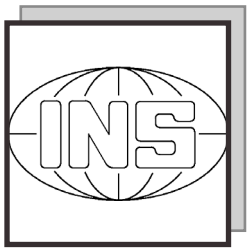
Universität Stuttgart

Schriftenreihe der Institute des Studiengangs Geodäsie und Geoinformatik

Technical Reports
Department of Geodesy and Geoinformatics



Quo vadis geodesia ... ?

Festschrift for Erik W. Grafarend

on the occasion of his 60th birthday

Part 1

Friedhelm Krumm,
Volker S. Schwarze
(Eds.)

ISSN 0933-2839 Report Nr. 1999.6-1

Friedhelm Krumm / Volker Siegfried Schwarze (Eds.)

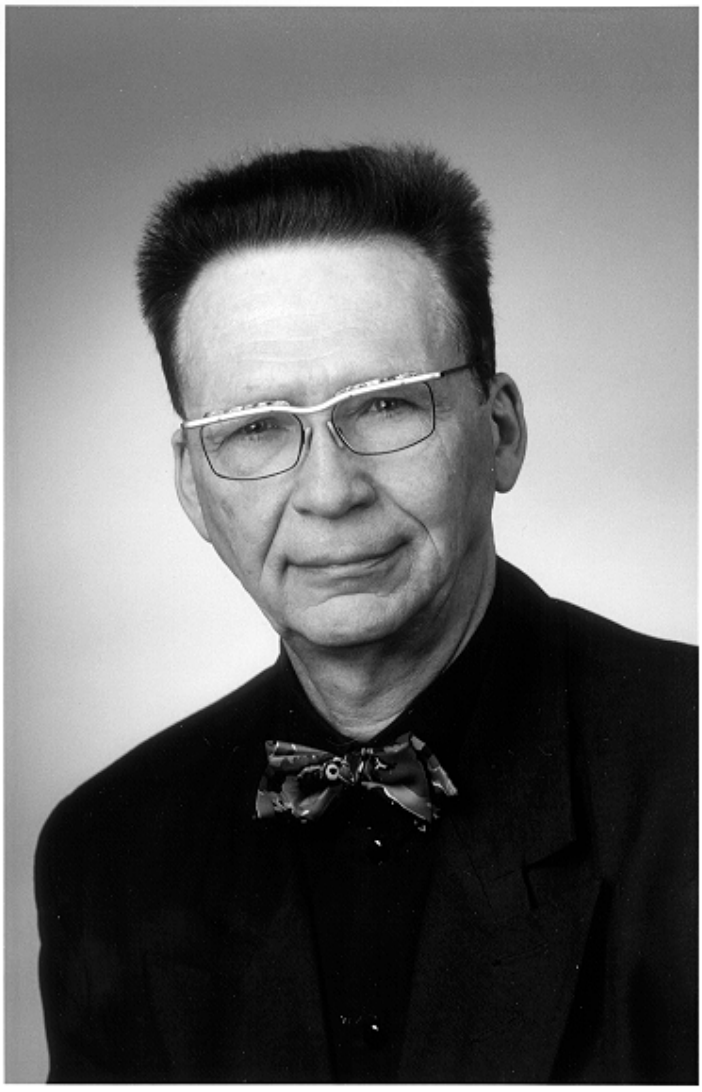
Quo vadis geodesia...?

Festschrift for Erik W. Grafarend

on the occasion of his 60th birthday

Part 1

Geodätisches Institut der Universität Stuttgart 1999



Foreword

This „Festschrift“ is dedicated to

Prof. Dr.-Ing. habil. Tekn. Dr. h.c. Dr.-Ing. E.h. Dr. h.c. Erik W. Grafarend

in honour of his 60th birthday. The two volumes of this „Festschrift“ cover a broad spectrum of geodetic research and mirror his scientific contributions and widespread scientific interests. Prof. Grafarend's scientific work spans from the rotation of a gyroscope to the rotation of the Earth as described by a deformable, viscoelastic body. It also covers geodetic sensors, from theodolites to inertial systems, from photogrammetric cameras to artificial Earth satellites.

Prof. Grafarend's major scientific contribution is towards a unified theory of Geodesy. In working with the gravitational field of the Earth he used the disciplines of Geodesy and Physics, of geometry space and gravity space, and beautifully showed the interrelationship between these models. This enjoyable interplay between these disciplines lead him to a unified approach towards geodesy: For example, the motion of a satellite, the trajectory of a light ray and the course of a plumbline can all be interpreted as geodesics - curves of minimal distance on some properly chosen space-time manifold. Prof. Grafarend also looked at the statistical nature of the geodetic measurement processes. It is therefore not surprising that he has written many important papers on adjustment theory and statistical inference.

No area of geodetic interest has been neglected by Prof. Grafarend to achieve a unified theory of geodesy. His scientific work in totality looks like the geodetic version of the „Glasperlenspiel“ by Hermann Hesse with its never resting „magister geodesiae“ Prof. Grafarend in its centre. In that sense these two volumes of papers written by internationally acknowledged scientists give a timely and representative overview over geodesy, as Prof. Grafarend sees it. It also becomes obvious how many people have been stimulated by his ideas and visions, and wish to express their respect and gratitude on the occasion of his 60th birthday.

The editors appreciate all the work by the authors of „Festschrift“.

Friedhelm Krumm Volker Schwarze

Stuttgart, October 1999

Welcome Address

On behalf of the directorate of the Universität Stuttgart it is my great pleasure to welcome all of you to the two-day symposium *Quo vadis geodesia...?* being held on occasion of the 60th Birthday of Prof. Dr.-Ing. habil. Tekn. Dr. h.c. Dr.-Ing. E.h. Dr. h.c. Erik W. Grafarend. Nobody will believe that Erik Grafarend has reached the age of *wise men*, who normally show silver hair and behave according to their age. He seems *ageless*, he stimulates, researches and lectures the basics of geodetic science at our university since 20 years without showing any loss of his young spirit. His always straightforward interests to introduce new theoretical concepts into the disciplines of statistical inference, mathematical and astronomical geodesy, cartographic map projections, and basic surveying computations have made him an national and international well-acknowledged and awarded scientist. We feel honored to have him as member in the professors' board, we feel also honored today by your participation in this symposium.

Erik Grafarend was born on Oct. 30, 1939 in Essen, Germany. After primary and secondary education he started his academic career studying Mining and Surveying at the Technical University Clausthal-Zellerfeld, Germany (1959-1964). He graduated 1964 as Diplom-Ingenieur (MSc) in Mining and Surveying. Spending three years as research associate in the Institute of Mining and Surveying of this university he obtained his doctoral degree (PhD) by submitting a thesis on *Azimuth Gyroscopes*. He realized very soon that natural sciences might be helpful to provide him deeper knowlegde and roots for his theory-oriented engineering concepts. Thus, he became again student studying Physics and graduated 1968 with the Diplom-Physiker (MSc in Physics). Prof. Helmut Wolf at Bonn university became his mentor in 1969, when Erik Grafarend was promoted being lecturer for geodesy and geophysics at the Institute of Theoretical Geodesy. After submitting a second thesis for qualifying himself to become professor (Habilitation) he was appointed 1970 Associate Professor at Bonn University. This thesis is well-known, here he wrote down the *Accuracy measures of a point manifold in multidimensional Euclidean spaces*. Due to his splendid publications he qualified himself very soon for a position as Full Professor - he accepted 1975 the offer of the University of the Federal Armed Forces Munich to take over the Chair of Astronomical and Physical Geodesy. He spent five very lifely and stimulating years in Munich, organised workshops and seminars to discuss and to deep statistical inference concepts, for example to prove robust statistics and to fully explore the potential of least squares prediction and collocation. Parallel to his Munich activities he became Adjunct Professor of the Technical University of Darmstadt lecturing the basics of satellite orbit computations.

Since 1980 he is with the Universität Stuttgart, he accepted the position being Full Professor and Head of the Geodetic Institute embedded in the Curriculum of Geodesy and Surveying. Besides his strong scientific interests he took over also the responsibility being Dean of the Faculty of Civil Engineering and Surveying (1985-1986) and served from 1995-1999 as Dean for Educational Affairs within the Curriculum of Geodesy and GeoInformatics.

Since the early seventies Erik Grafarend is heavily internationally engaged, he is one of the *pillars* of the International Association of Geodesy (IAG). Here he held several positions: Special Study Group member of many SSGs, SSG President, Commission President, etc. He realised 1980 that there is a need to launch a high level scientific journal, as result the *manuscripta geodaetica* was published. This journal is today one of the internationally recognized journals of our profession, its weight in citation index we gratefully acknowledge. But besides all these activities a real measure of Erik Grafarends output are the more than 230 publications, all of high level. Very often, we the readers of his publications, have difficulties to fully understand his definitions, sets and proofs. But as many of us know him, we give him the confidence that the final results of his complex mathematical boxes are okay. The validation by examples is often done by his staff members.

It seems to us obvious that he was awarded several times. This awarding series started 1971 when he got the Award of the National Science Foundation of Germany (DFG) for research on *The Geodetic*

Reference Surface, the IAG Bomford prize followed 1975 (during the General Assembly in Grenoble/France), also to mention the Senior K. and W.A. Heiskanen Award of The Ohio State University, Department of Geodetic Science, Columbus/Ohio. The National Academy of Sciences awarded him 1978 the Senior Scientist Award. As we realized in the first sentence of this welcome address he obtained three honorary doctoral degrees: the Tekn. Dr. h.c. by the Kungl. Tekniska Högskolan, Stockholm/Sweden (1989), the Dr.-Ing. E.h. by the Technical University of Darmstadt/Germany (1996), and the Dr. h.c. by the Technical University of Budapest/Hungary (1998). Most probably, this series of awards is still open.

What have his colleagues at the Universität Stuttgart done to keep such an outstanding scientist satisfied? Well, we provided him 1997 adequate room facilities, a dream he had from the mid eighties. Many of you may have wondered whether the furniture and the room furnishings are standard of our university. They are not. It is his personal note of room design and interior architectural taste. We further offered him the opportunity to share lectures and exercises in adjustment and statistical inference, an interest he was also looking for from his beginnings at this university. Thus, we feel having done quite a lot to offer him a good atmosphere for science and education, at least during the last few years. We hope, that the next five years Erik Grafarend will further contribute high level papers to geodesy and related disciplines. Let me close with the statement: " Erik, thank you very much for your contributions. Congratulations to your 60^{th} Birthday. We wish you and Ulrike a further interesting life. Please keep your shape and spirit young. "

Dieter Fritsch, Vice President Education
Universität Stuttgart

Contents Part 1

Contents Part 2

Spaceborne VLBI beyond 2000

József Ádám

After a long preparation period, the idea of space VLBI has became a reality. In February 12th, 1997 a dedicated radio telescope has been launched into Earth orbit by the Institute of Space and Astronautical Science (ISAS) from Japan and integrated in the ground-based VLBI networks for astrophysical studies.

A straightforward extension from the ground - based VLBI to space is called *space VLBI* (or orbiting VLBI), which uses radio telescopes in space. With launching of one or more space VLBI satellites, space VLBI observations will be available for astrometric, geodetic and geodynamic applications as well. *The space VLBI observables may be useful to improve the determination of the Earth's gravity field and in the unification and connection of reference frames inherent in the space VLBI technique.* The space VLBI missons offer and will provide new types of satellite observables (VLBI time delay, delay rate and differential VLBI tracking data) with high accuracy for these potential applications.

After the successful launch of the first space VLBI satellite, known as MUSES-B before launch, was renamed HALCA (which stands for Highly Advanced Laboratory for Communications and Astronomy, and is meant to sound like *haruka*, the Japanese word for „distant"). HALCA is the orbital element of the VLBI Space Observatory Programme (VSOP), a large international collaboration of space agencies and radio astronomy observatories which have combined resources to create the first dedicated space VLBI mission. Simultaneous observations with HALCA's 8 meter diameter radio telescope and ground-based radio telescopes synthesize a radio telescope that has an effective diameter over twice the size of the Earth, providing finer detail images at a given radio wavelength than can be obtained from the Earth. In its elliptical orbit, HALCA ranges as far as 21000 km from Earth's center, so that when it observes in conjunction with a ground-based telescope, a maximum baseline of 33000 km is achieved, yielding a resolution of 0,3 milliarcseconds at an operating frequency of 5 GHz. That's enough to see things the size of 10 light-years at the quasar's distance of 6.5×10^9 light-years (a threefold improvement over ground-based arrays operating at the same frequency). VSOP observations at 1.6 GHz (18 cm) and 5 GHz (6 cm) have yielded the highest resolution images ever made of extragalactic radio sources at these frequencies (*e.g. Day, 1997; Hirabayashi, 1999; Hirabayashi et al., 1998; Paragi et al., 1998*). Further details about the VSOP mission, including the current observation schedule, images from previous observations, and information about proposing for observations, are available from *http://www.vsop.isas.ac.jp.*

The potential geodetic-geodynamic applications of the space VLBI technique were first pointed out by *Fejes et al.(1986).* Based on the paper by *Dermanis and Grafarend (1981), J. Ádám* started the theoretical work as a Humboldt Fellow during his research stay at the Department of Geodetic Science, Stuttgart University in 1985 (*Adam, 1989*). In the following years several very detailed theoretical studies were carried out on this subject *(*see e.g. *Adam (1990), Kulkarni (1992), Kulkarni et al. (1991), Klatt (1995), Zheng (1992)).* After the theoretical investigations and simulation studies made by these experts in different institutions in the frame of the IAG Special Study Group 2.109 („Application of Space VLBI in the Field of Astrometry and Geodynamics"), see *Ádám (1995), Fejes (1992),* the next obvious step was to formulate a plan to experiment with real potential measurements, when available in order to prove the feasibility of space VLBI for geodesy and geodynamics. In response for the First VSOP Announcement of Opportunity, the GEDEX (Space VLBI Geodesy Demonstration Experiment) proposal was submitted by an international team (*I. Fejes as P.I. and J. Ádám, P. Charlot, S.Frey, N. Kawaguchi, Z.H. Qian, and H. Schuh* as co-I.-s). The VSOP Scientific Review Committee accepted the proposal in 1996 with assigned code V002. For a more detailed

description of the GEDEX concept is given by *Fejes et al. (1996)* and status reports are in *Fejes (1998) and Kulkarni et al. (1998).*

The future of spacebore VLBI beyond 2000 is promising. An other space VLBI satellite called RADIOASTRON in Russia is still planned to be launched, possibly in 2001. A preproject of NASA called ARISE (Advanced Radio Interferometry between Space and Earth) will be a mission consisting of one (or possibly two) 25-meter radio telescopes in highly elliptical Earth orbit. The telescope(s) would observe in conjunction with a large number of radio telescopes on the ground, using the technique of space VLBI, in order to obtain the highest resolution (10-microarcsecond) images of the most energetic astronomical phenomena in the universe. Details on ARISE are available from

http://arise. jpl. nasa.gov/arise/whatisarise/what-is-arise.html.

References used for this review and for the outlook are as follows.

References

Ádám J. (1989): A Possible Application of the Space VLBI Observations for Establishment of a new Connection of Reference Frames. (Proceedings of the Int. Summer School of Theoretical Geodesy on „Theory of Satellite Geodesy and Gravity Field Determination", Eds.: F.Sansó and R. Rummel, Assisi, Italy, May 23-June 3, 1988.) *Lecture Notes in Earth Sciences,* No. 25, pp. 459-475, Springer-Verlag, Berlin-Heidelberg, 1989. (Abstract in: *Veröff. Zentralinst. Physik Erde,* 102(1989), Part II., p. 300, Potsdam.)

Ádám J. (1990): Estimability of Geodetic Parameters from Space VLBI Observables. *Report No. 406,* Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio, July, 1990.

Ádám J. (1995): Report of IAG Special Study Group 2.109: Application of Space VLBI in the Field of Astrometry and Geodynamics. *Travaux de l'Association Internationale de Géodésie,* Tome 30, Ed. by P. Willis, pp. 114-143, Paris, 1995.

Ádám J. (1997): Application of Space VLBI in the field of Astrometry and Geodynamics: State-of-the-Art 1995. IAG/CSTG *Bulletin No. 14,* pp. 51-72, Münich, 1997.

Ádám J. and Mueller I.I. (1990): Merits of Space VLBI Missions for Geodynamics. (Paper presented at the 123rd Colloquium of the International Astronomical Union, Held in Greenbelt, Maryland, USA, April 24-27, 1990.) Abstract in: *Observatories in Earth Orbit and Beyond,* Ed. Y. Kondo, p. 507, Kluwer Academic Publishers, 1990.

ARISE-Advanced Radio Interferometry Between Space and Earth. *http://arise.jpl.nasa.gov/arise/whatisarise/what-is-arise.html*(1999).

Day, Ch. (1997): Radio Telescope in Space Maps Quasar Jet. *Physics Today*, p. 23, September 1997.

Dermanis A. and Grafarend E.W. (1981): Estimability analysis of geodetic, astrometric and geodynamical quantities in very long baseline interferometry. *Geophysical Journal of the Royal astronomical Society,* 64(1981), pp. 31-56.

Fejes I. (1992): Report of IAG Special Study Group. 2.109 - Application of Space VLBI in the Field of Astrometry and Geodynamics. *Travaux de L'Association Internationale de Géodésie,* Tome 29, Ed. by P. Willis, pp. 189-192, Paris, 1992.

Fejes I. (1994): Techniques and methods in Geodynamical Research: Space VLBI. *Acta Geod. Geoph. Hung.,* Vol. 29(3-4), pp. 443-456(1994).

Fejes I. (1999): The Status and Perspective of the Space VLBI Geodesy Demonstration Experiment (GEDEX). *Adv. Space Res.,* Vol. 23, No. 4, pp. 781-784, 1999.

Fejes I., Kawaguchi, N. and Mihály Sz. (1996): Space VLBI Geodesy: background of an experiment proposal. *Astrophysics and Space Science,* 239, pp. 275-280, 1996.

Fejes I., Almár I., Ádám J. and Mihály Sz. (1986): Space VLBI: Potential applications in Geodynamics. *Adv. Space Res.,* Vol.6., No.9, pp. 205-209, 1986.

Hirabayashi H. and 52 others (1998): Overview and Initial Results of the Very Long Baseline Interferometry Space Observatory Programme. *Science,* Vol.281, pp. 1825-1829, September 18, 1998.

Hirabayashi H. (1999): VSOP Results. Proceedings of the International Workshop on Geodetic Measurements by the collocation of Space Techniques on Earth (GEMSTONE), pp. 189-196, Koganei, Tokyo, Japan, January 25-28, 1999.

Klatt C.V. (1995): Space VLBI Orbit Reconstruction, Ph.D. Thesis, York University, North York, Ontario, Canada, August.

Kulkarni M.N. (1992): A Feasibility Study of Space VLBI for Geodesy and Geodynamics. *Report* No. 420, Dept. of Geodetic Science and Surveying, The Ohio State Univ., Columbus, Ohio, June.

Kulkarni M.N. and Ádám J. (1993): A sensitivity analysis for orbit determination of space VLBI satellites for interconnection of reference frames. *Acta Geod. Geoph. Mont. Hung.*, Vol. 28(3-4), pp. 441-456(1993).

Kulkarni M.N., Mueller, I.I., Schaffrin, B. and Ádám J. (1991): Sensitivity Analysis of Earth Rotation Parameters from Space VLBI. (Poster paper presented at the AGU Spring meeting Baltimore, Maryland, May 28-June 1, 1991.) Abstract in: *A supplement to EOS, AGU,* Spring Meeting 1991, Program and Abstract, p. 94, April 23, 1991.

Kulkarni M.N., Ádám J., Fejes I., Frey S., Kampes, B., Zheng,Y. and Hu,X (1998): Geodesy Demonstration Experiment (GEDEX) for Space VLBI: State of the Art and Software Development. In: Geodesy on the Move-Gravity, Geoid, Geodynamics and Antarctica, *IAG Symposia Vol.119,* pp. 383-388, Springer Verlag, Berlin-Heidelberg, 1998.

Lemoine F.G. and 14 others (1998): The Development of the Joint NASA GSFC and the National Imagery and Mapping Argency (NIMA) Geopotential Model EGM96. *NASA/TP-1998-206861,* July 1998.

Paragi Z., Frey S., Fejes I., Porcas R.W., Schilizzi R.T. and Venturi T. (1998): 1.6 GHz Space VLBI Observation of 3C446. Paper presented at the 32nd COSPAR Scientific Assembly, *„VSOP Results and the Future of Space VLBI",* Nagoya, Japan, July 12-19, 1998.

Sovers O.J., Fanselow J.L. and Jacobs Ch.S. (1998): Astrometry and geodesy with radio interferometry: experiments, models, results. *Reviews of Modern Physics,* Vol. 70, No.4,pp.1393-1454, October 1998.

Zheng Y. (1992): The Application of Space VLBI for Astrometry, Geodynamics and Space Physics (in Chinese). *Ph.D. Thesis,* Shanghai Observatory, Chinese Academy of Sciences, Shanghai, August.

Zheng Y. (1995): The Research for VLBI Software and Astrogeodynamics (in Chinese), Postdoctoral Subject Report, Nanjing University, China, June.

Zheng Y. (1997): SPVK User's Guide, Z.Z. Institute of Surveying & Mapping, China, May.

Robust Geodetic Parameter Estimation under Least Squares through Weighting on the Basis of the Mean Square Error

Francis W. O. Aduol

Abstract

A technique for the robust estimation of geodetic parameters under the least squares method when weights are specified through the use of the mean square error is presented. The mean square error is considered in the specification of observational weights instead of the conventional approach based on the observational variance. The practical application of the proposed approach is demonstrated through computational examples based on a geodetic network. The results indicate that the least squares estimation with observational weights based on the mean square error is relatively robust against outliers in the observational set, provided the network (or the system) under consideration has a good level of reliability, as to make the network (or system) stable under estimation.

Introduction

The classical approach in the estimation of geodetic parameters is through the least squares method within the framework of the Gauss-Markov model given as:

$$\tilde{y} = Ax + \varepsilon; \varepsilon \sim (0, \sigma_0^2 W^{-1}) \tag{1}$$

where y is an $n \times 1$ vector of observations, A is an $n \times m$ design matrix, x is an $m \times 1$ vector of unknown parameters, ε is an $n \times 1$ vector of observational errors, σ_0^2 is the variance of unit weight, and W is an $n \times n$ positive-definite weight matrix.

This estimation model is based on the assumption that the observational errors collected in the vector ε occur randomly and are distributed according to the normal distribution. with this assumption and under the least squares condition that $\varepsilon^T W \varepsilon$ be minimum, the following estimates may be obtained:

$$\begin{aligned}
&\hat{x} = (A^T WA)^{-1} A^T W\tilde{y} \\
&D(\hat{x}) = \Sigma_{\hat{x}\hat{x}} = \hat{\sigma}_0^2 (A^T WA)^{-1} \\
&\hat{y} = Ax = A(A^T WA)^{-1} A^T W\tilde{y} \\
&D(\hat{y}) = \Sigma_{\hat{y}\hat{y}} = A\Sigma_{\hat{x}\hat{x}} A^T = \hat{\sigma}_0^2 A(A^T WA)^{-1} A^T \\
&\hat{\varepsilon} = y - \hat{y} = y - A\hat{x} = y - A(A^T WA)^{-1} A^T Wy \\
&D(\hat{\varepsilon}) = \Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = \Sigma_{\tilde{y}\tilde{y}} - A\Sigma_{\hat{x}\hat{x}} \\
&\hat{\sigma}_0^2 = \hat{\varepsilon}^T W\hat{\varepsilon} / (n-m)
\end{aligned} \tag{2}$$

In the event however that the observational vector y may be contaminated with a bias parameter b (whereby the bias may be as a result of gross errors, or systematic errors, or a combination of both),

then the assumption $\varepsilon \sim (0, \sigma_0^2 W^{-1})$ gets invalidated, in that the errors on y, which now also comprise b can no longer be considered to be distributed according to the normal distribution. The consequence of this is that if the estimation of the unknown parameters still be performed according to the least squares condition, under the Gauss-Markov model as in (1), then the so obtained estimates will be biased as a result of b. To deal with this problem, two options come into consideration: (i) one performs the estimation under the least squares under the model (1) but seeks to identify and remove outliers (biased observations) from the observational dataset in what we may refer to as *outlier detection*, or (ii) one adopts estimation techniques that are robust with respect to the biases under *robust estimation*.

The propagation of outlier detection in geodesy and surveying was motivated by the works of W Baarda [**2**, **3**, **4**]. Today outlier isolation forms an integral component of any major geodetic data processing and analysis. However the detection and isolation of outliers within the framework of the Gauss-Markov model as specified in (1), still suffers from the tendency of the ordinary least squares method to spread out the effect of outliers among observations, thereby rendering the isolation of the outliers difficult, and sometimes altogether impossible. To cope with this problem, robust estimation techniques offer real alternatives.

The objective in robust estimation is to perform an estimation of the parameters from the observations in such a way that the estimates of the parameters so obtained are virtually unaffected by any biases or outliers that may be present in the observations. An extensive study of the application of robust estimation in geodesy is reported in [**5**]. Robust estimation techniques in the estimation of parameters in general were however brought to the fore through the works of P J Huber [**8**, **9**, **10**], while a further extensive treatment of the subject has been presented by [**7**]. The core of Huber's technique is the M-estimator, which is based on the maximum-likelihood method.

A general characteristic of the robust estimation techniques is that they restrict a range of observational error within which the observations may be accepted, and observations associated with observational error outside the specified range are 'cut off' from the estimation process within the process of 'winsorisation'. The problem in this approach however is that the decision on where the 'cut-off' point itself should be is rather subjective. As an alternative approach in robust estimation, a procedure for robust estimation based on iterative weighting of observations was suggested in [**1**]. This was an attempt at a procedure that would avoid excluding any observations from the estimation procedure, but include all the observations within the estimation procedure except with appropriate weighting.

In this presentation, we extend the concept of iterative weighting by considering it from the point of view of the observational weights based on the *mean square error* (MSE), and evaluate the effectiveness of the method through the computation of a practical network.

The Mean Square Error

Let us consider a parameter vector ξ, whose realisation (obtained through estimation or otherwise) is $\hat{\xi}$, then the mean square error of $\hat{\xi}$ is given as

$$M(\hat{\xi}) = E[(\hat{\xi} - \xi)(\hat{\xi} - \xi)^T] \tag{3}$$

In general we have that $E(\hat{\xi}) = \xi + \beta$ where β is a bias vector. Thus we may rewrite (3) as

$$\begin{aligned} M(\hat{\xi}) &= E[(\hat{\xi} - (E(\hat{\xi}) - \beta))(\hat{\xi} - (E(\hat{\xi}) - \beta))^T] \\ &= E[(\hat{\xi} - E(\hat{\xi}))(\hat{\xi}^T - E(\hat{\xi})^T)] + \beta\beta^T \end{aligned} \tag{4}$$

But we have that the dispersion $D(\hat{\xi})$ of $\hat{\xi}$ is given as

$$D(\hat{\xi}) = E[(\hat{\xi} - E(\hat{\xi}))(\hat{\xi}^T - E(\hat{\xi})^T)] \tag{5}$$

Thus

$$M(\hat{\xi}) = D(\hat{\xi}) + \beta\beta^T \tag{6}$$

(see e.g. [**11**] and [**6**]).
In the special case that $\beta = 0$, we have then that

$$M(\hat{\xi}) = D(\hat{\xi}) \tag{7}$$

From the fact that the mean square error incorporates the biases in the realisation of a parameter, the mean square error is a much more effective and efficient estimate of the quality of the parameter in the sense of *accuracy*. The dispersion on the other hand, respectively the variance, as is ordinarily known, gives the *precision* of the estimate or realisation, which however only becomes also a measure of accuracy in the special case when $\beta = 0$, in which case (7) obtains.

We have from (6) that in the special case that it is a single independent parameter being considered, the mean square error is given as: *mean-square-error = variance+bias²* .

The Estimation Model

In the event that the observation $\tilde{y}$ in (1) is contaminated with a bias b, then we have that

$$E(\tilde{y}) = y + b \tag{8}$$

where y is the 'true' value of the parameter.
But we have that

$$\tilde{y} = E(\tilde{y}) + \varepsilon; \varepsilon \sim (0, \Sigma_{\tilde{y}\tilde{y}}) \tag{9}$$

Then with (8) and (9) we have

$$\tilde{y} = y + b + \varepsilon \tag{10}$$

which with $v := b + \varepsilon$, becomes

$$\tilde{y} = y + v \tag{11}$$

For

$$y = Ax \tag{12}$$

we then have that $\tilde{y} = Ax + b + \varepsilon$ or

$$\tilde{y} = Ax + v, E(v) = E(b + \varepsilon) = b, M(\tilde{y}) = \Sigma_{\tilde{y}\tilde{y}} + bb^T \tag{13}$$

We adopt this as the model for the estimation of the parameters within the framework of least squares.

We note therefore from (13) that if we can estimate v such that

$$E(v) = E(b + \varepsilon) = b + E(\varepsilon) = b, \tag{14}$$

then we would have been able to obtain an unbiased estimate of x that is relatively free from the influence of the bias b.

In the conventional least squares approach, whereby the model is defined according to (1), if the model had a bias parameter as to be described according to (13), but with the stochastic part described

through $\varepsilon \sim (0, \Sigma_{\tilde{y}\tilde{y}} = \sigma_0^2 W^{-1})$, then the model would have not been appropriately specified, so that the parameters estimated with the model will be biased. We seek to overcome the bias effect in that we define the estimation model through (14) and weight the observations according to the mean square error (MSE), which already incorporates the bias effect. We propose then to define the weight W of the observations as

$$W = \sigma_0^2 M_{\tilde{y}\tilde{y}}^{-1} \tag{15}$$

in which we have taken $M_{\tilde{y}\tilde{y}} = M(\tilde{y})$.

If we assume independence of observations $\tilde{y}_i (i = 1, \ldots n)$, then we have that for an observation $\tilde{y}$, the mean square error may be given as

$$m_i = \sigma_i^2 + b_i^2 \tag{16}$$

for σ_i^2 and b_i being respectively the variance and bias of $\tilde{y}_i$. Then the weight of $\tilde{y}_i$ can now be defined as

$$w_i = \frac{\sigma_0^2}{m} \tag{17}$$

With the weights so defined, we notice that $M_{\tilde{y}\tilde{y}}^{-1}$ will exist due to the fact that $M_{\tilde{y}\tilde{y}}$ has been taken to be a diagonal matrix, and hence W according to (15) can be evaluated.

The question however is how does one evaluate the mean square error in the first place, when the bias b itself is in the first instance unknown, and must in any case be evaluated. We seek to deal with this problem in that we evaluate b iteratively and hence also W.

The Estimation Process

We begin the estimation process by assuming nominally that $b = 0$. With this, we notice that we will simply be having the Gauss-Markov model as described in (1). From this, the first estimates of b as 'residuals' will have been obtained. With the residuals v_i, a new value for m_i is obtained according to (16), however with σ_i being as originally set, since these are the original variances of the observations, which are assumed known a priori. With the new mean square error values, the estimation process is repeated. The process is repeated until convergence for the estimated parameters is achieved at the specified level of tolerance. In particular, since the main parameters being estimated are the unknown parameter vector x, the convergence of the x parameters would be more appropriately adopted as control for the iteration.

Through the iterative process, the mean square error of an observation is estimated for simultaneously as well and consequently the mean-square-error weight of the observations. The robustness of the procedure is thus contained in the mean-square-error weight, which is a much more comprehensive and realistic representation of the observational weights.

The Test Example

The test network

A two-dimensional network as shown in Fig. 1 was adopted for the test example. The network comprises 9 points, which are linked by distance observations. A single distance observation was considered to have been measured with a standard error of 3mm+0.5ppm; with this the eventual standard error for the mean distance adopted was then deduced from the number of individual measurements from which the particular mean distance is obtained. The network has a total of 30 distance measurements.

Experimental design

Four versions of the network were computed; these were designated as Net-0, Net-1, Net-2, and Net-4. The networks were specified according to the numbers of gross errors they contained as follows: Net-0 - no gross errors; Net-1 - one gross error; Net-2 - two gross errors; and Net-4 - four gross errors. The gross errors were simulated into the networks as given in Table 1.

Table 1: The simulated gross errors

Line	Error [metres]	Network
4-7	+0,780	1,2,4
2-3	-5,067	2,4
7-11	+0,355	4
3-4	-0,055	4

Each version of the network was then computed on the basis of both the ordinary least squares and the least squares method with mean-square-error weights as proposed here. The network was computed throughout in free-network mode.

Results

In the results presented below, X and Y are estimated point coordinates in metres; σ_X and σ_Y are estimated positional standard errors in metres; a and b are the major and minor axes of the positional error ellipse in metres, while φ is the orientation of the major axis of the ellipse in degrees taken with respect to the X axis.

Net-0

Table 2: Conventional least squares

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.186	3462429.367	0.0044	0.0059	0.0063	0.0039	155.4
2	5439065.854	3468259.524	0.0044	0.0053	0.0056	0.0040	150.5
3	5457025.454	3476522.257	0.0054	0.0059	0.0059	0.0054	170.3
4	5465079.259	3433374.141	0.0045	0.0064	0.0065	0.0042	163.0
5	5448374.040	3427727.520	0.0039	0.0043	0.0047	0.0034	143.1
6	5439527.166	3423319.106	0.0053	0.0066	0.0076	0.0036	145.4
7	5447601.042	3443324.505	0.0060	0.0038	0.0060	0.0037	9.8
8	5411104.687	3454335.360	0.0064	0.0092	0.0097	0.0056	158.0
11	5464986.157	3457965.548	0.0086	0.0053	0.0086	0.0053	179.8

Table 3: Robustified least squares (3 iterations)

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.186	3462429.370	0.0030	0.0042	0.0045	0.0026	155.4
2	5439065.857	3468259.522	0.0030	0.0043	0.0045	0.0026	157.1
3	5457025.455	3476522.260	0.0036	0.0040	0.0040	0.0036	3.4
4	5465079.258	3433374.142	0.0030	0.0041	0.0043	0.0028	161.1
5	5448374.040	3427727.518	0.0027	0.0032	0.0034	0.0024	146.8
6	5439527.168	3423319.103	0.0036	0.0049	0.0055	0.0025	147.8
7	5447601.039	3443324.504	0.0041	0.0028	0.0042	0.0027	12.9
8	5411104.688	3454335.360	0.0042	0.0060	0.0064	0.0036	156.0
11	5464986.155	3457965.548	0.0057	0.0035	0.0057	0.0035	2.1

Observations treated as containing gross errors in the adjustment: Nil

Net-1

Table 4: Conventional least squares

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.119	3462429.461	0.0616	0.0834	0.0882	0.0545	155.4
2	5439065.822	3468259.569	0.0625	0.0745	0.0796	0.0559	150.5
3	5457025.475	3476522.203	0.0767	0.0828	0.0830	0.0765	170.3
4	5465079.471	3433373.962	0.0631	0.0899	0.0922	0.0597	163.0
5	5448374.111	3427727.603	0.0556	0.0606	0.0666	0.0484	143.1
6	5439527.284	3423319.082	0.0741	0.0932	0.1078	0.0505	145.4
7	5447600.816	3443324.577	0.0841	0.0537	0.0848	0.0525	9.8
8	5411104.631	3454335.397	0.0898	0.1302	0.1367	0.0795	158.0
11	5464986.118	3457965.474	0.1209	0.0752	0.1209	0.0752	179.8

Table 5: Robustified least squares (6 iterations)

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.189	3462429.366	0.0033	0.0047	0.0051	0.0027	153.8
2	5439065.857	3468259.524	0.0031	0.0042	0.0044	0.0028	156.4
3	5457025.454	3476522.262	0.0038	0.0043	0.0043	0.0038	178.6
4	5465079.254	3433374.146	0.0042	0.0049	0.0056	0.0032	143.1
5	5448374.039	3427727.516	0.0031	0.0035	0.0036	0.0029	153.1
6	5439527.165	3423319.103	0.0043	0.0053	0.0061	0.0029	144.4
7	5447601.044	3443324.502	0.0055	0.0030	0.0055	0.0030	177.8
8	5411104.690	3454335.360	0.0045	0.0065	0.0069	0.0038	155.0
11	5464986.154	3457965.550	0.0061	0.0038	0.0061	0.0038	4.8

Observations treated as containing gross errors in the adjustment

Line	Gross error as isolated (metres)	Redundancy
4 – 7	-0.7952	100%

Net-2

Table 6: Conventional least squares

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.886	3462429.822	0.4272	0.5778	0.6114	0.3775	155.4
2	5439067.003	3468259.524	0.4332	0.5164	0.5515	0.3876	150.5
3	5457022.867	3476521.657	0.5314	0.5738	0.5750	0.5301	170.3
4	5465079.528	3433373.759	0.4374	0.6229	0.6391	0.4135	163.0
5	5448374.140	3427727.373	0.3855	0.4203	0.4615	0.3351	143.1
6	5439527.264	3423318.920	0.5135	0.6457	0.7470	0.3502	145.4
7	5447600.915	3443324.313	0.5827	0.3720	0.5880	0.3637	9.8
8	5411105.229	3454336.092	0.6223	0.9021	0.9475	0.5507	158.0
11	5464986.014	3457965.650	0.8379	0.5210	0.8379	0.5210	179.8

Table 7: Robustified least squares (8 iterations)

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.187	3462429.365	0.0037	0.0049	0.0052	0.0032	153.5
2	5439065.854	3468259.524	0.0038	0.0043	0.0045	0.0035	147.7
3	5457025.460	3476522.263	0.0061	0.0045	0.0062	0.0043	14.5
4	5465079.254	3433374.146	0.0043	0.0051	0.0058	0.0033	144.4
5	5448374.039	3427727.516	0.0030	0.0035	0.0036	0.0029	157.4
6	5439527.165	3423319.103	0.0043	0.0052	0.0061	0.0028	143.5
7	5447601.044	3443324.503	0.0054	0.0030	0.0054	0.0030	176.4
8	5411104.689	3454335.358	0.0046	0.0065	0.0068	0.0042	158.9
11	5464986.154	3457965.550	0.0061	0.0039	0.0061	0.0039	0.5

Observations treated as containing gross errors in the adjustment

Line	Gross error as isolated (metres)	Redundancy
4 – 7	-0.7951	100%
2 – 3	+5.0783	100%

Net-4

Table 8: Conventional least squares

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.864	3462429.832	0.4259	0.5760	0.6096	0.3764	155.4
2	5439066.998	3468259.726	0.4319	0.5149	0.5498	0.3864	150.5
3	5457022.859	3476521.637	0.5298	0.5721	0.5733	0.5284	170.3
4	5465079.501	3433373.799	0.4361	0.6210	0.6371	0.4122	163.0
5	5448374.131	3427727.363	0.3843	0.4190	0.4601	0.3341	143.1
6	5439527.268	3423318.877	0.5119	0.6437	0.7447	0.3491	145.4
7	5447600.832	3443324.287	0.5810	0.3709	0.5862	0.3625	9.8
8	5411105.206	3454336.094	0.6204	0.8993	0.9446	0.5491	158.0
11	5464986.187	3457965.713	0.8354	0.5194	0.8354	0.5194	179.8

Table 9: Robustified least squares (8 iterations)

Point	X	Y	σ_X	σ_Y	a	b	φ
1	5428972.136	3462429.354	0.0061	0.0066	0.0073	0.0052	142.2
2	5439065.800	3468259.517	0.0071	0.0058	0.0077	0.0049	148.5
3	5457025.431	3476522.214	0.0085	0.0089	0.0091	0.0083	29.9
4	5465079.216	3433374.153	0.0060	0.0082	0.0088	0.0051	152.9
5	5448374.005	3427727.512	0.0042	0.0060	0.0060	0.0042	2.0
6	5439527.134	3423319.092	0.0061	0.0074	0.0085	0.0046	145.3
7	5447600.999	3443324.497	0.0077	0.0055	0.0077	0.0055	176.9
8	5411104.642	3454335.337	0.0065	0.0085	0.0086	0.0063	163.0
11	5464986.482	3457965.652	0.0123	0.0083	0.0124	0.0081	167.4

Observations treated as containing gross errors in the adjustment

Line	Gross error as isolated (metres)	Redundancy
4 – 7	-0.7951	100%
2 – 3	+5.0783	100%
2 –11	+0.3184	99%
5 –11	+0.2669	100%

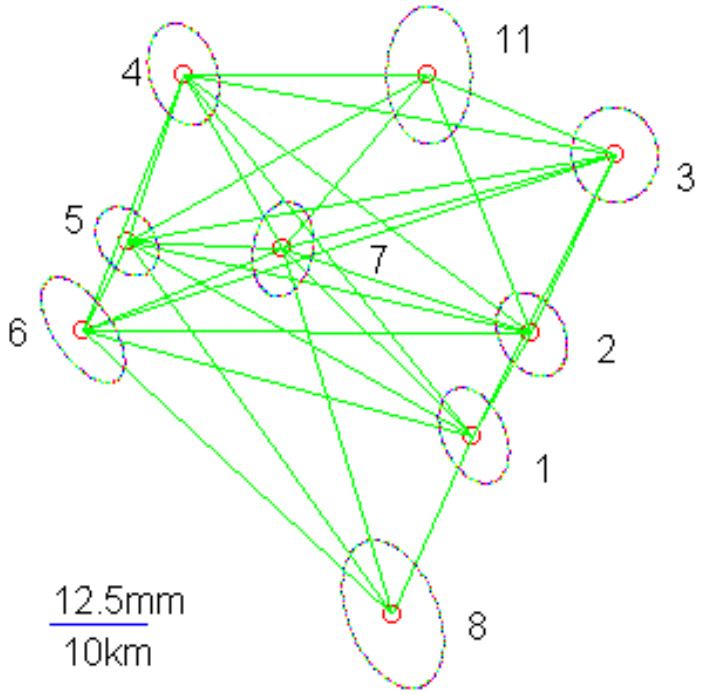


Fig. 1: Net 0 - Conventional least squares

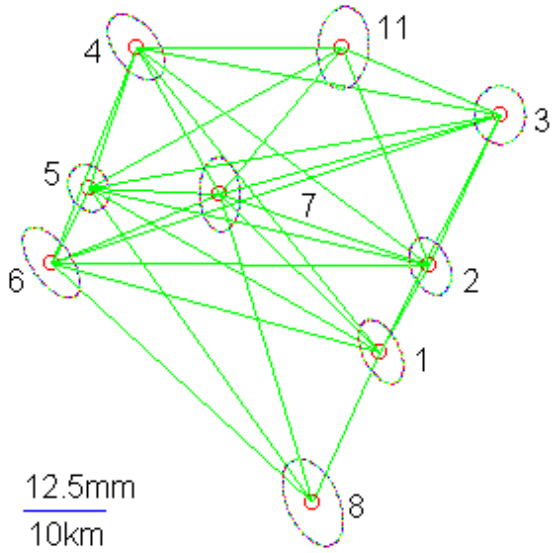


Fig. 3: Net 1 - Robustified least squares

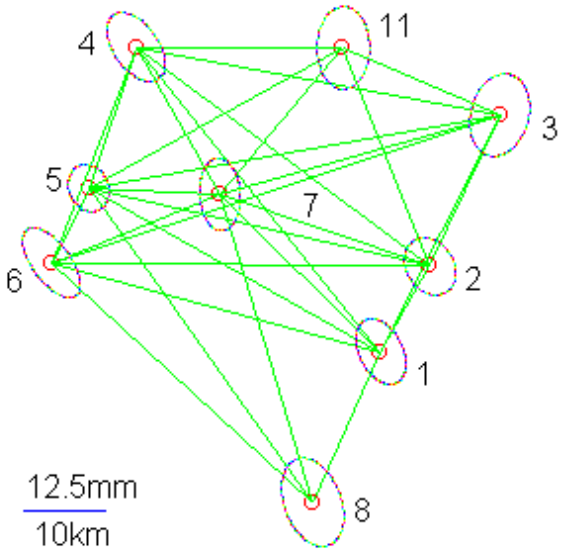


Fig. 4: Net 2 - Robustified least squares

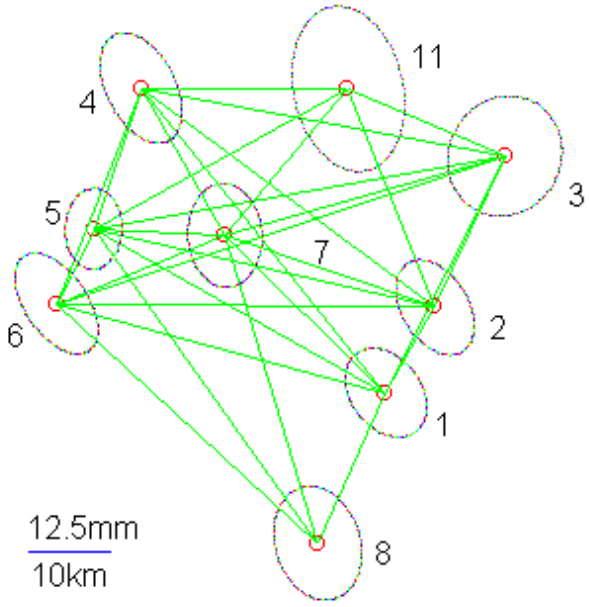


Fig. 5: Net 4 - Robustified least squares

Discussion

The technique described here, like most techniques for robust estimation and management of outliers in observations, depends considerably for is effectiveness on the reliability of the network. The technique is only able to isolate outliers and damp their effects on the estimation process through the fact that the bias-free observations are in a position to estimate effectively the unknown parameters and at the same time resist the influences from the outlying observations. This way, the effects of the outlying observations on the estimated parameters are rendered minimal.

If however the bias-free observations should be overwhelmed by the outlying observations, either through sheer numbers or through geometric distribution within the observational set, then an adequate solution of the estimates may be rendered difficult, or altogether impossible. For instance, in the present study, in the case with four gross errors in the network a converging solution was only obtained after eight iterations. However, although the results indicate that the estimated parameters have been obtained with relatively acceptable precision, the space of convergence of the parameters is biased, as can be ascertained through comparing the results in Table 9 with those in Table 3. This bias has been caused by the fact that the network was not sufficiently robust in configuration (i.e. in geometry, as well as observational type, number and quality) as to able to isolate the observations containing gross errors, which in the first place were rather 'unsuitably' distributed. The gross errors were here distributed such that out of the five network points, 2,3,4,7,11, connected with gross-error-contaminated observations three of the points, namely 3,4,7, were each connected with gross-error-contaminated observations. The result of this was that the gross errors in lines 3-4 and 7-11 could not adequately be isolated, and instead lines 2-11 and 5-11 were interpreted as the ones containing the gross errors.

In the cases with one and three gross errors, whose results are presented in Tables 5 and 7, the biases were effectively isolated, even though in this case point 7 was still connected by two gross-error-contaminated observations. The results for these two cases were found to be even more precise than

those from the ordinary gross-error-free least squares case presented in Table 2. In the initial case with no gross errors we notice from Table 3 that the results for the robustified least squares technique are considerably more precise than the case with ordinary least squares. Thus we have that even with observations that are effectively gross-error free one obtains more efficient estimates than with the ordinary least squares approach.

Conclusion

The results of this study demonstrate that the definition of the observational weights through the mean square error results in robustified least squares estimates. The technique tested was able to cope effectively with outliers in the observational set. The effectiveness of the technique however, as can be expected, is dependent on the reliability of the network, and especially on the particular observations contaminated with outliers. When the network reliability is sufficiently high, the technique of weighting observations on the basis of the mean square error instead of the variance can be relied on to yield fairly reliable estimates even with gross errors in the observational set. The computational process is rendered rather slower than in the case of weights based on variances, due to the fact that the mean square error has essentially to be determined iteratively.

Acknowledgement

This work was completed when the author was a visiting DAAD (German Academic Exchange Service) scholar at the Geodätisches Institut of the University Karlsruhe, with Prof Günter Schmitt as his host. The author is grateful to DAAD for this support and to Prof Schmitt for kindly hosting him while at the Geodätisches Institut. The test network adopted for the study was also kindly made available to the author by the Geodätisches Institut, Karlsruhe, and this assistance is also acknowledged. The author also wishes to thank Dr S M Musyoka, Department of Surveying, University of Nairobi, for his assistance in the preparation of the diagrams.

References

1. Aduol, F.W.O., 1994. Robust geodetic parameter estimation through iterative weighting. Survey Review, 32, 252: 359-367.
2. Baarda, W., 1967. Statistical concepts in geodesy. *Netherlands Geodetic Commission, publications on Geodesy, New Series*, Vol. 2, No. 4, Delft.
3. Baarda, W., 1968a. Statistics - a compass for the land surveyor. *Computing Centre of the Delft Geodetic Institute*.
4. Baarda, W., 1968b. A testing procedure for use in geodetic networks. *Netherlands Geodetic Commission, publications on Geodesy, New Series*, Vol. 2, No. 5, Delft.
5. Borutta, H. 1988. Robuste Schätzverfahren für geodätische Anwendungen. Schriftenreihe Studiengang Vermessungswesen Universität der Bundeswehr München, Heft 33. München.
6. Grafarend, E.W., Schaffrin, B., 1993. *Ausgleichungsrechnung in linearen Modellen*. Wissenschaftsverlag, Mannheim. Pp. 116 - 117.
7. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P., and Stahel, W.A., 1986. *Robust Statistics - the Approach based on Influence Functions*. John Wiley & Sons, New York.
8. Huber, P.J., 1964. Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35: 73-101.
9. Huber, P.J., 1972. Robust statistics - A review. *Annals of Mathematical Statistics*, 43: 1041-1067.
10. Huber, P.J., 1981. *Robust Statistics*. John Wiley & Sons, New York.
11. Toutenburg, H., 1992. *Lineare Modelle*. Physica Verlag, Heidelberg. Pp 35 - 36.

Care while using the NMEA 0183!

Alireza A. Ardalan and Joseph L. Awange

Abstract

The use of standard *NMEA 0183* for GPS and GLONASS receiver interface is supposed to provide smooth compatibility in data transfer from the receiver to the computer for processing. The present paper highlights the fact that although the Ashtech receiver is made to confirm to this *NMEA 0183* format, the different way in which the Ashtech Manual *(Ashtech 1997, P104)* defines the ellipsoidal height leads to different results as opposed to what would be obtained if the NMEA 0183's definition of the ellipsoidal height is used. The results with the Ashtech's definition indicates the computed means of the deviations dy and dz of the GPS values (with Selective Availability) to be better than those of GLONASS and combined GPS-GLONASS which is misleading. The difference in the results for the dz-component from the two different definition of the ellipsoidal height is in the magnitude of geoidal undulation.

1. Introduction

Nowadays all the GPS or GLONASS receivers are capable of providing position and velocity information in real time. Some applications however, requires that the data gathered by the receivers be stored externally and processed later as is the case in Geodetic applications. Post processing of the data requires that the receivers be able to transfer the stored data to the processing equipment i.e. the computers. This therefore means that the receivers must have communications ports and a specific protocol upon which the data are to be transferred.

To avoid incompatibility of equipment connection that would arise if receivers manufacturers were to device their own encoding procedures, data rate, signal level and message format e.t.c, the National Marine Electronics Association (*NMEA*) prepared in early 80's a standard procedure. Most GPS receivers manufactures therefore provide receivers with a data communications ports that conform to the *NMEA 0183* format *(R.B Langley 1995)*.

Examples of receivers that conform to the *NMEA 0183* format are *Trimble Navigation (see Trimble Navigation 1992)* and *Ashtech (Ashtech 1997)*. In using the *NMEA GGA* format for GPS and GLONASS data transfer, great care has to be observed in the way different receivers make use of the standard *NMEA GGA* format specifications. When using the standard *NMEA GGA* format, the user assumes that the receiver conforms to the NMEA CGA format (especially when stated) and that the redefinition of the *NMEA GGA* specifications is not valid. One expects to have the receiver's data transferred to the computer using the *NMEA GGA* format when stated and one then proceeds to compute the absolute point positioning for instance. The present paper considers the case in which the ellipsoidal height as presented in the standard *NMEA GGA* format *(NMEA 0183 1994)* has been defined in a different way in the Ashtech's receiver manual *(Ashtech 1997)*. To demonstrate the danger in neglecting such a redefinition, two experiments designated A and B based on Absolute Point Positioning using GPS and GLONASS data collected by an Ashtech receiver is performed. Insight on Absolute Point Positioning and other GPS and GLONASS positioning procedures are elaborately given in *(A. Leick 1995, A. Mathes 1998, G. Strang and K. Borre 1997)*. Experiment A considers purely the standard *NMEA GGA (NMEA 0183 1994)* format of ellipsoidal height's definition. The use of NMEA GGA is recom-

mended by the Ashtech receiver manual and the assumption in this Experiment is that no knowledge of the alternate definition of the ellipsoidal height by *Ashtech (1997, P104)* is known.

The experiment demonstrates the error that could be incurred by a user who is insensitive to the different way in which the Ashtech receiver adopts the standard specification as given in *NMEA 0183*. In Experiment B, the same GPS and GLONASS data used in Experiment A are used but with the ellipsoidal height as defined in the Ashtech manual *(Ashtech 1997, P104)*. An elaborate review of the NMEA 0183 is given by *R.B Langley (1995)*.

The present study was motivated by the fact that in one of our experiments we expected the combination of GPS and GLONASS to provide better results than those of GPS alone. This, as pointed out by *A. Mathes* (1998), is due to the fact that GPS observable are subject to Selective Availability (SA) while those of GLONASS are free from SA. The combination of both GPS and GLONASS was expected therefore to improve on the accuracy. The experiments we carried out to this effect using the Ashtech receiver however indicated the contrary. The results of the combination were worse! While those of GPS appeared to be better than both GLONASS alone and combined GPS-GLONASS. We thus had to investigate why this was the case. Our investigation led us to the fact that this behaviour of the results was due to the way in which the Ashtech receiver manual *(Ashtech 1997, P104)* redefined the ellipsoidal height in a different way from the standard *NMEA GGA* format. Care therefore must be taken on using *NMEA GGA* format. Presented in Section 2 are the procedures for Experiment A, Section 3 considers the procedures for Experiment B, while Section 4 concludes the results.

2. Experiment A

In this Experiment, the Ashtech GG24 receiver is used for the positioning of a single point i.e. Absolute Point Positioning. This receiver is capable of receiving signals from both GPS and GLONASS satellites. The observations were carried out in the second floor of the Institute of Navigation, Stuttgart University, using a receiver whose antenna is situated on top of the roof of the building. The data were collected in the *NMEA GGA* format in three different modes (i.e. GPS alone , GLONASS alone, and combined GPS-GLONASS). At each measuring mode, the data were collected for almost 20 minutes at 2 second intervals. A sample of *NMEA* records for the first 10 measurements in GLONASS alone mode is presented in *Table 2*. The aim of the experiment is to assess the suitability of the GLONASS as compared to GPS for point positioning and the advantages therein in combining the two.

Using the given known coordinates of the antenna as reference coordinates, the deviation in the computed antenna's position from the real value (reference coordintes) are computed first in ellipsoidal coordinates $\{\Delta\lambda, \Delta\phi, \Delta h\}$ and secondly in terms of 2+1 dimensional rectangular system $\{\Delta X, \Delta Y\}$-$\{\Delta Z\}$ according to following definitions:

(i) The orthogonal linear increments $\{\Delta X, \Delta Y\}$ are defined according to (2.1)-(2.2) in a plane which is tangent to the mean spherical radius, r, of the earth

$$\Delta Y = \frac{2\pi}{360} r\, \Delta\phi \tag{2.1}$$

$$\Delta X = \frac{2\pi}{360} r\, \Delta\lambda \cos\phi \tag{2.2}$$

where we assumed r=6380 (Km).

(ii) $\Delta Z = \Delta h$ is the difference between the measured and reference ellipsoidal height, which can be considered as an incremental vertical height difference above the $\{X, Y\}$ plane. h is ellipsoidal height

and neglecting the deflection of vertical it relates to orthometric height H and geoid undulation as follows

$$h = H + N \tag{2.3}$$

In *NMEA GGA* format H and N are given, therefore one can calculate h according to (2.3). *Table 3* shows a sample of the calculated ellipsoidal deviations $\{\Delta\lambda, \Delta\phi, \Delta h\}$ for the mode GLONASS alone, while *Table 4* depicts the deviations for the same case in terms of $\{\Delta X, \Delta Y\} - \{\Delta Z\}$.

Statistical inference of the time variations of $\{\Delta X, \Delta Y\} - \{\Delta Z\}$, in the three different measuring modes, namely GLONASS alone, GPS alone, and combined GLONASS-GPS, are as given in *Table 1*.

Figure 1-Figure 6 depicts the time variations of the $\{\Delta X, \Delta Y\} - \{\Delta Z\}$ in the three measuring modes, namely GLONASS alone, GPS alone, and combined GLONASS-GPS.

By comparing the results given in *Table 1*, we can see from the computed means that the GPS system is a bit more accurate than the GLONASS alone system or the combined GPS-GLONASS despite the availability of the Selective Availability which is astonishing. However, the results in both systems are biased.
From the results the following deductions could be drawn:

(i) In terms of the computed means of the x, y, z components, the combined case was worse as opposed to the stand alone cases. The best means were those of GLONASS in (dy), and GPS in (dx) and (dz) despite the presence of the S.A. The GPS results were however expected to be the worse.

(ii) In terms of *standard deviations* (dispersion of the data around the mean), the GLONASS results were better.

Table 1: Statistical information of the variation of the repeated positions measurements $\{\Delta X, \Delta Y\} - \{\Delta Z\}$ in the three measuring modes, namely GLONASS alone, GPS alone and combined GLONASS-GPS.

Statistical information	GLONASS alone (m)	GPS alone (m)	combined GPS and GLONASS (m)
Mean(dx)	14.1994	-9.2328	16.5384
Mean(dy)	-2.6517	5.0870	-5.9186
Mean(dz)	44.1784	35.3394	56.6020
STD(dx)	0.9251	9.4333	3.2719
STD(dy)	0.5468	17.7989	6.0557
STD(dz)	9.6975	28.9730	13.7515
Min(dx)	12.5907	-30.4415	2.2454
Min(dy)	-3.6635	-25.0245	-29.4971
Min(dz)	31.3290	-12.7310	-2.6710
max(dx)	15.6478	9.6925	23.1072
Max(dy)	-1.5107	43.0858	4.5580
Max(dz)	59.9890	97.2790	85.2690

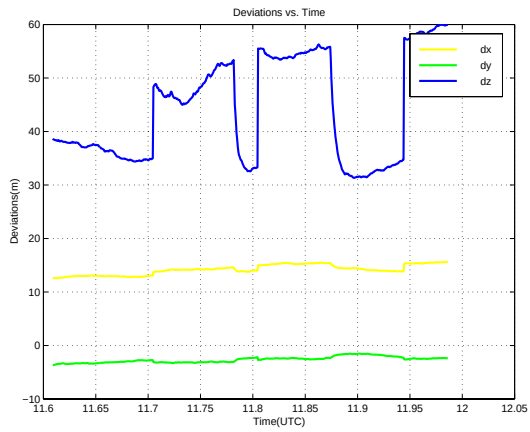


Figure 1: Deviation of measured coordinates from their known values versus time (UTC) when using GLONASS system alone.

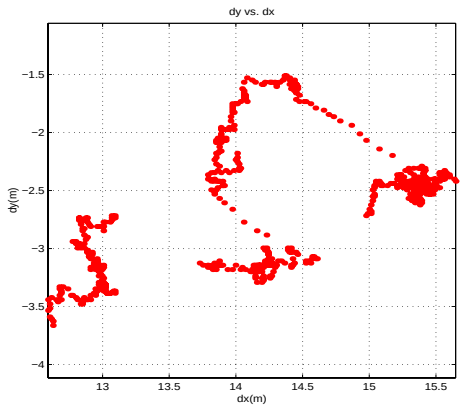


Figure 2: Variation of measured position minus its real value, projected into 2-D space of tangent plane, when using GLONASS system alone.

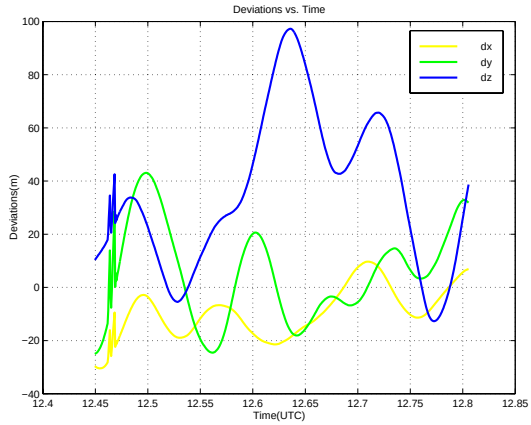


Figure 3: Deviation of measured coordinates from their known values versus time (UTC) when using GPS system alone.

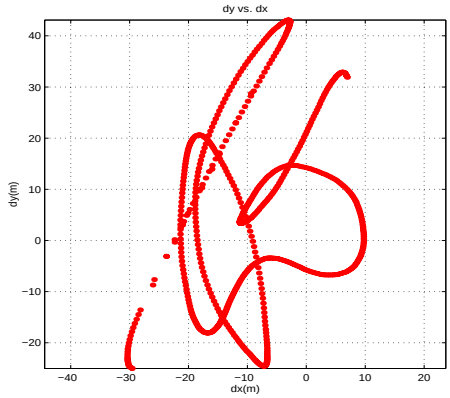


Figure 4: Variation of measured position minus its real value, projected into 2D space of tangent plane, when using GPS system alone.

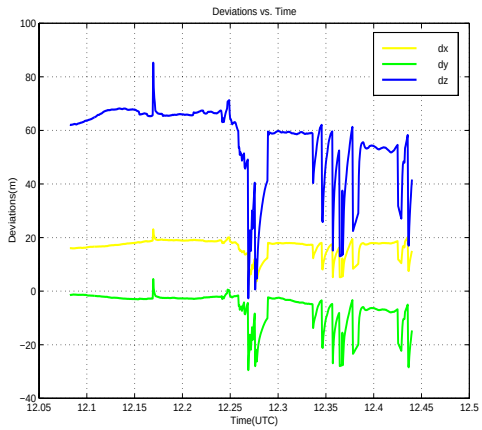


Figure 5: Deviation of measured coordinates from their known values versus time (UTC) when using combined GPS-GLONASS systems.

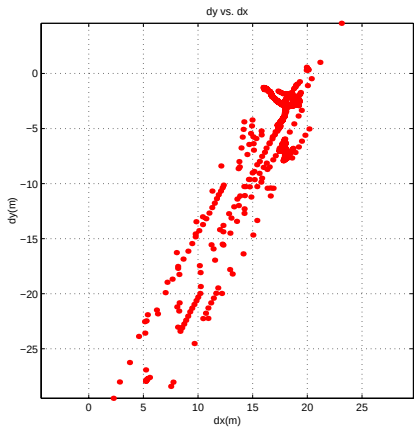


Figure 6: Variation of measured position minus its real value, projected into 2D space of tangent plane, when using combined GPS-GLONAS systems.

Table 2: A sample of NMEA GGA data recording format, first 10 records of GLONASS alone mode.

```
$GPGGA,113632.00,4846.99236,N,00910.52603,E,1,05,01.4,+00322.38,M,+0046.65,M,,*52
$GPGGA,113634.00,4846.99236,N,00910.52603,E,1,05,01.4,+00322.26,M,+0046.65,M,,*5B
$GPGGA,113636.00,4846.99238,N,00910.52603,E,1,05,01.4,+00322.22,M,+0046.65,M,,*53
$GPGGA,113638.00,4846.99239,N,00910.52602,E,1,05,01.4,+00322.18,M,+0046.65,M,,*54
$GPGGA,113640.00,4846.99240,N,00910.52601,E,1,05,01.4,+00322.09,M,+0046.65,M,,*56
$GPGGA,113642.00,4846.99243,N,00910.52600,E,1,05,01.4,+00322.09,M,+0046.65,M,,*56
$GPGGA,113644.00,4846.99244,N,00910.52601,E,1,05,01.4,+00322.11,M,+0046.65,M,,*5F
$GPGGA,113646.00,4846.99246,N,00910.52601,E,1,05,01.4,+00322.05,M,+0046.65,M,,*5A
$GPGGA,113648.00,4846.99248,N,00910.52600,E,1,05,01.4,+00321.95,M,+0046.65,M,,*51
```

Table 3: Deviation of the measured ellipsoidal coordinates from the real values, for first 10 measurements in GLONASS alone mode.

UTC hour	$\Delta\phi$ deg	$\Delta\lambda$ deg	Δh m
1.1608889E+1	1.2627385E-5	3.6634810E-4	3.8659000E+1
1.1609444E+1	1.2627385E-5	3.6634810E-4	3.8539000E+1
1.1610000E+1	1.2627385E-5	1.6263637E-4	3.8499000E+1
1.1610556E+1	1.2615156E-5	3.6078050E-4	3.8459000E+1
1.1611111E+1	1.2602927E-5	3.5892463E-4	3.8369000E+1
1.1611667E+1	1.2590699E-5	3.5335703E-4	3.8369000E+1
1.1612222E+1	1.2602927E-5	3.5150117E-4	3.8389000E+1
1.1612778E+1	1.2602927E-5	3.4778943E-4	3.8329000E+1
1.1613333E+1	1.2590699E-5	3.4407770E-4	3.8229000E+1

Table 4: Deviation of the measured coordinates from the real values in terms of the linear increments $\{\Delta X, \Delta Y\} - \{\Delta Z\}$, for first 10 records in GLONASS alone mode.

UTC (hour)	ΔX (m)	ΔY (m)	ΔZ (m)
1.1608889E+1	1.2627385E+1	-3.6634810E+0	3.8659000E+1
1.1609444E+1	1.2627385E+1	-3.6634810E+0	3.8539000E+1
1.1610000E+1	1.2627385E+1	-3.6263637E+0	3.8499000E+1
1.1610556E+1	1.2615156E+1	-3.6078050E+0	3.8459000E+1
1.1611111E+1	1.2602927E+1	-3.5892463E+0	3.8369000E+1
1.1611667E+1	1.2590699E+1	-3.5335703E+0	3.8369000E+1
1.1612222E+1	1.2602927E+1	-3.5150117E+0	3.8389000E+1
1.1612778E+1	1.2602927E+1	-3.4778943E+0	3.8329000E+1
1.1613333E+1	1.2590699E+1	-3.4407770E+0	3.8229000E+1
1.1613889E+1	1.2615156E+1	-3.4222183E+0	3.8309000E+1

3. Experiment B

Based on the second and third conclusions of Experiment A, This experiment was performed in order to determine the cause of the conclusions. The aim was to find out why the behaviour of Experiment A was contrary to the expectation. The answer to the behaviour of Experiment A was on the way the third component of the $\{\Delta x, \Delta y\}$-$\{\Delta z\}$ system was defined. While the NMEA GGA (*NMEA 0183* 1994) requires that the ellipsoidal height information be related to geoid, i.e. orthometric height H together with geoidal undulation N be recorded, the Ashtech system in using the NMEA GGA *(NMEA 0183 1994)* format records ellipsoidal height h and geoidal undulation N as in (2.4). Therefore, when ΔZ is corrected from the form

$$\Delta Z = \{h = H + N\} \tag{2.4}$$

to

$$\Delta Z = h \tag{2.5}$$

the expected results were obtained.
In this experiment, Experiment A was repeated with the ellipsoidal height defined as given by (2.5). The results are as presented in *Table 5*-
Table 7, and *Figure 7-Figure 12*.

Table 5: Statistical information of the variation of the repeated positions measurements $\{\Delta X, \Delta Y\} - \{\Delta Z\}$ in the three measuring modes, namely GLONASS alone, GPS alone and combined GLONASS-GPS.

Statistical information	GLONASS alone (m)	GPS alone (m)	combined GPS and GLONASS (m)
Mean(dx)	14.1994	-9.2328	16.5384
Mean(dy)	-2.6517	5.0870	-5.9186
Mean(dz)	-2.4716	-11.3106	9.9520
STD(dx)	0.9251	9.4333	3.2719
STD(dy)	0.5468	17.7989	6.0557
STD(dz)	9.6975	28.9730	13.7515
Min(dx)	12.5907	-30.4415	2.2454
Min(dy)	-3.6635	-25.0245	-29.4971
Min(dz)	-15.3210	-59.3810	-49.3210
Max(dx)	15.6478	9.6925	23.1072
Max(dy)	-1.5107	43.0858	4.5580
Max(dz)	13.3390	50.6290	38.6190

Table 6: Deviation of the measured ellipsoidal coordinates from the real values, for first 10 measurements in GLONASS alone mode.

UTC hour	Δ deg	Δ deg	Δh m
1.160888888888889E+1	-3.290000000077953E-5	1.721033333321742E-4	-7.990999999999985E+0
1.160944444444444E+1	-3.290000000077953E-5	1.721033333321742E-4	-8.110999999999990E+0
1.161000000000000E+1	-3.256666666828778E-5	1.721033333321742E-4	-8.150999999999954E+0
1.161055555555556E+1	-3.240000000204190E-5	1.719366666659283E-4	-8.190999999999974E+0
1.161111111111111E+1	-3.223333333579604E-5	1.717699999996824E-4	-8.281000000000006E+0
1.161166666666667E+1	-3.173333333705841E-5	1.716033333334366E-4	-8.281000000000006E+0
1.161222222222222E+1	-3.156666666370711E-5	1.717699999996824E-4	-8.260999999999967E+0
1.161277777777778E+1	-3.123333333121536E-5	1.717699999996824E-4	-8.320999999999970E+0
1.161333333333333E+1	-3.089999999872362E-5	1.716033333334366E-4	-8.420999999999992E+0
1.161388888888889E+1	-3.073333333247774E-5	1.719366666659283E-4	-8.341000000000008E+0

Table 7: Deviation of the measured coordinates from the real values in terms of the linear increments $\{\Delta X, \Delta Y\} - \{\Delta Z\}$, for first 10 records in GLONASS alone mode.

UTC (hour)	ΔX (m)	ΔY (m)	ΔZ (m)
1.160888888888889E+1	1.262738461818584E+1	-3.663481006607940E+0	-7.990999999999985E+0
1.160944444444444E+1	1.262738461818584E+1	-3.663481006607940E+0	-8.110999999999990E+0
1.161000000000000E+1	1.262738453431918E+1	-3.626363671275906E+0	-8.150999999999954E+0
1.161055555555556E+1	1.261515600130433E+1	-3.607805003609889E+0	-8.190999999999974E+0
1.161111111111111E+1	1.260292746837071E+1	-3.589246335943872E+0	-8.281000000000006E+0
1.161166666666667E+1	1.259069885189529E+1	-3.533570332945821E+0	-8.281000000000006E+0
1.161222222222222E+1	1.260292730096225E+1	-3.515011664488600E+0	-8.260999999999967E+0
1.161277777777778E+1	1.260292721725802E+1	-3.477894329156567E+0	-8.320999999999970E+0
1.161333333333333E+1	1.259069864283776E+1	-3.440776993824533E+0	-8.420999999999992E+0
1.161388888888889E+1	1.261515558237710E+1	-3.422218326158516E+0	-8.341000000000008E+0

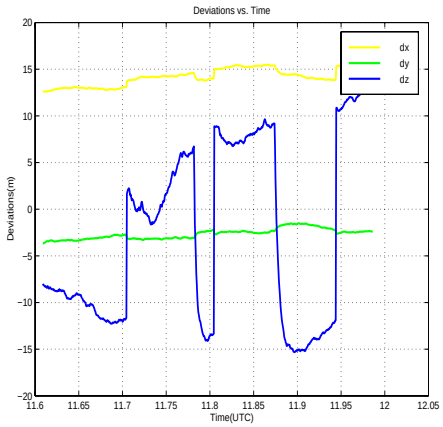


Figure 7: Deviation of measured coordinates from their known values versus time (UTC) when using GLONASS system alone.

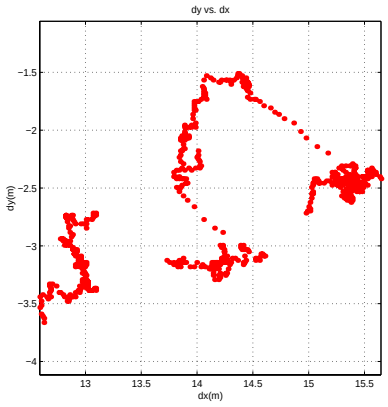


Figure 8: Variation of measured position minus its real value, projected into 2-D space of tangent plane, when using GLONASS system alone.

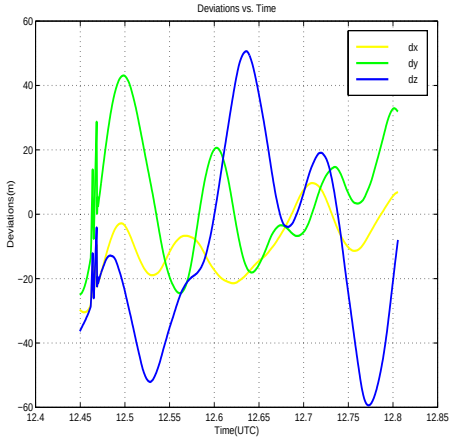


Figure 9: Deviation of measured coordinates from their known values versus time (UTC) when using GPS system alone.

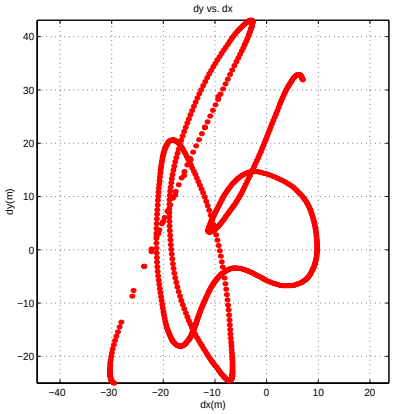


Figure 10: Variation of measured position minus its real value, projected into 2D space of tangent plane, when using GPS system alone.

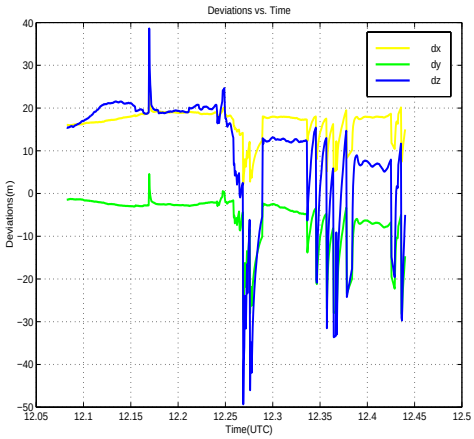


Figure 11: Deviation of measured coordinates from their known values versus time (UTC) when using combined GPS-GLONASS systems.

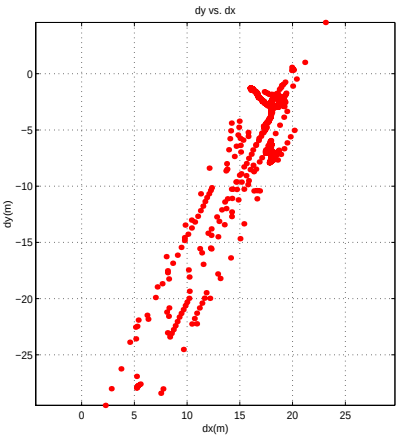


Figure 12: Variation of measured position minus its real value, projected into 2D space of tangent plane, when using combined GPS-GLONAS systems.

The following deductions can now be drawn for Experiment B:

By comparing the summary of the results given in *Table 5* we now see from the computed mean values that the GLONASS alone system is a bit more accurate than the GPS alone system because of the effect of the availability of the Selective Availability. The combined GPS-GLONASS, which is meant to improve on this effect of SA, is actually seen to be better than the GPS alone system. In summary we have:

(i) In terms of the computed means in x, y, and z components, the combined case was in between the stand-alone cases. The best means were those of GLONASS alone in (dy) and (dz), with the (dz) value of the combined GPS-GLONASS case being better than the GPS alone system.

(ii) In terms of *standard deviations* (dispersion of the data around the mean), the GLONASS results were better. This was followed by the combined GPS-GLONASS case.

4. Conclusions

The two Experiments above show that an error of the magnitude of the geoidal undulation will be incurred owing to the different definition of the NMEA GGA by the Ashtech receiver. Care needs therefore to be taken in using the NMEA GGA format! The calculations were carried out by a MATLAB program, which is included in the Appendix.

Acknowledgements

We kindly appreciate the courtesy of the Institute of Navigation, University of Stuttgart which allowed us the facilities to undertake the present study. We would like to give special thanks to Prof. Dr.-Ing. habil Erik W. Grafarend and Prof. Dr.-Ing. Alfred Kleusberg for agreeing to read and discuss the document. Their ideas were of great help.

References

Ashtech 1997: GG24™ OEM Board & Sensor, GPS + GLONASS™ Reference Manual. Nr. 630098, Revised Version B. Sunnyvale, CA USA 1997.

Langley, R.B (1995): NMEA 0183: A GPS Receiver Interface Standard. GPS World July 1995.

Leick, A (1995): GPS Satellite Surveying, John Wiley & Sons, U.S.A 1995

Mathes, A (1998): GPS und GLONASS als Teil eines Meßsystems in der Geodäsie am Beispiel des Systems HIGGINS. Dissertationen, DGK Heft. Nr.500 (1998)

NMEA 0183 (1994) Standard for Interfacing Marine Electronic Devices, Version 2.01, 1994

Strang, G and K Borre (1997) Linear Algebra, Geodesy, and GPS. Wellesky-Cambridge Press.

Trimble Navigation (1992): 4000SSE Geodetic System Surveyor, 4000SSE Geodetic Surveyor. Operational Manual Part Number 20576-00. Sunnyvale, CA.

Appendix: Matlab Program for Computing Absolute Point Positioning

```
%                       exe1.m Program
%               Effect of Combination of GPS and GLONASS in Positioning accuracy


        clear all
        format long e
    fid1=fopen('glo.g24');
%  fid1=fopen('gps.g24');
%  fid1=fopen('gpsglo.g24');
```

```
    nop=680  % number of GLONAS measurements
   %nop=642; %number of GPS measurements
   %nop=645  %number of combined GPS-GLONAS measurements
        utc(nop)=0;
        phi(nop)=0;
        lam(nop)=0;
        H(nop)=0;
        N(nop)=0;

        for i=1:nop
                line=fgetl(fid1);
                utc(i)=str2num(line(8:9));
                utc(i)=utc(i)+str2num(line(10:11))/60;
                utc(i)=utc(i)+str2num(line(12:16))/3600;
                phi(i)=str2num(line(18:19));
                phi(i)=phi(i)+str2num(line(20:27))/60;
                if (line(29) == 'S') | (line(29) == 's')
                        phi(i)=-phi(i);
                end
                lam(i)=str2num(line(31:33));
                lam(i)=lam(i)+str2num(line(34:41))/60;
                if line(43) == 'W' | line(43) == 'w'
                        lam(i)=360-lam(i);
                end

                H(i)=str2num(line(55:63));
                N(i)=str2num(line(67:74));
        end
        h=H+N;
        fclose(fid1)
        phi1=48.7832389;
        lam1=9.17526173;
        h1=330.371;
        r=6380000;

        dphi=phi-phi1;
        dlam=lam-lam1;
        dh=h-h1;
        dy=2*pi*r/360.*dphi;
        dx=2*pi*r/360.*dlam.*cos(phi.*pi./180);
        dz=dh;

        mdy=mean(dy);
        mdx=mean(dx);
        mdz=mean(dz);

        mindy=min(dy);
        mindx=min(dx);
        mindz=min(dz);

        maxdy=max(dy);
        maxdx=max(dx);
        maxdz=max(dz);

        stddy=std(dy);
        stddx=std(dx);
        stddz=std(dz);

        set(plot(utc,dx,'-y',utc,dy,'-g',utc,dz,'-b'),'LineWidth',2)
        title('Deviations vs. Time')
        xlabel('Time(UTC)')
        ylabel('Deviations(m)')
%        gtext('dx')
%        gtext('dy')
%        gtext('dz')
    legend('dx','dy','dz')
        grid
        print graph5.ps -depsc2
        pause

    clf
%        plot(dx,dy,'-r')
        set(plot(dx,dy,'.r'),'MarkerSize',15)
        title('dy vs. dx')
        xlabel('dx(m)')
        ylabel('dy(m)')
        grid

    set(gca,'AspectRatio',[1 1])
    %set(gca,'XTick',-100:20:100,'YTick',-100:20:100)
        print graph6.ps -depsc2


stat=[mdy;mdx;mdz;stddy;stddx;stddz;mindy;mindx;mindz;maxdy;maxdx;maxdz];
%save stat.txt stat  -ascii -double
```

Somigliana-Pizzetti Minimum Distance Telluroid Mapping

Alireza A. Ardalan

Abstract

A minimum distance mapping from the physical surface of the earth to the telluroid under the normal filed of *Somigliana-Pizzetti* is constructed. The point-wise minimum distance mapping under the constraint that actual gravity potential at the a point of physical surface of the earth be equal to normal potential of *Somigliana-Pizzetti* leads to a system of four nonlinear equations. The normal equations of minimum distance mapping are derived and solved via *Newton-Raphson iteration*. The problem of the existence and uniqueness of the solution is addressed. As a case study the quasi-geoid for the state Baden-Württemberg (Germany) is computed.

0. Introduction

We start with the definition of telluroid, after *E. Grafarend* (1978), as the best approximate representation of the surface of the earth. Given the geometry and potential field of the earth surface the telluroid can be completely defined as soon as we define a projection scheme. Telluroid mapping from the known surface of the earth has already been studied by the *A. Bode and E. Grafarend* (1982). They have presented an *isoparametric mapping* from the surface of the earth onto the telluroid under the influence of the spherical normal field including the centrifugal term. *Isoparametric mapping* is based on three assumptions/constrains, namely (i) $\lambda_p = \Lambda_P$, (ii) $\phi_p = \Phi_P$, and (iii) $w_p = W_P$. The triple coordinates $\{\lambda_p, \phi_p, w_p\}$ are representing the known longitude, latitude and gravity potential on the surface of the earth, and $\{\Lambda_P, \Phi_P, W_P\}$ are the normal counterpart of the same quantities on the telluroid. The first two constraints in *A. Bode and E. Grafarend* (1982) approach are referring to the definition of the mapping from the surface of the earth onto the telluroid (in their case *isoparametric mapping*). The present approach differs from approach proposed by *A. Bode and E. Grafarend* (1982) in (i) the mapping scheme and (ii) the normal field. Here we will use the *Somigliana-Pizzetti* field as the normal field and employ minimal distance criterion for mapping the surface of the earth onto the telluroid. *E. Grafarend* and *P. Lohse* (1991) have already used the minimum distance mapping to map the points on the physical surface of the earth onto the reference ellipsoid.

Paragraph one deals with the definition of two types of ellipsoidal coordinate systems which are used throughout the sequel. The set up of the variational equations of *Somigliana-Pizzetti minimum distance telluroid mapping* is dealt with in paragraph two. Paragraph three is devoted to our case study, i.e., "*quasi-geoid* for the state Baden-Württemberg".

1. Ellipsoidal Coordinates

Somigliana-Pizzetti field as the gravity field of a level ellipsoid can be most easily described in terms of any type of ellipsoidal coordinates which has an ellipsoid-of-revolution as one of it coordinate surfaces. For this reason here we will use *Jacobi spheroidal* coordinates $\{\lambda, \phi, u\}$ to present the *Somigliana-Pizzetti field*. *Jacobi spheroidal* coordinates $\{\lambda, \phi, u\}$ is one of the four variants of ellipsoidal coordi-

nates in which the *Laplace partial differential equation* separates. A résumé of basic geometry of *Jacobi spheroidal* coordinates $\{\lambda,\phi,u\}$ is given in *Definition 1-1*. Details on Jacobi spheroidal coordinates can be found in *P. Moon and D. Spencer* (1953, 1961), *N. Thong and E. Grafarend* (1989) for example. Besides since the GPS/GLONASS coordinate are normally given in terms of *Gauss* spheroidal coordinates $\{l,b,h\}$ *Definition1-2* is included which covers some basic properties of *Gauss* spheroidal coordinates $\{l,b,h\}$. *Definition 1-3* provides us with forward and backward transformation between *Gauss* and *Jacobi spheroidal coordinates* after *E. Grafarend, A. Ardalan, M. Sideris* (1999).

Definition 1-1: *Jacobi spheroidal* coordinates $\{\lambda,\phi,u\}$ in $\mathbb{R}^3$

(i) Conversion of Cartesian coordinates $\{x,y,z\}$ *into Jacobi spheroidal coordinates* $\{\lambda,\phi,u\}$

(a) Forward transformation from spheroidal coordinates $\{\lambda,\phi,u\}$ *to Cartesian coordinates* $\{x,y,z\}$

$$\begin{aligned} x &= \sqrt{u^2+\varepsilon^2}\cos\phi\cos\lambda \\ y &= \sqrt{u^2+\varepsilon^2}\cos\phi\sin\lambda \\ z &= u\sin\phi \end{aligned} \tag{1.1}$$

$\varepsilon := \sqrt{a^2-b^2}$ defines the absolute eccentricity.

(b) Backward transformation from Cartesian coordinates $\{x,y,z\}$ *to spheroidal coordinates* $\{\lambda,\phi,u\}$

$$l = \begin{cases} \arctan\dfrac{y}{x} & \text{for } x>0 \text{ and } y\geq 0 \\ \arctan\dfrac{y}{x}+\pi & \text{for } x<0 \text{ and } y\neq 0 \\ \arctan\dfrac{y}{x}+2\pi & \text{for } x>0 \text{ and } y<0 \\ \dfrac{\pi}{2} & \text{for } x=0 \text{ and } y>0 \\ 3\dfrac{\pi}{2} & \text{for } x=0 \text{ and } y<0 \end{cases} \tag{1.2}$$

$$\phi = (\operatorname{sgn} z)\arccos\left\{\frac{1}{2\varepsilon^2}[(x^2+y^2+z^2)+\varepsilon^2-\sqrt{(x^2+y^2+z^2+\varepsilon^2)^2-4\varepsilon^2(x^2+y^2)}]\right\}^{1/2} \tag{1.3}$$

$$u = \{\frac{1}{2}[x^2+y^2+z^2-\varepsilon^2+\sqrt{(x^2+y^2+z^2-\varepsilon^2)^2+4\varepsilon^2z^2}]\}^{1/2} \tag{1.4}$$

(ii) Jacobi matrix of forward transformation $\{x,y,z\}\mapsto\{\lambda,\phi,u\}$

$$J = \begin{bmatrix} -\sqrt{u^2+\varepsilon^2}\cos\phi\sin\lambda & -\sqrt{u^2+\varepsilon^2}\sin\phi\cos\lambda & u/\sqrt{u^2+\varepsilon^2}\cos\phi\cos\lambda \\ \sqrt{u^2+\varepsilon^2}\cos\phi\cos\lambda & -\sqrt{u^2+\varepsilon^2}\sin\phi\sin\lambda & u/\sqrt{u^2+\varepsilon^2}\cos\phi\sin\lambda \\ 0 & u\cos\phi & \sin\phi. \end{bmatrix} \tag{1.5}$$

(iii) Length element

$$dS^2 = [d\lambda,d\phi,du]J^*J\begin{bmatrix} d\lambda \\ d\phi \\ du \end{bmatrix} \tag{1.6}$$

(iv) Metric tensor

$$G := J^* J = \begin{bmatrix} (u^2+\varepsilon^2)\cos^2\phi & 0 & 0 \\ 0 & u^2+\varepsilon^2\sin^2\phi & 0 \\ 0 & 0 & (u^2+\varepsilon^2\sin^2\phi)/(u^2+\varepsilon^2) \end{bmatrix} := [g_{nm}] \quad \forall \quad n,m \in \{1,2,3\}$$

(1.7)

■

Definition1-2: *Gauss* spheroidal coordinates $\{l,b,h\}$ in $\mathbb{R}^3$

(i) Conversion of Cartesian coordinates $\{x,y,z\}$ into Gauss spheroidal coordinates $\{l,b,h\}$

(a) Forward transformation from spheroidal coordinates $\{l,b,h\}$ to Cartesian coordinates $\{x,y,z\}$

$$\begin{cases} x = [\dfrac{a}{\sqrt{1-e^2\sin^2 b}} + h]\cos b\cos l \\ y = [\dfrac{a}{\sqrt{1-e^2\sin^2 b}} + h]\cos b\sin l \\ z = [\dfrac{a(1-e^2)}{\sqrt{1-e^2\sin^2 b}} + h]\sin b \end{cases} \tag{1.8}$$

$e := \sqrt{a^2-b^2}\,/\,a$ defines the relative eccentricity.

(b) Backward transformation from Cartesian coordinates $\{x,y,z\}$ to spheroidal coordinates $\{l,b,h\}$

$$l = \begin{cases} \arctan\dfrac{y}{x} & \text{for } x>0 \text{ and } y \geq 0 \\ \arctan\dfrac{y}{x} + \pi & \text{for } x<0 \text{ and } y \neq 0 \\ \arctan\dfrac{y}{x} + 2\pi & \text{for } x>0 \text{ and } y<0 \\ \dfrac{\pi}{2} & \text{for } x=0 \text{ and } y>0 \\ 3\dfrac{\pi}{2} & \text{for } x=0 \text{ and } y<0 \end{cases} \tag{1.9}$$

$b(x,y,z)$, $h(x,y,z)$, can be derived either by *Newton* iteration *or* by solving a system of algebraic equations (*E. Grafarend and P. Lohse* (1991)), *or* by using closed formulae of *K. Borkowski* (1989), *H. Heikkinen* (1982) or *M. Paul* (1973), for instance.

(ii) Jacobi matrix of forward transformation $\{l,b,h\} \mapsto \{x,y,z\}$

$$J := \begin{bmatrix} x_l & x_b & x_h \\ y_l & y_b & y_h \\ z_l & z_b & z_h \end{bmatrix}, \tag{1.10}$$

subject to

$$x_l = D_l x = -[\frac{a}{\sqrt{1-e^2\sin^2 b}} + h]\cos b\sin l$$

$$x_b = D_b x = -[\frac{a}{\sqrt{1-e^2\sin^2 b}} + h]\sin b\cos l + \frac{ae^2\sin b\cos b}{(1-e^2\sin^2 b)^{3/2}}\cos b\cos l$$
$$x_h = D_h x = \cos b\cos l$$

$$y_l = D_l y = [\frac{a}{\sqrt{1-e^2\sin^2 b}} + h]\cos b\cos l$$
$$y_b = D_b y = -[\frac{a}{\sqrt{1-e^2\sin^2 b}} + h]\sin b\sin l + \frac{ae^2\sin b\cos b}{(1-e^2\sin^2 b)^{3/2}}\cos b\sin l$$
$$y_h = D_h y = \cos b\sin l$$

$$z_l = D_l z = 0$$
$$z_b = D_b z = [\frac{a(1-e^2)}{\sqrt{1-e^2\sin^2 b}} + h]\cos b + \frac{a(1-e^2)e^2\sin b\cos b}{(1-e^2\sin^2 b)^{3/2}}\sin b$$
$$z_h = D_h z = \sin b$$

(iii) Distance element "Metric of $\{\mathbb{R}^3, g_{kl}\}$*"*

$$ds^2 = [dl, db, dh]J^* J\begin{bmatrix} dl \\ db \\ dh \end{bmatrix} \tag{1.11}$$

(iv) Metric tensor

$$G := J^* J = \begin{bmatrix} (\frac{a}{\sqrt{1-e^2\sin^2 b}} + h)^2\cos^2 b & 0 & 0 \\ 0 & (\frac{a(1-e^2)}{(1-e^2\sin^2 b)^{3/2}} + h)^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.12}$$

■

Definition 1-3: *Forward* transformation of *Gauss* ellipsoidal coordinates into *Jacobi* spheroidal coordinates and vice versa

(i) Forward transformation equations $\{\lambda, \phi, u\} \mapsto \{l, b, h\}$

$$\lambda = l \tag{1.13}$$
$$\phi = \arctan(\sqrt{1-e^2}\tan b) \tag{1.14}$$
$$u = \frac{1}{\sqrt{1-e^2}}\cos b\left[\frac{a(1-e^2)}{(1-e^2\sin^2 b)^{1/2}} + h\right]\left[1 + (1-e^2)\tan^2 b\right]^{1/2} \tag{1.15}$$

(ii) Backward transformation equations $\{\lambda, \phi, u\} \mapsto \{l, b, h\}$

$$l = \lambda \tag{1.16}$$
$$b = \arctan(\frac{1}{\sqrt{1-e^2}}\tan\phi) \tag{1.17}$$
$$h = \sqrt{1-e^2}\,u\cos(\phi)[1 + \frac{1}{1-e^2}\tan^2\phi]^{1/2} - a(1-e^2)\left[1 - e^2\frac{\tan^2\phi}{1-e^2+\tan^2\phi}\right]^{-1/2} \tag{1.18}$$

■

2. Variational Equations of Somigliana-Pizzetti Minimum Distance Telluroid Mapping

We define the *Somigliana-Pizzetti minimum distance telluroid mapping*, as follows:

> Given the point $p(\mathbf{x})$ on the surface of the Earth $\mathcal{M}_h^2$, i.e. $p(\mathbf{x}) \in \mathcal{M}_h^2$, with potential value $w_p = w(\mathbf{x})$,
> find the point $P(\mathbf{X})$ such that;
> (i) the normal *Somigliana-Pizzetti* potential field $W_P = W(\mathbf{X})$ at point $P(\mathbf{X}) \in \mathcal{M}_H^2$ be equal to the actual potential at $p(\mathbf{x}) \in \mathcal{M}_h^2$,
> (ii) the point $P(\mathbf{X}) \in \mathcal{M}_H^2$ be at minimum (*Euclidean*) distance from the point $p(\mathbf{x}) \in \mathcal{M}_h^2$ .on the physical surface of the earth.

The surface $\mathcal{M}_H^2$ is called *Molodensky telluroid*, or specifically in our case the *Molodensky telluroid of Somigliana-Pizzetti type*. *Figure 2-1* shows the points $p(\mathbf{x})$ on the earth's surface and its minimum distance projection $P(\mathbf{X})$ onto the telluroid.

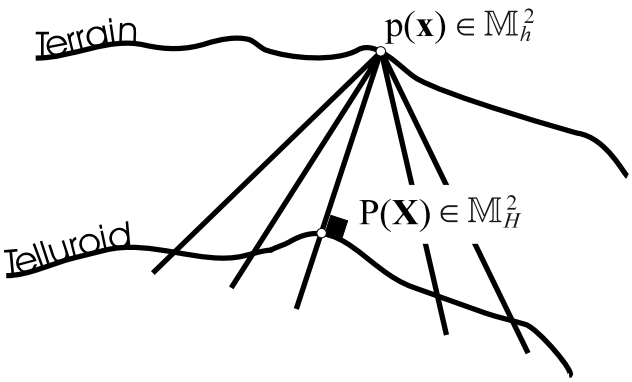


Figure 2-1: Point $p(\mathbf{x}) \in \mathcal{M}_h^2$ on the topographic surface and its minimum distance mapping onto the surface of type telluroid $\mathcal{M}_H^2$: Various projections, but one orthogonal projection of $\mathcal{M}_h^2$ onto $\mathcal{M}_H^2$.

In order to solve the original optimisation problem we proceed as following:

> Minimise the *Euclidean* distance between the point $p(\mathbf{x}) \in \mathcal{M}_h^2$ on the physical surface of the earth and the point $P(\mathbf{X}) \in \mathcal{M}_H^2$ on the *telluroid*,
>
> such that
>
> the normal potential $W_P = W(\mathbf{X})$ on the telluroid, according to *Somigliana-Pizzetti* normal field, be equal to the actual potential $w_p = w(\mathbf{x})$ at the topographic surface of the earth.

Analytically we formulate the optimisation problem by minimising the *constraint Lagrangean*:

$$
\begin{aligned}
&L(x_1, x_2, x_3, x_4) := \left\| \mathbf{x} - \mathbf{X} \right\|^2 + x_4 (W_P - w_p) \\
&= [x - X(x_1, x_2, x_3)]^2 + [y - Y(x_1, x_2, x_3)]^2 + [z - Z(x_1, x_2, x_3)]^2 + x_4 [W(x_1, x_2, x_3) - w_p] \qquad (2.1) \\
&= \min_{x_1, x_2, x_3, x_4}
\end{aligned}
$$

where $(x_1, x_2, x_3) = (\Lambda, \Phi, U)$ are *Jacobi spheroidal coordinates of the point* $P \sim \mathbf{X} \in \mathcal{M}_H^2$ on the telluroid, and x_4 is the unknown *Lagrange multiplier*.

Since the most suitable coordinate system to present the *Somigliana-Pizzetti field* is ellipsoidal coordinates, we formulate our minimisation problem in terms of *Jacobi spheroidal coordinates* $\{\lambda,\phi,u\}$.

Definition 2-1 presents the *Somigliana-Pizzetti* gravity potential field in terms of *Jacobi-spheroidal* coordinates $\{\lambda,\phi,u\}$. *Somigliana-Pizzetti* field has been developed by *P. Pizzetti* (1894) and *C. Somigliana* (1930) and recently extensively analysed by *E. Grafarend and A. Ardalan* (1999) in functional analytical terms.

Definition 2-1: *Somigliana-Pizzetti* field as developed by *P. Pizzetti* (1894), *C. Somigliana* (1930), and review by *E. Grafarend* and *A. Ardalan* (1999)

Somigliana-Pizzetti field as the gravity field of a rotational ellipsoid

$$\begin{aligned} W(\phi,u) &= \frac{GM}{\varepsilon}\operatorname{arc\,cot}(\frac{u}{\varepsilon}) \\ &+\frac{1}{6}\Omega^2 a^2 \frac{(3\frac{u^2}{\varepsilon^2}+1)\operatorname{arc\,cot}(\frac{u}{\varepsilon})-3\frac{u}{\varepsilon}}{(3\frac{b^2}{\varepsilon^2}+1)\operatorname{arc\,cot}(\frac{b}{\varepsilon})-3\frac{b}{\varepsilon}}(3\sin^2\phi-1) \\ &+\frac{1}{2}\Omega^2(u^2+\varepsilon^2)\cos^2\phi \end{aligned} \tag{2.2}$$

■

Using the forward transformation relations of $\{\lambda,\phi,u\}\mapsto\{x,y,z\}$ (see *Equation* (1.1) in *Definition 1-1*) the functional $L(x_1,x_2,x_3,x_4)$ can be written as

$$\begin{aligned} L(\Lambda_P,\Phi_{P2},U_P,\lambda) := &(x_p-\sqrt{{U_P}^2+\varepsilon^2}\cos\Phi_P\cos\Lambda_P)^2+(y_p-\sqrt{{U_P}^2+\varepsilon^2}\cos\Phi_P\sin\Lambda_P)^2+(z_p+U_P\sin\Phi_P)^2 \\ &+x_4(W(\Phi_P,U_P)-w_p) \end{aligned} \tag{2.3}$$

or

$$\begin{aligned} L(x_1,x_2,x_3,x_4) := &(x_p-\sqrt{{x_3}^2+\varepsilon^2}\cos x_2\cos x_1)^2+(y_p-\sqrt{{x_3}^2+\varepsilon^2}\cos x_2\sin x_1)^2+(z_p+x_3\sin x_2)^2 \\ &+x_4(W(x_2,x_3)-w_p) \end{aligned} \tag{2.4}$$

where $\{x_1,x_2,x_3\}$ are unknown *Jacobi spheroidal coordinates* of the point $P(\Lambda,\Phi,U)=P(x_1,x_2,x_3)$ on the telluroid ($P(\mathbf{X})\in\mathcal{M}_H^2$), $W(\Phi,U)=W(x_2,x_3)$ corresponds to *Somigliana-Pizetti* potential field at point $P(\Lambda,\Phi,U)\in\mathcal{M}_H^2$ according to (2.2), and w_p refers to actual gravity potential at point $p\{x,y,z\}$ on the surface of the earth.

The functional $L(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4)$ is minimal if and only if following two conditions hold:

$$
\text{(i)}\quad
\begin{cases}
f_1 := \dfrac{\partial L}{\partial x_1}(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4) = 2\sqrt{\hat{x}_3{}^2+\varepsilon^2}\cos\hat{x}_2(x_p\sin\hat{x}_1 - y_p\cos\hat{x}_1) = 0 \\
f_2 := \dfrac{\partial L}{\partial x_2}(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4) = 2(x_p - \sqrt{\hat{x}_3{}^2+\varepsilon^2}\cos\hat{x}_2\cos\hat{x}_1)\sqrt{\hat{x}_3{}^2+\varepsilon^2}\sin\hat{x}_2\cos\hat{x}_1 \\
\qquad\qquad +2(y_p - \sqrt{\hat{x}_3{}^2+\varepsilon^2}\cos\hat{x}_2\sin\hat{x}_1)\sqrt{\hat{x}_3{}^2+\varepsilon^2}\sin\hat{x}_2\sin\hat{x}_1 \\
\qquad\qquad -2(z_p - \hat{x}_3\sin\hat{x}_2)\hat{x}_3\cos\hat{x}_2 + \hat{x}_4(\dfrac{\partial W}{\partial x_2}(\hat{x}_2,\hat{x}_3)) = 0 \\
f_3 := \dfrac{\partial L}{\partial x_3}(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4) = -2\dfrac{(x_p - \sqrt{\hat{x}_3{}^2+\varepsilon^2}\cos\hat{x}_2\cos\hat{x}_1)\hat{x}_3\cos\hat{x}_2\cos\hat{x}_1}{\sqrt{\hat{x}_3{}^2+\varepsilon^2}} \\
\qquad\qquad -2\dfrac{(y_p - \sqrt{\hat{x}_3{}^2+\varepsilon^2}\cos\hat{x}_2\sin\hat{x}_1)\hat{x}_3\cos\hat{x}_2\sin\hat{x}_1}{\sqrt{\hat{x}_3{}^2+\varepsilon^2}} \\
\qquad\qquad -2(z_p - \hat{x}_3\sin\hat{x}_2)\sin\hat{x}_2 + \hat{x}_4\dfrac{\partial W}{\partial x_3}(\hat{x}_2,\hat{x}_3) = 0 \\
f_4 := \dfrac{\partial L}{\partial x_4}(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4) = W(\hat{x}_2,\hat{x}_3) - w_p = 0
\end{cases}
\tag{2.5}
$$

$$
\text{(ii)}\ \frac{\partial^2 L}{\partial x_i \partial x_j}(\hat{x}_1,\hat{x}_2,\hat{x}_3)\ \text{be positive-semi-definite for } i,j = 1,2,3 \tag{2.6}
$$

Partial derivatives $\partial W/\partial\phi$ and $\partial W/\partial u$ of (2.5) can be readily derive from (2.2) as follows

$$
\frac{\partial W}{\partial \phi} = \frac{\partial W}{\partial x_2} = a^2\Omega^2\frac{(3x_3{}^2+\varepsilon^2)\operatorname{arccot}(\frac{x_3}{\varepsilon}) - 3x_3\varepsilon}{[(3b^2+\varepsilon^2)\operatorname{arccot}(\frac{b}{\varepsilon}) - b\varepsilon]}\sin x_2\cos x_2 - \Omega^2(x_3{}^2+\varepsilon^2)\sin x_2\cos x_2 \tag{2.7}
$$

$$
\begin{aligned}
\frac{\partial W}{\partial u} = \frac{\partial W}{\partial x_3} = &-\frac{GM}{x_3{}^2+\varepsilon^2} \\
&-\frac{1}{3}\Omega^2 a^2\frac{\varepsilon(3x_3{}^2+2\varepsilon^2)+(-3x_3{}^3-3x_3\varepsilon^2)\operatorname{arccot}(\frac{x_3}{\varepsilon})}{(x_3{}^2+\varepsilon^2)[\operatorname{arccot}(\frac{b}{\varepsilon})\varepsilon^2+(-3\varepsilon+3\operatorname{arccot}(\frac{b}{\varepsilon})b)b]}(3\sin^2 x_2 - 1) \\
&+\Omega^2 x_3\cos^2 x_2
\end{aligned}
\tag{2.8}
$$

Equations (2.5) builds up the variational equations of the optimisation problem (2.1). System of equations (2.5) is a nonlinear system; the *B. Taylor* expansion of it reads

$$
\begin{aligned}
&\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0) + \frac{1}{1!}\mathbf{F}'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \\
&+\frac{1}{2!}\mathbf{F}''(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)\otimes(\mathbf{x}-\mathbf{x}_0) + \mathcal{O}_3\big((\mathbf{x}-\mathbf{x}_0)\otimes(\mathbf{x}-\mathbf{x}_0)\otimes(\mathbf{x}-\mathbf{x}_0)\big) \\
&= \mathbf{F}_0 + \mathbf{J}_0(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}\mathbf{H}_0(\mathbf{x}-\mathbf{x}_0)\otimes(\mathbf{x}-\mathbf{x}_0) + \mathcal{O}_3
\end{aligned}
\tag{2.9}
$$

where

$$\mathbf{F}=\begin{bmatrix} f_1(x_1,x_2,x_3,x_4)\\ f_2(x_1,x_2,x_3,x_4)\\ f_3(x_1,x_2,x_3,x_4)\\ f_4(x_1,x_2,x_3,x_4)\end{bmatrix},\ \mathbf{x}=\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\end{bmatrix},\ \mathbf{F}':=\mathbf{J}=\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} & \frac{\partial f_1}{\partial x_4}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4}\\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4}\\ \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4}\end{bmatrix},$$

$\mathbf{F}'':=\mathbf{H}=$

$$\begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 x_2} & \frac{\partial^2 f_1}{\partial x_1 x_3} & \frac{\partial^2 f_1}{\partial x_1 x_4} & \frac{\partial^2 f_1}{\partial x_2 x_1} & \frac{\partial^2 f_1}{\partial x_2^2} & \frac{\partial^2 f_1}{\partial x_2 x_3} & \frac{\partial^2 f_1}{\partial x_2 x_4} & \frac{\partial^2 f_1}{\partial x_3 x_1} & \frac{\partial^2 f_1}{\partial x_3 x_2} & \frac{\partial^2 f_1}{\partial x_3^2} & \frac{\partial^2 f_1}{\partial x_3 x_4} & \frac{\partial^2 f_1}{\partial x_4 x_1} & \frac{\partial^2 f_1}{\partial x_4 x_2} & \frac{\partial^2 f_1}{\partial x_4 x_3} & \frac{\partial^2 f_1}{\partial x_4^2}\\ \frac{\partial^2 f_2}{\partial x_1^2} & \frac{\partial^2 f_2}{\partial x_1 x_2} & \frac{\partial^2 f_2}{\partial x_1 x_3} & \frac{\partial^2 f_2}{\partial x_1 x_4} & \frac{\partial^2 f_2}{\partial x_2 x_1} & \frac{\partial^2 f_2}{\partial x_2^2} & \frac{\partial^2 f_2}{\partial x_2 x_3} & \frac{\partial^2 f_2}{\partial x_2 x_4} & \frac{\partial^2 f_2}{\partial x_3 x_1} & \frac{\partial^2 f_2}{\partial x_3 x_2} & \frac{\partial^2 f_2}{\partial x_3^2} & \frac{\partial^2 f_2}{\partial x_3 x_4} & \frac{\partial^2 f_2}{\partial x_4 x_1} & \frac{\partial^2 f_2}{\partial x_4 x_2} & \frac{\partial^2 f_2}{\partial x_4 x_3} & \frac{\partial^2 f_2}{\partial x_4^2}\\ \frac{\partial^2 f_3}{\partial x_1^2} & \frac{\partial^2 f_3}{\partial x_1 x_2} & \frac{\partial^2 f_3}{\partial x_1 x_3} & \frac{\partial^2 f_3}{\partial x_1 x_4} & \frac{\partial^2 f_3}{\partial x_2 x_1} & \frac{\partial^2 f_3}{\partial x_2^2} & \frac{\partial^2 f_3}{\partial x_2 x_3} & \frac{\partial^2 f_3}{\partial x_2 x_4} & \frac{\partial^2 f_3}{\partial x_3 x_1} & \frac{\partial^2 f_3}{\partial x_3 x_2} & \frac{\partial^2 f_3}{\partial x_3^2} & \frac{\partial^2 f_3}{\partial x_3 x_4} & \frac{\partial^2 f_3}{\partial x_4 x_1} & \frac{\partial^2 f_3}{\partial x_4 x_2} & \frac{\partial^2 f_3}{\partial x_4 x_3} & \frac{\partial^2 f_3}{\partial x_4^2}\\ \frac{\partial^2 f_4}{\partial x_1^2} & \frac{\partial^2 f_4}{\partial x_1 x_2} & \frac{\partial^2 f_4}{\partial x_1 x_3} & \frac{\partial^2 f_4}{\partial x_1 x_4} & \frac{\partial^2 f_4}{\partial x_2 x_1} & \frac{\partial^2 f_4}{\partial x_2^2} & \frac{\partial^2 f_4}{\partial x_2 x_3} & \frac{\partial^2 f_4}{\partial x_2 x_4} & \frac{\partial^2 f_4}{\partial x_3 x_1} & \frac{\partial^2 f_4}{\partial x_3 x_2} & \frac{\partial^2 f_4}{\partial x_3^2} & \frac{\partial^2 f_4}{\partial x_3 x_4} & \frac{\partial^2 f_4}{\partial x_4 x_1} & \frac{\partial^2 f_4}{\partial x_4 x_2} & \frac{\partial^2 f_4}{\partial x_4 x_3} & \frac{\partial^2 f_4}{\partial x_4^2}\end{bmatrix}$$

and $\otimes$ stands for *Kronecker* tensor product.

Newton iteration solution (*X. Chen et al.* (1997)) can be performed by the n-sequence

$$\mathbf{x}-\ \mathbf{x}_0=\Delta\mathbf{x}=\mathbf{J}_0^{-1}(\mathbf{F}-\mathbf{F}_0)=\left(\mathbf{J}(\mathbf{x}_0)\right)^{-1}(\mathbf{F}-\mathbf{F}_0) \tag{2.10}$$

$$\mathbf{x}-\ \mathbf{x}_0\ =\mathbf{J}^{-1}(\mathbf{F}-\mathbf{F}_0)$$

$$\Rightarrow\quad \mathbf{x}_1=\mathbf{x}_0+\mathbf{J}_0^{-1}(\mathbf{F}-\mathbf{F}_0) \tag{2.11}$$

$$\Rightarrow\quad \mathbf{x}_2\ =\mathbf{x}_1+\mathbf{J}_1^{-1}(\mathbf{F}-\mathbf{F}_1) \tag{2.12}$$

$$\Rightarrow\qquad \cdots\qquad \Rightarrow \mathbf{x}_n=\mathbf{x}_{n-1} \tag{2.13}$$

where Jacobean matrix of linearized form of the variational equations (2.5) reads as

$$\mathbf{J}:=\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} & \frac{\partial f_1}{\partial x_4}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4}\\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4}\\ \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4}\end{bmatrix} \tag{2.14}$$

$$\frac{\partial f_1}{\partial x_1}=2(x_3^{\,2}+\varepsilon^2)^{1/2}\cos x_2(x_p\cos x_1+y_p\sin x_1)$$

$$\frac{\partial f_1}{\partial x_2}=2(x_3^{\,2}+\varepsilon^2)^{1/2}\sin x_2(-x_p\sin x_1+y_p\cos x_1)$$

$$\frac{\partial f_1}{\partial x_3}=-2\frac{\cos x_2(-x_p\sin x_1+y_p\cos x_1)x_3}{(x_3^{\,2}+\varepsilon^2)^{1/2}}$$

$$\frac{\partial f_1}{\partial x_4}=0$$

$$\frac{\partial f_2}{\partial x_1}=-2(x_3^{\,2}+\varepsilon^2)^{1/2}\sin x_2(x_p\sin\ x_1-y_p\cos x_1)$$

$$\frac{\partial f_2}{\partial x_2}=2(x_3^{\,2}+\varepsilon^2)(\sin x_2)^2(\cos x_1)^2+2(x_p-(x_3^{\,2}+\varepsilon^2)^{1/2}\cos(x_2)\cos x_1)(x_3^{\,2}+\varepsilon^2)^{1/2}\cos x_2\cos x_1+2(x_3^{\,2}+\varepsilon^2)(\sin x_2)^2(\sin x_1)^2+2(y_p-(x_3^{\,2}+\varepsilon^2)^{1/2}\cos x_2\sin x_1)(x_3^{\,2}+\varepsilon^2)^{1/2}\cos x_2\sin x_1+2x_3^{\,2}(\cos x_2)^2+2(z_p-x_3\sin x_2)x_3\sin x_2+x_4(\Omega^2a^2((3x_3^{\,2}/\varepsilon^2+1)\operatorname{acot}(x_3/\varepsilon)-3x_3/\varepsilon)/((3b^2/\varepsilon^2+1)\operatorname{acot}(b/\varepsilon)$$

$$
\begin{aligned}
&-3b/\varepsilon)\cos(x_2)^2-\Omega^2 a^2((3x_3{}^2/\varepsilon^2+1)\operatorname{acot}(x_3/\ )-3x_3/\ )/((3b^2/\varepsilon^2+1)\operatorname{acot}(b/\ )-3b/\ )(\sin x_2)^2\\
&+\Omega^2(x_3{}^2+\varepsilon^2)(\sin x_2)^2-\Omega^2(x_3{}^2+\varepsilon^2)(\cos x_2)^2)\\
\frac{\partial f_2}{\partial x_3} &= -2\sin x_2(\cos x_1)^2\cos(x_2)x_3+2(x_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\cos x_1)/(x_3{}^2+{}^2)^{1/2}\sin x_2\cos x_1(x_3)\\
&-2\sin x_2(\sin x_1)^2\cos x_2(x_3)+2(y_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\sin x_1)/(x_3{}^2+{}^2)^{1/2}\sin x_2\sin x_1(x_3)\\
&+2x_3\cos x_2\sin x_2-2(z_p-x_3\sin x_2)\cos x_2+x_4(\Omega^2a^2(6x_3{}^2\operatorname{acot}(x_3/\ )\\
&-(3x_3{}^2/{}^2+1)/\ /(1+x_3{}^2/{}^2)-3/\ )/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )\sin x_2\cos x_2\\
&-2\Omega^2x_3\cos x_2\sin x_2)\\
\frac{\partial f_2}{\partial x_4} &= \Omega^2a^2((3x_3{}^2/{}^2+1)\operatorname{acot}(x_3/\ )-3x_3/\ )/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )\sin x_2\cos x_2\\
&-\Omega^2(x_3{}^2+{}^2)\cos x_2\sin x_2\\
\frac{\partial f_3}{\partial x_1} &= 2\frac{\cos x_2x_3(\sin x_1x_p-\cos x_1y_p)}{(x_3{}^2+\varepsilon^2)^{1/2}}\\
\frac{\partial f_3}{\partial x_2} &\quad 2\sin x_2(\cos x_1)^2\cos(x_2)x_3+2(x_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\cos x_1)/(x_3{}^2+{}^2)^{1/2}\sin x_2\cos(x_1)x_3\\
&-2\sin x_2(\sin x_1)^2\cos(x_2)x_3+2(y_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\sin x_1)/(x_3{}^2+{}^2)^{1/2}\sin x_2\sin(x_1)x_3\\
&+2x_3\cos x_2\sin x_2-2(z_p-x_3\sin x_2\cos x_2+x_4(\Omega^2a^2(6x_3/{}^2\operatorname{acot}(x_3/\ )-(3x_3{}^2/{}^2+1)/\ /(1\\
&+x_3{}^2/{}^2)-3/\ )/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )\sin x_2\cos x_2-2\Omega^2x_3\cos x_2\sin x_2)\\
\frac{\partial f_3}{\partial x_3} &= 2/(x_3{}^2+{}^2)(\cos x_2)^2(\cos x_1)^2x_3{}^2+2(x_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\cos x_1)/(x_3{}^2+{}^2)^{3/2}\cos x_2\cos(x_1)x_3{}^2\\
&-2(x_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\cos x_1)/(x_3{}^2+{}^2)^{1/2}\cos x_2\cos x_1+2/(x_3{}^2+{}^2)(\cos x_2)^2\sin(x_1)^2x_3{}^2\\
&+2(y_p-(x_3{}^2+{}^2)^{1/2}\cos x_2\sin x_1)/(x_3{}^2+{}^2)^{3/2}\cos x_2\sin x_1x_3{}^2-2(y_p\\
&-(x_3{}^2+{}^2)^{1/2}\cos x_2\sin x_1)/(x_3{}^2+{}^2)^{1/2}\cos x_2\sin x_1+2\sin x_2{}^2+x_4(2\mathrm{GM}/{}^4/(1+x_3{}^2/{}^2)^2x_3\\
&+1/6\Omega^2a^2(6/{}^2\operatorname{acot}(x_3/\ )-12x_3/{}^3/(1+x_3{}^2/{}^2)+2(3x_3{}^2/{}^2+1)/{}^3/(1+x_3{}^2/{}^2)^2x_3)\\
&/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )(3\sin(x_2)^2-1)+\Omega^2(\cos x_2)^2)\\
\frac{\partial f_3}{\partial x_4} &\quad \mathrm{GM}/{}^2/(1+x_3{}^2/{}^2)+1/6\Omega^2a^2(6x_3/{}^2\operatorname{acot}(x_3/\ )-(3x_3{}^2/{}^2+1)/\ /(1+x_3{}^2/{}^2)\\
&-3/\ )/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )(3(\sin x_2)^2-1)+\Omega^2x_3(\cos x_2)^2\\
\frac{\partial f_4}{\partial x_1} &= 0\\
\frac{\partial f_4}{\partial x_2} &= \Omega^2a^2((3x_3{}^2/{}^2+1)\operatorname{acot}(x_3/\ )-3x_3/\ )/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )\sin x_2\cos x_2\\
&-\Omega^2(x_3{}^2+{}^2)\cos(x_2)\sin(x_2)\\
\frac{\partial f_4}{\partial x_3} &\quad \mathrm{GM}/{}^2/(1+x_3{}^2/{}^2)+1/6\Omega^2a^2(6x_3/{}^2\operatorname{acot}(x_3/\ )-(3x_3{}^2/{}^2+1)/\ /(1+x_3{}^2/{}^2)-\\
&3/\ )/((3b^2/{}^2+1)\operatorname{acot}(b/\ )-3b/\ )(3(\sin x_2)^2-1)+\Omega^2x_3(\cos x_2)^2\\
\frac{\partial f_4}{\partial x_4} &= 0
\end{aligned}
$$

The solution set $(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4)$ derived from final step of *Newton iteration* (2.13) provides the necessary condition (2.5) for having a minimal solution. This extremal solution is minimal if condition (ii) of (2.6) be also satisfied. Indeed, we must show that *Hesse matrix* $\mathbf{H}_L$ is *positive semi-definite*, i.e., the characteristic polynomials of $|\mathbf{H}_L-\lambda\mathbf{I}|=0$ are all *non-negative*. The *Hesse matrix* $\mathbf{H}_L$ of second derivatives is given below.

$$\mathbf{H}_L = \frac{\partial^2 L}{\partial x_i \partial x_j} = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^{\,2}} & \frac{\partial^2 L}{\partial x_1 x_2} & \frac{\partial^2 L}{\partial x_1 x_3} \\ \frac{\partial^2 L}{\partial x_2 x_1} & \frac{\partial^2 L}{\partial x_2^{\,2}} & \frac{\partial^2 L}{\partial x_2 x_3} \\ \frac{\partial^2 L}{\partial x_3 x_1} & \frac{\partial^2 L}{\partial x_3 x_2} & \frac{\partial^2 L}{\partial x_3^{\,2}} \end{bmatrix} \tag{2.15}$$

$$\frac{\partial^2 L}{\partial x_1^{\,2}} = 2(x_3^{\,2} + \varepsilon^2)^{1/2} \cos x_2 (x_p \cos x_1 + y_p \sin x_1)$$

$$\frac{\partial^2 L}{\partial x_1 x_2} = 2(x_3^{\,2} + \varepsilon^2)^{1/2} \sin x_2 (-x_p \sin x_1 + y_p \cos x_1)$$

$$\frac{\partial^2 L}{\partial x_1 x_3} = -2x_3 \cos x_2 (-x_p \sin x_1 + y_p \cos x_1) / (x_3^{\,2} + \varepsilon^2)^{1/2}$$

$$\frac{\partial^2 L}{\partial x_2 x_1} = \frac{\partial^2 L}{\partial x_1 x_2}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial x_2^{\,2}} =& 2(x_3^{\,2}+\varepsilon^2)(\sin x_2)\, x_3^{\,2} (\cos x_1)^2 + 2(x_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \cos x_1)(x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \cos x_1 \\ &+2(x_3^{\,2}+\varepsilon^2)(\sin x_2)^2 (\sin x_1)^2 + 2(y_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \sin x_1)(x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \sin x_1 \\ &+2x_3^{\,2} \cos x_2^{\,2} + 2(z_p - x_3 \sin(x_2)\, x_3 \sin x_2 + x_4 (\Omega^2 a^2 ((3x_3^{\,2}/\varepsilon^2+1)\operatorname{acot}(x_3/\varepsilon) \\ &-3x_3/\varepsilon)/((3b^2/\varepsilon^2+1)\operatorname{acot}(b/\varepsilon) - 3b/\varepsilon)(\cos x_2)^2 - \Omega^2 a^2 ((3x_3^{\,2}/\varepsilon^2+1)\operatorname{acot}(x_3/\varepsilon) \\ &-3x_3/\varepsilon)/((3b^2/\varepsilon^2+1)\operatorname{acot}(b/\varepsilon) - 3b/\varepsilon)(\sin x_2)^2 + \Omega^2 (x_3^{\,2}+\varepsilon^2)(\sin x_2)^2 \\ &-\Omega^2 (x_3^{\,2}+\varepsilon^2)(\cos x_2)^2)\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial x_2 x_3} =& -2\cos x_2 (\cos x_1)^2 x_3 \sin x_2 \\ &+2(x_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \cos x_1)/(x_3^{\,2}+\varepsilon^2)^{1/2} \sin x_2 \cos(x_1)\, x_3 \\ &-2\cos x_2 (\sin x_1)^2 x_3 \sin x_2 + 2(y_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \sin x_1)/(x_3^{\,2} \\ &+\varepsilon^2)^{1/2} \sin x_2 \sin(x_1)\, x_3 + 2\sin(x_2)\, x_3 \cos x_2 - 2(z_p - x_3 \sin x_2)\cos x_2 \\ &+x_4 (\Omega^2 a^2 (6x_3/\varepsilon^{\,2} \operatorname{acot}(x_3/\varepsilon) - (3x_3^{\,2}/\varepsilon^2+1)/\varepsilon/(1+x_3^{\,2}/\mathrm{e}^2) \\ &-3/\varepsilon)/((3b^2/\varepsilon^2+1)\operatorname{acot}(b/\varepsilon) - 3b/\varepsilon)\sin x_2 \cos x_2 - 2\Omega^2 x_3 \cos x_2 \sin x_2)\end{aligned}$$

$$\frac{\partial^2 L}{\partial x_3 x_1} = \frac{\partial^2 L}{\partial x_1 x_3}$$

$$\frac{\partial^2 L}{\partial x_3 x_2} = \frac{\partial^2 L}{\partial x_2 x_3}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial x_3^{\,2}} =& 2/(x_3^{\,2}+\varepsilon^2)(\cos x_2)^2 (\cos x_1)^2 x_3^{\,2} + 2(x_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \cos x_1)/(x_3^{\,2}+\varepsilon^2)^{3/2} \cos x_2 \cos(x_1)\, x_3^{\,2} \\ &-2(x_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \cos x_1)/(x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \cos x_1 \\ &+2/(x_3^{\,2}+\varepsilon^2)(\cos x_2)^2 (\sin x_1)^2 x_3^{\,2} + 2(y_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \sin x_1)/(x_3^{\,2} \\ &+\varepsilon^2)^{3/2} \cos x_2 \sin(x_1)\, x_3^{\,2} - 2(y_p - (x_3^{\,2}+\varepsilon^2)^{1/2} \cos x_2 \sin x_1)/(x_3^{\,2} \\ &+\varepsilon^2)^{1/2} \cos x_2 \sin x_1 + 2(\sin x_2)^2 + x_4 (2\mathrm{GM}/\varepsilon^4/(1+x_3^{\,2}/\varepsilon^2)^2\, x_3 \\ &+1/6\,\Omega^2 a^2 (6/\varepsilon^2 \operatorname{acot}(x_3/\varepsilon) - 12x_3/\varepsilon^3/(1+x_3^{\,2}/\varepsilon^2) \\ &+2(3x_3^{\,2}/\varepsilon^2+1)/\varepsilon^3/(1+x_3^{\,2}/\varepsilon^2)^2\, x_3)/((3b^2/\varepsilon^2+1)\operatorname{acot}(b/\varepsilon) - 3b/\varepsilon)(3(\sin x_2)^2 - 1) \\ &+\Omega^2 (\cos x_2)^2)\end{aligned}$$

We proved the *positive-definteness* of the *Hesse-matrix* H_L of second derivatives by a numerical test.

3. Case study quasi-geoid of Baden-Württemberg

Next, we shall present the results of the minimum distance mapping of the *physical surface of the earth* $\mathbb{M}_h^2$ to the *Somigliana-Pizzetti telluroid* $\mathbb{M}_H^2$ for 157 GPS stations in the state Baden-Württemberg/Germany. *Table 1* shows the ten first GPS points of the GPS file of Baden-Württemberg. The coordinates are given in terms of *Gauss ellipsoidal coordinates* $\{l,b,h\}$ with respect to the GRS80 reference ellipsoid. This set of points constitute the Baden-Württemberg part (BWREF) of the German GPS network (DREF) which itself is part to European GPS network (EUREF).

Table 1: A part of the GPS file of the state Baden-Württemberg/Germany

Point ID Number	Longitude (l_p) (deg)	Latitude (b_p) (deg)	Ellipsoidal height (h_p) (m)	Geopotential Number (m^2/s^2)
621707001	8.623833676111	49.71395228806	218.6128	2148.9295328735
631805402	8.812993394167	49.60717789944	587.0355	5785.3299183738
632000110	9.065490941944	49.65066941389	590.7425	5811.1345798656
632107425	9.215170690833	49.65909479083	234.5042	2305.4354809625
632302808	9.575139530833	49.67395950667	379.2613	3733.6137953571
632400308	9.808090562778	49.64845024611	412.8049	4058.8965149575
641400308	8.159236366944	49.59699675556	349.7134	3434.9047294867
641600108	8.470342882778	49.59000986444	136.4937	1339.6896647405
641701308	8.628513233333	49.52709190306	143.9359	1415.1545995409
642100208	9.250862697222	49.55034041111	523.2340	5160.5154883848

The *Gauss ellipsoidal coordinates* $\{l,b,h\}$ of 157 GPS stations are converted to *Jacobi* ellipsoidal coordinates $\{\lambda,\phi,u\}$ according to forward transformation equations (1.13)-(1.15). *Table 2* presents the *Jacobi* ellipsoidal coordinates of the sample stations of *Table 1*.

Table 2: Transferred *Jacobi* ellipsoidal coordinates $\{\lambda,\phi,u\}_p$ of *Table 1*

Point ID Number	λ_p	ϕ_p	u_p
62230002	9.617432869327	49.60898966016	6356968.4504330
62230009	9.549475653403	49.60842348799	6356966.7368298
62230010	9.552776702951	49.61342803807	6356955.3539607
62230029	9.518509890067	49.66515664873	6356940.4330151
62230035	9.584526828221	49.60509389361	6356960.5466633
62230051	9.535064264303	49.62428053013	6356953.7623961
62230054	9.526483136723	49.63746430741	6356947.9810204
62230058	9.525534139567	49.64979432997	6356958.1832061
62230060	9.512823290742	49.65372897068	6356948.4890495
62230064	9.544545338174	49.62132133267	6356961.9798826

Newton Raphson iteration solution of the normal equations (2.5) led to point-wise *telluroid mapping* of all GPS stations in the state of *Baden-Württemberg*. A portion of the results for first ten GPS stations is presented in *Table 3*. Columns 2-4 are referring to *Jacobi* spheroidal coordinates of telluroid projection points. Column 5 presents the difference between u component of the GPS stations and their telluroid projection. Finally, column 6 shows the projection of $u_p - U_p$ along the unit vector $\mathbf{E}_u$. The geometrical height $H = (u - U_p)\sqrt{G_{33}}$ presents the separation between the surface of the earth and *Molodensky telluroid*, specifically the minimum distance mapping of the physical surface of the earth to the *Somigliana-Pizzetti* telluroid. If this height be considered as the height above the *reference ellipsoid*, by definition we have a presentation of the *quasi-geoid*. *Figure 4-1* is the result of the *minimum distance mapping* described here for *Baden-Württemberg* in the form of a *quasi-geoid* map.

Finally, the calculated quasi-geoid is compared with new *European Gravimetric Quasi-Geoid* (*EGG97*) (*H. Denker and W. Torge*, 1998). The summary of statistics of this comparison is given in Table 4.

Table 3: Telluroid mapping of the sample GPS stations of *Table 1*

Point ID Number	Λ_P	Φ_P	U_P	$u_p - U_p$	$(u-U_p)\sqrt{G_{33}}$ (m)
621707001	8.6238336761	49.619013350	6356923.6746	47.105907368	47.039684057
631805402	8.8129933941	49.512181451	6357295.0294	44.696132976	44.633028957
632000110	9.0654909419	49.555696232	6357297.6868	45.749522637	45.685046439
632107425	9.2151706908	49.564126141	6356939.7418	46.953250759	46.887093564
632302808	9.5751395308	49.578998873	6357085.8037	45.852377559	45.787813832
632400308	9.8080905627	49.553475871	6357118.3259	46.921681831	46.855543946
641400308	8.1592363669	49.501994892	6357054.1765	47.891587520	47.823938736
641600108	8.4703428827	49.495004292	6356840.8091	47.737670235	47.670215083
641701308	8.6285132333	49.432053186	6356848.6787	47.321327123	47.254288287
642100208	9.2508626972	49.455313888	6357231.6235	44.212039829	44.149472768

Table 4: Statistics of comparison between calculated height anomalies at 157 GPS station in Baden-Württemberg and EGG97

Statistics of $N_{EGG97} - \zeta$	(m)
mean	0.995013857205397
std	1.3227781145868
max	7.40296621419591
min	–0.921456458099193
number of sample points	157

4. Final Remarks and Conclusions

From a review of *Table 1 to Table 4* following conclusions can be made: (i) $\{\Lambda_P, \Phi_P\}$ of the telluroid point P is very close to $\{\lambda_p, \phi_p\}$ of point p on the surface of the earth. This reveals the fact that the minimum distance mapping of the physical surface of the earth to the *Somigliana-Pizzetti* telluroid is very close to the mapping along the coordinate line of u. (ii) The calculated quasi-geoid for GPS station based on minimum distance mapping of the physical surface of the earth to the *Somigliana-Pizzetti* deviates from *EGG97* by (0.995 ± 1.322778)(m) on average. This difference can be mainly associated to the interpolations process involved in providing the GPS stations with geopotential numbers. Indeed, since the present GPS stations of *Baden-Württemberg* are not identical with the first order level stations, where we have the geopotential numbers, such a interpolation is unavoidable. However, the present results, which are based on a very simple interpolation process, are indicating the minimum distance mapping of the physical surface of the earth to the *Somigliana-Pizzetti* telluroid as an optimal method in quasi-geoid calculations. This is especially valid if the GPS stations are occupied at the first order levelling stations, which we recommend for the future national GPS campaigns.

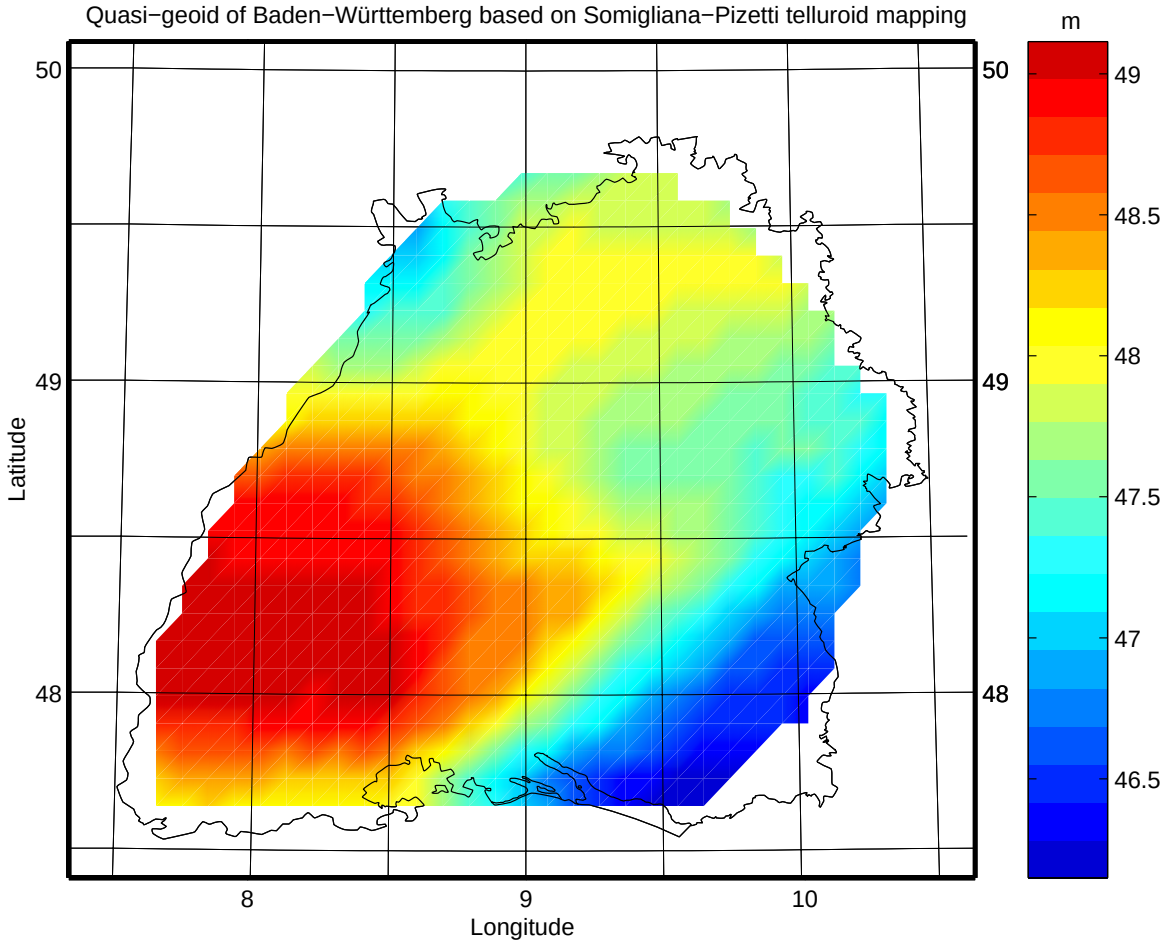


Figure 4-1: Quasi-geoid map of Baden-Württemberg, based on the minimum-distance mapping of the physical surface of the earth to the *Somigliana-Pizzetti* telluroid; variation: 46-49m.

References

BODE A. and E.W. GRAFAREND (1982): The telluroid mapping based on a normal gravity potential including the centrifugal term. Bolletino di Geodesia e Scienze Affini 41: 21-56.

BORKOWSKI K.M. (1989): Accurate algorithms to transform geocentric to geodetic coordinates. Bull. Géod. 63: 50-56.

CHEN X., Z. NASHED and L. QI (1997): Convergence of Newton's method for singular smooth and non-smooth equations using adaptive outer inverses. SIAM J. Optim. 7: 445-462.

DENKER H. and W. TORGE (1998): The European gravimetric quasi-geoid EGG97 -an IAG supported continental enterprise, In: R. Forsberg el al. (eds) IAG symp. proceed. 119: 249-254, Springer, Berlin-Heidelberg-New-York

GRAFAREND E.W. (1978): The definition of the telluroid. Bull. Geod. 52: 25-37.

GRAFAREND E.W. and P. LOHSE (1991): The minimal distance mapping of the topographic surface onto the (reference) ellipsoid of revolution. Manus. Geod. 16: 92-110.

GRAFAREND E.W. and A.A. ARDALAN (1999): World Geodetic Datum 2000 (WGD2000). Accepted for publication, Journal of Geodesy.

GRAFAREND E.W., R. SYFFUS and R. YOU (1995): Projective heights in geometry and gravity space. Allgemeine Vermessungs Nachrichten 382-403.

GRAFAREND E.W., A.A. ARDALAN and M. SIDERIS (1999): The Spheroidal Fixed-Free Two-Boundary Value Problem For Geoid Determination (The Spheroidal Bruns Transform). Accepted for publication, Journal of Geodesy.

HEIKKINEN M. (1982): Geschlossene Formeln zur Berechnung räumlicher geodätischer Koordinaten aus rechtwinkligen Koordinaten. Z.f. Verm. wesem 107: 207-211.

MOON P. and D.E. SPENCER (1953): Recent investigations of the separation of Laplace's equation. Ann. Math. Soc. Proc. 4: 302-307.
MOON P. and D.E. SPENCER (1961): Field theory handbook. Springer-Verlog, New York, Heidelberg, Berlin.
PAUL M.K. (1973): A note on computation of geodetic coordinates from geocentric (Cartesian) coordinates. Bull. Géod. 108: 135-139.
PIZZETTI P. (1894): Geodesia--Sulla espressione della gravita alla superficie del geoide, supposto ellissoidico. Atti Reale Accademia dei Lincei 3: 166-172.
SOMIGLIANA C. (1930): Geofisica--Sul campo gravitazionale esterno del geoide ellissoidico. Atti della Reale Academia Nazionale dei Lincei Rendiconti 6: 237-243.
THONG N.C. and E.W. GRAFAREND (1989): A spheroidal model of the terrestrial gravitational field. Manuscr. Geod. 14: 285-304.

Geodesy and Semantics
Progress by Graphs

Hans-Peter Bähr

1. Abstract

In Germany, basic geodetic research is coordinated by the German Geodetic Commission. Presently, semantic modeling is emerging from image analysis. In this respect, new challenges are identified besides well known problems which are put into a new context, like *terms*, *concepts*, *knowledge representation*, *isomorphism*. It has been shown that nets and graphs of all kind form suitable tools for knowledge representation and knowledge processing.

As an example, ERIK GRAFAREND is displayed in a semantic graph.

2. Research Coordination by the German Geodetic Commission (DGK)

„Quo vadis geodesia“ honors ERIK GRAFAREND and his contributions to our professional field. In addition to the many distinctions and responsibilities, he recently was elected Chairman of the Scientific Board of the German Geodetic Commission („Vorsitzender des Wissenschaftlichen Beirats der Deutschen Geodätischen Kommission“). This is a very demanding and important position since the „German Geodetic Commission“ (DGK) affiliated to the Bavarian Academy of Science, coordinates teaching and research in the German geodetic community. It has to be pointed out, that „Geodesy“ in German covers a somewhat broader understanding than as is conventionally the case within the English context. This means, that the DGK includes - besides the „core Geodesy“ - surveying, cartography, photogrammetry, remote sensing and land management. The fixed number of the 45 ordinary members integrated in the DGK are elected; they represent geodetic competence in the broad sense and try to focus diverging activities.

Positive results may be shown for research programs emerging from the DGK on long-term themes of the international scientific community. An example from the „core Geodesy“ is the „Forschungseinrichtung Satellitengeodäsie“ established in 1983 as a follow up to the scientific program SFB 78, which integrated academics as well as experts from federal administration and which led to the installation of the Geodetic Fundamental Station in Wettzell in 1972.

Within the DGK, particularly within the Scientific Board, challenging research themes are defined, thoroughly discussed and formulated for submission as long-term projects. The overall aim is scientific progress in the geodetic field through contribution from different areas and resulting synergetic effects.

Another activity in this respect is the joint program „Semantic Modeling“, funded by the German Science Foundation (DFG) since 1993. At the end of the 80´s, an idea emerged in the DGK to try something similar to the very successful “Satellite Geodesy Program” in the photogrammetric domain. Under the chairmanship of FRITZ ACKERMANN a group of a dozen experts from photogrammetry, cartography and computer science developed a program on „Semantic Modeling“. A first phase integrated research teams from 8 German universities, documented in W. FÖRSTNER/L. PLÜMER

(Eds., (1997)). The second phase is coming to an end in the year 2000, documented by a second SMATI workshop (SMATI 1999, Munich).

3. Limitations of Geometric Models

According to the famous definition of F. R. HELMERT (1880), Geodesy is "the science of the measurement and mapping the earth´s surface" (translation taken from TORGE 1980). This definition is basically given from a *geometrical* point of view. The realisation of this task requires primarily geometrical tools. However, today there is an overall tendency to incorporate more rigorously attributes in order to describe the earth like concepts, semantics, context and related tools.

The geometrical description of objects, at least from the conventional geodetic point of view, uses analytical models which were introduced by Greek philosophers and scientists. We have to stress the fact, that these models are very useful; but they are nonetheless *models* and may not fit to the „real world" as they describe reality only in a limited way. There is evidently no „point" in the real world, a fact, which causes some difficulties. To quote LAKOFF (1988):

"Theories may become so ingrained in our culture or in our intellectual life that we do not even recognise them as theories .....".

The real world is perceived by human beings as continuous while it is represented in digital image processing by discrete primitives. This contradiction is reflected by the complementary models in spatial and frequency domain, too.

Geometry may be considered as just an attribute of objects among others. In geodesy, however, including cartography and photogrammetry, geometric properties are given a prominent importance. On the other hand, interpretation of the surveyed objects has always played an important role in cartography and photogrammetry. Nevertheless, one was not aware of this condition, as object attributes (i. e. semantic features) were spontaneously added to the measured geometric parameters by the human.

This situation changes dramatically when the computer has to be trained to take over not only the geometric domain but also the semantics, for instance in image understanding. This step requires „modeling of semantics", which has proved to be a challenging task for the geodesist who exhibits a tendency to overestimate geometrical properties.

Rigorously spoken, it is not possible to separate geometry and semantics. Both features are essential for complete object description.

4. The Nature of Knowledge

When entering computer vision geometric and semantic features have to be modeled together in a much broader context. The more general concept of *knowledge* is taken, which is a very useful metaphor (LAKOFF,G.AND M. JOHNSON, 1980), when describing retrieval from pictorial information assisted by the computer.
There are several definitions of knowledge. KEITH DEVLIN (1991) formulates: „knowledge involves a mental state and a concept of truth". This is a very cautious attempt to describe the environment. MAKATO NAGAO (1990) forms an equation: „knowledge = cognition + logic". The latter definition is obviously more useful, as „cognition" is a more human oriented, general concept than „truth". Finally,

„logic“ includes a systematic model of knowledge which seems to be essential. Logic, order, rules, systematization etc. might be indispensable properties included in knowledge. We shall discuss this again in the next paragraphs.

Without losing the level of general acceptance we have to discriminate factual and procedural knowledge. Factual knowledge lists facts whereas procedural knowledge gives rules for action. In image analysis both contribute synergetically. We are going to show later that both domains must not be separated.

It is no wonder, that the two natures of „knowledge“ are evident in many different disciplines. In language science for instance *words* correspond to facts whereas *meaning* is based on rules within a context. In mathematics and computer science *declarative* algorithmic languages, like PROLOG are separated from *procedural* languages like FORTRAN. In philosophy, *representationistic* views (i. e. ARISTOTELES, FREGE) have to be discriminated from *instrumentalistic* approaches (like PLATON, WITTGENSTEIN), - see R.KELLER (1995), H.P. BÄHR (1998). Finally, in psychology, *male* is attributed to facts and *female* to rules.[1].

5. Knowledge Representation and Knowledge Processing by Graphs

What are the available tools to represent and to process knowledge? According to NAGAO this has to be based on logical rules. There are many alternatives but we think that graphs (networks) offer the best tool to structure knowledge. This is due to the twofold nature of knowledge, factual and procedural, as discussed in the previous section. Graphs may easily take *nodes* for the concepts (or facts) and the connecting *edges* for the context (or rules). Beside this, graphs allow the easy inclusion of topological features.

In image analysis, in language - or in whatever field where knowledge has to be represented and processed - facts, like objects or concepts are often overestimated or given too much importance in comparison to the interrelations which model the context like meaning or semantics.

In artificial intelligence, many solutions have been proposed for representing knowledge in nets (see H.KOCH et al. (1997)). Artificial neural networks try to simulate the process of learning in the human brain. They form an *implicit* representation of knowledge, i.e. knowledge representation and knowledge processing are elements of the same system. Information is introduced by the human operator during the analysis procedure. This is also true for Delaunay-Triangulation, another network tool which has shown good performance in image processing (K.-J. SCHILLING and TH. VÖGTLE (1996)). Graphs for implicit representation of knowledge follow a pragmatic approach. There is no rigorous modeling of the nets but just a heuristic approach.

This is not the case for *explicit* representation of knowledge in networks. Knowledge is thoroughly modeled a priori (according to NAGAO) in Semantic Nets or Markoff Random Fields to mention just two alternatives. Nevertheless, these tools allow not only knowledge representation but also knowledge processing.

Both, *implicit* representation or *explicit* representation of knowledge by graphs are adequate to serve the twofold nature of knowledge as described before. Besides the facts, they model the interrelations between concepts which contain the context. The context, however, determines the *meaning*, the *semantics* of both linguistic or visual features. To quote WITTGENSTEIN (1953): *„ for a large class of*

[1] *“Willst Du erfahren was sich ziemt, so frage nur bei edlen Frauen nach”* (Goethe, Torquato Tasso)

cases though not for all in which we employ the word „meaning" it can be defined thus: The meaning of a word is its use in the language". WITTGENSTEIN puts the word, the fact, in its individual context. The appropriate tool to do this in computer graphics is within a network.

Semantic networks (like in H. NIEMANN et al. (1990) and F.QUINT (1997)) are formally structured. This is not the case in neural nets that we humans build through continuous learning. D. HOFSTADTER (1979, see Fig. 1.) gives us an insight in a „tiny portion of the author´s semantic network". He groups his associations and relations between the main concepts of his book, GÖDEL, ESCHER and BACH. Figure 1 shows a small segment of the „tiny portion" given in his book directly associated with *Bach*. There is one cluster formed by *Bach*, *Goldberg*, *music*, *canons*, *fugues*, *musical offering*. Another group is composed by *semantic*, *language*, *sameness* and *isomorphisms*.

6. Sameness in Geodesy and in Semantics

In section 4 we stressed the fact that a semantic network reflects the design of a particular individual. Consequently, Figure 1 gives the associations of DOUGLAS HOFSTADTER. Nevertheless, we may follow him „more or less". It is most probable, that all possible readers of this text associate „Bach" with „Music". However, the association of „Music" and „Language" as given in Fig. 1 may be less common; it is part of the particular message DOUGLAS HOFSTADTER gives in his famous book on „GÖDEL, ESCHER, BACH: An Eternal Golden Braid". There are different reasons for accepting or not accepting an individual semantic network. As for the trivial case, concepts may be simply lacking - I am, for instance, not so sure that every reader has grasped the term „musical offering" („musikalisches Opfer"). It is obvious that *language* plays a most important role in semantic networks. This aspect will not be treated here in more detail (see H.-P. BÄHR and A.SCHWENDER (1996)).

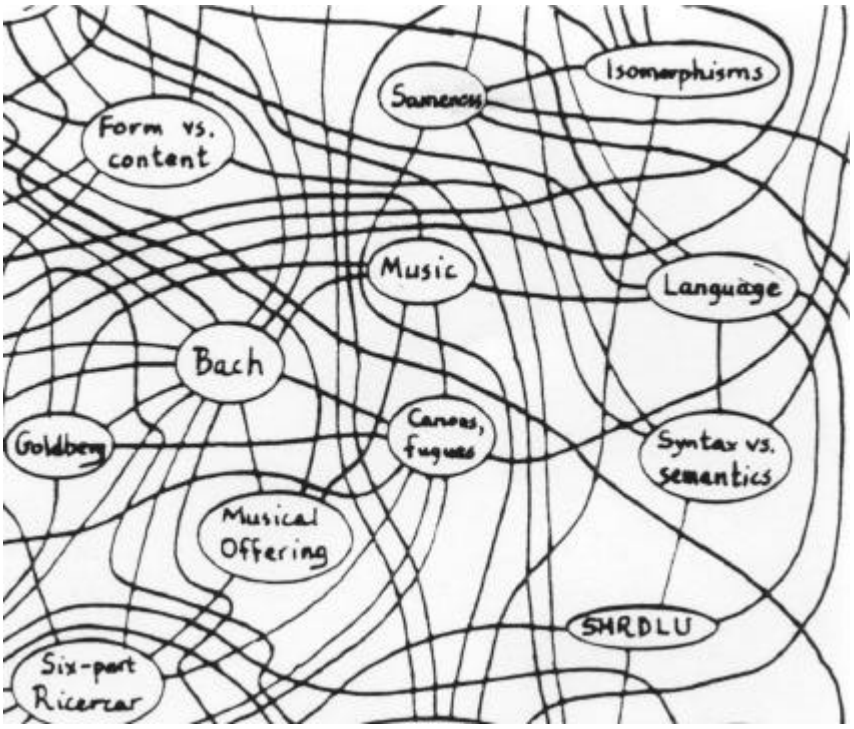


Fig. 1: A tiny portion of Hofstadter´s semantic network (HOFSTADTER 1979)

A fundamental question is how to define sameness in semantic networks. It cannot be done, obviously, just by "matching" the nodes, i.e. the concepts. As we showed earlier, the interrelations between the „facts", the context, has seriously to be taken into account.
Interestingly, in Geodesy we encounter basically the same problem. It corresponds for instance to the question: „Does a measurement fit"? In image processing the definition of „homogeneity" leads to the same problem. The assignment of a pattern of spectral signatures to a predefined class is of the same nature. A real world pattern will never fit exactly, same like individual semantic networks. Nevertheless, *sameness may be accepted under certain given conditions*.

This discussion leads to the concept of isomorphism. Isomorphism between two semantic networks does not only require sameness of facts, concepts or nodes in the graph, but also sameness of the "triggering pattern" as described by HOFSTADTER. This includes the interrelations between the nodes, which means, that beyond a certain level of detail, there will be no identical webs existing at all.

HOFSTADTER gives the example of spider nets, which never will be fully identical. Difference can be looked at in a local or in a global context. For given conditions, two nets may be accepted as identical locally or globally.

Geodesists are well prepared to contribute to the challenging discussion on sameness of semantics, as they are trained to evaluate fitness of data. However, one should know that fitness of semantics like in feature extraction from imagery, leads to much more complex problem compared to purely geometrical questions.

7. Conclusion: Erik Graph-Arend Displayed in a Semantic Graph

Figure 2 shows a „tiny portion" of the author´s semantic network focussed on ERIK GRAFAREND.

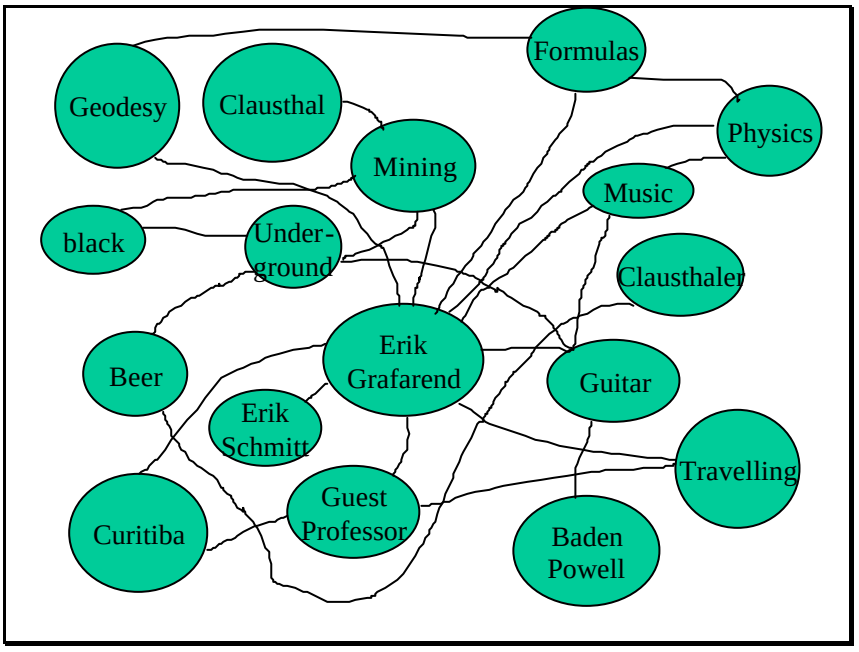


Fig. 2: Erik Grafarend in a semantic network

I did this performance, nota bene, spontaneously; there is obviously no hierarchy, no weighting, no completeness, just associations. But this is the way we are organised. Sometimes chaotic systems may

be superior to formally well-structured ones. We have to confess that this is in contradiction to what had been postulated for "knowledge" in section 4.
It is true what was said earlier, that any individual who knows ERIK GRAFAREND, will have his/her own background, that means he/she would design a very special network. Anyway, some of the concepts in the nodes will necessarily match like *geodesy*, *guitar* or *formulas*. Others, like *Curitiba* or *Erik Schmitt* are characteristic for the author´s web and show limited access from outside.

8. References

BÄHR, H.-P. and SCHWENDER, A. (1996): Linguistic Confusion in Semantic Modeling. In: KRAUS, K., WALDHÄUSL, P. (eds.): XVIII Congress of ISPRS, Vienna. International Archives of Photogrammetry and Remote Sensing, Vol. XXXI, Part B6, Commission VI, pp. 13-17.

BÄHR, H.-P. (1998b): From Data to Inference: Examples for Knowledge Representation in Image Understanding. Symposium Commission III ,ISPRS, Columbus/Ohio

BÄHR, H.-P. (1999): GIS-Introduction – Main Concepts. In: BÄHR, H.-P., VÖGTLE, Th. (eds.): GIS for Environmental Monitoring. Schweizerbart Verlag Stuttgart

DEVLIN, K. J.: Logic and Information. Cambridge University Press., 1995.

HELMERT, F.R.: Die mathematischen und physikalischen Teorien der Höheren Geodäsie. Teubner 1880

HOFSTADTER, D. R. (1979): Goedel, Escher, Bach: an Eternal Golden Braid. New York, Basic Books Inc.

KELLER, R. (1995): Zeichentheorie - zu einer Theorie semiotischen Wissens. Francke Verlag Tübingen und Basel

KOCH, H. et al. (1997): Knowledge Based Interpretation of Aerial Images and Maps Using a Digital Landscape Model as Partial Interpretation. In: SMATI 97, Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, Förstner, W. and Plümer, L. (Eds.), Birkhäuser

KUNZ, D. et al. (1997): A new Approach for Satellite Image Analysis by Means of a Semantic Network. In: SMATI 97, Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, Förstner, W. and Plümer, L. (Eds.), Birkhäuser

LAKOFF, G. (1988): Cognitive Semantics. In: U. Era, M. Santambrogio, P. Violi: Meaning and Mental Presentations. Indiana University Press, pp. 119 -154

LAKOFF, G., JOHNSON (1980), M.: Metaphors We Live By. The University of Chicago Press

LENK, H (1993).: Interpretationskonstrukte. Zur Kritik der interpretatorischen Vernunft. Suhrkamp

NAGAO, M. (1990): Knowledge and Inference. Academic Press

NIEMANN, H., SAGERER, G. (1990): ERNEST: A semantic network system for pattern understanding, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12 (9) 883-905.

QUINT, F. (1997): Kartengestützte Interpretation monokularer Luftbilder. Deutsche Geodätische Kommission, Serie C, Heft Nr. 477, München

SCHILLING, K.-J. e VÖGTLE, Th. (1996): Satellite Image Analysis Using Integrated Knowledge Processing. In: KRAUS, K., WALDHÄUSL, P. (eds.): XVIII Congress of ISPRS, Vienna. International Archives of Photogrammetry and Remote Sensing, Vol XXXI, Part B3, Commission III, pp. 752-757

SEGL, K. (1996): Integration von Form- und Spektralmerkmalen durch künstliche neuronale Netze bei der Satellitenbildklassifizierung. Deutsche Geodätische Kommission, Serie C, Heft Nr. 468, München

TAUER, W. and HUMBORG, G. (1992): Runoff Irrigation in the Sahel Zone. Technical Centre for Agricultural and Rural Cooperation, Margraf Weikersheim

TORGE, W. (1980): Geodesy. de Gruyter

WITTGENSTEIN, L. (1953): Philosophical Investigations. Basil Blackwell Oxford

Diffusion with space memory

Michele Caputo and Wolfango Plastino

Abstract

For a better resolution of the gravity values monitored on the surface of the Earth or underground is needed to analyze the time variation of the elevation at the measurement site. An important variation of the elevation is due to the effect of the pore filling of the ground caused by the migration processes of the underground water often associated to the ocean tides.
In order to obtain a better representation of the diffusion processes of fluids the Darcy's law has been modified introducing a general *time* memory formalism represented by fractional derivatives which imply a time filtering of the pressure gradient without singularities (Caputo, 1998a, 1999); a model which is particularly valid when considering the local phenomenology. In this note we introduce in Darcy's law the *space* fractional derivatives of the pressure which seems appropiate when considering a half space in order to represent the effect of the medium previously affected by the fluid.
We find the Green function for the general boundary and initial value problem. In particular, we discuss the initial value problem when the pressure and its space derivatives are nil on the boundary at any time while the pressure in the medium is constant at the initial time and also the problem when on the boundary the pressure is constant while its first and second order derivative are there nil at any time and the initial value of the pressure in the medium is nil.
Keywords: Porous media, Diffusion, Memory, Fractional derivative.

Glossary

k $(s^{-2}m^2)$	ratio of the fluid pressure to the fluid density [see Eq.(2)].
n	fractional order of differentation [see Eq.(3) and Eq.(4)] (dimensionless).
$p(x,t)$ $(kg\ s^{-2}m^{-1})$	fluid pressure.
$q(x,t)$ $(kg\ s^{-1}m^{-2})$	fluid mass flow rate in the porous medium.
r, R	radius of the inner and outer circles, respectively, of the integration path of Eq.(A1) shown in Fig.3.
t (s)	time.
x (m)	distance from the boundary plane.
α $(s\ m^n)$	coefficient of the Darcy's law modified [see Eq.(3)].
αk $(s^{-1}m^{2+n})$	pseudodiffusivity.
β (s)	coefficient of the classic Darcy's law [see Eq.(3)].
$\rho(x,t)$ $(kg\ m^{-3})$	fluid density.
ω, ε	imaginary and real parts in the plane of the integral in Eq.(A1).

1. Introduction.

In monitoring the local values of the gravity at measurement sites located on the surface of the Earth or underground the knowledge of the time variation of the elevation of the site has become specially important. The principal periodic variations of elevation are due to the solid Earth tide

but other important variations are the secular variations due to tectonic activity and those due to the indirect effect of the pore filling of the ground caused by the migration of underground water. The latter phenomenon, at some sites, is due to the tidal variation of the sea level in the near coast; the water load causes a migration of the fluids which is governed by the equations of diffusion.
Some data on the flow of fluids in rocks exhibit properties which may not be interpreted with the classic theory of the propagation of pressure and of fluids in porous media (Bell and Nur, 1978; Roeloffs, 1988) based on the classic Darcy's law which states that the flux is proportional to the pressure gradient.
Memory has been used previously in studying electromagnetic phenomena by (e.g., Graffi, 1936), diffusion (e.g., Kabala and Sposito, 1991; Hu and Cushman, 1994; Indelman and Abramovich, 1994) and biological phenomena (e.g., Volterra, 1930). In this note we shall use space memory represented by fractional order derivative operating on the pressure.
Classic cases of use of time fractional order derivatives as memory operators are those of energy dissipation in anelastic media (e.g., Caputo, 1969; Caputo an Mainardi, 1972; Bagley and Torvik, 1983, 1986; Körnig and Müller, 1989), of dispersion in dielectrics (e.g., Le Mehaute an Crepy, 1983; Jacquelin 1984, 1991; Pelton et al.; 1983; Caputo and Plastino, 1998) of population growth (e.g., Caputo, 1984) and of diffusion in financial (e.g., Mainardi et al., 1998; Caputo, 1998b) and hydrologic phenomena (e.g., Caputo, 1999).
The time derivative of fractional order used in the former cases is also presented and discussed (Caputo, 1969; Lucko and Gorenflo, 1998); in the present note, we shall use in the space domain.
Among other memory models developed in the research on the diffusion of fluids in rocks must be considered the use of the fractional derivative introduced in the Darcy's law operating on the flow as well as on the pressure gradient which imply a filtering of the pressure gradient without singularities (Caputo, 1998a).
The time fractional order derivative of the pressure represents the local variations and is particularly valid when considering local phenomena. In an infinite medium is more appropriate to introduce the space fractional order derivative instead of the time fractional derivative order to represent the effect of the medium previously affected by the fluid. Therefore, the flow is not directly related to the instantaneous pressure gradient in the measurement site but to the spatial fractional derivative i.e. to the pressure gradient investigated in the path from the starting point to the measurement site.
In this note we shall devote our attention particularly to find the Green function of the initial value and of the boundary value problems in a semi-infinite medium bounded by plane.
We will first find the general solution of the initial and boundary value problem; namely when the pressure is initially constant in the medium and nil with its first and second order derivatives at all times on the boundary.
Then we discuss separately the boundary value problem. Specifically we discuss the case when the pressure and its first and second order spatial derivatives are assigned on the boundary while, in the medium, is assigned the initial value of the pressure.

2. The model.

In order to find general solution of the problem, that is the pressure distribution in the porous media affected by space memory we begin setting the constitutive equations. The first equation is the classic continuity equation between the time variation of the density and the divergence of the flux

$$q_x + \rho_t = 0 \tag{1}$$

Another constitutive equation is that relating the pressure to the variation of the density from its undisturbed condition

$$p = k\rho \tag{2}$$

Successively, to take into account the observed deviations of the flow from those implied by the classic diffusion equation, we introduce, as follows, a space memory formalism in Darcy's law consistent with the flow dependence on the history of the pressure gradient.

$$q = \alpha \frac{\partial^{1+n}}{\partial x^{1+n}} p + \beta \frac{\partial}{\partial x} p \tag{3}$$

with $0 \leq n < 1$, where the definition of derivative of fractional order *1+n* is (Caputo, 1969)

$$\frac{\partial^{1+n}}{\partial x^{1+n}} p(x,t) = \left(\frac{1}{\Gamma(1-n)} \right) \int_0^x (x-v)^{-n} \left(\frac{\partial^2 p(x,v)}{\partial v^2} \right) dv$$

In the constitutive equation (3), for sake of generality (Caputo, 1999), the effect of the memory affects only the part of the pressure p with factor α, while the term with factor β represents the part of the pressure gradient not affected by the memory and behaving as in the classic Darcy's law.
Replacing p/k from Eq.(2) in Eq.(1) and taking into account the derivative respect to x variable in Eq.(2) we obtain a single equation in p

$$-\frac{p_t}{k} = \alpha \frac{\partial^{2+n}}{\partial x^{2+n}} p + \beta \frac{\partial^2}{\partial x^2} p \tag{4}$$

In order to solve Eq.(4) we take its Laplace Transform (LT) respect to x variable using the LT theorem (Caputo, 1969):

$$LT\left(\frac{\partial^{1-\alpha}}{\partial x^{1-\alpha}} p(x,t) \right) = -u^{\alpha} p(x,0) + u^{1-\alpha} LT\left(p(x,t)\right)$$

where u is the LT variable and obtain the equation

$$P_t + k\left[\alpha u^{2+n} + \beta u^2\right] P = \alpha k \left[u^{1+n} p(0,t) + u^n p_x(0,t) + u^{n-1} p_{xx}(0,t)\right]$$
$$+\beta k \left[u p(0,t) + p_x(0,t)\right] \tag{5}$$

where $P(u,t) = LT_{x,u} p(x,t)$. Proceeding to the solution, now we take place the LT of Eq.(5) respect to t variable and obtain

$$V(u,v) = \frac{P(u,0)}{v + k\left[\alpha u^{2+n} + \beta u^2\right]}$$

$$+\frac{\alpha k}{v + k\left[\alpha u^{2+n} + \beta u^2\right]} LT_{t,v}\left[u^{1+n} p(0,t) + u^n p_x(0,t) + u^{n-1} p_{xx}(0,t)\right]$$
$$+\frac{\beta k}{v + k\left[\alpha u^{2+n} + \beta u^2\right]} LT_{t,v}\left[u p(0,t) + p_x(0,t)\right] \tag{6}$$

where $V(u,v) = LT_{t,v} P(u,t)$ and $P(u,0) = LT_{x,u}(p(x,0))$. The solution p is then be obtained by inverting both LT. The inverse $LT_{t,v}$ of Eq.(6) is

$$P(u,t) = LT_{t,v}^{-1} V(u,v) = P(u,0) e^{-tku^2[\alpha u^n + \beta]}$$

$$+\alpha k e^{-tku^2[\alpha u^n + \beta]} *_t \left[u^{1+n} p(0,t) + u^n p_x(0,t) + u^{n-1} p_{xx}(0,t)\right]$$
$$+\beta k e^{-tku^2[\alpha u^n + \beta]} *_t \left[u p(0,t) + p_x(0,t)\right] \tag{7}$$

The inverse $LT_{x,u}$ of Eq.(7) gives finally

$$p(x,t) = LT_{x,u}^{-1}P(u,t) = p(x,0) *_x LT_{x,u}^{-1}\left(e^{-tku^2[\alpha u^n+\beta]}\right)$$
$$+\alpha k\left[p(0,t) *_t LT_{x,u}^{-1}\left(u^{1+n}e^{-tku^2[\alpha u^n+\beta]}\right) + p_x(0,t) *_t LT_{x,u}^{-1}\left(u^n e^{-tku^2[\alpha u^n+\beta]}\right)\right.$$
$$+p_{xx}(0,t) *_t LT_{x,u}^{-1}\left(u^{n-1}e^{-tku^2[\alpha u^n+\beta]}\right)$$
$$\left.+\beta k\left[p(0,t) *_t LT_{x,u}^{-1}\left(ue^{-tku^2[\alpha u^n+\beta]}\right) + p_x(0,t) *_t LT_{x,u}^{-1}\left(e^{-tku^2[\alpha u^n+\beta]}\right)\right]\right] \quad (8)$$

which gives the formal general solution of the problem and includes the boundary conditions *p(0,t), p_x(0,t), p_{xx}(0,t)* in terms of the 2^{nd}, 3^{rd}, 4^{th} lines and also the initial condition *p(x,0)* in terms of 1^{st} line.

2.1. The explicit pressure solution.

Using Eq.(A3) of the appendix with γ= *0, 1+n, n, n-1, 1* and substituting in Eq.(8) we obtain the solution reduced to simple integrations

$$p(x,t) = p(x,0) *_x \left(-\frac{1}{\pi}\int_0^\infty e^{-rx}e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}tsin\left(n\pi\right)\right)dr\right)$$
$$+\alpha k\left[p(0,t) *_t \left(-\frac{1}{\pi}\int_0^\infty e^{-rx}r^{1+n}e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}t\left(sin\left(n\pi\right)-(1+n)\pi\right)\right)dr\right)\right.$$
$$+p_x(0,t) *_t \left(-\frac{1}{\pi}\int_0^\infty e^{-rx}r^{n}e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}t\left(sin\left(n\pi\right)-n\pi\right)\right)dr\right)$$
$$\left.+p_{xx}(0,t) *_t \left(-\frac{1}{\pi}\int_0^\infty e^{-rx}r^{n-1}e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}t\left(sin\left(n\pi\right)-(n-1)\pi\right)\right)dr\right)\right]$$
$$+\beta k\left[p(0,t) *_t \left(-\frac{1}{\pi}\int_0^\infty e^{-rx}re^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}t\left(sin\left(n\pi\right)-\pi\right)\right)dr\right)\right.$$
$$\left.+p_x(0,t) *_t \left(-\frac{1}{\pi}\int_0^\infty e^{-rx}e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}tsin\left(n\pi\right)\right)dr\right)\right] \quad (9)$$

where the values of the integrals depend on the variables x and t. We note that in Eq.(9) we have two types of convolution, one relative to the time variable and one relative to the space variable. We note again that the first term in Eq.(9) takes into account the initial values in the medium while the other terms take into account the boundary values. The computation of the initial value term implies the convolution relative to the space variable only while the computation of the terms relative to the boundary values imply convolutions relative to the time variable only. The boundary values consist of the boundary values of the function and of its first and second order space derivatives.

2.2. The initial value problem.

We consider nil the pressure and its derivative respect to the x variable on the boundary for any t while the pressure in the medium has initial (t=*0*) constant value $C \neq 0$. The solution is then readily obtained from Eq.(9) considering only the first integral

$$p(x,t) = \frac{C}{\pi}\int_0^\infty \frac{1-e^{-rx}}{r}e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}sin\left(\alpha kr^{2+n}tsin\left(n\pi\right)\right)dr \quad (10)$$

In order to tentatively explore the effect of the space memory we will first assume β=*0*, which

excludes the portion of p following the classic Darcy's law in Eq.(3). The formula is then not difficult to compute for several values of x measuring the amplitude of the effect in units of C/π and measuring t, in all case considered, in units of αk (pseudodiffusivity). We considered for the curves shown in Fig.1 the values of *n=0.1, 0.2, 0.3, 0.4* which are sufficient to describe the dependence of the memory effect on the order of fractional derivation. The Fig.1 shows that the pressure at any point in the medium decreases during the time and the decrease diminishes with increasing of n. In the figure is also seen that at any given time the pressure increases with increasing distance from the boundary. The distances considered are in meters and cover a significant range of practical interest. It is easy to extend the range to distances of geodetic interest; however, we see that the effect of the memory is significant only in relatively short distances.

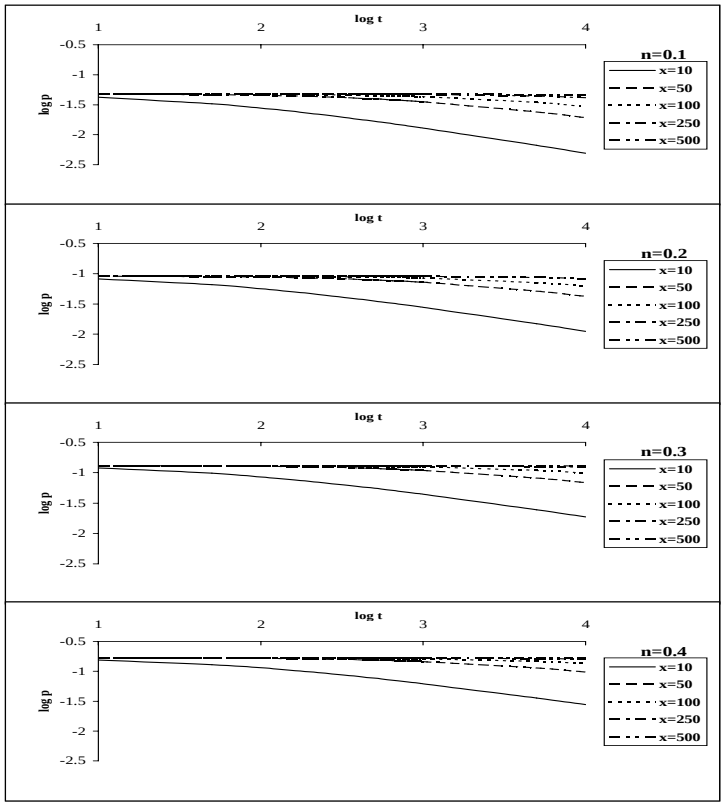


Fig.1. The initial value problem curves are related to the values of fractional derivative order *n=0.1, 0.2, 0.3, 0.4* and the distances from the boundary *x=10, 50, 100, 250, 500* meters. The amplitude is measured in units of C/π and the time in units of αk (pseudodiffusivity).

2.3. The boundary value problem.

In this case we consider nil the pressure for *t=0* for any x in the medium while the pressure on the boundary (*x=0*) is constant with value $C \neq 0$ and its derivatives respect to the x variable are nil for *x=0*. The solution is then readily obtained from Eq.(9) considering the second and fifth integral

$$p(x,t) = \alpha \left(\frac{C}{\pi} \int_0^\infty e^{-rx} \frac{1 - e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}}{r^{1-n}\left(\alpha r^n cos(n\pi)+\beta\right)} sin\left(\alpha k r^{2+n} t \left(sin\left(n\pi\right) - \left(1+n\right)\pi\right)\right) dr \right)$$

$$+\beta\left(\frac{C}{\pi}\int_0^{\infty} e^{-rx}\frac{1-e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)}}{r\left(\alpha r^n cos(n\pi)+\beta\right)} sin\left(\alpha k r^{2+n} t\left(sin\left(n\pi\right)-\pi\right)\right) dr\right) \tag{11}$$

Also in this case we exclude the portion of p following the classic Darcy's law in Eq.(3) and assume β=0. The amplitude is measured in units of $\alpha C/\pi$ and the time, in all case considered, in units of αk (pseudodiffusivity). The curves shown in Fig.2 are relatives to the values of n=0.1, 0.2, 0.3, 0.4. The Fig.2 shows that the pressure at any point at the boundary increases during the time and the increase diminishes with increasing of n. Besides, at any given time the pressure increases with decreasing distance from the boundary. The distances considered are in meters.

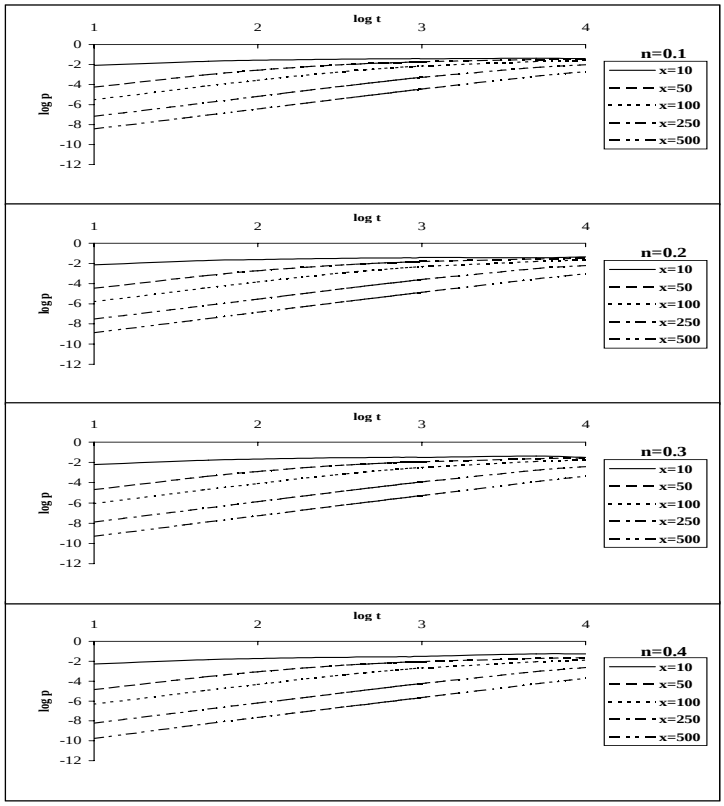


Fig.2. The boundary value problem curves are related to the values of fractional derivative order n=0.1, 0.2, 0.3, 0.4 and the distances from the boundary x=10, 50, 100, 250, 500 meters. The amplitude is measured in units of $\alpha C/\pi$ and the time in units of αk (pseudodiffusivity).

3. Conclusion.

The fluctuations in water level caused by Earth tides are not in complete agreement with the phases of the tides emphasizing that a memory mechanism could be the cause of this phenomenon. Particularly, the migration processes of the underground water near the coast are affected to this difference of phases. The Darcy's law modified by space derivative fractional order presented in this note may be a useful tool to describe the memory mechanism and to interpret part of the phenomenology also carachterized by anelasticity, inhomogeneity, anisotropy and of the medium.

Besides, we hope that the model of diffusion with memory in the space domain studied here be more usefull for applications to the study of the variations of the gravity field than that with memory in the time domain (Caputo, 1998a). Indeed, the latter seems more appropiate for

diffusion in layers of limited thickness, such as membranes or thin layers, while the former for diffusion in layers very thick such as in the case of water diffusion in thick layers of the Earth's crust which is of interest when studying the time variations of the gravity field.

Appendix A

We calculate the $LT_{x,u}^{-1}\left(u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}\right)$ of Eq.(8), where γ is real variable, integrating along the closed path shown in Fig.3 and taking the radius of the inner circle r to zero and that of the outer circle R to infinity.

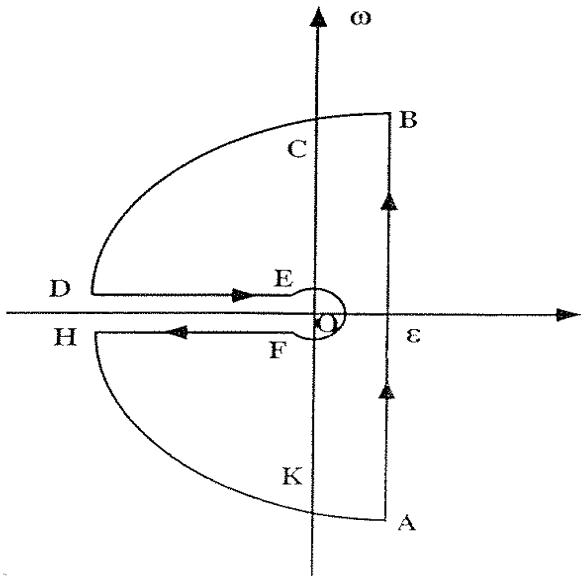


Fig.3. Path of the integration of Eq.(A1) in the complex plane. The path begins in A, follows the direction of the arrows, and return to A.

Inside the integration path there are no poles of the $LT_{x,u}^{-1}\left(u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}\right)$ because this has no poles in the negative complex plane of u and then the integral is therefore nil because the residuals are nil. The integrals along BC, CD, HK, KA are nil when the outer radius R of the path is infinite; the integral on EF is nil when the inner radius r of the path is nil (Caputo, 1969) and finally we may write

$$LT_{x,u}^{-1}\left(u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}\right)=\lim_{\omega\to\infty}\frac{1}{2\pi i}\left\{\int_{\varepsilon-i\omega}^{\varepsilon+i\omega}e^{ux}u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}du\right.$$

$$\left.+\int_{D}^{E}e^{ux}u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}du+\int_{F}^{H}e^{ux}u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}du\right\} \tag{A1}$$

where ε and ω are the real and imaginary parts in the plane of integration shown in Fig.3.
We assume
$u = r\ e^{i\vartheta} = r\ (cos\ \vartheta + i\ sin\ \vartheta)$
$u^{\gamma} = r\ e^{i\gamma\vartheta} = r^{\gamma}\ (cos\ \gamma\theta + i\ sin\ \gamma\vartheta)$
on $\vartheta = \pm\pi$, u = -r, du = - dr, where r is the modulus of u and noting that the integration on DE: $\vartheta = \pi$ and on FH: ϑ = -π. We may write Eq.(10) as

$$LT_{x,u}^{-1}\left(u^{\gamma}e^{-tku^2(\alpha u^n+\beta)}\right)=\frac{1}{2\pi i}\left[-\int_{\infty}^{0}e^{-rx}r^{\gamma}e^{i\gamma\pi}e^{-tkr^2(\alpha r^n(cos(n\pi)+isin(n\pi))+\beta)}dr\right.$$

$$- \int_0^\infty e^{-rx} r^\gamma e^{-i\gamma\pi} e^{-tkr^2(\alpha r^n (cos(n\pi) - isin(n\pi)) + \beta)} dr \Big] \tag{A2}$$

which may be simplified to

$$LT_{x,u}^{-1}\left(u^\gamma e^{-tku^2(\alpha u^n+\beta)}\right) = -\frac{1}{\pi}\int_0^\infty e^{-rx} r^\gamma e^{-tkr^2(\alpha r^n cos(n\pi)+\beta)} sin\left(\alpha k r^{2+n} t\left(sin\left(n\pi\right) - \gamma\pi\right)\right) dr \tag{A3}$$

Formula (A3) will be used to obtain the $LT_{x,u}^{-1} P(u,t)$.

References

Bagley, R.L., Torvik, P.J., *A theoretical basis for the application of fractional calculus to viscoelasticity*, Journal of Rheology, 27, 3, 201-210, 1983.

Bagley, R.L., Torvik, P.J., *On the fractional calculus model of viscoelastic behaviour*, Journal of Rheology, 30, 1, 133-155, 1986.

Bell, M.L., Nur, A., *Strenght changes due to reservoir-induced pore pressure and stresses and application to Lake Oroville*, Journal of Geophysical Research, 83, 4469-4483, 1978.

Caputo, M., *Elasticitá e dissipazione (Elasticity and anelastic dissipation)*, pp.150, Zanichelli Publisher, Bologna, 1969.

Caputo, M., Mainardi, F., *Linear models of dissipation in anaelsatic solids*, Rivista del Nuovo Cimento, 1, 2, 161-198, 1971.

Caputo, M., *A linear model for population growth in a limited habitat*, in La biogeografia delle isole, Atti Convegni Lincei, 62, 219-229, 1984.

Caputo, M., *3-dimensional physically consistent diffusion in anisotropic media with memory*, Rend.Mat. Acc. Lincei, s.9, v.9, 131-143, 1998a.

Caputo, M., *Evolutionary equilibrium equation between demand and offer*, in Atti Accad. Scienze Ferrara, 75, 167-198, 1998b.

Caputo, M., *Diffusion of fluids in porous media with memory*, Geothermics, 28, 1, 113-130, 1999.

Caputo, M., Plastino, W., *Rigorous time domain responses of polarizable media II*, Annali di Geofisica, 41, 3, 399-407, 1998.

Graffi, D., *Sopra alcuni fenomeni ereditari dell'elettrologia: I, II*, Rendiconti Istituto Scienze Lettere ed Arti, 2, 69, 128-181, 1936.

Hu, X., Cushman, H., *Non equilibrium statistical mechanical derivation of a non local Darcy's law for unsatured/saturated flow*, Stochastic Hydrology and Hydraulics, 8, 109-116, 1994.

Indelman, P., Abramovich, B., *Non-local properties of non-uniform averaged flows in heterogeneous media*, Water Resources Research, 30, 12, 3385-3393, 1994.

Jacquelin, J., *Use of fractional derivates to express the properties of energy storage phenomena in electrical networks*, Technical Report, Laboratoires de Marcoussis, 1984.

Jacquelin, J., *Synthése de circuits électriques equivalents a partir de measures d'impédences complexes*, 5éme Forum sur les impedences électrochimiques, 287-295, 1991.

Kabala, Z.J., Sposito, G., *A stochastic model of reactive solute transport with time-varying velocity in a heterogeneous acquifer*, Water Resorces Researches, 27, 3, 341-350, 1991.

Körnig, H., Müller, G., *Rheological models and interpretation of postglacial uplift*, Geophys. J.R. Astr. Soc., 98, 243-253, 1989.

Le Mehaute, A., Crépy, G., *Introduction to transfer motion in fractal media: the geometry of kinetics,* Solid State Ionic, 9&10, 17-30, 1983.

Lucko, Y., Gorenflo, R., *The inital value problem for some fractional differential equations with the Caputo derivatives*, Fachbereich Mathematik und Informatik, Serie A:Mathematik, Preprint No.A8-98, 1-23, 1998.

Mainardi, F., Paradisi, P., Gorenflo, R., *Probability Distributions Generated by the Fractional Diffusion Equations*, in Econophysics, Kluver Academic Publisher, 1998.

Pelton, W.H., Sill, W.R., Smith, B.D., *Interpretation of complex resistivity and dielectric data, Part 1*, Geophysical Transactions, 29, 4, 297-330,1983.

Roeloffs, E.A, *Fault stability changes induced beneath a reservoir with cyclic variations in water level*, Journal of Geophysical Research, 93, 2107-2124, 1988.

Volterra, V., *Leçons sur la théorie mathématique de la lutte pour la vie*, (Edited by M. Vrelot) Paris, 1930.

Procrustes Analysis and Geodetic Sciences

Fabio Crosilla

Abstract

Procrustes analysis is a well known technique to provide least squares matching of two or more factor loading matrices or for the multidimensional rotation and scaling of different matrix configurations. Applied at first as a useful tool in factor analysis, today it has become a popular method of shape analysis (Goodall 1991, Dryden and Mardia 1998).
This paper reviews the development of the most significant algorithms used in this particular field. Starting from the solution of the classical "orthogonal procrustes problem"(Schönemann 1966) a first extension including a scaling factor and a central dilation will be presented (Schönemann and Carroll 1970).
The solution of the "generalized orthogonal procrustes problem" to sets of more than two matrices will be then reported (Gower 1975, Ten Berge 1977).
Furthermore, "weighted procrustes analysis" will be considered for the cases in which the residuals of a matching procedure are differently weighted across columns (Lissitz et al. 1976) or across rows (Koschat and Swayne 1991) of a matrix configuration .
Finally, some possible applications of procrustes methods for point coordinates transformations in geodesy and photogrammetry will be mentioned. All this makes it possible to emphasize the capabilities of the method proposed.

1. Unweighted procrustes analysis

The so-called "orthogonal procrustes problem" (Schönemann, 1966) is the least squares problem that makes it possible to transform a given matrix A into a given matrix B by an orthogonal transformation matrix T in such a way to minimize the sum of squares of the residual matrix E = AT - B , that is

$$tr\ (E'\ E) = min,$$

under the orthogonal condition for matrix T, that is T'T = TT' = I. In order to satisfy the minimum condition, a Lagrangean function, defined as

$$F = tr\ (E'\ E) + tr\ [\ L\ (T'T - I)\ ],$$

where L is a matrix of Lagrangean multipliers, must be minimized, by setting the partial derivative of F with respect to T equal to zero, that is

$$\partial F/\ \partial T = (\ A'A + A'A)\ T - 2\ A'B + T\ (L + L') = 0 \qquad (1.1)$$

Putting A'A = P, A' B = S and (L+ L')/2 = Q, one can note that matrices P and Q are symmetric, so that multiplying (1.1) on the left by T', it results that

$$Q = T'S - T'\ P\ T = Q'.$$

Now, since T' PT is symmetric, T'S must be symmetric too; In this way

$$T'S = S'T. \qquad (1.2)$$

Multiplying (1.2) on the left by T

$$S = T\ S'T$$

and on the right by T'

$$T'\ S\ T' = S'$$

we finally have

$$SS' = TS'\ ST'. \qquad (1.3)$$

Matrices S'S and SS' are symmetric matrices, both of which can be transformed in a diagonal form by orthonormal matrices and both of which have the same eigenvalues, according to Schönemann et al. (1965). Equation (1.3) can be finally written as

$$W\ D_S\ W' = T\ V\ D_S\ V'T'$$

so that W= TV or $T = WV'$

Matrix T is the orthogonal matrix which satisfies the least squares principle defined by tr (E'E) = min. A first generalization to the Schönemann (1966) orthogonal procrustes problem was given by Schönemann and Carroll (1970) when a least squares method for fitting a given matrix A to another given matrix B under choice of an unknown rotation, an unknown translation and an unknown central dilation was presented. The chosen model is the following

$$B = c\ AT + J\ g' + E \qquad (1.4)$$

where J' = (1111...1), A, B and T have the same meaning as before, g is a vector of translation components, c is a scalar of central dilation and E is a matrix of residuals. To obtain the least squares solution for model (1.4), the Lagrangean function, written as

$$F = tr\ (E'\ E) + tr\ [\ L\ (T'T - I)\ ]$$

where

$$tr\ (E'\ E) = tr\ B'B + c^2\ tr\ T'A'\ AT + p\ g'g - 2\ c\ tr\ B'\ AT - 2tr\ B'J\ g' + 2\ c\ tr\ T'A'Jg'$$

and p is the number of rows of matrices A and B, must be differentiated with respect to the unknowns T, g and c. Once the derivatives are set at zero, it results that

$$\partial F/\ \partial T = 2\ c^2\ A'A\ T - 2\ c\ A'B + 2\ c\ A'\ J\ g' + TQ = 0 \qquad (1.5)$$

where Q = (L + L')/2 = Q' is an unknown symmetric matrix,

$$\partial F/\ \partial g = 2p\ g - 2\ B'J + 2c\ T'A'J = 0 \qquad (1.6)$$

$$\partial F/\ \partial c = 2\ c\ tr\ T'A'AT - 2\ tr\ B'AT + 2\ tr\ T'A'J\ g' = 0 \qquad (1.7)$$

From equation (1.6) it follows that

$$g = (B- c\ AT)'\ J/p \qquad (1.8)$$

and from (1.5), according to some considerations already done about symmetry of matrices T'A'AT and Q, it results that

$$T'A'B - T'A'Jg' = \text{symmetric matrix}$$

which can also be written, according to (1.8) ,as

$$T'A'B - T'A'(JJ'/p)(B - cAT) = \text{symm.}$$

so that

$$T'A'(I - JJ'/p)\, B = \text{symm.} \qquad (1.9)$$

because $cT'A'(JJ'/p)AT$ is symmetric. Equation (1.9) does not consider g and c and a solution for T can be found following the already mentioned procedure applied for the "orthogonal procrustes" problem.

Letting $S^* = A'(I - JJ'/p)\, B,$

matrix products $S^{*'}S^*$ and $S^*S^{*'}$ will be computed, and from their singular value decomposition

$$S^{*'}S^* = V D_S V'$$
$$S^*S^{*'} = W D_S W'$$

the solution for $T = WV'$ can be found. Substituting (1.8) in (1.7), equation (1.7) can be solved for the contraction factor c

$$c = \text{tr } T'A'(I - JJ'/p)\, B / \text{tr } A'(I - JJ'/p)\, A \qquad (1.10)$$

Inserting the result for c in equation (1.8) the value of the estimated g can be finally obtained.

One interesting thing to note is that the matrix of best fit $\hat{B}$ and that of the residuals $E = B - \hat{B}$ given respectively by

$$\hat{B} = cAT + Jg' = c\,AT + (JJ'/p)(B - c\,AT) = (JJ'/p)B + c(I - JJ'/p)\,AT$$

and by

$$E = B - \hat{B} = (I - JJ'/p)(B - c\,AT)$$

do not involve g. The fit is the same independent of the origin of both data set configurations. For this reason, from the practical point of view in terms of computation, the first recommended step to take consists of the calculation of the two column centered matrices

$$A^* = A - J\,k' \quad \text{and}$$
$$B^* = B - J\,h'$$

where $k = A'J/p$ and $h = B'J/p$.

Afterward one has to enter a standard orthogonal procrustes algorithm to obtain the transformation matrix T and the matrix A^*T. The final computations are related to the scalar c, given by

$$c = tr [(T' A^{*}') B^{*}] / tr A^{*}'A^{*}$$

and the matrix of best fit

$$\hat{B} = c (A^{*}T) + J h'.$$

To measure the least squares fit, the criterion function itself, that is

$$e= tr E'E = tr B'(I - JJ')/p B - (tr T'A' (I - JJ')/p B)^2 / tr A'(I- JJ')/p A$$

is commonly adopted in the literature. This is however not symmetric in the sense that a fit of A to B generates different values for the elements of e than a fit of B to A. In order to satisfy a symmetry condition, Lingoes and Schönemann (1974) defined at first a new symmetric measure of fit given by

$$e_S = e\ u^{-1/2}$$

where $$u = tr B' (I - JJ'/p) B/ tr A' (I - JJ'/p) A$$

which satisfy e_S (A,B) = e_S (B,A),and does not depend on the order of matching and u is a constant. As reported by Lingoes and Schönemann (1974) e_S depends upon the norm of the target matrix B. Such dependency must be avoided when the comparison of the fits for different target matrices is wanted. To achieve the desired scale invariance, the following measure was finally proposed

$$S = e / tr B'(I - JJ'/p) B$$

In this case the measure remains invariant for different orders of fitting and is 0 - 1 bounded.

2. Generalized Procrustes analysis

A generalization of the classical procrustes analysis was given by Gower (1975) and Ten Berge (1977) where the problem of best fitting more than two matrices was taken into account. Instead of considering the matching of all possible independent matrix pairs, the procrustes analysis is generalized in such a way that m matrices are simultaneously subjected to similarity transformations until a proper fit criterion is reached. The criterion adopted consists in the minimization of the sum of square distances between each point of the m ones $P_j^{(i)}$ (i = 1...m) belonging to the same cluster and their centroid G_j (j= 1...n), summed for all n clusters. The problem of rotating, translating and scaling m matrices ($m \geq 2$) toward a best least squares fit consists in finding orthonormal matrices T_i (i= 1...m), translation vectors t_i, and scale factors c_i, for which the function

$$S = tr \sum_{i<j}^{m} [(c_i A_i T_i + t_i) - (c_j A_j T_j + t_j)]' [(c_i A_i T_i + t_i) - (c_j A_j T_j + t_j)] \qquad (2.1)$$

is minimized (Gower 1975).
The Gower's method starts with an initial centering of the matrices, so that all column sums are zero and a successive scaling of each A_i by a general $w^{1/2}$ so that $\sum w\ tr\ A'_i A_i = m$.

In the following, with the symbol A_i, column centered and scaled matrices will be considered. Rotation matrices and scaling constants are adjusted in sequence. The solution of the rotation problem consists of finding orthonormal matrices T_i for which the function

$$f(T_1...T_m) = \sum_{i<j} \text{tr} (A_iT_i - A_jT_j)' (A_iT_i - A_jT_j)$$

is minimized or equivalently the function

$$g(T_1...T_m) = \sum_{i<j} \text{tr } T_i' A_i' A_jT_j$$

is maximized. Ten Berge (1977) suggests the following iterative procedure to solve the problem:

Step 1: Rotate A_1 to $\sum_{j=2}^{m} A_j$, thus yielding $A_1T_1^{(1)}$

Step 2: Rotate A_2 to $A_1T_1^{(1)} + \sum_{j=3}^{m} A_j$, thus yielding $A_2T_2^{(1)}$

Step m: Rotate A_m to $\sum_{j=1}^{m-1} A_jT_j^{(1)}$, thus yielding $A_mT_m^{(1)}$

Step m+1 : Rotate $A_1T_1^{(1)}$ to $\sum_{j=2}^{m} A_jT_j^{(1)}$, thus yielding $A_1T_1^{(2)}$

The procedure terminates when the combined effect of some steps does not raise g above a certain threshold value. The procedure will converge and g will be maximized if, and only if, all the matrix products

$$T_i'A_i'A_{i+1}$$

are symmetric and positive semi-definite . Proof of the theorem is reported in Ten Berge 1977, page 269. Let us now consider the problem of computation of the scaling constants c_i. Let it be $\sum \text{tr } A_i'A_i = m$ for which scaling constants $c_1....c_m$ are wanted to maximize

$$h(c_1....c_m) = \Sigma_{i<j} c_ic_j \text{ tr } A_i'A_j \qquad (2.2)$$

with the constraint

$$\Sigma c_i^2 \text{ tr } A_i'A_i = \quad \text{tr } A_i'A_i = m \qquad (2.3)$$

that satisfies the condition of maximizing (2.2) and minimizing the least squares function

$$\Sigma_{i<j} \text{tr } (c_iA_i - c_jA_j)' (c_iA_i - c_jA_j).$$

If we consider the particular rescaled matrix A_i^* as

$$A_i^* = (\text{tr } A'_i A_i)^{-1/2} A_i$$

it follows that tr $A_i^{*\prime} A_i^* = 1$, for i= 1...m. We look for m scalars d_i (i= 1..m), able to maximize

$$h^*(d_1 \ldots d_m) = \sum_{i<j} d_i d_j \operatorname{tr} A_i^{*\prime} A_j^*, \tag{2.4}$$

with the constraint

$$\sum d_i^2 \operatorname{tr} A_i^{*\prime} A_i^* = \sum d_i^2 = \sum c_i^2 \operatorname{tr} A_i' A_i = \sum \operatorname{tr} A_i' A_i = m \tag{2.5}$$

Now let the mxm matrix Y be written as

$$Y = \begin{bmatrix} \operatorname{tr} A_1' A_1 & \operatorname{tr} A_1' A_2 & \ldots. & \operatorname{tr} A_1' A_m \\ \ldots. & \ldots. & \ldots. & \ldots. \\ \operatorname{tr} A_m' A_1 & \operatorname{tr} A_m' A_2 & \ldots. & \operatorname{tr} A_m' A_m \end{bmatrix}$$

and $Y_D = \operatorname{diag}(Y)$, and $F = Y_D^{-1/2} Y Y_D^{-1/2}$.

Putting d_i in a vector d, than condition (2.4) can be written as

$$h^*(d) = 1/2\, d' (F - I) d \tag{2.6}$$

which must be maximized subject to

$$d'd = m.$$

Considering the singular value decomposition of matrix F, (F= PLP'), and letting z = P' d it follows that

$$h^*(d) = 1/2\, d' (F - I) d = 1/2\, d' (PLP' - I) d =$$

$$= 1/2\, (z'Lz - m) \leq 1/2\, (l_1\, z'z - m) = 1/2\, m\, (l_1 - 1)$$

where l_1 is the greatest eigenvalue of matrix L. Condition (2.6) is maximized when $d = m^{1/2} p_1$. In this case

$$h^*(m^{1/2} p_1) = 1/2\, m\, p_1' (F-I) p_1 = 1/2\, m\, p_1' (PLP' - I)\, p_1 =$$

$$= 1/2\, m (\, e_1' L\, e_1 - p_1' p_1) = 1/2\, m\, (l_1 - 1).$$

From this result and the constraint (2.5) it follows that

$$c_i = (\, m / \operatorname{tr} A_i' A_i)^{1/2}\, p_i \qquad (\text{Ten Berge (1977)}).$$

Up to now the so-called unweighted least squares solutions have been taken into account. This is appropriate when the residuals have equal variance and hence should be weighted equally. If one wishes to weight the residuals differently, a weighted least squares criterion is more appropriate.

3. Weighted procrustes analysis

Two different ways of weighting the residuals are usually applied in the procrustean literature: across columns or across rows. The corresponding least squares criteria are then:

$$\text{tr}(AT - B)' \, W_n^2 \, (AT - B) \tag{3.1}$$

$$\text{tr}(AT - B) \, W_p^2 \, (AT - B)' \tag{3.2}$$

where W_n and W_p are diagonal weight matrices containing information about the dispersion of the residuals.
To find an orthogonal matrix T minimizing (3.1) is easy, since it is equivalent to minimize (3.1) replacing B by W_n B and A by W_n A, respectively (Lissitz et al. 1976). The second problem is more difficult to solve. A very interesting algorithm was introduced by Koschat and Swayne (1991). The algorithm is based on the possibility of considering the general problem like a specific one for which it is simple to find a valid solution. If matrix A is characterized by orthogonal column vectors characterized by the same euclidean lenght l, the problem is to find an orthogonal matrix T that minimizes (3.2), that is

$$\begin{aligned} \text{tr} (B - AT) W_p^2 (B-AT)' &= \text{tr}(B \, W_p^2 \, B') - 2\text{tr}(ATW_p^2B') + \text{tr} \, (W_p^2T' \, A'AT) \\ &= \text{tr}(B \, W_p^2 \, B') - 2\text{tr} \, (W_p^2B'AT) + l^2 \, \text{tr} \, (W_p^2) \end{aligned}$$

or equivalently, that maximizes

$$\text{tr} \, (W_p^2 \, B' \, AT),$$

which is very similar to the solution of the classical unweighted least-squares problem, previously reported. Writing the singular value decomposition of $W_p^2B'A$ as U L V' the solution can be found as

$$T = VU' \tag{3.3}$$

In case the column vectors of A are not orthogonal with respect to each other, a connection with the case reported above can be made in the following way. Once the nxp matrices A and B are given, and the (n+p)xp matrices A* and B* are defined as

$$A^* = \begin{bmatrix} A \\ A^0 \end{bmatrix} \qquad\qquad B^* = \begin{bmatrix} B \\ B^0 \end{bmatrix}$$

one has to fix matrix A° so that

$$A^{*\prime} \, A^* = l^2 \, I_p, \text{ for some } l \, . \tag{3.4}$$

Of course the matrix A* satisfies the condition (3.4) if and only if A° satisfies

$$A^{\circ\prime}A^\circ = l^2 \, I_p - A'A.$$

In order to satisfy a positive-definite right hand-side of this equation, a sufficiently large l must be fixed. In this case an infinite number of solutions for A° are possible. Koschat and Swayne (1991) suggested the use of the Cholesky decomposition

$$A°' = \text{chol}\,(l^2 I_p - A'A)$$

where l^2 is set equal 1,1 times the largest eigenvalue of A'A. The pxp matrix B° can be chosen arbitrarily. The algorithm reported by Koschat and Swayne (1991), allows definition of a sequence (T_i, B_i*), and is of the form: for i = 2..., define the (n+p)xp matrix B_i* as

$$B_i^* = \begin{bmatrix} B \\ A^0 T_{i-1} \end{bmatrix}$$

and the corresponding weighted procrustes residuals

$$\text{tr}\,(B_i^* - A^*T)\, W_p^2\, (B_i^* - A^*T)' \qquad (3.5)$$

As A* is characterized by orthogonal columns of equal lenght, matrix T_i (i= 1,2,3…) reported in formula (3.5), can be computed at each iteration by formula (3.3). The solution method of this problem can be successfully applied in image analysis and in photogrammetry where, if matrices A and B contain the centred coordinates of corresponding control points on two different images or in two different reference systems of coordinate, it is desired to test whether the object described by A can be transformed into the object described by B through rotation and dilation along specified directions.
The problem may be formulated as a regression problem

$$B = AX + E$$

under the constraint

$$X = TK$$

where T'T = TT' = I and K is diagonal with positive values. The solution can be obtained by minimizing

$$\text{tr}\,(B - ATK)(B - ATK)' = \text{tr}\,(BK^{-1} - AT)K^2\,(BK^{-1} - AT)' \qquad (3.6)$$

If K is known this problem corresponds to the problem of minimizing (3.2). The algorithm just described can be successfully applied to find T. For a given T, the diagonal values in K are given as (Koschat and Swayne, 1991)

$$K_{ii} = \frac{(B')_i\,(AT)_i}{(AT)'_i\,(AT)_i}$$

where $(B')_i$ and $(AT)_i$ denote the i-th column vectors of the matrices B and AT. If K and T are unknown, an iterative algorithm can be used. This permits determination of a sequence (K_i, T_i) whose limit is the solution (K,T). Koschat and Swayne (1991) recommend to start by choosing K_1 to be the identity matrix.

4. Procrustes analysis and geodetic applications

Geodetic data analysis often requires the application of rescaling, rotation and translation procedures of different data matrix configurations.

It seems therefore very strange that up to now procrustes analysis has not been widely applied in the geodetic literature. With this technique linearization problems of non linear equations systems and iterative procedures of computation could be avoided, in general, with significant computational time saving and less analytical difficulties.

To the author's knowledge only a single geodetic application of the procrustes technique was done in the 1980s for the construction of an ideal variance-covariance matrix (criterion matrix) of a control net point coordinates.

In that case the solution of the "classical procrustes problem" by Schönemann (1966) was applied to compute an unknown rotation matrix T able to guarantee a least squares matching of the matrices AT and B, where A is a variance covariance eigenvector matrix of a control net point coordinate vector and B is an ideal pseudo eigenvector matrix for the same vector. This last matrix was created a priori by 2D rotations of the "essential" eigenvector component pairs of the net point coordinates in such a way to orient them to the greatest possible extent along a direction orthogonal to that of the movement predicted by the deformation model. See for instance Crosilla (1983a, 1985) for the basic methodology and Crosilla (1983b) for further numerical developments of this technique.

The procedure known in the literature as a generalization of the orthogonal procrustes problem, given by Schönemann and Carroll (1970), could be applied with success for the transformation problems solutions of point coordinates between different reference systems.
In geodesy it is a common practice to transform by similarity 3D coordinates related to WGS 84 reference ellipsoid into 3D coordinates of a local reference system. For this purpose it is necessary to know in advance the approximate values of the unknown transformation parameters. Sometimes it is not easy to fix some of these values, like, for instance, in close range photogrammetry where often rotation angles between the model and the absolute reference systems are difficult to identify in advance.

In these cases procrustes methods are powerful because they do not require the knowledge of a priori unknown parameters values; from the computational point of view they just require some products of matrices containing point coordinates in different reference frames and the eigenvalue-eigenvector decomposition of a 3x3 matrix.

Some first numerical results of coordinate transformations with procrustes seem really satisfactory when compared with the results obtained with the classical methods and are worth of more deep investigations. A paper in progress will report these results and some further considerations.

Promising results are also expected by using generalized procrustes analysis, by Gower (1975) and Ten Berge (1977), for the computation of the International Terrestrial Reference System. As is well known the ITRS is based on the idea that each individual set of coordinates obtained by space geodesy measurements is related to a particular reference system. According to the procrustes approach, to combine all these coordinates into a unique frame, it is necessary to transform each solution by a 7 parameter similarity to an unknown common system satisfying the (2.1) Gower (1975) least squares function, previously reported.

In the author's opinion the same model could be expanded to compute the International Terrestrial Reference Frame (ITRF96) recently introduced by Sillard, Altamini and Boucher (1998) where a 14-parameter similarity is proposed to transform station positions and velocities into a combined system of reference.

Finally, a recent book by Dryden and Mardia (1998), presents very interesting applications of procrustes techniques and related statistics for the definition of objects' size-and-shape and the study of their variations. These examples are very interesting and worth applying to the deformation analysis of geodetic networks carried out by the comparison of two or more network adjustment results, relating to measurements made at different times.

Conclusions

Procrustes analysis seems to be a very promising technique in geodetic applications where transformation problems between systems of reference often have to be solved.
With respect to the classical transformation methods of solution, procrustes procedures take advantage of the symmetrical property of two matrices obtained by simple products of the original ones containing the point coordinate values to be matched. Spectral decomposition of these matrix products makes it possible to then compute the transformation parameters without any approximate value of the unknown parameters and with less computational time.
Very stimulating applications might be possible for the International Terrestrial Reference System and Frame computations and for the analysis of a deformation network by repeated measurements made at different times.

References

Crosilla F. 1983a, "A criterion matrix for the second order design of control networks", Bull. geod. 57: 226-239

Crosilla F. 1983b "Procrustean transformation as a tool for the construction of a criterion matrix for control networks" Manusc. geod. 8: 343-370

Cosilla F. 1985 " A criterion matrix for deforming networks by multifactorial analysis techniques", in E. Grafarend F. Sansò eds. Optimization and design of geodetic networks, Springer Verlag, Berlin Heidelberg New York Tokyo

Dryden I.L. Mardia K.V., 1998 "Statistical Shape Analysis",J. Wiley, Chichester

Goodall C. 1991 " Procrustes methods in the statistical analysis of shape", Journal Royal Statist. Soc. B 53, N° 2, pp 285-339

Gower J.C. 1975 " Generalized procrustes analysis", Psychometrika Vol. 40, N°1, pp 33-51

Koschat M. A. Swayne D. F. 1991" A weighted procrustes criterion" , Psychometrika vol 56 N° 2 pp 229- 239

Lingoes J.C. Schönemann P.H. 1974 " Alternative measures of fit for the Schönemann- Carroll matrix fitting algorithm", Psychometrika Vol 39, N°4 pp 423-427

Lissitz R.W. Schönemann P.H. Lingoes J.C. 1976 "A solution to the weighted procrustes problem in which the transformation is in agreement with the loss function", Psychometrika vol 41 N° 4, pp 547-550

Schönemann P.H. Bock R.D. Tucker L.R. 1965 "Some notes on a theorem by Eckart and Young" Univ. of North Carolina Psychometric Laboratory Research Memorandum N°25

Schönemann P.H. 1966 " A generalized solution of the orthogonal procrustes problem", Psychometrika, vol 31, N°1, pp 1-10

Schönemann P.H. Carroll R.M. 1970 "Fitting one matrix to another under choice of a central dilation and a rigid motion", Psychometrika, vol 35, N°2 pp245-255

Sillard P., Altamini Z., Boucher C. 1998 " The ITRF 96 realization and its associated Velocity field", Geophysical research letters, vol 25 N° 17, pp 3223-3226

Ten Berge J.M.F. 1977" Ortogonal procrustes rotation for two or more matrices",Psychometrika vol 42, N° 2, pp 267-276

On the maintenance of a proper reference frame for VLBI and GPS global networks

Athanasios Dermanis

1. Introduction

The use of an appropriate terrestrial reference frame in order to describe the position of points on the earth, as well as its temporal variation, is a problem of both theoretical and practical importance. This is the zero order optimal design problem in the terminology of Grafarend (1974), extended from the space to the space-time domain.

A terrestrial reference frame consists of a particular point O, its origin and a triad of orthonormal vectors $\vec{\mathbf{e}}=[\vec{e}_1\ \vec{e}_2\ \vec{e}_3]$, and it is related to a quasi-intertial orthonormal frame $\vec{\mathbf{e}}^I=[\vec{e}_1^I\ \vec{e}_3^I\ \vec{e}_2^I]$ with origin at the geocenter C. A point P of the earth has position vector $\vec{x}=\vec{OP}=x^1\vec{e}_1+x^2\vec{e}_2+x^3\vec{e}_3=\vec{\mathbf{e}}\mathbf{x}$ and terrestrial coordinates $\mathbf{x}=[x^1x^2x^3]^T$; its inertial position vector is $\vec{x}_I\equiv\vec{CP}=x_I^1\vec{e}_1^I+x_I^2\vec{e}_2^I+x_I^3\vec{e}_3^I=\vec{\mathbf{e}}^I\mathbf{x}_I$ and it has inertial coordinates $\mathbf{x}_I=[x_I^1x_I^2x_I^2]^T$. To relate the two frames, we need the displacement vector $\vec{d}\equiv\vec{CO}=\vec{\mathbf{e}}^I\mathbf{d}_I=\vec{\mathbf{e}}\,\mathbf{d}$ and the components Q_i^k of the elements of the terrestrial triad with respect to the inertial ones, i.e. $\vec{e}_i=\sum_k\vec{e}_k^IQ_i^k$, or $\vec{\mathbf{e}}=\vec{\mathbf{e}}^I\mathbf{Q}$ in matrix form. The matrix $\mathbf{Q}$ is a proper orthogonal matrix ($\mathbf{Q}^{-1}=\mathbf{Q}^T$, $|\mathbf{Q}|=+1$) as a consequence of the orthonormality and common orientation (right-handed) of $\vec{\mathbf{e}}$ and $\vec{\mathbf{e}}^I$. The relation between the two frames is expressed by

$$\vec{x}_I=\vec{\mathbf{e}}^I\mathbf{x}_I\equiv\vec{CP}=\vec{CO}+\vec{OP}=\vec{d}+\vec{x}=\vec{\mathbf{e}}^I\mathbf{d}_I+\vec{\mathbf{e}}\mathbf{x}=\vec{\mathbf{e}}^I\mathbf{d}_I+\vec{\mathbf{e}}^I\mathbf{Q}\mathbf{x}\,, \tag{1}$$

or in component form

$$\mathbf{x}_I=\mathbf{d}_I+\mathbf{Q}\mathbf{x}\,,\qquad \mathbf{x}=\mathbf{Q}^T(\mathbf{x}_I-\mathbf{d}_I)=\mathbf{Q}^T\mathbf{x}_I-\mathbf{d}\,. \tag{2}$$

The motion of any earth point in inertial space described by the function $\mathbf{x}_I(t)$, where t denotes time, should be determined by observations. The corresponding motion $\mathbf{x}(t)$, with respect to a terrestrial frame, depends in addition on the more or less arbitrary choice of the terrestrial frame, i.e. of the functions $\mathbf{Q}(t)$ and $\mathbf{d}(t)$. If the earth was rigid (or in applications where rigidity is a valid approximation) there are choices of $\mathbf{Q}(t)$ and $\mathbf{d}(t)$, such that the terrestrial coordinates $\mathbf{x}=\mathbf{Q}^T(t)\mathbf{x}_I(t)-\mathbf{d}(t)$ are constant. For a deformable earth, where deformations are known to be small, it is reasonable to establish a terrestrial frame in a way that the temporal variations of $\mathbf{x}(t)=\mathbf{Q}^T(t)\mathbf{x}_I(t)-\mathbf{d}(t)$ appear to be "as small as possible", i.e., such that the largest part of the inertial motions $\mathbf{x}_I(t)$ is absorbed by the rotation and position of the terrestrial frame. The optimal choice of the terrestrial frame depends directly on the specific optimality criterion, which gives concrete mathematical meaning to the loose expression "as small as possible".

The solution of the problem requires, apart from the choice of the optimality criterion, the knowledge of the motion $\mathbf{x}_I(t)$ with respect to the inertial frame, of every point of the earth. Operationally, this is possible only for points on the surface of the earth, while the motion of internal points has to be deduced from theoretical arguments.

The general form of an optimality criterion is

$$\int_{t_0}^{t_F}\int_{E} F\left(x(t),\frac{d\mathbf{x}}{dt}(t)\right)d\mathbf{x}\,dt=\min\,, \tag{3}$$

where F is an appropriate known function and integration is carried out over the earth E and the time interval $[t_0,t_F]$ for which observational data are available.

A more modest problem is the maintenance of a reference frame for a set of discrete points P_i, $i=1,\ldots,n$, which are the positions of observation stations distributed all over the world and engaged in a collective analysis of the acquired data.

Formerly, the problem of frame definition was solved in a discrete way, corresponding to discrete data $\mathbf{x}_i(t_k)=\mathbf{x}(P_i,t_k)$, collected in repeated campaigns over short time intervals, which could efficiently considered as "instantaneous" data corresponding to a single epoch t_k. The most popular approach starts with a more or less arbitrary frame definition at the initial epoch t_0 and then fits the coordinates of each epoch t_k to those already for the previous epoch t_{k-1}, by applying the optimality criterion

$$\sum_{i=1}^{n}[\mathbf{x}_i(t_k)-\mathbf{x}_i(t_{k-1})]^T[\mathbf{x}_i(t_k)-\mathbf{x}_i(t_{k-1})]=\min\ . \tag{4}$$

Setting

$$\mathbf{x}(t_k)=\begin{bmatrix}\mathbf{x}_1(t_k)\\ \vdots\\ \mathbf{x}_n(t_k)\end{bmatrix}, \tag{5}$$

and introducing the notation $\mathbf{x}^0=\mathbf{x}(t_{k-1})$, $\mathbf{x}=\mathbf{x}(t_k)$, $\delta\mathbf{x}=\mathbf{x}-\mathbf{x}^0$, the optimality criterion $\delta\mathbf{x}^T\delta\mathbf{x}=\min$ can be incorporated in a linearized adjustment model $\mathbf{l}=\mathbf{A}\delta\mathbf{x}+\mathbf{v}$, where linearization is based on the approximate values $\mathbf{x}^0=\mathbf{x}(t_{k-1})$, by introducing a set of *inner constraints* $\mathbf{E}\delta\mathbf{x}=\mathbf{0}$, where the rows of $\mathbf{E}$ form a basis for the null space $N(\mathbf{A})=\{\mathbf{x}|\mathbf{A}\mathbf{x}=\mathbf{0}\}$ of the matrix $\mathbf{A}$.

Nowadays, observation stations are engaged in continuous data collection, operating as permanent stations within the framework of international organizations, such as the International GPS Service (IGS) and provide the means for the establishment of an official International Terrestrial Reference System (ITRS) with the care of the International Earth Rotation Service (IERS).

Remark:
The term reference frame used here corresponds to the term ITRS used by IERS, which preserves the term International Terrestrial Reference Frame (ITRF) to the set of stations engaged in the realization of the ITRS.

With the availability of continuously available data the stepwise approach of the past is no more desirable or practical from the implementation point of view. Instead, we propose to visualize the data (which are discrete but with a very high rate of repetition) as continuous and to seek a time-continuous solution to the reference frame choice problem. Such a solution may be eventually discretized or implemented in a discrete approximation. Furthermore, the optimality criterion to be introduced for the frame choice in the discrete point network, must coincide with (or at least attempt to imitate) the optimality criterion introduced for the whole earth on the basis of theoretical considerations.

We will distinguish between two types of networks, which we call for the sake of convenience VLBI- and GPS-type networks. In VLBI-type networks where observations are translation-invariant, the position of the geocenter C cannot be determined and the origin of the frame O must be also introduced. Thus we must determine both functions $\mathbf{Q}(t)$ and $\mathbf{d}(t)$. In GPS-type networks observations are linked to the geocenter through the use of satellite orbits, in which case $O\equiv C$ is already known and only $\mathbf{Q}(t)$ should be determined.

In Dermanis (1995) we introduced a methodology for the solution to the space-time datum problem, which considered also scale transformations. Here we will restrict ourselves to rigid transformations, since scale is provided for both VLBI- and GPS-type networks, within the framework of a non-relativistic approach, through the assumption that mean of the readings of a set of reference clocks does not accelerate or decelerate with respect to Newtonian time. Distance (and thus scale) is entering the problem only implicitly through the observation of time intervals. On the other hand we will generalize the approach by looking into alternative optimality criteria and also by introducing "masses" or "weights" m_i, for each station P_i, which may reflect either a measure of the quality of station data, or the degree of participation of the station to the optimality criterion, in relation to the part of earth masses closest to the particular point.

2. Transformation from a preliminary reference frame to an optimal one

The basic idea of our approach is to make use of the fact that the available observations can very well determine the shape of the network, but not the additional information of its orientation (and position) with respect to a reference frame, which is contained in a set of network coordinates $\mathbf{x}$. At any single epoch t there exist an infinite number of coordinates $\mathbf{x}(t)$ which give rise to the same shape for the network.
If $\mathbf{x}'(t)$ and $\mathbf{x}(t)$ are two coordinate sets which both correspond to the "observed" shape at epoch t, there exists a rigid transformation between the two, defined point-wise by

$$\mathbf{x}'_i(t)=\mathbf{R}\big(\boldsymbol{\theta}(t)\big)\mathbf{x}_i(t)+\mathbf{b}(t)\,, \tag{6}$$

where $\mathbf{R}$ is a proper orthogonal matrix. This means that if a preliminary solution $\mathbf{x}(t)$ is available, we can switch to an optimal solution $\mathbf{x}'(t)$, by applying an optimization principle and determining the optimal six functions $\boldsymbol{\theta}(t)=[\theta_1(t)\,\theta_2(t)\,\theta_3(t)]^T$ and $\mathbf{b}(t)=[b_1(t)\,b_2(t)\,b_3(t)]^T$ which transform to the coordinates $\mathbf{x}'(t)$ satisfying the optimality criterion. But such a solution is always available, because a reference frame must be introduced for the analysis of the data which lead to the coordinate estimates $\mathbf{x}(t)$ at every epoch t. The only requirement is that the function $\mathbf{x}(t)$ is a smooth one, i.e. continuous with continuous derivatives up to a certain order. The reference frame for $\mathbf{x}(t)$ may be introduced during the adjustment of the observations, by a set of minimal constraints, which define a frame without any influence on the shape of the network.
The original known rotation $\mathbf{Q}(t)$ and displacement $\mathbf{d}(t)$ in $\mathbf{x}=\mathbf{Q}^T\mathbf{x}_I-\mathbf{d}$ must be combined with the optimal relative rotation $\mathbf{R}(t)$ and relative displacement $\mathbf{b}(t)$ in $\mathbf{x}'=\mathbf{R}\mathbf{x}+\mathbf{b}$, in order to obtain the final ones $\mathbf{Q}'(t)$ and $\mathbf{d}'(t)$ in $\mathbf{x}'=\mathbf{Q}'^T\mathbf{x}_I-\mathbf{d}'$ by means of

$$\mathbf{x}'=\mathbf{R}\mathbf{x}+\mathbf{b}=\mathbf{R}(\mathbf{Q}^T\mathbf{x}_I-\mathbf{d})+\mathbf{b}=\mathbf{R}\mathbf{Q}^T\mathbf{x}_I-\mathbf{R}\mathbf{d}+\mathbf{b}=\mathbf{Q}'^T\mathbf{x}_I-\mathbf{d}' \quad\Rightarrow\quad \mathbf{Q}'=\mathbf{Q}\mathbf{R}^T \;\&\; \mathbf{d}'=\mathbf{R}\mathbf{d}-\mathbf{b}\,. \tag{7}$$

3. Optimal solutions of minimum energy and minimum length (geodesics)

A particular optimality criterion is based on the instantaneous (relative to the terrestrial frame) kinetic energy of the earth $T(t)=\frac{1}{2}\int_E|\mathbf{v}|^2\,dm$, where $\mathbf{v}=\frac{d\mathbf{x}}{dt}$ is the velocity, which is integrated over a time interval $[t_0,t_F]$ and the resulting total energy is minimized

$$\int_{t_0}^{t_F}T_E(t)\,dt=\frac{1}{2}\int_{t_0}^{t_F}\int_E\mathbf{v}^T\mathbf{v}\,dm\,dt=\min\,. \tag{8}$$

The discrete analog for a network of points P_i, $i=1,\ldots,n$, each of which has optimal coordinates $\mathbf{x}'_i$ and it is assigned a mass m_i, takes the form

$$\int_0^t T_{\mathrm{N}}(\mathrm{t})\,dt=\tfrac{1}{2}\int_0^t\left(\frac{d\mathbf{x}'}{dt}\right)^T\mathbf{M}\frac{d\mathbf{x}'}{dt}\,dt=\min\,, \tag{9}$$

where

$$T_N(t)=\frac{1}{2}\left(\frac{d\mathbf{x}'}{dt}\right)^T\mathbf{M}\frac{d\mathbf{x}'}{dt}=\frac{1}{2}\sum_{i=1}^n m_i\left(\frac{d\mathbf{x}_i'}{dt}\right)^T\frac{d\mathbf{x}_i'}{dt}=\frac{1}{2}\sum_{i=1}^n m_i\,\mathbf{v}_i'^T\mathbf{v}_i'=\frac{1}{2}\sum_{i=1}^n m_i|\mathbf{v}_i'|^2\,,\quad (M_{ik}=\delta_{ik}m_i), \tag{10}$$

is the kinetic energy of the network.
The minimization principle (9) is in fact equivalent to the minimization principle

$$\int_{t_0}^{t_F}\sqrt{\left(\frac{d\mathbf{x}'}{dt}\right)^T\mathbf{M}\frac{d\mathbf{x}'}{dt}}\,dt=\int_{t_0}^{t_F}\sqrt{d\mathbf{x}'^T\mathbf{M}d\mathbf{x}'}=\int_{t_0}^{t_F}ds=s|_{t_0}^{t_F}=\min \tag{11}$$

which is leading to a solution $\mathbf{x}'(t)$ which is a geodesic curve (curve of minimum length $s|_{t_0}^{t_F}$) in the network coordinate space X where any set of network coordinates $\mathbf{x}$ belong, $\mathbf{x}\in X$. Distance is measured by an element of length ds defined by $ds^2=d\mathbf{x}'^T\mathbf{M}d\mathbf{x}'=\sum_{i=1}^n m_i d\mathbf{x}_i'^T d\mathbf{x}_i'$. This means that the "distance" between two network coordinate sets $\mathbf{x}'$ and $\mathbf{x}''$ is measured by

$$\rho(\mathbf{x}',\mathbf{x}'')=\|\mathbf{x}'-\mathbf{x}''\|=\sqrt{\sum_{i=1}^n m_i(\mathbf{x}_i'-\mathbf{x}_i'')^T(\mathbf{x}_i'-\mathbf{x}_i'')}=\sqrt{\sum_{i=1}^n m_i d_i^2}\,, \tag{12}$$

where $d_i=\|\mathbf{x}_i'-\mathbf{x}_i''\|$ is the usual euclidean distance between the two positions of point P_i. The minimization problem (10) is a standard problem of the calculus of variations. Its solution $\mathbf{x}'(t)$, described by means of the curvilinear coordinates

$$\mathbf{u}(s)=\begin{bmatrix}\boldsymbol{\theta}(s)\\ \mathbf{b}(s)\\ t(s)\end{bmatrix} \tag{13}$$

expressed as functions of arc length s, satisfies the Euler-Lagrange differential equations

$$\frac{\partial L}{\partial\boldsymbol{\theta}}-\frac{d}{ds}\left(\frac{\partial L}{\partial\dot{\boldsymbol{\theta}}}\right)=\mathbf{0}\,,\qquad \frac{\partial L}{\partial\mathbf{b}}-\frac{d}{ds}\left(\frac{\partial L}{\partial\dot{\mathbf{b}}}\right)=\mathbf{0}\,,\qquad \frac{\partial L}{\partial t}-\frac{d}{ds}\left(\frac{\partial L}{\partial\dot{t}}\right)=0\,,\qquad \left(\dot{\boldsymbol{\theta}}=\frac{d\boldsymbol{\theta}}{ds},\ \dot{\mathbf{b}}=\frac{d\mathbf{b}}{ds},\ \dot{t}=\frac{dt}{ds}\right). \tag{14}$$

The derivation of the explicit form of the Euler-Lagrange equations has been carried out in Dermanis (1995), for the special case $\mathbf{M}=\mathbf{I}$, but they can be easily generalized to the present case of varying point masses m_i. Additionally the geodesic differential equations, corresponding to the optimality criterion (11) have been derived, yielding (as expected) identical results.

The resulting equations are

$$[(\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt})\times]\left(\mathbf{h}_x+\mathbf{C}_x\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}\right)+\left[\frac{d\mathbf{h}_x}{dt}+\frac{d\mathbf{C}_x}{dt}\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+\mathbf{C}_x\mathbf{R}^T\left(\frac{d\boldsymbol{\Omega}}{dt}\frac{d\boldsymbol{\theta}}{dt}+\boldsymbol{\Omega}\frac{d^2\boldsymbol{\theta}}{dt^2}\right)\right]-\frac{\ddot{s}}{\dot{s}}\left(\mathbf{h}_x+\mathbf{C}_x\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}\right)=\mathbf{0} \tag{15}$$

$$\frac{d^2\mathbf{b}}{dt^2}-\frac{\ddot{s}}{\dot{s}}\frac{d\mathbf{b}}{dt}=\mathbf{0} \tag{16}$$

$$\left(\frac{d\mathbf{x}}{dt}\right)^T\frac{d^2\mathbf{x}}{dt^2}+\mathbf{h}_x^T\mathbf{R}^T\left(\frac{d\boldsymbol{\Omega}}{dt}\frac{d\boldsymbol{\theta}}{dt}+\boldsymbol{\Omega}\frac{d^2\boldsymbol{\theta}}{dt^2}\right)-\frac{1}{2}\left(\frac{d\boldsymbol{\theta}}{dt}\right)^T\boldsymbol{\Omega}^T\mathbf{R}\frac{d\mathbf{C}_x}{dt}\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}-\frac{\ddot{s}}{\dot{s}}\left[\left(\frac{d\mathbf{x}}{dt}\right)^T\frac{d\mathbf{x}}{dt}+\mathbf{h}_x^T\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}\right]=0 \tag{17}$$

where $\mathbf{C}_x$ is the moment of inertia matrix of the network with respect to the initial reference frame

$$\mathbf{C}_x=-\sum_{i=1}^{n}m_i[\mathbf{x}_i\times]^2=\sum_{i=1}^{n}m_i[(\mathbf{x}_i^T\mathbf{x}_i)-\mathbf{x}_i\mathbf{x}_i^T]=\mathbf{x}^T\mathbf{M}\mathbf{x}-\sum_{i=1}^{n}m_i\mathbf{x}_i\mathbf{x}_i^T\ , \tag{18}$$

$\mathbf{h}_x$ is its relative angular momentum vector of the network with respect to the initial reference frame

$$\mathbf{h}_x=\sum_{i=1}^{n}m_i[\mathbf{x}_i\times]\dot{\mathbf{x}}_i \tag{19}$$

and $\boldsymbol{\Omega}=\boldsymbol{\Omega}(\boldsymbol{\theta})$ is a matrix defined by

$$[\boldsymbol{\omega}_k\times]\equiv\frac{\partial\mathbf{R}}{\partial\theta_k}\mathbf{R}^T\ ,\qquad \boldsymbol{\Omega}=[\boldsymbol{\omega}_1\boldsymbol{\omega}_2\boldsymbol{\omega}_3]\ . \tag{20}$$

Remark: We make repeated use of the notation

$$\mathbf{a}=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\qquad\Rightarrow\qquad[\mathbf{a}\times]\equiv\begin{bmatrix}0&-a_3&a_2\\a_3&0&-a_1\\-a_2&a_1&0\end{bmatrix} \tag{21}$$

and of the properties ($\mathbf{Q}^{-1}=\mathbf{Q}^T$)

$$[\mathbf{a}\times]\mathbf{b}=-[\mathbf{b}\times]\mathbf{a}\ ,\qquad[(\mathbf{Q}\mathbf{a})\times]=\mathbf{Q}[\mathbf{a}\times]\mathbf{Q}^T\ ,\qquad[\mathbf{a}\times][\mathbf{b}\times]=\mathbf{b}\mathbf{a}^T-(\mathbf{a}^T\mathbf{b})\mathbf{I}\ . \tag{22}$$

We have also assumed that reference frame $\mathbf{x}(t)$ has been chosen in a way that $\bar{\mathbf{x}}=\mathbf{0}$, where $\bar{\mathbf{x}}$ are the coordinates of the center of mass of the network, defined by

$$\bar{\mathbf{x}}=\frac{1}{m}\sum_{i=1}^{n}m_i\mathbf{x}_i\ ,\qquad m=\sum_{i=1}^{n}m_i\ . \tag{23}$$

(It is always possible to switch from any reference frame $\mathbf{x}_0(t)$ to a "centered" one $\mathbf{x}(t)$ with $\bar{\mathbf{x}}=\mathbf{0}$, using $\mathbf{x}_i(t)=\mathbf{x}_{0i}(t)-\bar{\mathbf{x}}_0$).

Of the seven equations (15), (16), (17), the last one can be solved for the factor $\frac{\ddot{s}}{\dot{s}}$ (s been length and dots denoting differentiation with respect to time) which replaced in the first two, will yield a system of six non-linear differential equations for the six unknown functions $\boldsymbol{\theta}(t)=[\theta_1(t)\,\theta_2(t)\,\theta_3(t)]^T$ and $\mathbf{b}(t)=[b_1(t)\,b_2(t)\,b_3(t)]^T$. The resulting equations are very complicated and they can be solved only by numerical methods. Furthermore any solution yields a frame definition where the "curve" $\mathbf{x}'(t)$ is the closest between its end points $\mathbf{x}'(t_0)$ and $\mathbf{x}'(t_F)$, but not necessarily the shortest possible. To arrive at a truly optimal solution, which by the way is not unique, we must select the optimal among all initial (or boundary) values which are necessary for obtaining a specific solution of the relevant differential equations.

We will not pursue this matter any further, but we will follow a different approach motivated by the methods used in the theoretical study of earth rotations.

4. Tisserand axes

The rotation of the earth is governed by the differential equation $\frac{d\vec{h}}{dt}=\vec{l}$, where $\vec{h}=\int_E \vec{x}\times\vec{v}\,dm$ is the angular momentum, $\vec{l}=\int_E \vec{x}\times\vec{a}\,dm$ is the acting torque, $\vec{v}=\frac{d\vec{x}}{dt}$ are the velocities of earth points and $\vec{a}$ the corresponding acting accelerations. The rotational equation which in the inertial frame becomes simply $\frac{d\mathbf{h}_I}{dt}=\mathbf{l}_I$, obtains a more complicated form when expressed with respect to the terrestrial frame. Differentiation of $\vec{\mathbf{e}}=\vec{\mathbf{e}}^I\mathbf{Q}$, yields $\frac{d\vec{\mathbf{e}}}{dt}=\vec{\mathbf{e}}^I\frac{d\mathbf{Q}}{dt}=\vec{\mathbf{e}}\mathbf{Q}^T\frac{d\mathbf{Q}}{dt}\equiv\vec{\mathbf{e}}[\boldsymbol{\omega}\times]$, where $\vec{\omega}=\vec{\mathbf{e}}\boldsymbol{\omega}$ is the instantaneous rotation vector of the earth, so that $\vec{v}=\frac{d\vec{x}}{dt}=\frac{d}{dt}(\vec{\mathbf{e}}\mathbf{x})=\vec{\mathbf{e}}(\frac{d\mathbf{x}}{dt}+[\boldsymbol{\omega}\times]\mathbf{x})$ and similarly $\frac{d\vec{h}}{dt}=\vec{\mathbf{e}}(\frac{d\mathbf{h}}{dt}+[\boldsymbol{\omega}\times]\mathbf{h})=\vec{l}=\vec{\mathbf{e}}\,\mathbf{l}$. The components of $\vec{h}=\vec{\mathbf{e}}\mathbf{h}$ in the terrestrial system become

$$\mathbf{h}=\int_E[\mathbf{x}\times]\left(\frac{d\mathbf{x}}{dt}+[\boldsymbol{\omega}\times]\mathbf{x}\right)dm=\left(-\int_E[\mathbf{x}\times][\mathbf{x}\times]dm\right)\boldsymbol{\omega}+\int_E[\mathbf{x}\times]\frac{d\mathbf{x}}{dt}\,dm\equiv\mathbf{C}\boldsymbol{\omega}+\mathbf{h}_R \tag{24}$$

where $\mathbf{C}$ is the inertia matrix and $\mathbf{h}_R$ the relative angular momentum of the earth. Replacing $\mathbf{h}=\mathbf{C}\boldsymbol{\omega}+\mathbf{h}_R$ in the rotational equations $\frac{d\mathbf{h}}{dt}+[\boldsymbol{\omega}\times]\mathbf{h}=\mathbf{l}$ yields the *Liouville equations*

$$\mathbf{C}\frac{d\boldsymbol{\omega}}{dt}+\frac{d\mathbf{C}}{dt}\boldsymbol{\omega}+\frac{d\mathbf{h}_R}{dt}+[\boldsymbol{\omega}\times](\mathbf{C}\boldsymbol{\omega}+\mathbf{h}_R)=\mathbf{l}\,. \tag{25}$$

The choice of the terrestrial frame in the study of earth rotation is dictated by the need to simplify the analytical work involved in solving the Liouville equations.
Two choices are under consideration (Munk & MacDonald, 1960, ch. 3.2, p. 10): the *principal axes* or *figure axes*, defined so that $\mathbf{C}$ becomes diagonal, and the *Tisserant axes* for which the relative angular momentum vanishes, $\mathbf{h}_R=\mathbf{0}$. The first choice is more appropriate for the theory of rotation of a rigid earth but it has a serious shortcoming when an elastic earth model is used: as a consequence of rotational elastic deformation the third (polar) axis of figure intersecting the earth at a point F, undergoes a diurnal rotation around the corresponding position F_0 of the rigid earth model with a radius of $F_0F=60\,\mathrm{m}$, while F_0 undergoes a rotation around the position O of third Tisserand axis, with radius of only $OF_0=2\,\mathrm{m}$ and a Chandler period of about 430 days (Moritz and Mueller, 1987, ch. 3.3.1). For this reason the Tisserand axes are the preferred ones for the description of the rotation of the deformable earth. Furthermore the Tisserand axes have the advantage that they minimize the relative to the terrestrial frame kinetic energy of the earth, i.e., $T_E=\int_E|\mathbf{v}|^2\,dm=\min$ (Moritz and Mueller, 1987, ch. 3.1). Both choices of figure and Tisserand axes, cannot determine a displacement but only the rotation from an initial arbitrary reference frame. In theory they are both considered to be geocentric.
The figure axes are uniquely defined for any body that has no axis of symmetry. They are therefore well defined for the real earth, but not for an ellipsoidal model-earth where only one direction (that of symmetry) coincides with one figure axes and the position of the other two must be arbitrarily chosen.
On the contrary the Tisserant axes are not uniquely defined. Indeed if $\mathbf{x}$ are coordinates with respect to a set of Tisserand axes and we consider a new set of axes defined by the transformation $\tilde{\mathbf{x}}=\mathbf{S}\mathbf{x}$, where $\mathbf{S}$ is a time-independent orthogonal matrix then

$$\tilde{\mathbf{h}}_R=\int_E[\tilde{\mathbf{x}}\times]\tilde{\mathbf{v}}\,dm=\int_E[(\mathbf{S}\mathbf{x})\times]\mathbf{S}\mathbf{v}\,dm=\int_E\mathbf{S}[\mathbf{x}\times]\mathbf{S}^T\mathbf{S}\mathbf{v}\,dm=\mathbf{S}\int_E[\mathbf{x}\times]\mathbf{v}\,dm=\mathbf{S}\mathbf{h}_R=\mathbf{S}\mathbf{0}=\mathbf{0} \tag{26}$$

and the $\tilde{\mathbf{x}}$ axes are also Tisserand axes. To choose a particular set of Tisserand axes we must fix their position $\mathbf{x}(t_0)$ at an initial epoch t_0.

For a discrete network of mass points we may define a set of "Tisserand" axes by setting the corresponding relative momentum equal to zero

$$\mathbf{h}_{x'} \equiv \sum_i m_i [\mathbf{x}'_i \times] \frac{d\mathbf{x}'_i}{dt} = \mathbf{0} \tag{27}$$

and try to find the transformation parameters $\boldsymbol{\theta}(t)$, $\mathbf{b}(t)$ which convert coordinates $\mathbf{x}(t)$ in an originally available reference frame into "Tisserand" coordinates $\mathbf{x}'(t) = \mathbf{R}\big(\boldsymbol{\theta}(t)\big)\mathbf{x}(t) + \mathbf{b}(t)$.

Setting

$$[\boldsymbol{\omega}_k \times] = \frac{\partial \mathbf{R}}{\partial \theta_k}\mathbf{R}^T, \qquad \frac{\partial \mathbf{R}}{\partial \theta_k} = [\boldsymbol{\omega}_k \times]\mathbf{R} \tag{28}$$

we have

$$\frac{d\mathbf{R}}{dt} = \frac{\partial \mathbf{R}}{\partial \theta_1}\frac{d\theta_1}{dt} + \frac{\partial \mathbf{R}}{\partial \theta_2}\frac{d\theta_2}{dt} + \frac{\partial \mathbf{R}}{\partial \theta_3}\frac{d\theta_3}{dt} = \frac{d\theta_1}{dt}[\boldsymbol{\omega}_1 \times]\mathbf{R} + \frac{d\theta_2}{dt}[\boldsymbol{\omega}_2 \times]\mathbf{R} + \frac{d\theta_3}{dt}[\boldsymbol{\omega}_3 \times]\mathbf{R} \tag{29}$$

$$\frac{d\mathbf{x}'_i}{dt} = \frac{d\mathbf{R}}{dt}\mathbf{x}_i + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} = \frac{d\theta_1}{dt}[\boldsymbol{\omega}_1 \times]\mathbf{R}\mathbf{x}_i + \frac{d\theta_2}{dt}[\boldsymbol{\omega}_2 \times]\mathbf{R}\mathbf{x}_i + \frac{d\theta_3}{dt}[\boldsymbol{\omega}_3 \times]\mathbf{R}\mathbf{x}_i + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} =$$

$$= -\frac{d\theta_1}{dt}[(\mathbf{R}\mathbf{x}_i)\times]\boldsymbol{\omega}_1 - \frac{d\theta_2}{dt}[(\mathbf{R}\mathbf{x}_i)\times]\boldsymbol{\omega}_2 - \frac{d\theta_3}{dt}[(\mathbf{R}\mathbf{x}_i)\times]\boldsymbol{\omega}_3 + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} =$$

$$= -[(\mathbf{R}\mathbf{x}_i)\times]\left(\frac{d\theta_1}{dt}\boldsymbol{\omega}_1 + \frac{d\theta_2}{dt}\boldsymbol{\omega}_2 + \frac{d\theta_3}{dt}\boldsymbol{\omega}_3\right) + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} =$$

$$= -[(\mathbf{R}\mathbf{x}_i)\times][\boldsymbol{\omega}_1\boldsymbol{\omega}_2\boldsymbol{\omega}_3]\begin{bmatrix} \frac{d\theta_1}{dt} \\ \frac{d\theta_2}{dt} \\ \frac{d\theta_3}{dt} \end{bmatrix} + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} \equiv -[(\mathbf{R}\mathbf{x}_i)\times]\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt} + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} =$$

$$= -\mathbf{R}[\mathbf{x}_i \times]\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt} + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt} \tag{30}$$

Therefore, the vanishing relative momentum becomes

$$\mathbf{h}_x = \sum_i m_i [\mathbf{x}'_i \times]\frac{d\mathbf{x}'_i}{dt} = \sum_i m_i \Big(\mathbf{R}[\mathbf{x}_i \times]\mathbf{R}^T + [\mathbf{b}\times]\Big)\left(-\mathbf{R}[\mathbf{x}_i \times]\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt} + \mathbf{R}\frac{d\mathbf{x}_i}{dt} + \frac{d\mathbf{b}}{dt}\right) =$$

$$=\sum_i m_i\left(-\mathbf{R}[\mathbf{x}_i\times][\mathbf{x}_i\times]\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+\mathbf{R}[\mathbf{x}_i\times]\frac{d\mathbf{x}_i}{dt}+\mathbf{R}[\mathbf{x}_i\times]\mathbf{R}^T\frac{d\mathbf{b}}{dt}\right)+$$

$$+\sum_i m_i\left(-[\mathbf{b}\times]\mathbf{R}[\mathbf{x}_i\times]\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+[\mathbf{b}\times]\mathbf{R}\frac{d\mathbf{x}_i}{dt}+[\mathbf{b}\times]\frac{d\mathbf{b}}{dt}\right)=$$

$$=-\mathbf{R}\left(\sum_i m_i[\mathbf{x}_i\times][\mathbf{x}_i\times]\right)\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+\mathbf{R}\sum_i m_i[\mathbf{x}_i\times]\frac{d\mathbf{x}_i}{dt}+\mathbf{R}\left(\sum_i m_i[\mathbf{x}_i\times]\right)\mathbf{R}^T\frac{d\mathbf{b}}{dt}-$$

$$-[\mathbf{b}\times]\mathbf{R}\left(\sum_i m_i[\mathbf{x}_i\times]\right)\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+[\mathbf{b}\times]\mathbf{R}\sum_i m_i\frac{d\mathbf{x}_i}{dt}+\left(\sum_i m_i\right)[\mathbf{b}\times]\frac{d\mathbf{b}}{dt}=$$

$$=-\mathbf{R}\mathbf{C}_x\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+\mathbf{R}\mathbf{h}_x+m\mathbf{R}[\bar{\mathbf{x}}\times]\mathbf{R}^T\frac{d\mathbf{b}}{dt}-m[\mathbf{b}\times]\mathbf{R}[\bar{\mathbf{x}}\times]\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+m[\mathbf{b}\times]\mathbf{R}\frac{d\bar{\mathbf{x}}}{dt}+m[\mathbf{b}\times]\frac{d\mathbf{b}}{dt}=\mathbf{0}\,. \quad (31)$$

Under the feasible assumption that $\bar{\mathbf{x}}=\mathbf{0}$, and the consequent $\frac{d\bar{\mathbf{x}}}{dt}=\mathbf{0}$, the last equation simplifies to

$$\mathbf{h}_{x'}=-\mathbf{R}\mathbf{C}_x\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+\mathbf{R}\mathbf{h}_x+m[\mathbf{b}\times]\frac{d\mathbf{b}}{dt}=\mathbf{0} \quad (32)$$

These are three equations in six unknowns, which is an underdetermined system. This reflects the fact that the Tisserant principle $\mathbf{h}_{x'}=\mathbf{0}$ can determine the orientation but not the position of the (geocentric) Tisserant axes with respect to the original working frame. For GPS-type networks where the original frame is already geocentric, we have $\mathbf{b}(t)=\mathbf{0}$, by definition. For VLBI-type networks we will set again $\mathbf{b}(t)=\mathbf{0}$ to obtain the orientation of a Tisserand frame parallel to the geocentric Tisserant frame, with the same origin as the original frame. We need a separate optimization principle for the determination of an optimal origin of the network, since the position of the geocenter remains undeterminable from the available data.

With the choice $\mathbf{b}(t)=\mathbf{0}$, the three transformation parameters $\boldsymbol{\theta}(t)$ to "Tisserand" coordinates should be determined from the solution of the three differential equations $\mathbf{h}_{x'}=-\mathbf{R}\mathbf{C}_x\mathbf{R}^T\boldsymbol{\Omega}\frac{d\boldsymbol{\theta}}{dt}+\mathbf{R}\mathbf{h}_x=\mathbf{0}$, which under the additional assumption that $|\boldsymbol{\Omega}|\neq 0$ take the form

$$\frac{d\boldsymbol{\theta}}{dt}=\boldsymbol{\Omega}^{-1}\mathbf{R}\mathbf{C}_x^{-1}\mathbf{h}_x\,. \quad (33)$$

These can be integrated to obtain a solution

$$\boldsymbol{\theta}(t)=\boldsymbol{\theta}(t_0)+\int_{t_0}^{t}\boldsymbol{\Omega}^{-1}\big(\boldsymbol{\theta}(\tau)\big)\mathbf{R}\big(\boldsymbol{\theta}(\tau)\big)\mathbf{C}_x^{-1}(\tau)\,\mathbf{h}_x(\tau)\,d\tau \quad (34)$$

which depends on the chosen initial value $\boldsymbol{\theta}(t_0)$ which determines one out of all the possible Tisserand frames. For example if $\boldsymbol{\theta}(t_0)=\mathbf{0}$ is chosen, the Tisserand axes will coincide with the original axes at the initial epoch, since $\mathbf{x}'(t_0)=\mathbf{R}\big(\boldsymbol{\theta}(t_0)\big)\mathbf{x}(t_0)=\mathbf{R}(\mathbf{0})\mathbf{x}(t_0)=\mathbf{I}\,\mathbf{x}(t_0)=\mathbf{x}(t_0)$.

The explicit form of the differential equations (33) depends on the chosen parametrization of the rotation matrix $\mathbf{R}$ in terms of three parameters $\boldsymbol{\theta}$.

5. Equivalence of Tisserant axes to a space-time generalization of Meissl's inner constraints

A different type of solution can be based on the extension of the well known concept of inner constraints, introduced by Meissl (1965, 1969). At any epoch t, when the network has coordinates $\mathbf{x}(t)$ in an original reference frame, the set of all coordinates $\mathbf{x}'(t)=\mathbf{R}\big(\boldsymbol{\theta}(t)\big)\mathbf{x}(t)+\mathbf{b}(t)$, resulting as the parameters $\boldsymbol{\theta}(t)$ and $\mathbf{b}(t)$ take any possible value, form a manifold, i.e. a "curved" subspace M_t of the network coordinate space X. In fact M_t is the set of all network coordinates which give rise to the same network shape as the one defined by $\mathbf{x}(t)$. Obviously $\mathbf{x}'(t)\in\mathsf{M}_t$ for any particular epoch t. The idea is now to impose on the curve $\mathbf{x}'(t)$ to be such that its velocity $\frac{d\mathbf{x}'}{dt}(t)$ is perpendicular to the manifold M_t, or more precisely to the "flat" space, which is tangent to the (cutved) manifold M_t at the point $\mathbf{x}'(t)$. Since the parameters $\boldsymbol{\theta}(t)$ and $\mathbf{b}(t)$ comprise a set of curvilinear coordinates for M_t, the tangent space is the set of all liner combinations of the vectors tangent to the coordinate curves, namely $\frac{\partial\mathbf{x}'}{\partial\theta_1}$, $\frac{\partial\mathbf{x}'}{\partial\theta_2}$, $\frac{\partial\mathbf{x}'}{\partial\theta_3}$, $\frac{\partial\mathbf{x}'}{\partial b_1}$, $\frac{\partial\mathbf{x}'}{\partial b_2}$, $\frac{\partial\mathbf{x}'}{\partial b_3}$. The orthogonality conditions $\frac{d\mathbf{x}'}{dt}\perp\frac{\partial\mathbf{x}'}{\partial\theta_k}$, $\frac{d\mathbf{x}'}{dt}\perp\frac{\partial\mathbf{x}'}{\partial b_k}$ take the form $\left(\frac{d\mathbf{x}'}{dt}\right)^T\mathbf{M}\frac{\partial\mathbf{x}'}{\partial\theta_k}=0$, $\left(\frac{d\mathbf{x}'}{dt}\right)^T\mathbf{M}\frac{\partial\mathbf{x}'}{\partial b_k}=0$, $k=1,2,3$, or in compact matrix notation

$$\left(\frac{\partial\mathbf{x}'}{\partial\boldsymbol{\theta}}\right)^T\mathbf{M}\frac{d\mathbf{x}'}{dt}=\sum_{i=1}^{n}m_i\left(\frac{\partial\mathbf{x}'_i}{\partial\boldsymbol{\theta}}\right)^T\frac{d\mathbf{x}'_i}{dt}=\mathbf{0}\,,\qquad\left(\frac{\partial\mathbf{x}'}{\partial\mathbf{b}}\right)^T\mathbf{M}\frac{d\mathbf{x}'}{dt}=\sum_{i=1}^{n}m_i\left(\frac{\partial\mathbf{x}'_i}{\partial\mathbf{b}}\right)^T\frac{d\mathbf{x}'_i}{dt}=\mathbf{0}\tag{35}$$

Replacing $\frac{\partial\mathbf{x}'_i}{\partial\boldsymbol{\theta}}=\mathbf{R}[\mathbf{x}_i\times]\mathbf{R}^T\boldsymbol{\Omega}$, $\frac{\partial\mathbf{x}'_i}{\partial\mathbf{b}}=\mathbf{I}$, and $\frac{d\mathbf{x}'_i}{dt}$ from (30), implementing the usual assumption that $\bar{\mathbf{x}}=\mathbf{0}$, $\frac{d\bar{\mathbf{x}}}{dt}=\mathbf{0}$, we arrive at

$$\boldsymbol{\Omega}^T\mathbf{R}\left(\mathbf{C}_x\mathbf{R}^T\frac{d\boldsymbol{\theta}}{dt}-\mathbf{h}_x\right)=\mathbf{0}\,,\qquad\frac{d\mathbf{b}}{dt}=\mathbf{0}\,.\tag{36}$$

The first one of (36) is equivalent to (33) and therefore the *inner constraint* or *Meissl frame* is a Tisserant frame! The second of (36) yields $\mathbf{b}=\text{constant}$ and provides a solution to the origin determination for VLBI-type networks: If we chosse $\mathbf{b}=\mathbf{0}$, this means that, in relation to the assumption $\bar{\mathbf{x}}=\mathbf{0}$, the network origin should remain at the "center of mass" of the network.
Any solution of (36) satisfies the geodesic or minimum energy equations (15), (16) and (17). Thus the Tisserand or Meissl frame solution $\mathbf{x}'(t)$ is a geodesic and even more it is a geodesic of minimum possible length among all geodesics, a property which follows from the fact that $\mathbf{x}'(t_0)\perp\mathsf{M}_{t_0}$ and $\mathbf{x}'(t_F)\perp\mathsf{M}_{t_F}$.
In order to see how the present solution is related to Meissl's concept of inner constraints, we must use $\frac{\partial\mathbf{x}'_i}{\partial\boldsymbol{\theta}}=[(\mathbf{R}\mathbf{x}_i)\times]\boldsymbol{\Omega}=[(\mathbf{x}'_i-\mathbf{b})\times]\boldsymbol{\Omega}$, $\frac{\partial\mathbf{x}'_i}{\partial\mathbf{b}}=\mathbf{I}$, $\bar{\mathbf{x}}=\mathbf{0}$ and $\frac{d\bar{\mathbf{x}}}{dt}=\mathbf{0}$, in order to rewrite the orthogonality conditions (35) in the form

$$\sum_i m_i[\mathbf{x}_i\times]\frac{d\mathbf{x}_i}{dt}=\mathbf{0}\,,\qquad\sum_i\frac{d\mathbf{x}_i}{dt}=\mathbf{0}\,.\tag{37}$$

Assume that the solution $\mathbf{x}'(t^0)$ has been determined at some epoch t^0 and we want to determine the solution at a slightly later epoch $t=t^0+\Delta t$, i.e. $\mathbf{x}'(t)=\mathbf{x}'(t^0+\Delta t)\approx\mathbf{x}'(t^0)+\frac{d\mathbf{x}'}{dt}(t^0)\,\Delta t$, using $\mathbf{x}'(t^0)=\mathbf{x}'^0$ as

a starting approximate value. If $\mathbf{x}_i'$ is replaced by $\mathbf{x}_i^0$, $\frac{d\mathbf{x}_i'}{dt}$ is approximated by $\frac{\Delta\mathbf{x}_i}{\Delta t}=\frac{\mathbf{x}_i-\mathbf{x}_i^0}{\Delta t}$, and we choose $m_i=1$, equations (37) are converted to the well known inner constraints:

$$\sum_i [\mathbf{x}_i^0\times]\Delta\mathbf{x}_i=\mathbf{0}\,,\qquad \sum_i \Delta\mathbf{x}_i=\mathbf{0}\,. \tag{38}$$

6. An illustrative example

A particular choice of rotation parameters is

$$\mathbf{R}(\boldsymbol{\theta})=\mathbf{R}(\theta_1,\theta_2,\theta_3)=\mathbf{R}_3(\theta_3)\mathbf{R}_2(\theta_2)\mathbf{R}_1(\theta_1) \tag{39}$$

yielding

$$[\boldsymbol{\omega}_1\times]=\frac{\partial\mathbf{R}}{\partial\theta_1}\mathbf{R}^T=-\mathbf{R}_3(\theta_3)\mathbf{R}_2(\theta_2)\mathbf{R}_1(\theta_1)[\mathbf{i}_1\times]\mathbf{R}^T=-\mathbf{R}[\mathbf{i}_1\times]\mathbf{R}^T \tag{40}$$

$$[\boldsymbol{\omega}_2\times]=\frac{\partial\mathbf{R}}{\partial\theta_2}\mathbf{R}^T=-\mathbf{R}_3(\theta_3)\mathbf{R}_2(\theta_2)[\mathbf{i}_2\times]\mathbf{R}_1(\theta_1)\mathbf{R}^T=-\mathbf{R}\mathbf{R}_1(-\theta_1)[\mathbf{i}_2\times]\mathbf{R}_1(\theta_1)\mathbf{R}^T \tag{41}$$

$$[\boldsymbol{\omega}_3\times]=\frac{\partial\mathbf{R}}{\partial\theta_3}\mathbf{R}^T=-\mathbf{R}_3(\theta_3)[\mathbf{i}_3\times]\mathbf{R}_2(\theta_2)\mathbf{R}_1(\theta_1)\mathbf{R}^T=-\mathbf{R}\mathbf{R}_1(-\theta_1)\mathbf{R}_2(-\theta_2)[\mathbf{i}_3\times]\mathbf{R}_2(\theta_2)\mathbf{R}_1(\theta_1)\mathbf{R}^T \tag{42}$$

$$\boldsymbol{\omega}_1=-\mathbf{R}\mathbf{i}_1=-\mathbf{R}_3(\theta_3)\mathbf{R}_2(\theta_2)\mathbf{i}_1 \tag{43}$$

$$\boldsymbol{\omega}_2=-\mathbf{R}\mathbf{R}_1(-\theta_1)\mathbf{i}_2=-\mathbf{R}_3(\theta_3)\mathbf{i}_2 \tag{44}$$

$$\boldsymbol{\omega}_3=-\mathbf{R}\mathbf{R}_1(-\theta_1)\mathbf{R}_2(-\theta_2)\mathbf{i}_3=-\mathbf{i}_3 \tag{45}$$

We may set

$$\mathbf{Q}=[\mathbf{q}_1\mathbf{q}_2\mathbf{q}_3]\equiv-\mathbf{R}^T\boldsymbol{\Omega}=-\mathbf{R}^T[\boldsymbol{\omega}_1\boldsymbol{\omega}_2\boldsymbol{\omega}_3]\,,\qquad \mathbf{q}_k=-\mathbf{R}^T\boldsymbol{\omega}_k \tag{46}$$

$$\mathbf{q}_1=\mathbf{i}_1 \tag{47}$$

$$\mathbf{q}_2=\mathbf{R}_1(-\theta_1)\mathbf{i}_2=\begin{bmatrix}0\\ \cos\theta_1\\ \sin\theta_1\end{bmatrix} \tag{48}$$

$$\mathbf{q}_3=\mathbf{R}_1(-\theta_1)\mathbf{R}_2(-\theta_2)\mathbf{i}_3=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta_1 & -\sin\theta_1\\ 0 & \sin\theta_1 & \cos\theta_1\end{bmatrix}\begin{bmatrix}\sin\theta_2\\ 0\\ \cos\theta_2\end{bmatrix}=\begin{bmatrix}\sin\theta_2\\ -\sin\theta_1\cos\theta_2\\ \cos\theta_1\cos\theta_2\end{bmatrix} \tag{49}$$

$$\mathbf{Q}\dot{\boldsymbol{\theta}}=\mathbf{C}_z^{-1}\mathbf{h}_z \tag{50}$$

or setting

$$\mathbf{C}_z^{-1}\mathbf{h}_z\equiv\mathbf{c}=\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix} \tag{51}$$

$$\begin{bmatrix} 1 & 0 & \sin\theta_2 \\ 0 & \cos\theta_1 & -\sin\theta_1\cos\theta_2 \\ 0 & \sin\theta_1 & \cos\theta_1\cos\theta_2 \end{bmatrix}\begin{bmatrix} \dot\theta_1 \\ \dot\theta_2 \\ \dot\theta_3 \end{bmatrix}=\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \tag{52}$$

$$\dot\theta_1+\sin\theta_2\dot\theta_3=c_1 \tag{53}$$

$$\cos\theta_1\dot\theta_2-\sin\theta_1\cos\theta_2\dot\theta_3=c_2 \tag{54}$$

$$\sin\theta_1\dot\theta_2+\cos\theta_1\cos\theta_2\dot\theta_3=c_3 \tag{55}$$

Inversion of the matrix **Q** gives

$$\mathbf{Q}^{-1}=\begin{bmatrix} 1 & 0 & \sin\theta_2 \\ 0 & \cos\theta_1 & -\sin\theta_1\cos\theta_2 \\ 0 & \sin\theta_1 & \cos\theta_1\cos\theta_2 \end{bmatrix}^{-1}=\begin{bmatrix} 1 & \sin\theta_1\tan\theta_2 & -\cos\theta_1\tan\theta_2 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\dfrac{\sin\theta_1}{\cos\theta_2} & \dfrac{\cos\theta_1}{\cos\theta_2} \end{bmatrix} \tag{56}$$

and the differential equations become

$$\dot{\boldsymbol\theta}=\mathbf{Q}^{-1}\mathbf{C}_z^{-1}\mathbf{h}_z=\mathbf{Q}^{-1}\mathbf{c} \tag{57}$$

or explicitly

$$\dot\theta_1=c_1+\sin\theta_1\tan\theta_2 c_2-\cos\theta_1\tan\theta_2 c_3 \tag{58}$$

$$\dot\theta_2=\cos\theta_1 c_2+\sin\theta_1 c_3 \tag{59}$$

$$\dot\theta_3=-\frac{\sin\theta_1}{\cos\theta_2}c_2+\frac{\cos\theta_1}{\cos\theta_2}c_3 \tag{60}$$

Acknowledgement

This work has been completed during the author's stay at the Geodetic Institute of the University of Stuttgart, with the financial support of the Alexander von Humboldt Foundation, which is gratefully acknowledged.

References

Dermanis, A. (1995): *The Nonlinear and Space-time Geodetic Datum Problem*. Presented at the International Meeting "Mathematische Methoden der Geodäsie", Mathematisches Forschungsinstitut Oberwolfach, 1-7 Oct. 1995.

Grafarend, E. (1974). Optimization of Geodetic Networks. *Bollettino di Geodesia e Scienze Affini*, 33, 4, 351-406.

Meissl, P. (1965): Über die Innere Genauigkeit dreidimensionaler Punkthaufen. *Zeitschrift für Vermessungswesen*, 90, 4, 109-118.

Meissl, P. (1969): Zusammenfassung und Ausbau der inneren Fehlertheorie eines Punkthaufens. *Deutsche Geodätische Kommission*, Reihe A, Nr. 61, 8-21.

Moritz, H. & I.I. Mueller (1987): *Earth Rotation. Theory and Observation*. Ungar, New York.

Munk, W.H. & G.J.F. MacDonald (1960): *The Rotation of the Earth. A Geophysical Discussion*. Cambridge University Press.

From Elliptic Arc Length to Gauss-Krüger Coordinates by Analytical Continuation

Egon Dorrer

Prologue

The majority of contemporary geodesists[1] considers Eric Grafarend as a most remarkable and outstanding scientist as well as brilliant scholar in the field of geodesy. His strong opinions, founded upon a thorough understanding of mathematical reasoning as applied to geodetic science, his clear views on essentials and needs for a science oriented university education with emphasis on fundamentals, and his openness to other fields and different cultures has gained him many friends throughout the world. In fact, Eric Grafarend's creative and productive power is enormous and can hardly be equaled. His background easily has enabled him investigating new and old problems of our common profession from a purely theoretical, highly mathematical point of view. This is in contrast to the majority of geodesists who consider geodesy as engineering science rather than (geo-) science *per se*, and which should be primarily oriented towards practical applications. Although, partly for that reason, most of Eric Grafarend's scientific publications are beyond the comprehension for those *other* geodesists, I nevertheless consider his work essential for a theoretically sound deepening of our profession.

Considering myself one of those *other* geodesists, I would not even think of attempting to write something that could come close to the quality of the level of Eric Grafarend's publications in terms of mathematical rigor and degree of abstraction. Let alone, I would not succeed anyhow. Yet, both of us seem to have a few convictions in common that encourage me to a contribution dedicated to him on occasion of his 60th birthday even knowing it would not meet his high scientific standards. In my scientific endeavor one of my peculiarities always has been a certain desire, if not longing, for search and investigation of novel ideas and unconventional methods or techniques, even though nothing spectacular would have to be expected from the final results. It is somewhat strange that my main interests were concentrated not so much on the outcome of a certain study but often rather more on the way leading to a solution of the problem. I think Eric Grafarend's way of living for science is not too far off from such an attitude – one level higher, of course. My modest contribution will then not be entirely in vain.

1 Introduction

It remains one of the mysteries in geodesy why most of the differential geometric relations on the spheroid (ellipsoid of revolution) always have been developed into truncated power series for numerical computations. Although understandable from a historical point of view when all calculations had to be carried out by hand, there is no reason why this should be done the same way today with computers. A typical example is the computation of Gauss-Krüger coordinates or UTM-coordinates (see,

[1] For reason of simplicity and in accordance with Eric Grafarend's understanding, in this paper the definition of *geodesy* is adopted according to the European view, i.e. encompassing the entire spectrum of fields of expertise of *surveying engineering* (a new term is *geomatics*), even though the author, who does not entirely agree with this, has difficulties in finding his own specialization (photogrammetry and remote sensing) properly represented under this name.

e.g., [Hubeny, 1953]). Virtually all existing software routines employ algorithms derived solely from incomplete power series that were developed ages ago and for regional use only, and nobody asks anymore if there existed more general, universal and mathematically sound algorithms. For the same reason I never really could understand why in geodesy hardly any complex numbers are used and why practically all geodesists – exceptions, of course, prove the rule – prefer to circumvent complex arithmetic despite their claims that conformal mappings on the spheroid are essential. It is a fact, however, that conformal mappings such as the Gauss-Krüger projection are based on and easiest represented by complex numbers, and algorithms written in computer languages containing complex arithmetic turn out to be rather short, effective and transparent.

When we notice departures from this line then mostly those originating from non-geodesists or outright *outsiders*. E.g., in [Klotz, 1993] efforts are undertaken to extend truncated power series from local to global by recursive definitions; the treatise [Lee, 1976] elaborates on conformal projections based on elliptic functions and integrals; and in [Gerstl, 1984] numerical evaluations of (complex) elliptic integrals are performed by Landen transformations. Seemingly unnoticed in the geodetic community and despite their innovative character, these research studies have remained in a somewhat dormant state. In a way, this is rather unfortunate because of the knowledge we are carelessly throwing away which, on the other hand, would be of considerable value and help for a better understanding of and insight in a truly geodetic matter. Really disillusioning is, in my opinion, that the mathematical relevance of transitions from real to complex by *analytic continuation,* since inherently having practical consequences, are rarely understood by many geodesists. E.g., who knows that any valid, in this case real, mathematical formulation for the arc length along the meridian of a spheroid as function of the (real) isometric latitude, immediately yields (conformal) Gauss-Krüger coordinates if the quantities used are extended to the complex domain.

The author remembers with horror the lectures on "Landesvermessung" when his teacher, with a relatively high degree of dilettantism, tried to explain both nature and background of the Gauss-Krüger projection and derive their mathematical relations. Instead of having kept to the simple essentials, the matter submerged into a sea of obscurity, and it would take the author many years of own search until he became sufficiently confident in comprehending the subject. The following paragraphs are to present the findings and results of the author's work as an outsider to the whole matter. By virtue of his understanding of belonging to an engineering science, main emphasis will ultimately be placed upon practical applicability rather than theoretical rigor. This leads to the presentation of not only general formulations and algorithms derived therefrom but also genuine yet simple and immediately applicable computer programs. Although not new in mathematical literature, the numerical evaluation of elliptic integrals of the second and third kind, essentially defining the arc length on the meridian of a spheroid, will be based entirely on the highly convergent Landen transformation. Whether the geodetic community finally will appreciate this or not remains to be seen. While this topic forms the kernel of the paper, the transition to Gauss-Krüger coordinates represents but a mere extension from real to complex numbers without modification of the algorithms.

2 Elliptic Integrals

The radius of curvature μ of the spheroid with semi-major axis $a = 1$ in the direction of the meridian at a point of geographic latitude φ is given by

$$\mu = \mu(k,\varphi) = \frac{1-k^2}{(1-k^2\sin^2\varphi)^{3/2}} \tag{1}$$

where k is the (first) numerical eccentricity of the meridian (often denoted e or ε in geodetic literature). An element of length $d\zeta$ along the meridian is then given by

$$d\zeta \;= \mu(k,\varphi)\; d\varphi\;. \tag{2}$$

Hence by integration we get

$$\zeta \;=\; \int_0^{\varphi} \mu(k,\varphi)\, d\varphi \;=\; \int_0^{\varphi} \frac{1-k^2}{(1-k^2\sin^2\varphi)^{3/2}}\, d\varphi \tag{3}$$

for the arc length normalized to $a=1$ (denoted *arc latitude* here). If, instead, we use the *reduced latitude* τ defined by

$$\tan\tau \;= \sqrt{1-k^2}\,\tan\varphi \;=\; k'\tan\varphi, \tag{4}$$

where k' is termed complementary modulus, then arc latitude is given by

$$\zeta \;=\; \int_0^{\tau} \sqrt{1-k^2\cos^2\tau}\, d\tau \;= \tag{5a}$$

$$= \int_0^{\frac{\pi}{2}-\tau} \sqrt{1-k^2\sin^2(\frac{\pi}{2}-\tau)}\; d(-\tau) \;= \tag{5b}$$

$$= \int_{\tau}^{\frac{\pi}{2}} \sqrt{1-k^2\sin^2\tau}\, d\tau\,. \tag{5c}$$

Equation (5b) is equivalent to Legendre's normal (incomplete) elliptic integral of the second kind [Korn et al., 1968] defined by

$$E_{(2)}(k,\theta) \;=\; \int_0^{\theta} \sqrt{1-k^2\sin^2\theta}\; d\theta \tag{6}$$

thus yielding, together with (5c), the simple relation

$$\boxed{\begin{aligned} \zeta \;&=\; E_{(2)}(k,\tfrac{\pi}{2}) - E_{(2)}(k,\tau) \;= \\ &=\; C_{(2)}(k) - E_{(2)}(k,\tau) \end{aligned}} \tag{7}$$

for arc latitude as function of reduced latitude, viz. as difference between complete and incomplete elliptic integral of the second kind.

Equation (3) is proportional to Legendre's normal (incomplete) elliptic integral of the third kind, generally defined by

$$E_{(3)}(n,k,\theta) \;=\; \int_0^{\theta} \frac{d\theta}{(1-n\sin^2\theta)\sqrt{1-k^2\sin^2\theta}}\,, \tag{8}$$

yet specialized to the case $n = k^2$, viz.

$$\dot{E}_{(3)}(k,\theta) = \int_0^{\theta} \frac{d\theta}{(1-k^2\sin^2\theta)^{3/2}} \tag{9}$$

thus giving rise to the relation

$$\boxed{\zeta = (1-k^2)\cdot\dot{E}_{(3)}(k,\varphi)} \tag{10}$$

for arc latitude as function of geographic latitude. Since the special elliptic integral (9) of the third kind can be expressed in terms of the elliptic integral of the second kind [Korn et al., 1968] according to

$$\dot{E}_{(3)}(k,\theta) = \frac{1}{k'^2}\left(E_{(2)}(k,\theta) - \frac{k^2\sin\theta\cos\theta}{\sqrt{1-k^2\sin^2\theta}}\right), \tag{11}$$

equation (10) may be rewritten in the final form

$$\boxed{\zeta = E_{(2)}(k,\varphi) - \frac{k^2\sin\varphi\cos\varphi}{\sqrt{1-k^2\sin^2\varphi}}}\,. \tag{12}$$

In order to complete this topic, the elliptic integral of the first kind is defined by

$$E_{(1)}(k,\theta) = \int_0^{\theta} \frac{d\theta}{\sqrt{1-k^2\sin^2\theta}}\,. \tag{13}$$

It follows from the foregoing that the arc latitude for a meridian can directly be determined from the elliptic integral of the second kind. This is relevant not only from a formal and theoretical point of view, shown by the equations (7) and (12), but also from a practical, application oriented one, even though elliptic integrals cannot be solved in closed form. The reason for this statement becomes immediately evident if properly defined software function routines are available for numerical evaluations of these integrals. To the author's knowledge such functions have never been explicitly utilized in geodetic works. Formulations and algorithms for the meridian arc length are exclusively based on expansions into power series in terms of the modulus, either truncated after a few terms whenever a certain accuracy level has been reached, or as recursive algorithm [Klotz, 1993]. The latter approach is global rather than local and should therefore be preferred, even though convergence problems may arise in case of long arcs and/or large values of the modulus. Since the Earth's spheroidal eccentricity is small no serious practical problems need to be expected.

The following paragraph deals with yet another numerical approach for the evaluation of elliptic integrals which, though well known in other disciplines where elliptic integrals frequently occur, has obviously not been adopted in geodetic literature. This approach is peculiar to elliptic integrals and elliptic functions and works only there.

3 Landen Transformation

The numerical evaluation of elliptic integrals is favorably attained via a sequence of so-called Landen transformations [Tricomi, 1948]. These are second-order periodic transformations causing the

modulus k to converge quadratically towards zero without effecting certain normal forms of the integrals. The transformations have been employed with advantage in [Bulirsch, 1965] for the evaluation of real elliptic integrals. [Gerstl, 1984] succeeded in generalizing these elegant algorithms to complex integrals, thus showing, for the first time, their practical use for calculating (conformal) Gauss-Krüger coordinates.

The principle of Landen's transformation is best understood if we consider it for the complete elliptic integral of the first kind, viz.

$$C_{(1)}(k) \;=\; E_{(1)}(k,\frac{\pi}{2}) \;=\; \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1-k^2\sin^2\theta}} \;. \tag{14}$$

By choosing different moduli [Korn et al., 1968], a series of interrelations, or transformations, of this integral exist. In particular, by virtue of the relation [Erdelyi et al., 1953]

$$C_{(1)}(k) \;=\; \frac{2}{1+k'}\cdot C_{(1)}\left(\frac{1-k'}{1+k'}\right) \tag{15}$$

a recursive process can be initiated with successively decreasing moduli to one whose magnitude is negligible, i.e. when

$$C_{(1)}(k=0) \;=\; \frac{\pi}{2} \;. \tag{16}$$

Let two successive and corresponding moduli k_{n+1}, k_n' be such that

$$k_{n+1} \;=\; \frac{1-k_n'}{1+k_n'} \;. \tag{17}$$

Thus the step from n to $n+1$ decreases the modulus, and by iterating the process we can descend from a given modulus down to zero. With $k_0 = k$ we obtain successively

$$\begin{aligned} C_{(1)}(k) = C_{(1)}(k_0) &= \frac{2}{1+k_0'} C_{(1)}\left(\frac{1-k_0'}{1+k_0'}\right) = \\ &= \frac{2}{1+k_0'} C_{(1)}(k_1) = \\ &= \frac{2}{1+k_0'}\frac{2}{1+k_1'} C_{(1)}(k_2) = \\ &= \\ &\;\;\vdots \\ &= \prod_{n=0}^{\infty} \frac{2}{1+k_n'}\frac{\pi}{2} \end{aligned} \tag{18}$$

This process, denoted Descending Landen Transformation [Abramowitz et al.,1965], can be conveniently exploited for a recursive algorithm that may be written as follows

$$C_{(1)}(k) = \begin{cases} \dfrac{\pi}{2}, & \text{if } k = 0 \\ \dfrac{2}{1+k'} \cdot C_{(1)}\left(\dfrac{1-k'}{1+k'}\right) & \text{else} \end{cases} . \tag{19}$$

```
[0]   K←CELI1 k;k1
[1]    →(~k=0)/2 ◊ K←○÷2 ◊ →0
[2]    K←(2÷1+k1)×CELI1(1-k1)÷1+k1←○○k
```

Figure 1. Recursive *APL2*-function for complete elliptic integral of the first kind

A realization of this algorithm as recursive *APL2*-function is shown in Fig.1. The rapid convergence of the process is demonstrated in Tab.1 for three very different values of the modulus.

Table 1. Decreasing modulus by and convergence of Descending Landen Transformation of complete elliptic integral of the first kind for three different moduli

k	0.999	0.5	0.08169683121517
0	0.914406543055135	0.0717967697244909	0.0016741848008124
1	0.423693166734227	0.0012920262399948	0.0000007007246689
2	0.049424882188761	0.0000004173332996	0.0000000000001227
3	0.000611451806208	0.0000000000000435	0
4	0.000000009346835	0	
5	0.000000000000002		
6	0		
*C*1	4.49559639584214	1.6857503548126	1.57342723266909

Legendre's incomplete elliptic integral of the first kind (13) can be obtained in analogous fashion solely by expansion to the integral's argument θ. This means generalizing the transformation relation (15) by including the argument. The new relation

$$E_{(1)}(k,\theta) = \frac{E_{(1)}\left(\dfrac{1-k'}{1+k'}, \theta + \arctan k' \tan\theta\right)}{1+k'} \tag{20}$$

(see [Erdelyi et al., 1953]) is capable of initiating a recursive process similar to (18) if, in addition to the iteration statement (17) for two successive moduli, two successive arguments are constrained to

$$\tan(\theta_{n+1} - \theta_n) = k_n' \tan\theta_n \qquad (\theta_{n+1} > \theta_n) . \tag{21}$$

Thus the step from n to n+1 decreases the modulus but increases the amplitude. The iteration process will be terminated if the magnitude of the final modulus is negligible, i.e. when

$$E_{(1)}(k = 0, \theta) = \theta . \tag{22}$$

```
[0] F←k ELI1 p;k1;q
[1]  →(~k=0)/2 ◊ F←p ◊ →0
```

```
[2]  F←(((1-k1)*1+k1)ELI1(0⌊0.5-(q-2×p)*01)+q←p+¯30k1×30p)*1+k1←00k
```

Figure 2. Recursive *APL2*-function for elliptic integral of the first kind

Equation (20) is part of a recursive algorithm analogous to (19) which is realized as *APL2*-function in Fig.2, yet completed to guarantee increasing amplitudes.

Transformations for the incomplete elliptic integral of the second kind (6) also are given in [Erdelyi et al., 1953]. The one particularly useful for numerical computations is the relation

$$E_{(2)}(k,\theta) = \frac{1+k'}{2}\left(E_{(2)}(\dot{k},\dot{\theta}) + \dot{k}\sin\dot{\theta}\right) - k' E_{(1)}(k,\theta) \qquad (23)$$

$$\text{where} \quad \dot{k} = \frac{1-k'}{1+k'} \quad \text{and} \quad \dot{\theta} = \theta + \arctan k' \tan\theta \ .$$

A realization of the corresponding recursive algorithm as *APL2*-function is shown in Fig. 3.

```
[0] E←k ELI2 p;k1;k¯;p¯
[1]  →(~k=0)/2 ◊ E←p ◊ →0
[2]  k¯←(1-k1)*1+k1←00k ◊ p¯←(0⌊0.5-(p¯-2×p)*01)+p¯←p+¯30k1×30p
[3]  E←(((k¯ ELI2 p¯)+k¯×10p¯)×(1+k1)*2)-k1×k ELI1 p
```

Figure 3. Recursive *APL2*-function for elliptic integral of the second kind

Finally, Fig.4 exhibits an *APL2*-function for elliptic arc latitude according to (12). The convergence behavior for the incomplete elliptic integrals is identical to that of the complete elliptic integral of the first kind previously discussed in detail. The numerical correctness of all these algorithms may be

```
[0]    a←k ELARC p
[1]     a←(k ELI2 p)-(k×k×102×p)*2×00k×10p
```

Figure 4. *APL2*-function for elliptic arc length

confirmed by calculating the arc length (in m) for the four geographic latitudes 30°, 45°, 60° and 90° on Bessel's spheroid (a = `6377397.155`, k = `0.08169683121517`):

```
      a×k ELARC 2 rad 30 45 60 90
3319786.50954331 4984439.26547085 6653376.12061161 10000855.7644355
```

Theoretically, these results ought to be accurate to 14 or 15 figures.

4 From Real Arc Length to Complex Gauss-Krüger

Since Legendre's elliptic integrals are valid also for complex arguments, they can be utilized for the determination of Gauss-Krüger coordinates. This stems from the fact that Gauss-Krüger coordinates belong to a conformal projection that preserves the scale on the principal meridian. Thus by analytic (or complex) continuation of the meridian arc length, i.e. by expanding one-dimensional arc length to the two-dimensional spheroidal surface, relation (12), inherently containing the elliptic integral of the second kind, will now, as analytic function, offer a conformal transformation between a complex Gauss-Krüger variable and a complex latitude variable yet to be defined. Unfortunately, as is known,

geographic latitude and longitude (φ,λ) do not possess properties specific for isometric (isothermal) coordinates required for conformal projections (see e.g. [Hubeny, 1953]). Only after transformation of φ into a so-called *isometric latitude* ψ, which can be carried out in closed form by the relation [Klotz, 1993]

$$\begin{aligned} \psi &= \operatorname{arsinh}\tan\varphi - k\tanh k\sin\varphi = \\ &= \operatorname{lam}(k,\varphi) \end{aligned} \quad , \tag{24}$$

are isometric coordinates obtained by the pair (ψ,λ). The relation on the first line in (24) is a slight, yet numerically more stable, modification of the one given in [Klotz, 1993]. The function $\operatorname{lam}(k,\varphi)$, taken from [Lee, 1976], denotes the *Lambertian* of φ for a certain elliptic eccentricity k, and is customarily called the *inverse Gudermannian* in Englisch literature and denoted by $\operatorname{gd}^{-1}\varphi$ for the sphere following Cayley's use of $\operatorname{gd} u$ to denote the *Gudermannian*. The latter is here denoted by $\operatorname{lam}^{-1} u$ for the sphere, i.e. $\operatorname{lam}^{-1}(k,\psi)$ as inverse within the context of (24). It cannot be represented in closed form but must be evaluated by iteration [Klotz, 1993].

The isometric coordinate pair (ψ,λ) can now conveniently defined as complex variable

$$\Lambda = \lambda + j\psi \tag{25}$$

denoted *complex longitude* or *Mercator variable* [Gerstl, 1984]. A conformal mapping represented by Λ and the identity as analytical function of Λ has uniform scale on the equator. Application of the inverse Lambertian to $j\overline{\Lambda} = \psi + j\lambda$ offers the definition of a *complex latitude*

$$\Phi = \operatorname{lam}^{-1}(k, j\overline{\Lambda}) \tag{26}$$

as analytic continuation of geographical latitude from the initial meridian into the spheroid's surface. Finally, the *complex arc length* $Z = \xi + j\eta$ as analytic continuation of (real) arc length ζ can be determined from relation (12), represented as function *elarc* in Fig.4

$$Z = \operatorname{elarc}(k,\Phi) . \tag{27}$$

Z coincides with the arc length on the initial meridian, and the conformal mapping of the spheroid offered by Z is identical to Gauss-Krüger's projection. Real part and imaginary part of Z define Gauss-Krüger coordinates. Hence, starting from complex longitude, Gauss-Krüger coordinates can be obtained directly from the compound relation

$$\xi + j\eta = \operatorname{elarc}(k,\operatorname{lam}^{-1}(k,\operatorname{lam}(k,\varphi) + j\lambda)) . \tag{28}$$

In the context of this treatise, relation (28) practically is in closed form. Its compactness and numerical generality can be considered unequaled by any other algorithm. Simply by inverting (28) we are able to determine geographic coordinates from Gauss-Krüger coordinates, viz.

$$\begin{aligned} \psi + j\lambda &= \operatorname{lam}(k,\operatorname{elarc}^{-1}(k,\xi + j\eta)) \\ \varphi &= \operatorname{lam}^{-1}(k,\psi) \end{aligned} \tag{29}$$

In order to demonstrate the validity of relation (28) a numerical example taken from [Klotz, 1993] has been calculated. With the geographic coordinates 52° and 30° for latitude and longitude on the Inter-

national ellipsoid, corresponding Gauss-Krüger coordinates are obtained by calling the *APL2*-function *GK* representing (28). Without going into details, the obtained result

```
      (a k)GK 2 rad 52 30
6200529.35513598J2033568.76509429
```

agrees with [Klotz, 1993] exactly to the 5 decimal figures given there.

The functions and procedures derived or stated in this paper can, of course, be easily applied to transformations of Gauss-Krüger coordinates from one meridian system to another, to manipulations of UTM-coordinates, etc. The relations investigated and used are based on sound mathematical grounds and stable, globally applicable numerical processes. Appreciation by the geodetic community in one way or another is hoped.

5 Literature and References

Abramowitz et al. (eds), (1965): Handbook of Mathematical Functions, Dover Publ., New York.

Bulirsch, R., (1965): Numerical calculation of elliptic integrals and elliptic functions. Num.Math. 7, 78-90.

Erdelyi et.al. (1953): Higher Transcendental Functions.Vol.2. McGraw-Hill, New York.

Gerstl, M. (1984): Die Gauss-Krügersche Abbildung des Erdellipsoides mit direkter Berechnung der elliptischen Integrale durch Landentransformation. DGK Reihe C, Heft. Nr. 296, München.

Hubeny, K. (1953): Isotherme Koordinatensysteme und konforme Abbildungen des Rotationsellipsoids. Österr. Verein f. Vermessungswesen, Sonderheft 13, Wien, 208 S.

Klotz, J. (1993): Eine analytische Lösung der Gauss-Krüger-Abbildung. ZfV 3/1993, 106-116.

Korn et al., (1968): Mathematical Handbook for Scientists and Engineers, McGraw-Hill, New York

Lee, L.P. (1976): Conformal Projections Based on Elliptic Functions. CARTOGRAPHICA, Monograph No. 16, York University, Toronto, Canada, 128 pages.

Tricomi, F., (1948): Elliptische Funktionen. Akademische Verlagsgesellschaft.

Tests of two forms of Stokes's integral using a synthetic gravity field based on spherical harmonics

Will E Featherstone

Abstract

Two gravimetric models of the geoid over Western Australia have been constructed using modified forms of Stokes's formula. The input data are synthetic gravity anomalies which have been generated by an artificial extension of the EGM96 global geopotential model to spherical harmonic degree and order 2700. This provides self-consistent sets of gravity anomalies and geoid heights, which are used as control on the effectiveness of a deterministically modified Stokes's kernel in relation to the common remove-restore technique with the spherical Stokes's kernel. The improved fit of the geoid model that uses a modification to allow for the neglect of the truncation error term and adapt its filtering properties indicates that the widely used remove-compute-restore approach is less appropriate for gravimetric geoid computation in the high-frequency band over Western Australia.

1 Introduction

In 1849, G. G. Stokes published a solution to the geodetic boundary value problem, which requires a global integration of gravity anomalies over the whole Earth to compute the separation (N) between the geoid and reference ellipsoid (Stokes, 1849). However, the incomplete global coverage and availability of accurate gravity measurements has precluded an exact determination of the geoid using Stokes's formula. Instead, an approximate solution is used in practice, where only gravity data in and close to the computation area are used. This approach is also attractive due to the increase in computational efficiency that is offered by working with a smaller integration area.

In 1958, M. S. Molodensky (cited in Molodensky *et al.*, 1962) proposed a modification to Stokes's formula to reduce the truncation error that results when gravity data are used over a limited area. However, Molodensky's modification did not receive a great deal of attention in practical geoid computations at that time because of the contemporaneous availability of low-frequency global gravity field information, derived from the analysis of the artificial Earth satellite orbits. These global geopotential models, expressed in terms of fully normalised spherical harmonics, are now routinely used in conjunction with terrestrial gravity data via a truncated form of Stokes's integral (eg. Vincent and Marsh, 1973; Sideris and She, 1995).

Assuming that the global geopotential model is a perfect fit to the low-degree terrestrial gravity field, this combined approach reduces the magnitude of the truncation error. This is because its Fourier series expansion begins at a higher degree, where the truncation coefficients are smaller in magnitude and the geopotential coefficients are expected to converge (cf. Grafarend and Engels, 1994). Another advantage of this combined solution is that it reduces the impact of the spherical approximation inherent to the derivation of Stokes's integral (eg. Heiskanen and Moritz, 1967); the reason being that most of the geoid's power is contained in the low-frequency band.

A formal description of the combination of a global geopotential model with terrestrial gravity

data has been proposed by Vaníček and Sjöberg (1991), which they refer to as the generalised Stokes scheme for geoid computation. Importantly, this satisfies a solution to the geodetic boundary value problem when formulated for a higher than second-degree reference model (Martinec and Vaníček, 1997). In this generalised scheme, the low-frequency geoid undulations, computed from a global geopotential model (N_M), are extended into the high frequencies by a global integration of complementary high-frequency terrestrial gravity anomalies (Δg^M). This is written as

$$N = N_M + \kappa \int_0^{2\pi} \int_0^{\pi} S^M(\cos\psi)\ \Delta g^M\ \sin\psi\ d\psi\ d\alpha \tag{1}$$

where $\kappa = R/4\pi\gamma$, R is the spherical Earth radius, γ is normal gravity evaluated on the surface of the reference ellipsoid as required by Bruns's formula (eg. Heiskanen and Moritz, 1967), ψ and α are the coordinates of spherical distance and azimuth angle about the computation point, respectively, and $S^M(\cos\psi)$ is the spheroidal form of Stokes's kernel, which is implicit to the generalised scheme, and has the series expansion

$$S^M(\cos\psi) = \sum_{n=M+1}^{\infty} \frac{2n+1}{n-1}\ P_n(\cos\psi) \tag{2}$$

where $P_n(\cos\psi)$ is the n-th degree Legendre polynomial.
In Eq. (1), the low-frequency component of the geoid undulation (N^M) can be computed from the spherical harmonic coefficients that represent the global geopotential model according to

$$N_M = \frac{GM}{r\gamma} \sum_{n=2}^{M} \left(\frac{a}{r}\right)^n \sum_{m=0}^{n} (\delta\overline{C}_{nm} \cos m\lambda + \overline{S}_{nm} \sin m\lambda)\overline{P}_{nm}(\cos\theta) \tag{3}$$

The corresponding high-frequency gravity anomalies (Δg^M) are evaluated by subtracting the same spherical harmonic degrees of the same global geopotential model from the terrestrial gravity anomalies (Δg) according to

$$\Delta g^M = \Delta g - \frac{GM}{r^2} \sum_{n=2}^{M} \left(\frac{a}{r}\right)^n (n-1) \sum_{m=0}^{n} (\delta\overline{C}_{nm} \cos m\lambda + \overline{S}_{nm} \sin m\lambda)\overline{P}_{nm}(\cos\theta) \tag{4}$$

In Eqs. (3) and (4), GM is the product of the Newtonian gravitational constant and mass of the solid Earth, oceans and atmosphere, a is the equatorial radius of the geocentric reference ellipsoid, (r, θ, λ) are the geocentric polar coordinates of each computation point, $\delta\overline{C}_{nm}$ and $\overline{S}_{nm}$ are the fully normalised geopotential coefficients of degree n and order m, which have been reduced by the even zonal harmonics of the reference ellipsoid, and $\overline{P}_{nm}(\cos\theta)$ are the fully normalised associated Legendre functions. It is assumed that the zero and first degree harmonic terms are inadmissible (eg. Heiskanen and Moritz, 1967).
The degree of spheroid (M) used for the generalised Stokes scheme can be chosen as the maximum degree of global geopotential model available, which is usually $M_{max} = 360$. However, there are more important considerations than simply taking the maximum degree of expansion available (eg. Featherstone, 1992). Firstly, the $M_{max} = 360$ models are constructed from both satellite-derived and terrestrial gravity data. Therefore, in many regional geoid computations, the same terrestrial gravity data are used twice in Eq. (1). Clearly, this introduces the correlation of errors between these data, which are rarely accounted for nor even acknowledged by most authors.
Another consideration is the leakage of low-frequency errors from the terrestrial gravity data into the combined solution for the geoid, much of which can be filtered out by the spheroidal kernel in Eq. (2) (Vaníček and Featherstone, 1998). This is considered to be a desirable scenario, because the low-frequency geopotential coefficients are currently the best source of this information, whereas terrestrial gravity anomalies are subject to low-frequency errors. Therefore, choosing

the degree of spheroid at, say, $M = 20$ (Vaníček and Kleusberg, 1987), which is probably the limit of the reliable resolution of the satellite-derived geopotential coefficients (notwithstanding resonant terms), avoids the correlations and reduces the leakage of terrestrial gravity anomaly errors.

2 Reduction of the Approximation Error

When high-frequency terrestrial gravity anomalies are used over a limited area, the generalised Stokes scheme becomes subject to a truncation error. Accordingly, there is an adjustment of Eq. (1) that involves limiting the integration domain to a spherical cap, bound by the spherical distance ψ_0 ($0 < \psi_0 < \pi$), which yields the approximation

$$\hat{N} \simeq N_M + \kappa \int_0^{2\pi} \int_0^{\psi_0} S^M(\cos\psi)\ \Delta g^M\ \sin\psi\ d\psi\ d\alpha \tag{5}$$

with a corresponding truncation error of

$$\delta N = \kappa \int_0^{2\pi} \int_{\psi_0}^{\pi} S^M(\cos\psi)\ \Delta g^M\ \sin\psi\ d\psi\ d\alpha \tag{6}$$

such that $N = \hat{N} + \delta N$. This truncation error can be expressed as a series expansion (eg. Vaníček and Featherstone, 1998) by

$$\delta N = 2\pi\kappa \sum_{n=M+1}^{\infty} Q_n^M(\psi_o)\ \Delta g_n \tag{7}$$

where the truncation coefficients

$$Q_n^M(\psi_o) = \int_{\psi_o}^{\pi} S^M(\cos\psi)\, P_n(\cos\psi)\ \sin\psi\, d\psi \tag{8}$$

can be evaluated using the algorithms of Paul (1973), and the n-th degree surface spherical harmonic of the gravity anomaly can be evaluated from the global geopotential model

$$\Delta g_n = \frac{GM}{r^2}\left(\frac{a}{r}\right)^n (n-1) \sum_{m=0}^{n} (\delta\overline{C}_{nm}\cos m\lambda + \overline{S}_{nm}\sin m\lambda)\overline{P}_{nm}(\cos\theta) \tag{9}$$

Therefore, the truncation error terms can be computed in the region $M \leq n \leq M_{max}$. If this is done, the truncation error then reduces to

$$\delta N = 2\pi\kappa \sum_{n=M_{max}+1}^{\infty} Q_n^M(\psi_o)\ \Delta g_n \tag{10}$$

However, if $\Delta g^M \neq 0$ ($2 \leq n \leq M$), the start of the series expansions in Eqs. (7) and (10) no longer hold, then there is a leakage of any low-frequency errors in the gravity data into the low-frequency geoid solution (when the integration is performed over a limited area; Vaníček and Featherstone, 1998). This is a direct consequence of the approximation of the generalised Stokes integral (Eq. 5), or any other gravity field convolution integral. Since Δg_n only depend on the physical properties of the Earth, it remains necessary to seek a modification of Stokes's integral that reduces the magnitude of the truncation error.

However, the common remove-compute-restore technique for the combined solution for the geoid (eg. Torge, 1991) makes no attempt to modify the integration kernel and thus reduce the truncation error or adapt its filtering properties. Instead, this scheme uses the spherical kernel as originally introduced by Stokes, which is

$$S(\cos\psi) = \sum_{n=2}^{\infty} \frac{2n+1}{n-1}\ P_n(\cos\psi) \tag{11}$$

Moreover, the remove-compute-restore approach generally uses the maximum degree (usually $M_{max} = 360$) of a global geopotential model to compute the residual gravity anomalies (Eq. 4). Accordingly, there is a disparity between the degree of the geopotential model and Stokes's kernel. The combined solution for the geoid in the remove-compute-restore scheme is thus written as

$$\hat{N}_1 \simeq N_{M_{max}} + \kappa \int_0^{2\pi} \int_0^{\psi_0} S(\cos\psi)\ \Delta g^{M_{max}}\ \sin\psi\ d\psi\ d\alpha \tag{12}$$

where the terms $N_{M_{max}}$ and $\Delta g^{M_{max}}$ are computed from the maximum available degree and order of a global geopotential model. In this combined solution for the geoid, little attempt has been made to reduce the truncation error or adapt the filtering properties of the spherical Stokes's kernel (Eq. 11). Admittedly, the truncation error has been reduced a great deal in the region ($2 \leq n \leq M_{max}$), if and only if the global geopotential model is a good fit to the terrestrial gravity anomalies over the area of interest. Conversely, this is at the expense of allowing any errors in the terrestrial gravity anomalies to propagate, virtually unattenuated, into the combined solution (Vaníček and Featherstone, 1998).

Accordingly, it remains preferable to apply a modification to the truncated form of the generalised Stokes integral (Eq. 5) or the truncated form of the spherical Stokes integral in the remove-compute-restore scheme (Eq. 12) to further reduce the errors associated with these approximations. Since Molodensky's pioneering work, several other authors have proposed modifications to Stokes's (1849) integral. These have been based on different criteria and can be broadly classified as deterministic modifications (eg. Molodensky *et al.* 1962; Wong and Gore 1969; Meissl 1971; Heck and Grüninger 1987; Vaníček and Kleusberg 1987; Vaníček and Sjöberg 1991; Featherstone *et al.* 1998) and stochastic modifications (eg. Wenzel 1982; Sjöberg 1991; Vaníček and Sjöberg 1991). The stochastic modifications, whilst offering an optimal combination (in a least-squares sense) of the data types together with a minimisation of the truncation error, require reliable error estimates of the input data. However, the error characteristics of the terrestrial gravity data are generally unknown, which renders the stochastic modifications of limited practical use. Therefore, the deterministic kernel modifications will have to be relied upon in the interim.

The deterministic kernel modifications can be further divided into two categories: modifications that reduce the truncation error according to some prescribed norm, and modifications that improve the rate of convergence of the series expansion of the truncation error. The modification scheme proposed by Featherstone *et al.* (1998) uses a combination of these, where the rate of convergence of the series expansion of an already-reduced truncation error is accelerated through a combination of the approaches proposed by Vaníček and Kleusberg (1987) and Meissl (1971). Essentially, this modification sets the Vaníček and Kleusberg (1987) kernel to zero at the truncation radius (ψ_0). Alternatively, the truncation radius can be chosen such that it coincides with a zero point of the Vaníček and Kleusberg (1987) kernel. This kernel modification can be written as

$$S_L^M(\cos\psi) = S^M(\cos\psi) - S^M(\cos\psi_0) - \sum_{k=2}^{L} \frac{2k+1}{2}\, t_k(\psi_0)\, [P_k(\cos\psi) - P_k(\cos\psi_0)] \tag{13}$$

where the modification coefficients $t_k(\psi_0)$ are computed from the solution of the following linear system of $L-1$ equations

$$\sum_{k=2}^{L} \frac{2k+1}{2}\, t_k(\psi_0)\, e_{nk}(\psi_0) = Q_n^M(\psi_0) \tag{14}$$

with

$$e_{nk}(\psi_0) = \int_{\psi_0}^{\pi} P_n(\cos\psi)\, P_k(\cos\psi) \sin\psi\, d\psi \tag{15}$$

which can be evaluated using the recursive algorithms of Paul (1973). The degree of this kernel modification (L) can be chosen to be greater than, equal to or less than the degree of the geopotential model (M) in the generalised Stokes formula (Eq. 5). However, if $L > M$, additional terms arise that account for this disparate combination and should be computed or their omission acknowledged.

The combined solution for the geoid considered in this study attempts to reach a compromise of the above two schemes, based on considerations of the data availability, their expected reliability and a reduction of the truncation error through the above deterministic modification of the generalised Stokes kernel. This compromise approach was used to compute the recent Australian gravimetric geoid model, AUSGeoid98 (Johnston and Featherstone, 1998). Mathematically, this is formalised as

$$\hat{N}_2 \simeq N_{M_{max}} + \kappa \int_0^{2\pi} \int_0^{\psi_0} S_L^M(\cos\psi)\ \Delta g^{M_{max}}\ \sin\psi\ d\psi\ d\alpha \tag{16}$$

where all terms have been defined earlier.

This utilises the maximum available expansion of the global geopotential model in conjunction with a low-degree deterministic kernel modification. This approach aims at reducing the truncation error so that it can be ignored, whilst relying more on the low-degree satellite solution by filtering a proportion of the low-frequency errors from the terrestrial gravity data. Empirical studies by Featherstone (1992) indicate that the modified kernels become numerically unstable for large L and small ψ_0, which enforces a low degree of kernel modification when a small integration radius is used. For simplicity, the degree of kernel modification is chosen equal to the degree of spheroid used in the generalised scheme (ie. $L = M = 20$). The integration radius was chosen to be $\psi_0 = 1°$, since this value was empirically selected for AUSGeoid98 (Johnston and Featherstone, 1998).

It is argued that this offers a geoid solution that is superior to the current remove-compute-restore approach because of its further reduction of the truncation error and adaption of the filtering properties of the kernel. However, it is also important to acknowledge the deficiencies of this attempted compromise, which are the reliance on the high-frequencies in the global geopotential model (which can contain 80% noise; eg. Lemoine *et al.*, 1998) and the correlations between the terrestrial gravity anomalies in the region $20 \leq M \leq 360$.

3 Tests with a synthetic gravity field in Western Australia

In order to compare the validity of the compromise in Eq. (16) and the remove-compute-restore technique (Eq. 12), a synthetic gravity field has been used. The expectation is that by using an error-free, self-consistent set of geoid heights and gravity anomalies, the effectiveness of each combined solution for the geoid can be determined. The approach is as follows: the synthetic gravity anomalies are reduced by the complete expansion of the global geopotential model, these used to compute the geoid according to Eqs. (12) and (16), then these results compared with the synthetic geoid heights. The approach that yields the closest fit to the synthetic geoid is assumed to deliver the better data combination.

In addition, the use of a synthetic gravity field avoids the assumptions and approximations introduced by the treatment of the topography and its density variations. This test is considered preferable to the 'conventional' comparison of gravimetric geoid solutions with the discrete geometrical control afforded by ellipsoidal heights and geodetic levelling. This is because the synthetic field has been generated so that it is uncontaminated by errors in these control data.

3.1 Construction of the synthetic field

The EGM96 global geopotential model (Lemoine *et al.*, 1998), complete to $M_{max} = 360$, has been artificially extended into the higher frequencies to construct the synthetic gravity field over

Western Australia. This is similar to the approach of Tziavos (1996), who used a $M_{max} = 360$ geopotential model to generate self-consistent geoid heights and gravity anomalies to test fast Fourier transform (FFT) based techniques. However, the latter only allowed an evaluation in this frequency band and thus prevented a determination of the performance in the higher frequencies and an assessment of the effect of neglecting the truncation error. In order to construct the synthetic gravity field in the higher frequencies, EGM96 has been artificially extended to spherical harmonic degree and order 2700 by artifically creating geopotential coefficients in the region $361 \leq n \leq 2700$ (cf. Holmes *et al.*, 1998). This upper limit was chosen to be commensurate with a spatial resolution of 4' by 4' and is also the point beyond which the fully normalised associated Legendre polynomials start to become numerically unstable.
The fully normalised EGM96 coefficients in the region $361 \leq n \leq 2700$ were generated by recycling the EGM96 coefficients from the orders in degree 360. To ensure that the degree variance of the synthetic gravity field continued to follow a Kaula-type rule in this extended region, a the artifical coefficients ($\overline{C}^{*}_{nm}$ and $\overline{S}^{*}_{nm}$) were scaled by $(b/r)^{n-360}$, where b is the semi-minor axis length of the reference ellipsoid. From Eq. (3), the synthetic geoid heights are given by

$$N_{syn} = \frac{GM}{r\gamma} \sum_{n=2}^{360} \left(\frac{a}{r}\right)^n \sum_{m=0}^{n} (\delta\overline{C}^{EGM96}_{nm} \cos m\lambda + \overline{S}^{EGM96}_{nm} \sin m\lambda)\overline{P}_{nm}(\cos\theta) + \frac{GM}{r\gamma} \sum_{n=361}^{2700} \left(\frac{a}{r}\right)^n \sum_{m=0}^{n} (\overline{C}^{*}_{nm} \cos m\lambda + \overline{S}^{*}_{nm} \sin m\lambda)\overline{P}_{nm}(\cos\theta) \quad . \tag{17}$$

From Eq. (4), the synthetic gravity anomalies are given by

$$\Delta g_{syn} = \frac{GM}{r^2} \sum_{n=2}^{360} \left(\frac{a}{r}\right)^n (n-1) \sum_{m=0}^{n} (\delta\overline{C}^{EGM96}_{nm} \cos m\lambda + \overline{S}^{EGM96}_{nm} \sin m\lambda)\overline{P}_{nm}(\cos\theta) + \frac{GM}{r^2} \sum_{n=361}^{2700} \left(\frac{a}{r}\right)^n (n-1) \sum_{m=0}^{n} (\overline{C}^{*}_{nm} \cos m\lambda + \overline{S}^{*}_{nm} \sin m\lambda)\overline{P}_{nm}(\cos\theta) \quad . \tag{18}$$

This synthetic field was relatively easy to implement in the existing computer programs for Eqs. (3) and (4). However, its computation becomes quite time consuming for the high degree components. As such, it is likely that the very high-frequency components of a synthetic gravity field will have to be constructed using alternative means, which are currently under investigation.

3.2 Geoid computation via the 1D-FFT technique

In the mid 1980s, the fast Fourier transform (FFT) technique began to find wide-spread use in gravimetric geoid computation because of its efficient evaluation of convolution integrals when compared to quadrature-based numerical integration. For many years, the planar, two-dimensional FFT was used (eg. Schwarz *et al.*, 1990). Strang van Hees (1990) then introduced the spherical, two-dimensional FFT. However, both of these FFT approaches are subject to approximation errors, the most notable of which is the simplification of Stokes's kernel. Therefore, Forsberg and Sideris (1993) proposed the spherical, multi-band FFT, which reduces the impact of the simplified kernel. Haagmans *et al.* (1993) then refined this approach to give the spherical, one-dimensional FFT, which requires no simplification of Stokes's kernel. For this reason, the 1D-FFT has been used in this investigation so that the exact kernels in Eqs. (11) and (13) can be used without the need for a simplification of the kernel.
Another consideration is that remove-compute-restore determinations of the geoid over a region using the FFT often convolve the whole rectangular grid of gravity anomalies with the spherical Stokes kernel (eg. Sideris and She, 1995). Therefore, this implementation is tested in this study, where in Eq. (12) the spherical integration radius (ψ_0) is replaced by the whole gravity data rectangle. Conversely, quadrature-based geoid determinations using numerical integration of

gravity anomalies over a spherical integration radius about each computation point. Therefore, each approach results in a different truncation error due to the neglect of the residual gravity anomalies in the remote zones outside each integration domain.
In order to make the 1D-FFT approach closely mimic quadrature-based numerical integration over a spherical cap, two adaptions of the 1D-FFT approach have been made (Featherstone and Sideris, 1998). The first is the limitation of the integration to a spherical cap by setting the kernel to zero outside the truncation radius (ψ_0) before transformation to the frequency domain. The modified kernel (Eq. 13) was implemented by evaluating it before transformation to the frequency domain. Comparisons with quadrature-based numerical integration software (Featherstone, 1992) were used to verify these adaptions. This approach was used for the computation of AUSGeoid98 (Johnston and Featherstone, 1998), since it allows an efficient evaluation of Eq. (16).

3.3 Comparison of Geoid Results with the Synthetic Model

Equations (17) and (18) were used to construct two, self-consistent 4' by 4' grids of geoid heights and gravity anomalies, respectively, over the region $-11° \leq \phi \leq -37°$ and $112° \leq \lambda \leq 131°$, which covers almost all of the state of Western Australia. These are shown in Figures 1a and 1b and their statistical properties summarised in Table 1. Table 1 also shows the statistical properties of the high-frequency synthetic gravity field, where the $M_{max} = 360$ expansion of EGM96 has been subtracted (cf. Eq. 4).

		max.	min.	mean	st. dev.	rms
total synth. geoid heights	$2 \leq n \leq 2700$	54.979	-40.905	-4.603	22.660	23.123
resid. synth. geoid heights	$361 \leq n \leq 2700$	1.060	-1.061	0.000	0.208	0.208
synth. gravity anomalies	$2 \leq n \leq 2700$	130.459	-188.572	-7.544	34.497	35.312
synth. gravity anomalies	$361 \leq n \leq 2700$	112.531	-122.314	-0.008	21.085	21.085

Table 1. Statistical properties of the synthetic geoid heights (metres) and gravity anomalies (mGal).

The synthetic geoid heights (Eq. 17) were used as control on the tests and the synthetic high-frequency gravity anomalies (Eq. 18; $361 \leq n \leq 2700$) input to the 1D-FFT geoid computation software's implementations of Eqs. (12) and (16). An integration radius of $\psi_0 = 1°$ was used in Eq. (16), since this was the value used in the computation of AUSGeoid98 (Johnston and Featherstone, 1998). No cap radius was specified in Eq. (12) so that the entire gravity data area was used for every geoid computation point. This approach was taken since it replicates the most common FFT-based implementation of the remove-compute-restore technique (eg. Sideris adn She, 1995). The results of the two 1D-FFT geoid computations were compared with the control grid of synthetic geoid heights over the region $-12° \leq \phi \leq -36°$ and $114° \leq \lambda \leq 129°$. This smaller area was chosen so as to eliminate the edge effect associated with the $\psi_0 = 1°$ integration radius. It should be pointed out that this edge effect affects the whole computation area when the cap-radius is unlimited. Nevertheless, the comparisons are conducted over the same area. Table 2 shows a statistical summary of the differences between the control grid of synthetic geoid heights and the results from the 1D-FFT implementations of Eqs. (12) and (16). Figures 1c and 1d show images of these differences, respectively.

		max.	min.	mean	st. dev.	rms
remove-compute-restore	Eq. 12 (ψ_0=1°, $S(\cos\psi)$)	0.058	-0.041	0.008	0.011	0.013
compromise approach	Eq. 16 (ψ_0=π, $S_{20}^{20}(\cos\psi)$)	0.035	-0.035	0.000	0.008	0.008

Table 2. The statistics of the differences between the synthetic control geoid heights and the geoid heights computed from Eqs. (12) and (16) (units in metres).

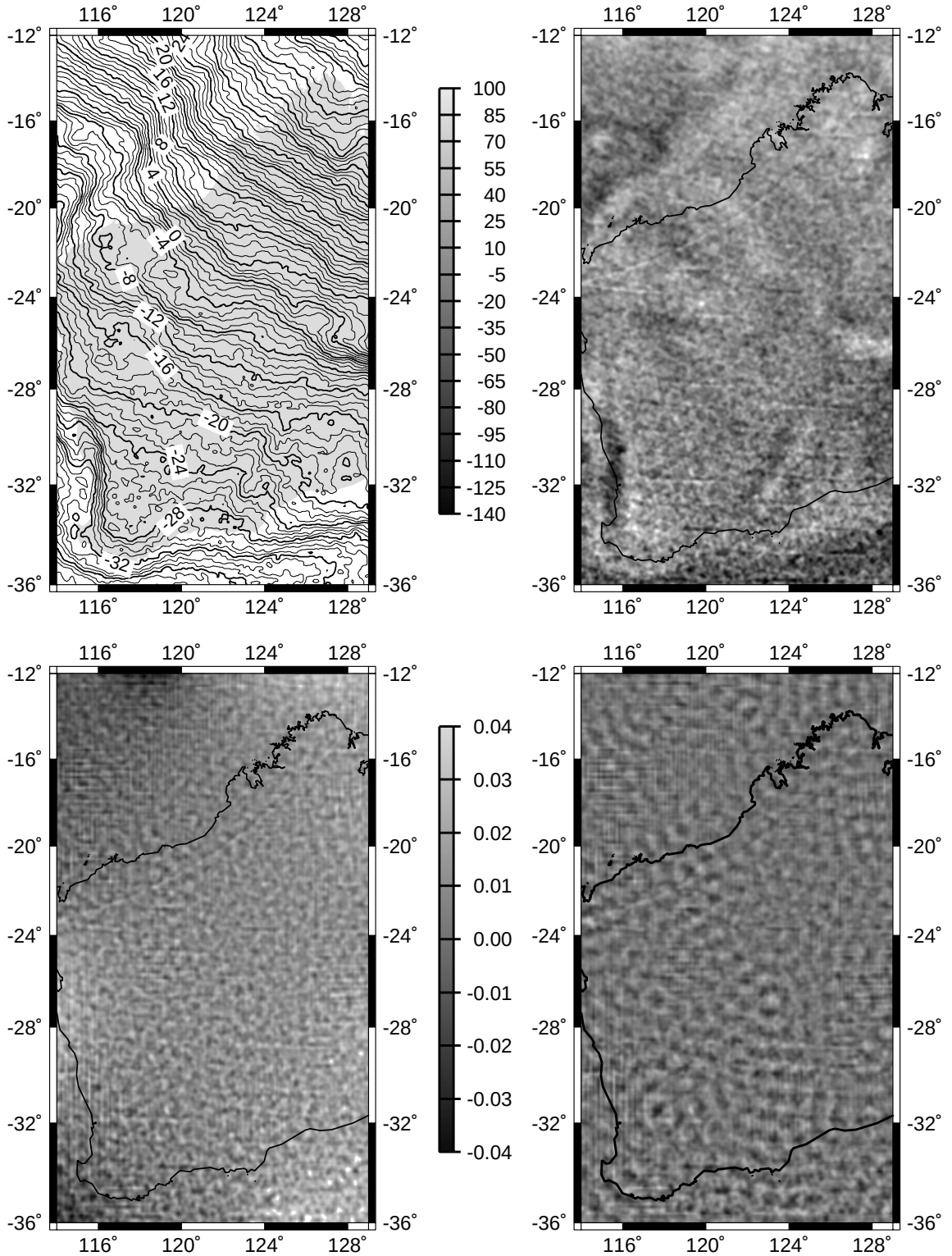


Figure 1. (a) The synthetic geoid heights (m) for $2 \leq n \leq 2700$, (b) The synthetic gravity anomalies (mGal) for $2 \leq n \leq 2700$, (c) The difference (m) between synthetic geoid heights and geoid heights computed from the remove-compute-restore technique (Eq. 12), (d) The difference (m) between synthetic geoid heights and geoid heights computed from the compromise approach (Eq. 16); [Mercator's projection].

4 Discussion, Conclusion and Recommendation

Prior to any discussion, it is essential to point out that the comparisons in Table 2 and between Figures 1c and 1d only consider the effect on the geoid of the neglect of the truncation error and the adaption of the filtering properties by the modified kernel in the high-frequency band ($361 \leq n \leq 2700$). This is because the EGM96 global geopotential model has been used both to construct the synthetic gravity field and produce the residual gravity anomalies in Eq. (4). Accordingly, the filtering and propagation of low-frequency gravity data errors cannot be tested. Future work will introduce low-frequency synthetic data errors in order to study the filtering effects of the kernels in these bands (cf. Vaníček and Featherstone, 1998). Also, using only the high-frequency components has dispensed with the correlations between the data which occur in practice, when using a high-degree, combined global geopotential model.
Nevertheless, the following can be concluded from this band-width-limited study: The improvement offered by the compromised approach in Eq. (16) over the remove-compute-restore approach (Eq. 12) is clearly shown in Table 2. The compromised approach delivers a closer fit to the control grid of geoid heights than does the remove-compute-restore approach. Therefore, the use of the $L = 20$ deterministically modified integration kernel (Eq. 13) over a spherical cap $\psi_0 = 1^\circ$ offers an improvement over the remove-compute-restore technique using the whole computation area. This indicates that the use of a theoretically more appropriate data combination yields better results than simply using more data in the combined solution for the geoid. This is principally because the truncation error has been reduced in size by the kernel modification, thus permitting its neglect, and the filtering properties of the modified kernel lead to a more accurate recovery of the high-frequency geoid undulations. However, due to the considerations described earlier, further work is necessary to quantify their relative effect in other frequency bands so as to replicate the situation in practical geoid computations.

Acknowledgments

I would like to thank the US National Imagery and Mapping Authority, and the US National Aeronautics and Space Administration for providing the EGM96 coefficients and Simon Holmes, a graduate student in the School of Spatial Sciences, for computing the synthetic gravity field.

References

Featherstone WE (1992) A GPS controlled gravimetric determination of the geoid of the British Isles. D.Phil thesis, Oxford University, England.

Featherstone WE, Evans JD, Olliver JG (1998) A Meissl-modified Vaníček and Kleusberg kernel to reduce the truncation error in gravimetric geoid computations. Journal of Geodesy 72(3): 154-160.

Featherstone WE, Sideris MG (1998) Modified kernels in spectral geoid determination: first results from Western Australia, in: Forsberg *et al.* (eds), Geodesy on the Move, Springer, Berlin, 188-193.

Forsberg R, Sideris MG (1993) Geoid computations by the multi-band spherical FFT approach. Bulletin Géodésique 18: 82-90.

Grafarend EW, Engels J (1994) The coinvergent series expansion of the gravitational field of a star-shaped body. manuscripta geodaetica 19: 18-30.

Haagmans RR, de Min E, van Gelderen M (1993) Fast evaluation of convolution integrals on the sphere using 1D-FFT, and a comparison with existing methods for Stokes's integral. manuscripta geodaetica 18(5): 227-241.

Heck B, Grüninger W (1987) Modification of Stokes's integral formula by combining two classical approaches. IUGG General Assembly, Vancouver, Canada.

Heiskanen WH, Moritz H (1967) Physical Geodesy. WH Freeman and Co., San Francisco, USA.

Holmes SA, Featherstone WE, Evans JD (1998) Towards a synthetic Earth gravity model, paper presented to the University of New South Wales Annual Research Seminar, Sydney, November.

Johnston GM, Featherstone WE (1998) AUSGEOID98 computation and validation: exposing the hidden dimension, proceedings of the 39th Australian Surveyors Congress, Launceston, 105-116.

Lemoine FG, Kenyon SC, Factor JK, Trimmer RG, Pavlis NK, Chinn DS, Cox CM, Klosko SM, Luthcke SB, Torrence MH, Wang YM, Williamson RG, Pavlis EC, Rapp RH, Olson TR (1998) The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96, NASA/TP-1998-206861, National Aeronautics and Space Administration, Maryland, USA.

Martinec Z, Vaníček P (1997) Formulation of the boundary-value problem for geoid determination with a higher-degree reference field. Geophysical Journal International, 126(1): 219-228.

Meissl P (1971) Preparations for the numerical evaluation of second-order Molodensky-type formulas. OSU Report 163, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA.

Molodensky MS, Eremeev VF, Yurkina MI (1962) Methods for Study of the External Gravitational Field and Figure of the Earth. Israeli Programme for the Translation of Scientific Publications, Jerusalem, Israel.

Paul MK (1973) A method of evaluating the truncation error coefficients for geoidal height, Bulletin Géodésique 47: 413-425.

Schwarz KP, Sideris MG, Forsberg R (1990) The use of FFT techniques in physical geodesy. Geophysical Journal International 100: 485-514.

Sideris MG, She BB (1995) A new, high-resolution geoid for Canada and part of the US by the 1D-FFT method. Bulletin Géodésique 69(2): 92-108.

Sjöberg LE (1991) Refined least squares modification of Stokes's formula. manuscripta geodaetica 16: 367-375.

Stokes GG (1849) On the variation of gravity on the surface of the Earth. Transactions of the Cambridge Philosophical Society 8: 672-695.

Strang van Hees GL (1990) Stokes's formula using fast Fourier techniques. manuscripta geodaetica 15: 235-239.

Torge W (1991) Geodesy (second edition), de Gruyter, Berlin.

Tziavos IN (1996) Comparisons of spectral techniques for geoid computations over large areas. Journal of Geodesy, 70(6): 357-373.

Vaníček P, Kleusberg A (1987) The Canadian geoid — Stokesian approach. manuscripta geodaetica 12(3): 86-98.

Vaníček P, Sjöberg LE (1991) Reformulation of Stokes's theory for higher than second-degree reference field and modification of integration kernels. Journal of Geophysical Research 96(B4): 6529-6539.

Vaníček P, Featherstone WE (1998) Performance of three types of Stokes's kernel in the combined solution for the geoid. Journal of Geodesy 72(11): 684-697.

Vincent S, Marsh JG (1973) Global detailed gravimetric geoid. in: Vies G (ed) Proceedings of the International Symposium on the use of Artificial Earth Satellites for Geodesy and Geodynamics, Athens, Greece, 825-855.

Wenzel H-G (1982) Geoid computation by least-squares spectral combination using integral kernels, Proceedeings of the International Association of Geodesy General Meeting, Tokyo, Japan, 438-453.

Wong L, Gore R (1969) Accuracy of geoid heights from modified Stokes kernels. Geophysical Journal of the Royal Astronomical Society 18: 81-91.

A Metric for Covariance Matrices

Wolfgang Förstner and Boudewijn Moonen

Diese Sätze führen dahin, die Theorie der krummen Flächen aus einem neuen Gesichtspunkt zu betrachten, wo sich der Untersuchung ein weites noch ganz unangebautes Feld öffnet ... so begreift man, dass zweierlei wesentlich verschiedene Relationen zu unterscheiden sind, theils nemlich solche, die eine bestimmte Form der Fläche im Raume voraussetzen, theils solche, welche von den verschiedenen Formen ... unabhängig sind. Die letzteren sind es, wovon hier die Rede ist ... man sieht aber leicht, dass eben dahin die Betrachtung der auf der Fläche construirten Figuren, ..., die Verbindung der Punkte durch kürzeste Linien[*] *u. dgl. gehört. Alle solche Untersuchungen müssen davon ausgehen, dass die Natur der krummen Fläche an sich durch den Ausdruck eines unbestimmten Linearelements in der Form* $\sqrt{(Edp^2 + 2Fdpdq + Gdq^2)}$ *gegeben ist ...*

Carl Friedrich Gauss

Abstract

The paper presents a metric for positive definite covariance matrices. It is a natural expression involving traces and joint eigenvalues of the matrices. It is shown to be the distance coming from a canonical invariant Riemannian metric on the space $Sym^+(n, \mathbb{R})$ of real symmetric positive definite matrices
In contrast to known measures, collected e. g. in Grafarend 1972, the metric is invariant under affine transformations and inversion. It can be used for evaluating covariance matrices or for optimization of measurement designs.

Keywords: Covariance matrices, metric, Lie groups, Riemannian manifolds, exponential mapping, symmetric spaces

1 Background

The optimization of geodetic networks is a classical problem that has gained large attention in the 70s.
1972 E. W. Grafarend put together the current knowledge of network design, datum transformations and artificial covariance matrices using covariance functions in his classical monograph [**?**]; see also [**?**]. One critical part was the development of a suitable measure for comparing two covariance matrices. Grafarend listed a dozen measures. Assuming a completely isotropic network, represented by a unit matrix as covariance matrix, the measures depended on the eigenvalues of the covariance matrix.
1983, 11 years later, at the Aalborg workshop on 'Survey Control Networks' Schmidt [**?**] used these measures for finding optimal networks. The visualization of the error ellipses for a single point, leading to the same deviation from an ideal covariance structure revealed deficiencies of these measures, as e. g. the trace of the eigenvalues of the covariance matrix as quality measure

[*] emphasized by the authors

would allow a totally flat error ellipse to be as good as a circular ellipse, more even, as good as the flat error ellipse rotated by 90^0.
Based on an information theoretic point of view, where the information of a Gaussian variable increases with $\ln \sigma^2$, the first author guessed the squared sum $d^2 = \sum_i \ln^2 \lambda_i$ of the logarithms of the eigenvalues to be a better measure, as deviations in both directions would be punished the same amount if measured in percent, i. e. relative to the given variances. He formulated the conjecture that the distance measure d would be a metric. Only in case d would be a metric, comparing two covariance matrices $\mathbf{A}$ and $\mathbf{B}$ with covariance matrix $\mathbf{C}$ would allow to state one of the two to be better than the other. Extensive simulations by K. Ballein [**?**] substantiated this conjecture as no case was found where the triangle inequality was violated.
1995, 12 years later, taking up the problem within image processing, the first author proved the validity of the conjecture for 2×2-matrices [**?**]. For this case the measure already had been proposed by V. V. Kavrajski [**?**] for evaluating the isotropy of map projections. However, the proof could not be generalized to higher dimensions. Using classical results from linear algebra and differential geometry the second author proved the distance d to be a metric for general positive definite symmetric matrices. An extended proof can be found in [**?**].
This paper states the problem and presents the two proofs for 2×2-matrices and for the general case. Giving two proofs for $n = 2$ may be justified by the two very different approaches to the problem.

2 Motivation

Comparing covariance matrices is a basic task in mensuration design. The idea, going back to Baarda 1973 [**?**] is to compare the variances of arbitrary functions $f = \mathbf{e}^T\mathbf{x}$ on one hand determined with a given covariance matrix $\mathbf{C}$ and on the other hand determined with a reference or criterion matrix $\mathbf{H}$.
One requirement would be the variance $\sigma_f^{2(C)}$ of f when calculated with $\mathbf{C}$ to be always smaller than the variance $\sigma_f^{2(H)}$ of f when calculated with $\mathbf{H}$. This means:

$$\mathbf{e}^T\mathbf{C}\mathbf{e} \le \mathbf{e}^T\mathbf{H}\mathbf{e} \quad \text{for all } \mathbf{e} \neq \mathbf{0}$$

or the Raleigh ratio

$$0 \le \lambda(\mathbf{e}) = \frac{\mathbf{e}^T\mathbf{C}\mathbf{e}}{\mathbf{e}^T\mathbf{H}\mathbf{e}} \le 1 \quad \text{for all } \mathbf{e} \neq \mathbf{0}.$$

The maximum λ from $1/2\partial\lambda(\mathbf{e})/\partial\mathbf{e} = \mathbf{0} \leftrightarrow \lambda\mathbf{H}\mathbf{e} - \mathbf{C}\mathbf{e} = (\lambda\mathbf{H} - \mathbf{C})\mathbf{e} = \mathbf{0}$ results in the maximum eigenvalue $\lambda_{\max}(\mathbf{C}\mathbf{H}^{-1})$ from the generalized eigenvalue problem

$$|\lambda\mathbf{H} - \mathbf{C}| = 0\,, \tag{1}$$

Observe that $\lambda\mathbf{e}^T\mathbf{H}\mathbf{e} - \mathbf{e}^T\mathbf{C}\mathbf{e} = \mathbf{e}^T(\lambda\mathbf{H} - \mathbf{C})\mathbf{e} = 0$ for $\mathbf{e} \neq \mathbf{0}$ only is fulfilled if (**??**) holds. The eigenvalues of (**??**) are non-negative if the two matrices are positive semidefinite.
This suggests the eigenvalues of $\mathbf{C}\mathbf{H}^{-1}$ to capture the difference in form of $\mathbf{C}$ and $\mathbf{H}$ completely. The requirement $\lambda_{\max} \le 1$ can be visualized by stating that the (error) ellipsoid $\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} = c$ lies completely within the (error) ellipsoid $\mathbf{x}^T\mathbf{H}^{-1}\mathbf{x} = c$.
The statistical interpretation of the ellipses results from the assumption, motivated by the principle of maximum entropy, that the stochastical variables are normally distributed, thus having density:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{\Sigma}}} \mathrm{e}^{-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x}}$$

with covariance matrix $\mathbf{\Sigma}$. Isolines of constant density are ellipses with semiaxes proportional to the square roots of the eigenvalues of $\mathbf{\Sigma}$. The ratio

$$\lambda_{\max} = \max_{\mathbf{e}} \frac{\sigma_f^{2(C)}}{\sigma_f^{2(H)}}$$

thus gives the worst case for the ratio of the variances when calculated with the covariances $\mathbf{C}$ and $\mathbf{H}$ respectively.
Instead of requiring the worst precision to be better than a specification one also could require the covariance matrix $\mathbf{C}$ to be closest to $\mathbf{H}$ in some sense. Let us for a moment assume $\mathbf{H} = \mathbf{I}$. Simple examples for measuring the difference in form of $\mathbf{C}$ compared to $\mathbf{I}$ are the trace

$$\operatorname{tr} \mathbf{C} = \sum_{i=1}^{n} \lambda_i(\mathbf{C}) \tag{2}$$

or the determinant

$$\det \mathbf{C} = \prod_{i=1}^{n} \lambda_i(\mathbf{C})\,. \tag{3}$$

These classical measures are invariant with respect to rotations (**??**) or affine transformations (**??**) of the coordinate system. Visualizing covariance matrices of equal trace or determinant can use the eigenvalues. Restricting to $n = 2$ in a 2D-coordinate system (λ_1, λ_2) covariance matrices of equal trace $\operatorname{tr}(\mathbf{C}) = c_{\mathrm{tr}}$ are characterized by the straight line $\lambda_1 = c_{\mathrm{tr}} - \lambda_2$ or $\sigma_1^2 = c_{\mathrm{tr}} - \sigma_2^2$. Covariance matrices of equal determinant $\det(\mathbf{C}) = c_{\det}$ are determined by the hyperbola $\lambda_1 = c_{\det}/\lambda_2$ or $\sigma_1^2 = c_{\det}/\sigma_2^2$. Obviously in both cases covariance matrices with very flat form of the corresponding error ellipse $\mathbf{e}^T\mathbf{C}\mathbf{e} = c$ are allowed. E. g., if one requires $c_{\mathrm{tr}} = 2$ then the pair $(0.02, 1.98)$ with a ratio of semiaxes $\sqrt{1.98/0.02} = 7$ is evaluated as being similar to the unit circle. The determinant measure is even more unfavourable. When requiring $c_{\det} = 1$ even a pair $(0.02, 50.0)$ with ratio of semiaxes 50 is called similar to the unit circle.
However, it would be desirable that the similarity between two covariance matrices reflects the deviation in variance in both directions according to the *ratio* of the variances. Thus deviations in variance by a factor f should be evaluated equally as a deviation by a factor $1/f$, of course a factor $f = 1$ indicating no difference. Thus other measures capturing the anisotropy, such as $(1 - \lambda_1)^2 + (1 - \lambda_2)^2$, not being invariant to inversion, cannot be used.
The conditions can be fulfilled by using the sum of the squared logarithms of the eigenvalues. Thus we propose the distance measure

$$\boxed{d(\mathbf{A}, \mathbf{B}) = \sqrt{\sum_{i=1}^{n} \ln^2 \lambda_i(\mathbf{A}, \mathbf{B})}} \tag{4}$$

between symmetric positive definite matrices $\mathbf{A}$ and $\mathbf{B}$, with the eigenvalues $\lambda_i(\mathbf{A}, \mathbf{B})$ from $|\lambda\mathbf{A} - \mathbf{B}| = 0$. The logarithm guarantees, that deviations are measured as factors, whereas the squaring guarantees factors f and $1/f$ being evaluated equally. Summing squares is done in close resemblance with the Euclidean metric.

This note wants to discuss the properties of d:

- d is invariant with respect to affine transformations of the coordinate system.
- d is invariant with respect to an inversion of the matrices.
- It is claimed that d is a *metric*. Thus

(i) positivity: $d(\mathbf{A},\mathbf{B}) \geq 0$, and $d(\mathbf{A},\mathbf{B}) = 0$ only if $\mathbf{A} = \mathbf{B}$.

(ii) symmetry: $d(\mathbf{A},\mathbf{B}) = d(\mathbf{B},\mathbf{A})$,

(iii) triangle inequality: $d(\mathbf{A},\mathbf{B}) + d(\mathbf{A},\mathbf{C}) \geq d(\mathbf{B},\mathbf{C})$.

The proof for $n = 2$ is given in the next Section. The proof for general n is sketched in the subsequent Sections **??** – **??**.

3 Invariance Properties

3.1 Affine Transformations

Assume the $n \times n$ matrix $\mathbf{X}$ to be regular. Then the distance $d(\overline{\mathbf{A}}, \overline{\mathbf{B}})$ of the transformed matrices

$$\overline{\mathbf{A}} = \mathbf{X}\mathbf{A}\mathbf{X}^T \quad \overline{\mathbf{B}} = \mathbf{X}\mathbf{B}\mathbf{X}^T$$

is invariant w. r. t. $\mathbf{X}$.
PROOF: We immediately obtain:

$$\begin{aligned}
\lambda(\overline{\mathbf{A}}, \overline{\mathbf{B}}) = \lambda(\mathbf{X}\mathbf{A}\mathbf{X}^T, \mathbf{X}\mathbf{B}\mathbf{X}^T) &= \lambda(\mathbf{X}\mathbf{A}\mathbf{X}^T(\mathbf{X}\mathbf{B}\mathbf{X}^T)^{-1}) \\
= \lambda(\mathbf{X}\mathbf{A}\mathbf{X}^T(\mathbf{X}^T)^{-1}\mathbf{B}^{-1}\mathbf{X}^{-1}) &= \lambda(\mathbf{X}\mathbf{A}\mathbf{B}^{-1}\mathbf{X}^{-1}) \\
= \lambda(\mathbf{A}\mathbf{B}^{-1}) &= \lambda(\mathbf{A},\mathbf{B})\,.
\end{aligned}$$

Comment: $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$ can be interpreted as covariance matrices of $\mathbf{y} = \mathbf{X}\mathbf{x}$ in case $\mathbf{A}$ and $\mathbf{B}$ are the covariance matrices of $\mathbf{x}$. Changing coordinate system does not change the evaluation of covariance matrices. Obviously, this invariance only relies on the properties of the eigenvalues, and actually was the basis for Baarda's evaluation scheme using so-called S-transformations.

3.2 Inversion

The distance is invariant under inversion of the matrices.
PROOF: We obtain

$$\begin{aligned}
d^2(\mathbf{A}^{-1},\mathbf{B}^{-1}) = d^2(\mathbf{A}^{-1}\mathbf{B}) &= \sum_{i=1}^{n} \left(\ln \lambda_i(\mathbf{A}^{-1}\mathbf{B})\right)^2 \\
= \sum_{i=1}^{n} \left(\ln[\lambda_i^{-1}(\mathbf{A}\mathbf{B}^{-1})]\right)^2 &= \sum_{i=1}^{n} \left(-\ln \lambda_i(\mathbf{A}\mathbf{B}^{-1})\right)^2 \\
= \sum_{i=1}^{n} \left(\ln \lambda_i(\mathbf{A}\mathbf{B}^{-1})\right)^2 &= d^2(\mathbf{A}\mathbf{B}^{-1}) \\
= d^2(\mathbf{A},\mathbf{B})\,. &
\end{aligned}$$

Comment: $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ can be interpreted as weight matrices of $\mathbf{x}$ if one chooses $\sigma_o^2 = 1$. Here essential use is made of the property $\lambda(\mathbf{A}) = \lambda^{-1}(\mathbf{A}^{-1})$. The proof shows, that also individual *inversions* of eigenvalues do not change the value of distance measure, as required.

3.3 d is a Distance Measure

We show that d is a distance measure, thus the first two criteria for a metric hold in general.

ad 1 $d \geq 0$ is trivial from the definition of d, keeping in mind, that the eigenvalues all are positive. Proof of $d = 0 \leftrightarrow \mathbf{A} = \mathbf{B}$:

$\leftarrow$: If $\mathbf{A} = \mathbf{B}$ then $d = 0$.
PROOF: From $\lambda(\mathbf{A}\mathbf{B}^{-1}) = \lambda(\mathbf{I})$ follows $\lambda_i = 1$ for all i, thus $d = 0$.

$\rightarrow$: If $d = 0$ then $\mathbf{A} = \mathbf{B}$.
PROOF: From $d = 0$ follows $\lambda_i(\mathbf{A}\mathbf{B}^{-1}) = 1$ for all i, thus $\mathbf{A}\mathbf{B}^{-1} = \mathbf{I}$ from which $\mathbf{A} = \mathbf{B}$ follows.

ad 2 As $(\mathbf{A}\mathbf{B}^{-1})^{-1} = \mathbf{B}\mathbf{A}^{-1}$ symmetry follows from the inversion invariance.

3.4 Triangle Inequality

For d providing a metric on the symmetric positive definite matrices the triangle inequality must hold.
Assume three $n \times n$ matrices with the following structure:

- The first matrix is the unit matrix:
$$\mathbf{A} = \mathbf{I}\,.$$
- The second matrix is a diagonal matrix with entries e^{b_i} thus
$$\mathbf{B} = Diag(\mathrm{e}^{b_i})\,.$$
- The third matrix is a general matrix with eigenvalues e^{c_i} and modal matrix $\mathbf{R}$, thus
$$\mathbf{C} = \mathbf{R}Diag(\mathrm{e}^{c_i})\mathbf{R}^T\,.$$

This setup can be chosen without loss of generality, as $\mathbf{A}$ and $\mathbf{B}$ can be orthogonalized simultaneously [**?**].
The triangle inequality can be written in the following form and reveals three terms

$$s \doteq d(\mathbf{A},\mathbf{B}) + d(\mathbf{A},\mathbf{C}) - d(\mathbf{B},\mathbf{C}) = d_c + d_b - d_a \geq 0\,. \tag{5}$$

The idea of the proof is the following:

(i) We first use the fact that d_b and d_c are independent on the rotation $\mathbf{R}$.
$$s(\mathbf{R}) = d_c + d_b - d_a(\mathbf{R})\,.$$

(ii) In case $\mathbf{R} = \mathbf{I}$ then the correctness of (**??**) results from the triangle inequality in $\mathbb{R}^n$. This even holds for any permutation $P(i)$ of the indices i of the eigenvalues λ_i of $\mathbf{B}\mathbf{C}^{-1}$. There exists a permutation P_{max} for which d_a is maximum, thus $s(\mathbf{R})$ is a minimum.

(iii) We then want to show, and this is the crucial part, that any rotation $\mathbf{R} \neq \mathbf{I}$ leads to a decrease of $d_a(\mathbf{R})$, thus to an increase of $s(\mathbf{R})$ keeping the triangle inequality to hold.

3.4.1 Distances d_c and d_b

The distances d_c and d_b are given by

$$d_c^2 = \sum_{i=1}^{n} b_i^2 \quad , \quad d_b^2 = \sum_{i=1}^{n} c_i^2\,.$$

The special definition of the matrices $\mathbf{B}$ and $\mathbf{C}$ now shows to be useful. The last expression results from the fact that the eigenvalues of $\mathbf{C}\mathbf{A}^{-1} = \mathbf{C}$ are independent on rotations $\mathbf{R}$.

3.4.2 Triangle Inequality for No Rotation

In case of no rotations the eigenvalues of $\mathbf{BC}^{-1}$ are $\mathrm{e}^{b_i}/\mathrm{e}^{c_i}$. Therefore the distance d_a yields

$$d_a^2 = \sum_{i=1}^{n} \left(\ln \frac{\mathrm{e}^{b_i}}{\mathrm{e}^{c_i}} \right)^2 = \sum_{i=1}^{n} (b_i - c_i)^2 \,.$$

With the vectors $\mathbf{b} = (b_i)$ and $\mathbf{c} = (c_i)$ the triangle inequality in $\mathbb{R}^n$ yields

$$|\mathbf{c}| + |\mathbf{b}| - |\mathbf{b} - \mathbf{c}| \geq 0$$

or

$$s = \sqrt{\sum_{i=1}^{n} c_i^2} + \sqrt{\sum_{i=1}^{n} b_i^2} - \sqrt{\sum_{i=1}^{n} (b_i - c_i)^2} \geq 0 \,.$$

holds.
For any permutation $P(i)$ we also get

$$s = \sqrt{\sum_{i=1}^{n} c_{P(i)}^2} + \sqrt{\sum_{i=1}^{n} b_i^2} - \sqrt{\sum_{i=1}^{n} (b_i - c_{P(i)})^2} \geq 0 \,. \tag{6}$$

which guarantees that there is a permutation $P_{max}(i)$ for which s in (**??**) is minimum.

3.4.3 d is Metric for 2×2-Matrices

We now want to show that the triangle inequality holds for 2×2 matrices. Thus we only need to show that $s(\mathbf{R}(\phi))$ is monotonous with ϕ in $[0..\pi/2]$, or equivalently that $d_a(\mathbf{R})$ is monotonous.

We assume (observe the change of notation in the entries b_i and c_i of the matrices)

$$\mathbf{B} = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} \qquad , \qquad b_1 > 0 \;,\; b_2 > 0 \,. \tag{7}$$

With

$$x = \sin \phi \tag{8}$$

the rotation $\mathbf{R}(x) = \mathbf{R}(\phi)$ is represented as

$$\mathbf{R}(x) = \begin{pmatrix} \sqrt{1 - x^2} & x \\ -x & \sqrt{1 - x^2} \end{pmatrix} ,$$

thus only values $x \in [0, 1]$ need to be investigated.
With the diagonal matrix $Diag(c_1, c_2)$, containing the positive eigenvalues

$$c_1 > 0, \quad c_2 > 0 \,, \tag{9}$$

this leads to the general matrix

$$\mathbf{C} = \mathbf{R} \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \mathbf{R}^T = \begin{pmatrix} c_1 x^2 + c_2(1 - x^2) & -x\sqrt{1 - x^2}(c_2 - c_1) \\ -x\sqrt{1 - x^2}(c_2 - c_1) & c_1(1 - x^2) + c_2 x^2 \end{pmatrix} .$$

The eigenvalues of $\mathbf{CB}^{-1}$ are (from Maple)

$$\lambda_1 = \frac{u(x) + \sqrt{v(x)}}{2 b_1 b_2} \geq 0 \qquad , \qquad \lambda_2 = \frac{u(x) - \sqrt{v(x)}}{2 b_1 b_2} \geq 0$$

with the discriminant

$$v(x) = u(x)^2 - w \geq 0$$

and

$$u(x) = (b_1\, c_1 + c_2\, b_2)(1 - x^2) + (b_1\, c_2 + b_2\, c_1)x^2 \geq 0\,, \tag{10}$$

$$w = 4b_1 b_2 c_1 c_2 \geq 0\,, \tag{11}$$

the last inequality holding due to (**??**)(**??**). The distance

$$d_a(x) = \sqrt{\ln^2 \lambda_1(x) + \ln^2 \lambda_2(x)}\,,$$

which is dependent on x, has first derivative

$$\begin{aligned}\frac{\partial d_a(x)}{\partial x} &= 2\frac{x(b_2 - b_1)(c_2 - c_1)}{v(x)\, d_a(x)} \left(\ln \frac{u(x) - \sqrt{v(x)}}{2b_1 b_2} - \ln \frac{u(x) + \sqrt{v(x)}}{2b_1 b_2} \right) \\ &= 2\frac{x(b_2 - b_1)(c_2 - c_1) \ln \dfrac{u(x) - \sqrt{v(x)}}{u(x) + \sqrt{v(x)}}}{v(x)\, d_a(x)}\,. \end{aligned} \tag{12}$$

For fixed b_1, b_2, c_1 and c_2 this expression does not change sign in $x \in [0, 1]$. This is because the discriminant $v(x) = u^2(x) - w(x)$ (cf. (**??**)) is always positive, due to

$$\begin{aligned} v(0) &= (b_1 c_1 - b_2 c_2)^2 \geq 0 \\ v(1) &= (b_2 c_1 - b_1 c_2)^2 \geq 0 \\ \frac{\partial v(x)}{\partial x} &= -4x(b_2 - b_1)(c_2 - c_1)\, u(x) \end{aligned}$$

with $u(x) \geq 0$ (cf. (**??**)) thus $v(x)$ being monotonous. Furthermore, $v(x)$ is always smaller than u^2 (cf. (**??**), (**??**)), thus the logarithmic expression always negative. As the triangle equation is fulfilled at the extremes of the interval $[0, 1]$ it is fulfilled for all x, thus for all rotations.

Comment: When substituting $x = \sin\phi$ (cf. (**??**)) the first derivative (**??**) is of the form $\partial d_a(\phi)/\partial\phi = \sin\phi\, f(\phi)$ with a symmetric function $f(\phi) = f(-\phi)$. Thus the derivative is skew symmetric w. r. t. $(0, 0)$, indication d_a to be symmetric $d_a(\phi) = d_a(-\phi)$, which is to be expected.

3.4.4 d is Claimed to be a Metric for $n \times n$-Matrices

The proof of the metric properties of d for 2×2 matrices suggests that in the general case of $n \times n$ matrices any rotation away from the worst permutation of the indices (cf. (**??**)) results in an increase of the value s. The proof for the case $n = 2$ can be used to show, that, starting with the worst permutation of the indices, any single rotation around one of the axes leads to a monotonous change of s. Therefore, for proving the case of general n, there would have to be shown that any combination of *two* rotations away from the worst permutation leads to a monotonous change of s allowing to reach any permutation by a rotation while increasing $s(\mathbf{R})$. Completing this line of proof has not been achieved so far.

4 Restating the Problem

In der Kürze liegt die Würze.

Deutscher Volksmund

Let

$$M(n,\mathbb{R}) := \{\, \mathbf{A} = (a_{ij}) \mid 1 \le i,j \le n \ ,\ a_{ij} \in \mathbb{R} \,\}$$

be the space of real $n \times n$-matrices, and let

$$S^+ := Sym^+(n,\mathbb{R}) := \left\{\, \mathbf{A} \in M(n,\mathbb{R}) \;\middle|\; \mathbf{A} = \mathbf{A}^T \ ,\ \mathbf{A} > 0 \,\right\}$$

be the subspace of real, symmetric, positive definite matrices. Recall that any symmetric matrix $\mathbf{A}$ can be substituted into a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ which gives a symmetric matrix $f(\mathbf{A})$ commuting with all matrices commuting with $\mathbf{A}$. In particular, a symmetric matrix $\mathbf{A}$ has an exponential $\exp(\mathbf{A})$, and a symmetric positive definite matrix A has a logarithm $\ln(\mathbf{A})$, and these assignments are inverse to each other. A symmetric positive definite matrix $\mathbf{A}$ also has a unique square root $\sqrt{\mathbf{A}}$ which is of the same type. Define, for $\mathbf{A}, \mathbf{B} \in Sym^+(n,\mathbb{R})$, their distance $d(\mathbf{A},\mathbf{B}) \ge 0$ by

$$\boxed{d^2(\mathbf{A},\mathbf{B}) := \operatorname{tr}\left(\ln^2\left(\sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}}\right)\right)}, \tag{13}$$

where tr denotes the trace. In particular, this shows that

$$d(\mathbf{A},\mathbf{B}) \ge 0 \qquad , \qquad d(\mathbf{A},\mathbf{B}) = 0 \iff \mathbf{A} = \mathbf{B}\,.$$

In more down-to-earth terms:

$$d(\mathbf{A},\mathbf{B}) = \sqrt{\sum_{i=1}^{n} \ln^2 \lambda_i(\mathbf{A},\mathbf{B})}, \tag{14}$$

where $\lambda_1(\mathbf{A},\mathbf{B}), \ldots, \lambda_n(\mathbf{A},\mathbf{B})$ are the *joint eigenvalues* of $\mathbf{A}$ and $\mathbf{B}$, i.e. the roots of the equation

$$\det(\lambda\mathbf{A} - \mathbf{B}) = 0.$$

This is the proposal of [**?**], i.e. of equation (**??**) above. (To see why these two definitions coincide, note that

$$\lambda\mathbf{A} - \mathbf{B} = \sqrt{\mathbf{A}}\,(\lambda\mathbf{E} - \sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}})\,\sqrt{\mathbf{A}}\,,$$

so that the joint eigenvalues $\lambda_i(\mathbf{A},\mathbf{B})$ are just the eigenvalues of the real symmetric positive definite matrix $\sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}}$; in particular, they are positive real numbers and so the definition (**??**) makes sense.) The equation (**??**) shows that d is invariant under congruence transformations with $\mathbf{X} \in GL(n,\mathbb{R})$, where $GL(n,\mathbb{R})$ is the group of regular linear transformations of $\mathbb{R}^n$:

$$\forall\, \mathbf{X} \in GL(n,\mathbb{R}) \ : \quad d(\mathbf{A},\mathbf{B}) = d(\mathbf{X}\mathbf{A}\mathbf{X}^T, \mathbf{X}\mathbf{B}\mathbf{X}^T) \tag{15}$$

(since $\det(\lambda\mathbf{A} - \mathbf{B})$ and $\det\left(\mathbf{X}(\lambda\mathbf{A} - \mathbf{B})\mathbf{X}^T\right)$ have the same roots); this is not easily seen from definition (**??**). It also shows that

$$d(\mathbf{A},\mathbf{B}) = d(\mathbf{B},\mathbf{A}) \ ,\ d(\mathbf{A},\mathbf{B}) = d(\mathbf{A}^{-1},\mathbf{B}^{-1}).$$

5 The results

One then has

Theorem 1. *The map d defines a distance on the space $Sym^+(n,\mathbb{R})$, i.e there holds*

(i) Positivity: $d(\mathbf{A},\mathbf{B}) \geq 0$, and $d(\mathbf{A},\mathbf{B}) = 0 \iff \mathbf{A} = \mathbf{B}$

(ii) Symmetry: $d(\mathbf{A},\mathbf{B}) = d(\mathbf{B},\mathbf{A})$

(iii) Triangle inequality: $d(\mathbf{A},\mathbf{C}) \leq d(\mathbf{A},\mathbf{B}) + d(\mathbf{B},\mathbf{C})$

for all $\mathbf{A},\mathbf{B},\mathbf{C} \in Sym^+(n,\mathbb{R})$. Moreover, d has the following invariances:

(iv) It is invariant under congruence transformations, i.e.

$$d(\mathbf{A},\mathbf{B}) = d(\mathbf{X}\mathbf{A}\mathbf{X}^T, \mathbf{X}\mathbf{B}\mathbf{X}^T)$$

for all $\mathbf{A},\mathbf{B} \in Sym^+(n,\mathbb{R})$, $\mathbf{X} \in GL(n,\mathbb{R})$

(v) It is invariant under inversion, i.e.

$$d(\mathbf{A},\mathbf{B}) = d(\mathbf{A}^{-1}, \mathbf{B}^{-1})$$

for all $\mathbf{A},\mathbf{B} \in Sym^+(n,\mathbb{R})$

The same conclusions hold for the space $SSym(n,\mathbb{R})$ of real symmetric positive definite matrices of determinant one, when one replaces the general linear group $GL(n,\mathbb{R})$ with the special linear group $SL(n,\mathbb{R})$, the $n\times n$–matrices of determinant one, and the space of real symmetric matrices $Sym(n,\mathbb{R})$ with the space $Sym_0(n,\mathbb{R})$ of real symmetric traceless matrices.

Remark 1. We use here the terminology "distance" in contrast to the standard terminology "metric" in order to avoid confusion with the notion of "Riemannian metric", which is going to play a rôle soon.

The case $n = 2$ is already interesting; see Remark **??** below.
All the properties except property (iv), the triangle inequality, are more or less obvious from the definition (see above), but the triangle inequality is not. In fact, the theorem will be the consequence of a more general theorem as follows.
The most important geometric way distances arise is as associated distances to Riemannian metrics on manifolds; the Riemannian metric, as an infinitesimal description of length is used to define the length of paths by integration, and the distance between two points then arises as the greatest lower bound on the length of paths joining the two points. More precisely, if M is a differentiable manifold (in what follows, "differentiable" will always mean "infinitely many times differentiable", i.e. of class $\mathcal{C}^\infty$), a *Riemannian metric* is the assignment to any $p \in M$ of a Euclidean scalar product $\langle - \,|\, - \rangle_p$ in the tangent space T_pM depending differentiably on p. Technically, it is a differentiable positive definite section of the second symmetric power S^2T^*M of the cotangent bundle, or a positive definite symmetric 2-tensor. In classical terms, it is given in local coordinates (U,x) as the "square of the line element" or "first fundamental form"

$$\mathrm{d}s^2 = g_{ij}(x)\mathrm{d}x^i\mathrm{d}x^j \tag{16}$$

(Einstein summation convention: repeated lower and upper indices are summed over). Here the g_{ij} are differentiable functions (the *metric coefficients)* subjected to the transformation rule

$$g_{ij}(x) = g_{kl}(y(x))\frac{\partial y^k}{\partial x^i}\frac{\partial y^l}{\partial x^j}\,.$$

A differentiable manifold together with a Riemannian metric is called a *Riemannian manifold*. Given a piecewise differentiable path $c : [a, b] \longrightarrow M$ in M, its *length* $L[c]$ is defined to be

$$L[c] := \int_a^b \|\dot{c}(t)\|_{c(t)} \, \mathrm{d}t \,,$$

where for $X \in T_pM$ we have $\|X\|_p := \sqrt{\langle X \mid X \rangle}_p$, the Euclidean norm associated to the scalar product in T_pM given by the Riemannian metric. In local coordinates

$$L[c] = \int_a^b \sqrt{g_{ij}(c(t)) \, \dot{c}^i(t) \dot{c}^j(t)} \, \mathrm{d}t \,.$$

Given p, $q \in M$, the distance $d(p, q)$ associated to a given Riemannian metric then is defined to be

$$d(p, q) := \inf_c L[c] \;, \tag{17}$$

the infimum running over all piecewise differentiable paths c joining p to q. This defines indeed a distance:

Proposition. *The distance defined by* (**??**) *on a connected Riemannian manifold is a metric in the sense of metric spaces, i.e. defines a map* $d : M \times M \longrightarrow \mathbb{R}$ *satisfying*

(i) $d(p, q) \geq 0$, $d(p, q) = 0 \iff p = q$

(ii) $d(p, q) = d(q, p)$

(iii) $d(p, r) \leq d(p, q) + d(q, r)$.

An indication of proof will be given in the next section.
So the central issue here is the fact that a Riemannian metric is the differential substrate of a distance and, in turn, defines a distance by integration. This is the most important way of constructing distances, which is the fundamental discovery of Gauss and Riemann. In our case, this paradigm is realized in the following way.
The space $Sym^+(n, \mathbb{R})$ is a differentiable manifold of dimension $n(n+1)/2$, more specifically, it is an open cone in the vector space

$$Sym(n, \mathbb{R}) := \left\{ \, \mathbf{A} \in M(n, \mathbb{R}) \;\middle|\; \mathbf{A} = \mathbf{A}^T \, \right\}$$

of all real symmetric $n \times n$-matrices. Thus the tangent space $T_{\mathbf{A}} Sym^+(n, \mathbb{R})$ to $Sym^+(n, \mathbb{R})$ at a point $\mathbf{A} \in Sym^+(n, \mathbb{R})$ is just given as

$$T_{\mathbf{A}} Sym^+(n, \mathbb{R}) = Sym(n, \mathbb{R}).$$

The tangent space $T_{\mathbf{A}} SSym^+(n, \mathbb{R})$ to $SSym(n, \mathbb{R})$ at a point $\mathbf{A} \in SSym(n, \mathbb{R})$ is just given as

$$T_{\mathbf{A}} SSym(n, \mathbb{R}) = Sym_0(n, \mathbb{R}) := \{ \, \mathbf{X} \in Sym(n, \mathbb{R}) \;\mid\; \mathrm{tr}\,(\mathbf{X}) = 0 \, \} \,,$$

the space of traceless symmetric matrices.
Now note that there is a natural action of $GL(n, \mathbb{R})$ on S^+, namely, as already referred to above, by congruence transformations: $\mathbf{X} \in GL(n, \mathbb{R})$ acts via $\mathbf{A} \mapsto \mathbf{X}\mathbf{A}\mathbf{X}^T$. If one regards $\mathbf{A}$ as the matrix corresponding to a bilinear form with respect to a given basis, this action represents a change of basis. This action is transitive, and the isotropy subgroup at $\mathbf{E} \in S^+$ is just the orthogonal group $O(n)$:

$$\left\{ \, \mathbf{X} \in GL(n, \mathbb{R}) \;\middle|\; \mathbf{X}\mathbf{E}\mathbf{X}^T = \mathbf{E} \, \right\} = \left\{ \, \mathbf{X} \in GL(n, \mathbb{R}) \;\middle|\; \mathbf{X}\mathbf{X}^T = \mathbf{E} \, \right\} = O(n),$$

and so S^+ can be identified with the homogeneous space $GL(n,\mathbb{R})/O(n)$ upon which $GL(n,\mathbb{R})$ acts by left translations (its geometric significance being that its points parametrize the possible scalar products in $\mathbb{R}^n$) [†] .
In general, a *homogeneous space* is a differentiable manifold with a *transitive* action of a Lie group G, whence it has a representation as a quotient $M = G/H$ with H a closed subgroup of G. The most natural Riemannian metrics in this case are then those for which the group G acts by isometries, or, in other words, which are invariant under the action of G; e.g. the classical geometries – the Euclidean, the elliptic, and the hyperbolic geometry – arise in this manner. Looking out for these metrics, Theorem **??** will be a consequence of the following theorem :

Theorem 2. *(i) The Riemannian metrics g on $Sym^+(n,\mathbb{R})$ invariant under congruence transformations with matrices $\mathbf{X} \in GL(n,\mathbb{R})$ are in one-to-one-correspondence with positive definite quadratic forms Q on $T_{\mathbf{E}}Sym^+(n,\mathbb{R}) = Sym(n,\mathbb{R})$ invariant under conjugation with orthogonal matrices, the correspondence being given by*

$$g_{\mathbf{A}}(\mathbf{X},\mathbf{Y}) \;=\; \mathbf{B}(\sqrt{\mathbf{A}^{-1}}\mathbf{X}\sqrt{\mathbf{A}^{-1}}, \sqrt{\mathbf{A}^{-1}}\mathbf{Y}\sqrt{\mathbf{A}^{-1}})\,,$$

where $\mathbf{A} \in Sym^+(n,\mathbb{R})$, $\mathbf{X},\mathbf{Y} \in Sym(n,\mathbb{R}) = T_{\mathbf{A}}Sym^+(n,\mathbb{R})$, and $\mathbf{B}$ is the symmetric positive bilinear form corresponding to Q

(ii) The corresponding distance d_Q is invariant under congruence transformations and inversion, i.e. satisfies

$$d_Q(\mathbf{A},\mathbf{B}) = d_Q(\mathbf{X}\mathbf{A}\mathbf{X}^T, \mathbf{X}\mathbf{B}\mathbf{X}^T)$$

and

$$d_Q(\mathbf{A},\mathbf{B}) = d_Q(\mathbf{A}^{-1}, \mathbf{B}^{-1})$$

for all $\mathbf{A},\mathbf{B} \in Sym^+(n,\mathbb{R})$, $\mathbf{X} \in GL(n,\mathbb{R})$, and is given by

$$d_Q^2(\mathbf{A},\mathbf{B}) = \frac{1}{4}Q\big(\ln\big(\sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}}\big)\big)\,. \tag{18}$$

*(iii) In particular, the distance in Theorem **??** is given by the invariant Riemannian metric corresponding to the canonical non–degenerate bilinear form* [‡]

$$\forall\,\mathbf{X},\mathbf{Y} \in M(n,\mathbb{R}) : \quad B(\mathbf{X},\mathbf{Y}) \;:=\; 4\ \mathrm{tr}\,(\mathbf{X}\mathbf{Y})$$

on $M(n,\mathbb{R})$, restricted to $Sym(n,\mathbb{R}) = T_{\mathbf{E}}Sym(n,\mathbb{R})^+$, i.e. to the quadratic form

$$\forall\,\mathbf{X} \in Sym(n,\mathbb{R}) : \quad Q(\mathbf{X}) := 4\ \mathrm{tr}\,\big(\mathbf{X}^2\big)$$

on $Sym(n,\mathbb{R})$. As a Riemannian metric it is, in classical notation,

$$\mathrm{d}s^2 = 4\ \mathrm{tr}\,\Big(\big(\sqrt{\mathbf{X}^{-1}}\,\mathrm{d}\mathbf{X}\,\sqrt{\mathbf{X}^{-1}}\big)^2\Big) = 4\ \mathrm{tr}\,\Big(\big(\mathbf{X}^{-1}\,\mathrm{d}\mathbf{X}\big)^2\Big) \tag{19}$$

where $\mathbf{X} = (X_{ij})$ is the matrix of the natural coordinates on $Sym^+(n,\mathbb{R})$ and $\mathrm{d}\mathbf{X} = (dX_{ij})$ is a matrix of differentials.

The same conclusions hold for the space $SSym(n,\mathbb{R})$ of real symmetric positive definite matrices of determinant one, when one replaces the general linear group $GL(n,\mathbb{R})$ with the special linear group $SL(n,\mathbb{R})$, the $n\times n$–matrices of determinant one, and the space of real symmetric matrices $Sym(n,\mathbb{R})$ with the space $Sym_0(n,\mathbb{R})$ of real symmetric traceless matrices.

[†] A concrete map $p : G \longrightarrow S^+$ achieving this identification is given by $p(\mathbf{A}) := \mathbf{A}\mathbf{A}^T$; it is surjective and satisfies $p(\mathbf{X}\mathbf{A}) = \mathbf{X}p(\mathbf{A})\mathbf{X}^T$. The fact that p is an identification map is then equivalent to the *polar decomposition*: any regular matrix can be uniquely written as the product of a positive definite symmetric matrix and an orthogonal matrix. This generalizes the representation $z = re^{i\varphi}$ of a nonzero complex number z.

[‡] the famous Cartan-Killing-form of Lie group theory

Remark 2. Although the expression (**??**) appears to be explicit in the coordinates, it seems to be of no use for analyzing the properties of the corresponding Riemannian metric, since the operations of inverting, squaring, and taking the trace gives, in the general case, untractable expressions. In particular, it apparently is of no help in deriving the expression (**??**) for the associated distance by direct elementary means.
There is, however, one interesting case where it can be checked to give a very classical expression; this is the case $n = 2$. In this case, one has

$$SSym^+(2,\mathbb{R}) = \left\{ \begin{pmatrix} x & y \\ y & z \end{pmatrix} \;\middle|\; x, y, z \in \mathbb{R}\,,\, xz - y^2 = 1\,,\, x > 0 \right\}$$

a hyperboloid in 3-space. This is classically known as a candidate for a model of the hyperbolic plane.In fact, in this case, one may show by explicit computation that the metric (**??**) restricts to the classical hyperbolic metric, and that the corresponding distance just gives one of the classical formulas for the hyperbolic distance. For details, see [**?**].

Of course, the next question is which invariant metrics there are. Also this question can be answered:

Addendum. *(i) The positive definite quadratic forms Q on $Sym(n,\mathbb{R})$ invariant under conjugation with orthogonal matrices are of the form*

$$Q(\mathbf{X}) = \alpha\ \mathrm{tr}\left(\mathbf{X}^2\right) + \beta\big(\ \mathrm{tr}\,(\mathbf{X})\ \big)^2\,, \qquad \alpha > 0, \beta > -\frac{\alpha}{n}$$

(ii) The positive definite quadratic forms Q on $Sym_0(n,\mathbb{R})$ invariant under conjugation with orthogonal matrices are unique up to a positive scalar and hence of the form

$$Q(\mathbf{X}) = \alpha\ \mathrm{tr}\left(\mathbf{X}^2\right), \qquad \alpha > 0.$$

In particular, the Riemannian metric (**??**) corresponds to the case $\alpha = 1$, $\beta = 0$. Since from the point of this classification all these metrics stand on an equal footing, it would be interesting to know by which naturality requirements this choice can be singled out.

6 The proofs

To put this result into proper perspective and to cut a long story short, let us very briefly summarize why Theorem **??**, and consequently Theorem **??**, are true. First, however, we indicate a proof of the Proposition above, since it is on this Proposition that our approach to the triangle equality for the distance defined by (**??**) is based.
The fact that $d(p,q) \geq 0$ and the symmetry of d are immediate from the definitions. There remains to show $d(p,q) = 0 \implies p = q$ and the triangle inequality.
For given $p \in M$, choose a coordinate neighbourhood $U \cong \mathbb{R}^n$ around p such that p corresponds to $0 \in \mathbb{R}^n$. We then have the expression (**??**) for the given metric in U. Moreover, we have in U the standard Euclidean metric

$$\mathrm{d}s_E^2 := \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j = \sum_{i=1}^{n} (dx^i)^2\,.$$

Let $\|-\|$ denote the norm belonging to the given Riemannian metric in U and $|-|$ the norm given by the standard Euclidean metric. For $r > 0$ let

$$\overline{\mathbb{B}}(p;r) := \{\, x \in \mathbb{R}^n \mid |x| \leq r \,\}$$

be the standard closed ball with radius r around $p = 0$, and

$$\mathbb{S}(p;r) := \{\, x \in \mathbb{R}^n \mid |x| = r \,\}$$

its boundary, the sphere of radius r around $p \in \mathbb{R}^n$.
As a continuous function $U \times \mathbb{R}^n \longrightarrow \mathbb{R}$ the norm $\|-\|$ takes its minimum $m > 0$ and its maximum $M > 0$ on the compact set $\overline{\mathbb{B}}(p;1) \times \mathbb{S}(p;1)$. It follows that we have

$$\forall\, q \in \overline{\mathbb{B}}(p;1)\,,\, X \in \mathbb{R}^n \,:\quad m|X| \leq \|X\|_q \leq M|X|$$

by homogeneity of the norm, and so by integrating and taking the infimum

$$\forall\, q \in \overline{\mathbb{B}}(p;1) \,:\quad m d_E(p,q) \leq d(p,q) \leq M d_E(p,q) \tag{20}$$

where $d_E(p,q) = |q-p|$ is the Euclidean distance. If $q \notin \overline{\mathbb{B}}(p;1)$, then any path c joining p to q meets the boundary $\mathbb{S}(p;1)$ in some point r, from which follows $L[c] \geq L[c'] \geq d(p,r) \geq m$ – where c' denotes the part of c joining p to r for the first time, say – whence $d(p,q) \geq m$. In other words, if $d(p,q) < m$ we have $q \in \overline{\mathbb{B}}(p;1)$, where we can apply (**??**). If now $d(p,q) = 0$, then surely $d(p,q) < m$, and then by (**??**) $m d_E(p,q) \leq d(p,q) = 0$, whence $d_E(p,q) = 0$, which implies $p = q$.
For the triangle inequality, let c be a path joining p to q and d a path joining q to r. Let $c * d$ be the composite path joining p to r. Then $L[c*d] = L[c] + L[d]$. Taking the infimum on the left hand side over all paths joining p to r gives $d(p,r) \leq L[c] + L[d]$. Taking on the right hand side first the infimum over all paths joining p to q and subsequently over all the paths joining q to r then gives $d(p,r) \leq d(p,q) + d(q,r)$, which is the triangle inequality.

Remark 3. In particular, (**??**) shows that the metric topology induced by the distance d on a connected Riemannian manifold coincides with the given manifold topology.

Now to the proof of Theorem **??**. Recall the terminology of [**?**], Chapter X: Let G be a Lie group with Lie algebra $\mathfrak{g}$, $H \subseteq G$ a closed Lie subgroup corresponding to the Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. Let M be the homogeneous space $M = G/H$. Then G operates as a symmetry group on M by left translations. M has the distinguished point $o = eH = H$ corresponding to the coset of the unit element $e \in G$ with tangent space $T_oM = \mathfrak{g}/\mathfrak{h}$. This homogeneous space is called *reductive* if $\mathfrak{g}$ splits as a direct vector space sum $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ for a linear subspace $\mathfrak{m} \subseteq \mathfrak{g}$ such that $\mathfrak{m}$ is invariant under the adjoint action $\mathrm{Ad} : H \longrightarrow GL(\mathfrak{g})$. Then canonically $T_oM = \mathfrak{m}$. In our situation, $G = GL(n,\mathbb{R})$, $H = O(n)$. Then $\mathfrak{g} = M(n,\mathbb{R})$, the full $n \times n$-matrices, and $\mathfrak{h} = Asym(n,\mathbb{R})$, the antisymmetric matrices. As is well known,

$$M(n,\mathbb{R}) = Asym(n,\mathbb{R}) \oplus Sym(n,\mathbb{R})\,,$$

since any matrix $\mathbf{X}$ splits into the sum of its antisymmetric and symmetric part via

$$X = \frac{X - X^T}{2} + \frac{X + X^T}{2}\,.$$

The adjoint action of $\mathbf{O} \in O(n)$ on $M(n,\mathbb{R})$ is given by $\mathbf{X} \mapsto \mathbf{OXO}^{-1} = \mathbf{OXO}^T$ and clearly preserves $Sym(n,\mathbb{R})$. So $Sym^+(n,\mathbb{R})$ is a reductive homogeneous space.
We now have the following facts from the general theory:
a) On a reductive homogeneous space there is a distinguished connection invariant under the action of G, called the *natural torsion free connection* in [**?**]. It is uniquely characterized by the following properties ([**?**], Chapter X, Theorem 2.10)

– It is G–invariant

– Its geodesics through $o \in M$ are the orbits of o under the one-parameter subgroups of G, i.e. of the form $t \mapsto \exp(tX) \cdot o$ for some $X \in \mathfrak{g}$, where $\exp : \mathfrak{g} \longrightarrow G$ is the exponential mapping of Lie group theory

– It is torsion free

In particular, with this connection M becomes an *affine locally symmetric space*, i.e. the geodesic symmetries at a point of M given by inflection in the geodesics locally preserve the connection (*loc. cit* Chapter XI, Theorem 1.1). If M is simply connected, M is even an *affine symmetric space*, i.e. the geodesic symmetries extend to globally defined transformations of M preserving the connection (*loc. cit.*, Chapter XI, Theorem 1.2). By homogeneity, these are determined by the geodesic symmetry s at o. In our case $M = Sym^+(n, \mathbb{R})$, M is even contractible, hence simply connected, and so with the natural torsion free connection an affine symmetric space. We have $o = \mathbf{E}$, the $n \times n$ unit matrix. For $G = GL(n, \mathbb{R})$, the exponential mapping of Lie group theory is given by the "naive matrix exponential" $e^{\mathbf{X}} = \sum_{k=0}^{\infty} t^k \mathbf{X}^k / k!$. So the geodesics are $t \mapsto \exp(t\mathbf{X})\, \mathbf{E} \exp(t\mathbf{X})^T = e^{2t\mathbf{X}}$, where $\mathbf{X} \in Sym(n, \mathbb{R})$, and s is given by $s(\mathbf{X}) = \mathbf{X}^{-1}$.
b) The Riemannian metrics g on M invariant under the action of G are in one-to-one-correspondence with positive definite quadratic forms Q on $\mathfrak{m}$ invariant under the adjoint action of H (*loc. cit, Chapter X*, Corollary 3.2), the correspondence being given by

$$\forall\, X \in T_oM = \mathfrak{m} : \quad g_o(X, X) = Q(X)\,.$$

This is intuitively obvious, since we can translate o to any point of M by operating on it with an element $g \in G$.
c) All G-invariant Riemannian metrics on M (there may be none) have the natural torsion free connection as their Levi-Cività connection (*loc. cit*, Chapter XI, Theorem 3.3). In particular, such a metric makes M into a Riemannian (locally) symmetric space, i.e. the geodesic symmetries are isometries, and the exponential map of Riemannian Geometry at o, $\mathrm{Exp}_o : T_oM = \mathfrak{m} \longrightarrow M$ is given by the exponential map of Lie group theory for G:

$$\forall\, X \in \mathfrak{m} : \quad \mathrm{Exp}_o(tX) = \exp(tX) \cdot o\,.$$

Collecting these results, we now can come to terms with formula (**??**). First we see that Part (i) of Theorem **??** is a standard result in the theory of homogeneous spaces. Furthermore, S^+, being a Riemannian symmetric space with the metric (**??**), is complete (*loc. cit*, Chapter XI, Theorem 6.4), the exponential mapping $\mathrm{Exp}_{\mathbf{E}}$ of Riemannian geometry is related to the exponential mapping $\exp : \mathrm{S} \longrightarrow \mathrm{S}^+$, $S = T_{\mathbf{E}}S^+$ from Lie theory and the matrix exponential $e^{\mathbf{X}}$ via $\mathrm{Exp}_{\mathbf{E}}(\mathbf{X}) = \exp(2\mathbf{X}) = e^{2\mathbf{X}}$ and is a diffeomorphism [§] .
Having reached this point, here is the showdown. Since, by general theory, the Riemannian exponential mapping is a radial isometry, we get for the square of the distance d_Q:

$$d_Q^2(\mathbf{A}, \mathbf{B}) = d_Q^2(\mathbf{E}, \sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}})$$

since d_Q is invariant under congruences by (**??**),

$$= Q(\frac{1}{2}\exp^{-1}(\sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}}))$$

[§] The fact that the naive matrix exponential is a diffeomorphism, whence S^+ is complete, can be seen by elementary means in the case under consideration. The main point is that it coincides with the exponential mapping coming from Riemannian Geometry (up to scaling with a factor of 2).

since $\mathrm{Exp}_{\mathbf{E}}$ is a radial isometry ,

$$= \frac{1}{4} Q(\ln(\sqrt{\mathbf{A}^{-1}}\mathbf{B}\sqrt{\mathbf{A}^{-1}}))\,,$$

and this is just equation (**??**). In particular, from this one directly reads off that the distance is invariant under inversion, as claimed. Of course, the invariances in question are for the particular case corresponding to (**??**) read off easily from the classical form (**??**) of the Riemannian metric. On the other hand, we see that the invariance under inversion comes from the structural facts that S^+ is a symmetric space, and that the geodesic symmetry at E, which on general grounds must be an isometry, is just given by matrix inversion (see a) above).
One should add that these arguments are general and pertain to the situation of a symmetric space of the non–compact type; for this, see [**?**].
The representation of the orthogonal group $O(n)$ on the symmetric matrices by conjugation is not irreducible, but decomposes as

$$Sym(n,\mathbb{R}) = Sym_0(n,\mathbb{R}) \oplus \mathfrak{d}(n,\mathbb{R}), \tag{21}$$

where $\mathfrak{d}(n,\mathbb{R})$ are the scalar diagonal matrices. It is easy to see that both summands are invariant under conjugation with orthogonal matrices, and it can be shown that both parts are irreducible representations of $O(n)$. From this it is standard to derive the Addendum. In the geometric framework of symmetric spaces, this describes the decomposition of the holonomy representation and correspondingly the canonical DE RHAM–decomposition

$$Sym^+(n,\mathbb{R}) \cong SSym(n,\mathbb{R}) \times \mathbb{R}^+$$

of the symmetric space $Sym(n,\mathbb{R})$ into irreducible factors. This is a direct product of Riemannian manifolds, i.e. the metric on the product is just the product of the metrics on the individual factors, that is given by the Pythagorean description. Thus it suffices to classify the invariant metrics on the individual factors, which accounts for the Addendum.
Thus, it transpires that the theorems above follow from the basics of Lie group theory and Differential Geometry and so should be clear to the experts. The main results upon which it is based appeared originally in the literature in [**?**]. All in all, it follows in a quite straightforward manner from the albeit rather elaborated machinery of modern Differential Geometry and the theory of symmetric spaces. In conclusion, it might therefore be still interesting to give a more elementary derivation of the result, as was done above in the case $n = 2$.
As a general reference for Differential Geometry and the theory of symmetric spaces I recommend [**?**], [**?**] (which, however, make quite a terse reading). A detailed exposition [**?**] covering all the necessary prerequisites is under construction; the purpose of this paper is to introduce the non–experts to all the basic notions of Differential Geometry and to expand the brief arguments just sketched.

References

[1] BAARDA, W. (1973): *S-Transformations and Criterion Matrices*, Band 5 der Reihe 1. Netherlands Geodetic Commission, 1973.

[2] BALLEIN, K., 1985, Untersuchung der Dreiecksungleichung beim Vergleich von Kovarianzmatrizen, persönliche Mitteilung, 1985

[3] FÖRSTNER, W., A Metric for Comparing Symmetric Positive Definite Matrices, *Note, August 1995*

[4] FUKUNAGA, K., *Introduction to Statistical Pattern Recognition*, Academic Press, 1972

[5] Grafarend, E. W. (1972): *Genauigkeitsmasse geodätischer Netze*. DGK A 73, Bayerische Akademie der Wissenschaften, München, 1972.

[6] Grafarend, E. W. (1974): Optimization of geodetic networks. *Bolletino di geodesia e Scienze Affini*, 33:351–406, 1974.

[7] Grafarend, E. W. and Niermann, A. (1994): Beste echte Zylinderabbildungen, *Kartographische Nachrichten*, 3:103–107, 1984.

[8] Kavrajski, V. V., *Ausgewählte Werke*, Bd. I: Allgemeine Theorie der kartographischen Abbildungen, Bd. 2. Kegel- und Zylinderabbildungen (russ.), GVSMP, Moskau, 1958, zitiert in [**?**]

[9] Kobayashi, S. & Nomizu, K., *Foundations of Differential Geometry*, Vol. I, Interscience Publishers 1963

[10] Kobayashi, S. & Nomizu, K., *Foundations of Differential Geometry*, Vol. II, Interscience Publishers 1969

[11] Moonen, B, *Notes on Differential Geometry*, `ftp://ftp.uni-bonn.cd/pub/staff/moonen/symmc.ps.gz`

[12] Nomizu, K., Invariant affine connections on homogeneous spaces, *Amer. J. Math, 76 (1954)*, 33-65

[13] Schmitt, G. (1983): Optimization of Control Networks – State of the Art. In: Borre, K.; Welsch, W. M. (Eds.), *Proc. Survey Control Networks*, pages 373–380, Aalborg University Centre, 1983.

Earth Rotation as a Geodesic Flow, a challenge beyond 2000 ?

Erwin Groten

Abstract

The title can be interpreted in different ways; the question may be treated alternatively. In other words, what part of energy is dissipated in such a way that energy is lost, as in case of coastal tides along shelf areas, so that the path of the pole is no longer a geodesic? And what part is preserved, as in tidal dissipation within the (closed) earth-moon-system where part of the energy is transferred within the closed system to the accelerated path of the moon from the decelerated earth? For instance, the role of the anelastic mantle was in detail discussed by Molodensky and Groten (1998).

In spite of increasing knowledge on global earth parameters it is still impossible to model completely earth rotation so that prediction would be possible for various applications in space science. The transition from celestial to terrestrial reference frames is a perpetual problem in a satellite geodesy where centimeter solutions are of practical interest. In spite of a lot of progress in recent time, we still rely on empirical approaches based on world-wide networks. Also in the interpretation of earth rotation observations a variety of unexplained phenomena prevails. F. Stacey's mentioning of a „never ending" story in his geophysical textbook still holds beyond the year 2000. We know that our present official nutation formulas and the precession constant are no longer up-to date now but precise celestial space systems are now independently defined and implemented and purely conventional terrestrial systems of ITRF-type represent more an artificial model than the actual earth. Insofar the geophysical interpretation of earth rotation data is mainly affected by those modelling effects. Then care is necessary, also beyond the year 2000, when IAU in 2001 updates present official formulas.

If IERS and WGAS of IAU could agree on the removal of existing inconsistencies, the way would be open for a new fundament system in a pure relativistic frame in 2001 and also for a new consistent GRS 2001 of SC-3 in IAG. This would indeed be progress along a geodesic.

1. Introduction

Everybody knows that I was always fascinated by Erik's work on exterior calculus, related to geodesia intrinseca. So when he showed me an excellent book on mathematics by Cushman and Bates (1997), I realized in a nice way the distinct formulation of geophysical problems in terms of mathematics, of physics using a limited number of parameters and the actual world. For the rotating earth the model of a Euler top etc. illustrates in an obvious way the problems we usually face with exact formulations of questions in earth rotation and geodynamics.

In talking about integrability and integrable systems we go "back to the roots" of E. Grafarend's work. He pointed out to me quite an interesting new discussion (Cushman and Bates 1997) of that topic. Erik treated this topic in relation to heights, deflections of the vertical and other classical "integrable" or "non-integrable" quantities. The idea to relate it to earth rotation is straight forward and leads to surprising results, as far as the motion of the pole is concerned. Unfortunately, its beauty gets lost when we face dissipative systems. Nevertheless, the mathematical beauty stands for itself and is as fascinating as in case of classical integrable systems as those in exterior or intrinsic geodesy or calculus.

Earth rotation studies, in particular polar motion research where we consider the rotation parameters of a deformable earth model from a terrestrial view point – i.e. in an earth-fixed frame of reference -, is often called a „never ending story“ because with each answer to questions we are confronted with several new open questions.

The rotation of an elastic or even anelastic earth in a relativistic frame is one of the most complex and intricate problems of relativistic hydrodynamics and still basically unresolved if contemporary accuracies of better than $\pm 10^{-10}$, as now possible and realistic, are asked for.

Even in a non-relativistic frame, the problem of a heterogeneous fluid outer core surrounding a stratified inner core and surrounded by an anelastic inhomogeneous mantle with a rather irregular ocean on top has never been solved as far as the frequencies from subdiurnal up to 18.6 years are concerned.

If we focus on particular aspects, such as core-mantle coupling at the core-mantle boundary (CMB), the variety of aspects becomes evident; starting from electromagnetic coupling and related „bumps“ which were long-time ago first contemplated by my good old friend S.K. Runcorn up to topographic coupling in connection with non-hydrostatic flattening at CMB, the unresolved physics behind all this is almost endless.

2. Various aspects of earth rotation research

Basically, the formulation of earth rotation in an earth fixed frame by classical means is done in terms of Liouville's equation; for instance, for the Chandlerian motion we thus get (Sekiguchi, 1994)

$$\begin{aligned} \dot{x}_1/\sigma + x_1 &= \psi_1 \\ \dot{x}_2/\sigma - x_2 &= \psi_2 \\ LOD/86400 &= \psi_3 \end{aligned}$$

with $\sigma = 2\pi/435$ days (Chandlerian wobble frequency), Ψ_i (i=1,2,3) = excitation functions components, LOD = length of day, $\dot{x}_i = dx_i/dt$ and x_i (i=1,2) polar motion components, as usually defined. Using Euler-angles we find geodesic solutions in terms of Legendre series.

The above equation is derived for uniform (mean) motion from its general form

$$\frac{d}{dt}\left[\mathbf{I}(t)\omega + \mathbf{h}(t)\right] + \omega \times \mathbf{I}(t)\omega + \mathbf{h}(t) = \mathbf{L} \qquad \text{where} \qquad \mathbf{H} = \mathbf{I}(t)\omega + \mathbf{h}(t)$$

with

$\mathbf{I}$ = inertia tensor, $\mathbf{H}$ = angular momentum (AM); $\mathbf{h}$ =AM due to motion wrt x_i;
$\mathbf{L} = \dot{\mathbf{H}}$ = torque (dot indicates time derivatives), t = time
ω = angular velocity, x_i = body fixed axes, wrt = with respect to.

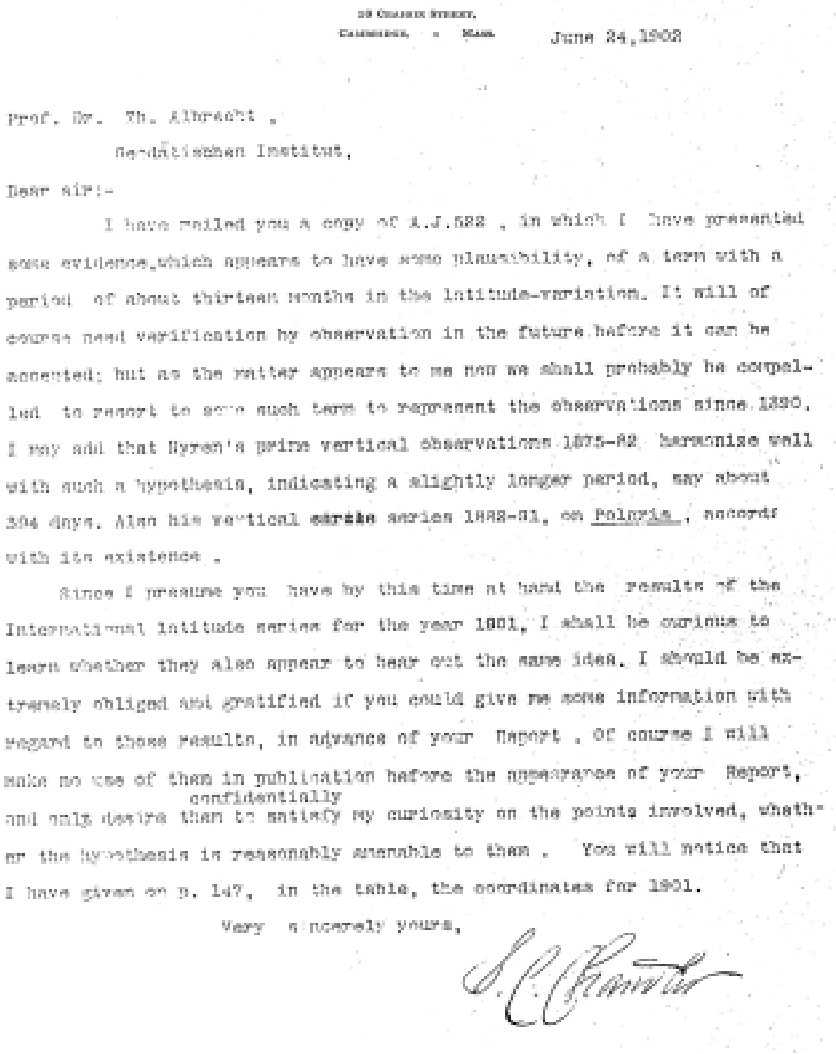
Solutions of Liouville equations for earth rotation in form of geodesics were the reason for the title of this paper suggested by Erik Grafarend. Nevertheless, it can also be interpreted in quite different ways, even in relativistic frames.

However, with the introduction of VLBI on the one side and superconducting gravimeters on the other side attenuation became a major topic which was already known from tidal (secular) friction down to seismic frequencies in terms of wave attenuation and associated quality factors Q_i which were assumed to be more or less frequency-dependent. There is still a lack of reconciliation and agreement of quality factors derived from various types of observations, e.g. at FCN frequency, but even at tidal

frequencies the impact of attenuation became obvious which affects the unified "geodesic" concepts and attenuation itself became a significant tool of geophysical investigation. The parallelism of (1) a rotating body (like the earth), (2) motion of bodies around each other (as in case of the earth-moon or satellite systems) and (3) associated tides is always fascinating but has its limits. Nevertheless, by interrelating Love and load Love numbers we may even extend that principle to load tides and I have learnt a lot on that from Peter Varga. Together with my collaborators we used superconducting gravimetry and VLBI data mainly to constrain theoretical models.

From a mathematical viewpoint, the deficient knowledge of structure parameters for the earth (density, temperature, pressure, quality factor Q for different frequencies, anelasticity parameters etc.) is one reason for the failure of exactly modelling the earth's rotation, so that the question of non-linear „ill-posed" problems becomes serious, as even small deviations in the input lead to large errors in the results of inverse (or direct) problem formulations. S.M. Molodensky adopted the name „pathological" vibrations for such ill-posed solutions in Hadamard's sense.

Consequently, the observational approach dominated earth rotation and the beginning of related investigations is indeed fully characterised by such approaches where, e.g., S. Chandler's letter to the Geodetic Institute in Potsdam is an excellent demonstration; see Fig. 1.



28 College Street,
Cambridge, Mass. June 24,1902

Prof. Dr. Th. Albrecht ,
Royal Prussian Institut,

Dear Sir:-

I have mailed you a copy of A.J.532 , in which I have presented some evidence,which appears to have some plausibility, of a term with a period of about thirteen months in the latitude-variation. It will of course need verification by observation in the future before it can be accepted; but as the matter appears to me now we shall probably be compelled to resort to some such term to represent the observations since 1890. I may add that Nyren's prime vertical observations 1875-82 harmonise well with such a hypothesis, indicating a slightly longer period, say about 304 days. Also his vertical ~~circle~~ series 1883-91, on Polaris , accords with its existence .

Since I presume you have by this time at hand the results of the International latitude series for the year 1901, I shall be curious to learn whether they also appear to bear out the same idea. I should be extremely obliged and gratified if you could give me some information with regard to those results, in advance of your Report , of course I will make no use of them in publication before the appearance of your Report, confidentially and only desire them to satisfy my curiosity on the points involved, whether the hypothesis is reasonably amenable to them . You will notice that I have given on p. 147, in the table, the coordinates for 1901.

Very sincerely yours,

S. C. Chandler

Fig. 1: Letter of S. Chandler to Th. Albrecht (Potsdam) after he had recovered the Chandler wobble (courtesy Prof. Jochmann, Berlin).

We might continue to point out other rotational components in polar motion such as the Nearly Diurnal Free Wobble (NDFW) related to free core nutation which appeared as a resonance phenomenon in the daily frequency band of earth tide observations where M.S. Molodensky started in 1957 a discussion of fluid outer core model alternatives parallel to Fedorov's polar motion consideration of related astronomical observations.

The empirical aspects of present definitions in earth rotation is obvious in the controversial definition of the Celestial Ephemeris Pole (CEP) which is presently reconsidered as only forced daily motion is subtracted from the earth rotation axis motion and free daily motion (because it cannot yet be modelled) is absent. Consequently, CEP still contains (contrary to its popular definition) daily perturbations such as NDFW.

For investigations, as we presently do at IPGD, in view of sub-diurnal interpretations we are hampered by the unsatisfying definition of CEP. Basically, this and other similar deficiencies of definitions lead to imprecise separation between nutation-precession and polar motion in the classical sense.

Until now we have still ignored the atmosphere, but if we look deeper into the subdiurnal or, generally speaking, high-frequency part of polar motion the interaction of ocean and atmosphere becomes one of the leading parts in generating functions.

Recently, even in long-periodic components the effects of „El Nino" and „La Nina" implied significant earth-rotation perturbations associated with „high" and „low" sea level variations in equatorial latitudes of the Pacific Ocean, also associated with low and high temperature and consequent density effects.

It is by no means clear, to what extent such climatic variations associated with strong wind and thunderstorms regionally affect long-term earth motion besides Milankovitch-cycles which originate from the motion of the earth in the ecliptic. B.F. Chao has recently discussed the climatic effects of variations of the obliquity, defining the angle between CEP and the ecliptic, as well as the influence of water reservoir water level variations and (global) earth rotation.

Even though monitoring the earth's rotation by VLBI and similar techniques now leads to relative accuracy of about $\pm 10^{-9}$ we still suffer in the interpretation of such data from the incomplete separation of plate tectonic motion at the VLBI-observatories from polar motion as deduced from such data.

Even completely independent techniques, such as inertial systems using INS-laser ring technology where, based on the Sagnac-effect, absolute motion is derived for such an INS-station, are effected by this deficient separation between plate tectonics and polar motion.

Consequently, purist people like H. Eichhorn in Gainsville/Florida always questioned the possibility to define a global rotation of a deformable earth but rather insisted in an individual rotation vector for every earth surface element at any observatory site.

With new possibilities to derive temporal variations of harmonic coefficients of the earth's gravitational potential, such as C_2^1, S_2^1 , the motion of the principal axis of inertial became an observable quantity but this axis is not identical with and difficult to relate to the conventional terrestrial pole of the IERS system.

In speaking of a „geodesic flow" instead of a „geodetic flow" of earth rotation we introduce basically a relativistic thinking. To model the motion of the earth in absolute space or in terms of a relativistic frame is still more controversial than ever before. One of the reasons for the intricacy is the lack of clear definitions and the deficiencies of implementing related corrections and reductions.

Take a simple example: Until 1998 it was clear what we mean by luni-solar and planetary precession. With ± 10^{-9} astrometry the planetary effects in luni-solar precession became significant as planetary attraction can no longer be ignored besides luni-solar attraction. Take another example:

Free motion due to ocean tides in the definition of Universal Time is significant and it therefore illogical to ignore free motion in CEP. "„Shifting" the CEP definition by including free motion (as far as it can be modelled) is simply a matter of definition but free motion is not constant in time but varying. Third case: Ignoring secular corrections in the transition of one dynamic time (such as TDT) to another, such as TDB, did not seriously affect results in the past; we may define theoretical „times" in a way similar to aberration where, we also leave out certain terms, by definition. However, if the „physical" meaning is not fully compatible with the abstract definition, difficulties arise wherever a step higher in accuracy is demanded. Similar problems arise with „geodesic precession" in case of defining a „non-rotating frame of reference" according to B. Guinot and others.

3. Outlook

As a result, IAU is now in a process of clarification in order to end up in the year 2001 with a relativistic frame of reference to which motion can be related exactly. The limits, however, are obvious as the definition of an inertial system is even not completely satisfying: an unaccelerated system has a clear dynamic meaning but its kinematical behavior is not defined: and we are close to the „Big Bang"-problem where the expanding world model does not answer all questions, as far as the kinematics are concerned.

So let us finally return to classical physics in an attempt to solve geophysical questions related to „generating functions". There are so many open questions in classical physics that, besides the ± 10^{-9} domain, still a lot of progress has to be made in classical physics applications. The basic formulae applied by us are shown in Table 1.

Table 1: The work from 1963 to 1998 in dealing with earth rotation at IPGD

1963: Hough equation:

$$\left(\mu\nabla^2 - \frac{\partial^2}{\partial t^2}\right)\nabla^2 p + 4\Omega^2 \frac{\partial^2}{\partial t^2}\ \frac{\partial^2 p}{\partial z} = 0 \qquad (1)$$

$\partial/\partial t$ = Eulerian derivative; t = time; p = pressure

1973: Euler's equation

$$p[\partial q/\partial t + (q\cdot\nabla)q] = -\nabla P - p\nabla\phi \qquad (2)$$

q = velocity in an inertial system, ϕ = Potential of external forces
$-\nabla P$ = force per unit volume; Ω = spin, μ = shear modulus.

1983: Navier's equation (without rotation; Stokesian form):

$$\frac{\partial q}{\partial t} + (q\cdot\nabla)q = -\nabla p + \nu\nabla^2 q \qquad (3)$$

$\nu = \eta/p$; p = density; η = viscosity (dynamical shear)

1993: Poincare's equation:

$$\nabla^2 p - \frac{4\Omega^2}{\omega_i^2} \frac{\partial^2 p}{\partial t^2} = 0 \tag{4}$$

1996: Boussinesq's equation:

$$\rho \frac{Dq}{Dt} = -\nabla P + \rho g \tag{5}$$

with g = gravity

1998:

$$\frac{\partial \omega_i}{\partial t} = \nabla x(q \times \omega_i) + \upsilon \nabla^2 \omega_i \tag{6}$$

with $\omega_i = \nabla x q$.

Liouville equation, as given above.

To a certain extent, the situation in earth rotation studies appears similar to the study of gravimetric problems, e.g., the determination of big „G", where observational intricacies lead to various assumptions and complicated theoretical explanations where, at the end, observational difficulties finally may primarily explain the controversial points. Take, e.g., the varying results obtained recently for big „G", i.e. Newton's Gravitational Constant (Schwarz et al., 1998, Kestenbaum, 1998).

Similarly, we still discuss frequency modulation of polar motion without finding appropriate models which could explain temporarily varying periods. As far as the Chandler period is concerned it may indeed be assumed that the Chandlerian period is not at all a free period but rather a conglomeration of forced vibrations around a period of 1.2 years. Who knows? In case of our results with sub-diurnal varying periods around the tidal bands oceanic effects of time-varying periods could be caused by non-linear effects in the ocean-atmosphere interaction at very high frequencies. But as in case of big „G" some people may agree with the Birmingham physicist C. Speake (see also Speake (1988) and Groten (1988)) who was quoted in Science, vol. 282, 18 Dec 1998 on page 2181 „Nobody gives a damn about Big G" also the earth rotation study will be a never ending story, so it will continue well beyond the year 2000.

The final answer to Erik's question in the title has certainly to be postponed but it is sure that we do not fully understand energy dissipation so that the deviations of pole path from a geodesic cannot be described in detail. The non-conservative forces for ocean-sea interaction may be partly understood for well surveyed systems such as the zonal winds in Antarctic regions, even for El-Nino-events but, e.g., for the short-period (sub-diurnal) phenomena investigated by us we are far from understanding what part is actually dissipative in the sense of non-conservative, then leading to deviations from an geodesic as a pole path.
If we follow Dziewonski, Kanamori etc. in assuming that part of the origin of Chandler wobble is due to (large) earthquakes there is a similar dissipation problem as the exact mechanism is unknown in which way energy due to seismic moments is radiated to polar motion, to heat and other forms of dissipation.

References:

Cushman, RH, LM Bates (1997) Global aspects of classical integrable systems, Birkhäuser-Verlag Basel, 427 p

Groten, E. (1988) Simulations for Studying the Yukawa-term. ZfV 8, 366-373

Kestenbaum, D. (1998) Gravity Measurements close in on Big G. Science, Vol 282, 2180-2181

Molodensky, S.M., E. Groten (1998) On the dynamical effects of an inhomogeneous liquid core in the theory of nutation. Journal of Geodesy 72: 385-403

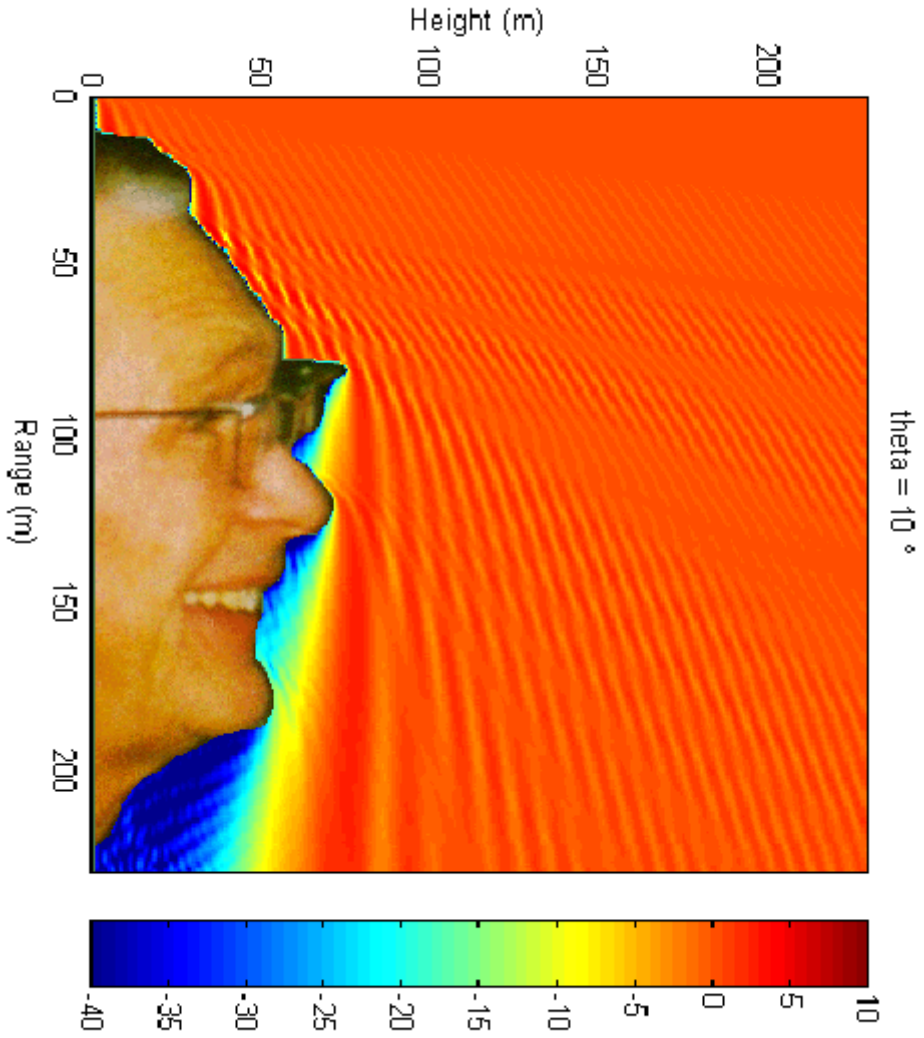
Schwarz, J.P., D.S. Robertson, T.M. Niebauer, J.E. Faller (1998) A free-fall determination of the Newtonian constant of gravity. Science, Vol. 282, 2230-2234

Sekiguchi, N. (1994) Theory of the excitation of the Earth's polar motion. Sendai Astron. Raportoi, 419, p. 226, Sendai Univ. Japan

Speake, C.C. and T.J. Quinn (1988) Search for a Short-Range, Isospin-Coupling Component of the Fifth Force with Use of a Beam Balance. Physical Review Letters, Vol 61, 12, 1340-1343

Propagation Modelling of GPS Signals

Bruce M Hannah, Kurt Kubik and Rodney A Walker



Simple Forward Propagation of 60 MHz Signal over Profile of Prof. Erik W. Grafarend

ABSTRACT

It has been previously shown [1] that the Parabolic Equation (PE) technique is well suited to solving GPS propagation problems over large domains. This paper presents the results of a modified model, which includes the effects of backscatter and provides time-domain representation of the propagated GPS signal. Results are presented for various difficult GPS propagation environments, including the determination of the Multipath Channel Impulse Response (MCIR) from the PE.

INTRODUCTION

The problem of signal multipath, in precision GPS applications, continues to remain largely unsolved. Traditionally two basic approaches have been taken to attempt to mitigate multipath effects in precision applications; special antenna designs; and specialised receiver architectures.

Special antenna designs, to mitigate multipath, are choke ring antennas and antenna ground planes. The fundamental aim of these designs is to physically limit the energy in unwanted signals, based on the direction of arrival at the antenna. A ground plane attenuates signals arriving at negative incident angles, whilst a choke ring antenna attenuates signals arriving below about 10°. For many environments these special antennas provide adequate performance, but for a truly ubiquitous sensor, it cannot be assumed that all multipath will arrive below 10° (see Figure 1). In the mining environment (or in any urban environment) a large variety of reflected signals will arrive from angles above 10° (side-wall reflections), thus reducing the effect of using such an antenna in these environments.

Another form of antenna mitigation useful for precision applications are beam-forming antennas. These antennas effectively form a narrow high-gain beam at the satellite signal of interest. Whilst these systems provide excellent multipath rejection in all environments, they are generally large and difficult to use in any practical system, since precise knowledge of the antenna attitude is required.

Special receiver architectures to minimise multipath effects essentially involve specialised correlator designs. Past developments in this area have included early-late slope (ELS)[3], narrow correlation[4], strobe correlation[5] and the Multipath Estimating Delay-Lock Loop MEDLL[6]. These approaches have all achieved some improvement in mitigation of the error effects of multipath. However, Weill[7] examined the theoretical limits for mitigation of code-phase multipath and found that one estimator for mitigating multipath can be claimed as optimal in a certain sense. Known as the minimum-mean-square error (MMSE) estimator, the multipath parameters are treated as random variables and the observed signal is used to construct a conditional probability density for the parameter values. In view of the disparity between the performance of the MMSE and what appears to be the state of the art, there seems to be room for more rigorous approaches.

The PE modelling technique is a linear time invariant model, however with the use of Fourier Time synthesis techniques, time dependence is able to be re-instated into the model. The resultant model is known as the PE-based Time Analysis (PETA) model.

The model is a wide-angle PE implementation optimised for GPS propagation. Multipath is characterised by amplitude, time delay, phase, and phase rate-of-change relative to the direct line-of-sight signal[2]. This model provides complete decomposition of the complex electromagnetic field components into these multipath parameters, through the simulation of the Multipath Channel Impulse Response. Results are presented for a variety of terrain features.

A discussion is made of the aims of this project whereby a complete environmental and receiver model is developed. This model encompasses the transmission of the signal from the satellite, its interaction with localised terrestrial terrain, and the manner in which the receiver correlators interpret the signal. This modelling strategy provides a tool that will assist in determining any relationships between multipath propagation behaviour and its effect in the receiver. These types of effects need to be determined before innovative mitigation techniques can be considered.

Without doubt, the progress in GPS multipath mitigation research has been significant and will continue. As the theoretical limits of correlation performance are approached there is justification for investigation of alternative receiver-based mitigation strategies. The PE propagation model, in conjunction with GPS receiver models, will form the basis for a comprehensive multipath analysis tool, necessary for extensive investigation of multipath mitigation techniques.

MULTIPATH MODELLING

The determination of propagation behaviour is important in the understanding of GPS multipath errors. The superposition of delayed replicas of the direct ranging signal leads to distortion of the received signal at the GPS antenna, and results in ranging errors of varying magnitude. In trying to develop an understanding of the impact these multipath signals have on the receiver, it is necessary to characterise the multipath signal. Multipath can be characterised by the three key parameters mentioned earlier; namely relative time delay, relative amplitude, and phase upon reflection. Together these parameters form the Multipath Channel Impulse Response (MCIR).

In this work it is shown how with the use of a PE propagation model and Fourier time-synthesis, the MCIR can be determined for various environments. Typical multipath propagation mechanisms are shown in Figure 1. To realistically model the propagation environment, we must not only deal with reflection, but also diffractive effects.

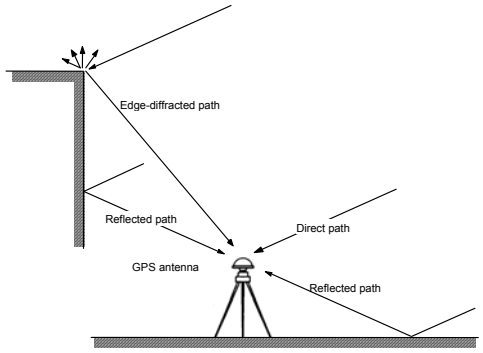


Figure 1 – Typical GPS antenna environment

The PE method is a full-wave solution to Maxwell's equations and hence provides the basis for the development of a multipath analysis tool.

PARABOLIC EQUATION MODELLING

The starting point for the development of an electromagnetic parabolic equation model is with the Helmholtz wave equation (1), for a field component, ψ, with assumed time dependence $e^{-j\omega t}$ [8].

$$\nabla^2 \psi + k^2 n^2 \psi = 0 \tag{1}$$

where $k = \frac{2\pi}{\lambda}$ is the free space wave-number,

and $n = \frac{k}{k_0}$ is the refractive index.

We make use of cylindrical coordinates with assumed far-field invariance in azimuth, and remove the rapid phase variation through the reduced function u, with

$$\psi(x,z) \approx u(x,z)e^{jkx} \tag{2}$$

This yields the simplified elliptic equation,

$$\frac{\partial^2 u}{\partial r^2} + 2jk_0 \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} + k_0{}^2\left(n^2 - 1\right)u = 0 \tag{3}$$

By defining the operators

$$P = \frac{\partial}{\partial r} \quad \text{and} \quad Q = \sqrt{n^2 + \frac{1}{k_0{}^2}\frac{\partial^2}{\partial z^2}}$$

we factorise equation (3) and select only the outgoing wave component. The result is a one-way, two-dimensional parabolic equation, which, for n equal to 1 (free-space propagation), is given by[9];

$$\frac{\partial u}{\partial r} + jk_0\left(1 - \sqrt{1 + \frac{1}{k_0{}^2}\frac{\partial^2}{\partial z^2}}\right)u = 0 \tag{4}$$

This equation is exact, within the limits imposed by the far-field approximation, and is evolutionary in range, allowing solution by an efficient Fourier transform based stepping technique[10]. The solution at a range-step (Δx) is given by

$$u(x + \Delta x, z) = F^{-1}\left[e^{jk\Delta x\left(\sqrt{1-\frac{p^2}{k^2}}-1\right)} F\left[u(x,z)\right]\right] \tag{5}$$

where p is the vertical wave-number and is related to k by $p = k\sin\theta$, with θ, the propagation angle relative to the horizontal. The *p-domain* defines the angular spectrum of the field, and together with the *z-domain*, they form a Fourier transform pair.

As can be seen from equation (5) we simply need to define some initial field condition at $x=0$, and march the solution out in range. For the case of GPS signal propagation, the field close to the Earth is essentially a uniform plane-wave. Thus we define our initial, or starting field condition, as a combination of incident and ground reflected Transverse Electric (TE) plane waves, and can write

$$E_i = E_0 e^{-jk(z\sin\theta - xcoz\theta)} + E_0 e^{-jk(z\sin\theta - xcoz\theta)}\Big|_{x=0} \tag{6}$$

Where E_0 is the incident field amplitude.

BOUNDARY-SHIFT TECHNIQUE FOR ARBITRARY TERRAIN

The boundary-shift technique[11], for handling arbitrary terrain within the PE code, involves the shifting of the field array (aperture) either up or down to account for the shift in the boundary position, and thus satisfy the boundary conditions. The field aperture immediately to the left of any obstructing terrain is stored then shifted down according to the height of the terrain element. The lower elements, those that would propagate into the terrain, are discarded and zeros inserted at the top of the array to maintain the correct number of elements. This modified field array is then propagated to the next array, with the Fourier-step technique. At negative terrain transitions, the reverse procedure is applied. The array is shifted up by the corresponding height, with the top elements discarded, and zeros inserted at the element positions where the field is obscured by the terrain. The result of the boundary shifting technique is simply a restructuring of the domain representation to that of a field propagating over a plane earth while accounting for diffractive effects over terrain.

IMPLEMENTATION OF BACKSCATTER

In the development of the PE, it was necessary to assume that the field was outgoing only. This one-way restriction can be lifted by using a store and forward method of back-propagating field components. The steps for implementation of a two-way PE model derived from a one-way model are as follows;

- The field is propagated with the one-way PE model in the forward ($+x$) direction
- The field components that will propagate into terrain (potential back-scatterers), are identified.
- These field values and indexes to their positions within the domain are stored for later use.

- The terrain profile and domain are mirrored vertically such that the one-way implementation can again be used without modifying the existing PE model code.
- The one-way PE is then used to propagate the stored field values that are added into the model as initial field conditions of the back-propagation.
- The field components of the forward and back implementations are then added to provide the resultant full field.

The use of this technique is justified by image theory, where the components at a vertical interface would travel to an image of the domain mirrored vertically about the vertical reflector. In addition, the method is complementary to the boundary-shift technique, where the down-shifted components normally discarded, are stored for use as the initial field values for a two-way PE implementation.

TIME-DOMAIN VIA FOURIER SYNTHESIS

The solution of the time-dependent field equation can be obtained by the Fourier transformation of the PE field solution[12], namely

$$u(x,z,t) = \int_{-\infty}^{\infty} S(f)\,u(x,z,f)\,e^{j2\pi ft}\,df \tag{7}$$

where *S(f)* is the spectrum of a source pulse and *u(x,z,f)* is the spatial transfer function derived from the PE modelling process. This integral is evaluated using Fast Fourier Transform (FFT) techniques at the spatial point of interest in the model domain, i.e. the antenna location. For this work we have chosen as our GPS time source, a *sinc* pulse of 1 nanosecond duration, modulated at the GPS *L1* frequency. The MCIR is the output of the PE Time analysis (PETA), and is given as a time series of delayed, and attenuated source pulses. The complex field in terms of the MCIR at a spatial point (*x,z*), is given by the addition of the decomposed plane waves

$$\psi_{PETA}(x,z) = e^{j2\pi ft_0} + \sum_{i=1}^{M} \alpha_i e^{j(2\pi ft_i + \phi_i)} \tag{8}$$

Here the first term represents the line-of-sight signal with a propagation time of t_0, from an arbitrary domain incident boundary at, $x=0$. The summation term represents, the M multipath signals, where α_i and t_i represent respectively, the i^{th} relative multipath amplitude and time of arrival. The phase term, ϕ_i, is the resultant phase shift due to the boundary reflection(s) for the i^{th} multipath signal. This equation can be normalised by assuming zero reference phase for the LOS signal. This normalisation is simply a change from absolute time delay, as presented by the PETA, to relative time delay, and is given by,

$$\psi_{PETA}(x,z) = 1 + \sum_{i=1}^{M} \alpha_i e^{j(2\pi f\tau_i + \phi_i)} \tag{9}$$

where τ_i is the time delay relative to the LOS signal.

DOMAIN REPRESENTATION AND PERFORMANCE

The propagation domain is represented by a two-dimensional plane that is specified by the azimuthal direction to the satellite, the maximum height, and the maximum range to be modelled. The antenna can be located at any point on the plane, above the terrain. Terrain information will be input from Digital Terrain Models (DTMs). The model domain is depicted in Figure 2.

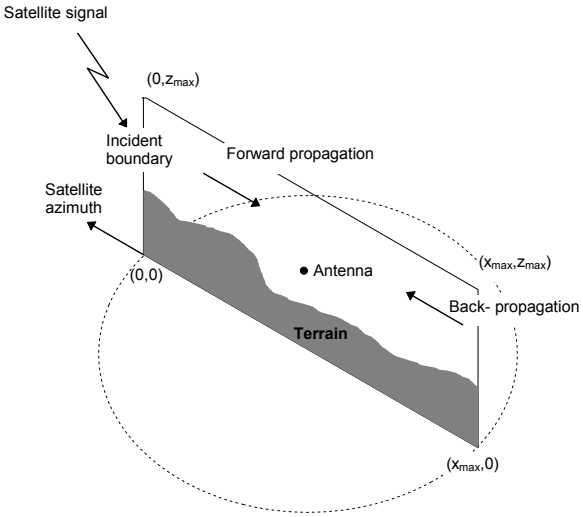


Figure 2 - Model Domain

The definitions of forward and back-propagation are relative to the directions specified in Figure 2.

Model simulation times for single frequency field values, with forward propagation only, is given by the proportionality

$$T_{PE}^{1-way} \propto k\theta A \tag{10}$$

where, k is the wavenumber, θ is the propagation angle, and A is the area of the domain plane. With inclusion of back-scatter this increases to

$$T_{PE}^{2-way} = (L+1)T_{PE}^{1-way} \tag{11}$$

for L back-scatterers. For the PETA the simulation time is

$$T_{PETA} = \frac{2\tau_{win}}{\tau_{pulse}} T_{PE}^{2-way} \tag{12}$$

where τ_{win} is the width of the time analysis window, and τ_{pulse} is the source pulse width.

GPS PROPAGATION RESULTS

Having established the basis for the modelling technique, results for several multipath environments are presented. These modelling results are based on the L1 GPS frequency of 1.575 GHz.

Validation of C/No

A comparison was made of predicted C/No against measured C/No[9]. A test site was chosen and data was recorded at 1 second epochs. The terrain, over which the satellite signal had propagated, was a single small building. The results are shown in Figure 3.

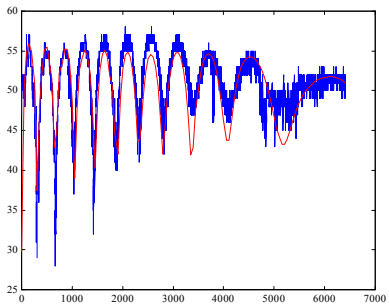


Figure 3 - Comparison of C/No Predicted vs. Observed

Figure 3 shows quite good agreement with measured results. These measurements were made in a relatively simple environment where the satellite signal was reflected from the roof-top of a large metal shed. The fading period and depth of fades agree well. Although this satellite was chosen because it's azimuth variation was small over the observation period there was a finite change which will affect the validity of the terrain profile used in the model.

Forward Specular Reflection Analysis

As a reference problem, a simple forward specular reflection problem is examined. The geometry of this problem is depicted in Figure 4. In this multipath situation we have the direct LOS signal (L) and a single multipath signal (R) arriving at the antenna.

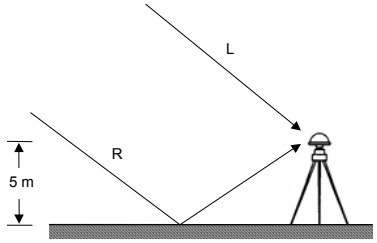


Figure 4 – Geometry of Forward Specular Reflection

A GPS satellite, rising in elevation from 5 degrees to 10 degrees, is modelled. Figure 5 and Figure 6 show, respectively, the calculated PE field, and the PETA result for a propagation angle of 8 degrees.

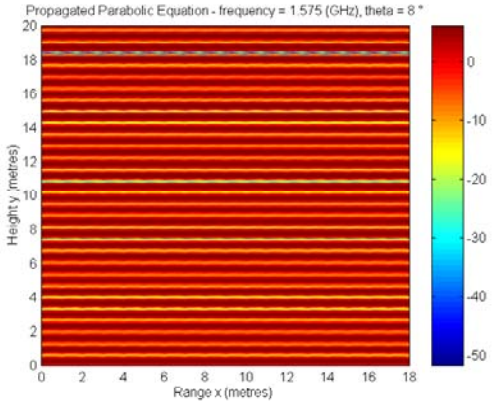


Figure 5 – Computed Field for Forward Scatter problem

This plot of the field strength shows the classical interference region pattern, with constructive and destructive interference clearly evident as a function of height. The computed field shown in this figure is for the full-space field and does not take into account the antenna gain pattern.

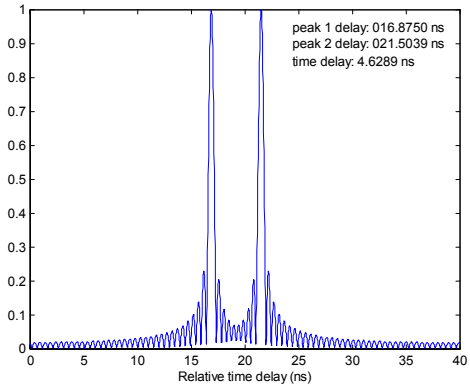


Figure 6 – Time analysis of Forward Scatter Problem

The time-domain analysis clearly shows the LOS and the multipath signals. Each of the multipath parameters is extracted from the PETA results, and using equation 9, we can reconstruct the total field. Figure 7 shows a comparison of the PETA estimated field compared to the full field solution as given by the PE propagation model.

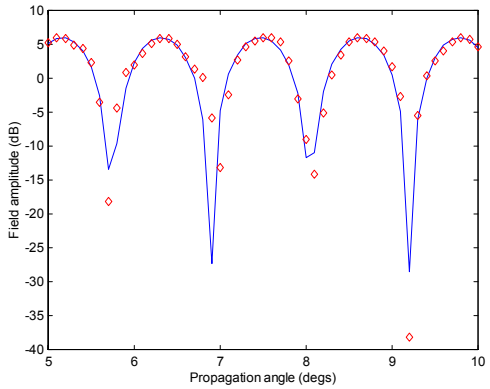


Figure 7 – Comparison PETA field vs. PE field

This figure again shows the classical fading pattern for a single multipath reflection. The results from the time-domain reconstruction are in good agreement with the full field result.

Forward Diffraction Analysis

Another propagation mode to consider is forward diffraction. A GPS satellite is modelled rising over a terrain obstuction, from an initial elevation angle of 5° to a final angle of 15°. In this case we can expect diffraction effects to dominate. The geometry of the problem is depicted in Figure 8, where the LOS signal exists, but *n* diffracted signals may also exist.

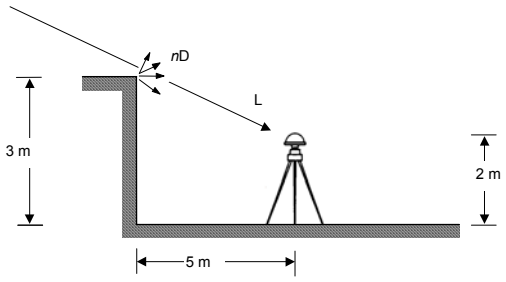


Figure 8 – Geometry of Forward Diffraction Problem

An instantaneous plot of the PE field, for a propagation angle of 10°, is given in Figure 9. The incident shadow boundary (ISB) can be seen (the interface between incident and diffraction illumination as described in [1]) as the -6dB boundary. Typical diffractive effects can be seen below the ISB (25-30dB attenuation of LOS signal) and forward scattering is seen above the ISB (c.f. Figure 5).

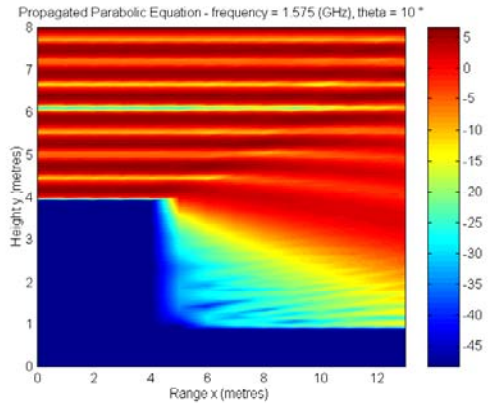


Figure 9 – Computed Field for Forward Diffraction Problem

Diffractive effects are evident in the region below the ISB. The diffractive effect on C/No is presented in Figure 10.

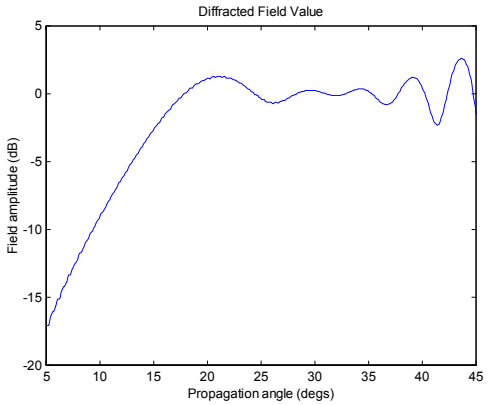


Figure 10 – Diffraction Effect on C/No

The computed field, for a satellite elevation change of 5° to 15°, shows a classical diffraction response. The field strength slowly rises as the antenna comes out of the shadow region, overshooting the incident field strength and oscillating about the 0dB LOS level. At approximately 35° the diffractive effect is almost zero and ground reflection interference starts to dominate. The ideal response for this scenario is for no field until 11° satellite elevation (when the satellite and antenna are LOS) and then a flat 0dB signal. The deviations from this ideal, seen in Figure 10, represent multipath from diffraction and ground reflections (forward scattering).

Figure 11 presents a comparison of the delay of the diffracted signal (upper plot) to that of the unobstructed line-of-sight (lower plot). This clearly indicates the additional path length caused by the diffraction of the signal around the terrain edge. If a receiver has a dynamic range of better than 20 dB then it is able to acquire and maintain track of the diffracted signal. At 5° the diffracted path delay is 0.2 ns representing approximately a 6 cm range error. The convergence of the two lines indicates that the diffractive effect (a function of elevation angle) eventually reduces to zero and the result is the same as that for the LOS case.

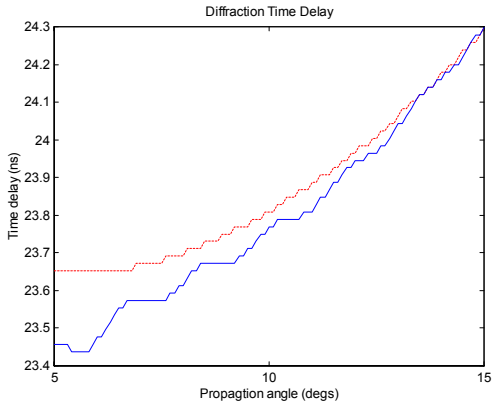


Figure 11 – Time delay profiles for Diffraction Problem

Stepped Back-Scatter Problem

Finally a much more complicated problem is examined. This is an example of an environment that can easily be found in urban environments. The geometry of the problem is seen in Figure 12.

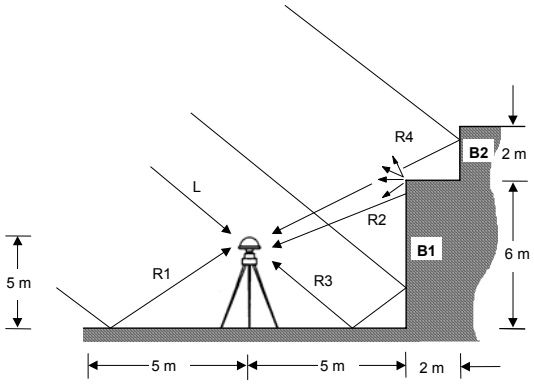


Figure 12 – Geometry of Stepped Back-Scatter Problem

This problem is referred to as a stepped backscatter example, where in addition to the forward scatter, signals are also scattered in the reverse propagation direction, from two distinct interfaces. Figure 13 shows the time-domain results for a 5° propagation angle.

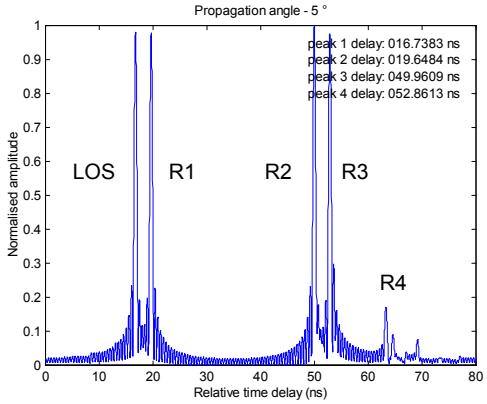


Figure 13 – Time Domain of Stepped Back-Scatter Problem (5°)

Here the LOS signal, forward scatter (R1), and two additional back-scattered multipath signals (R2 and R3) reflected from interface *B1* (see Figure 12) are seen. The first of these back-scattered multipath signals is identified as a backscatter from above (R2), that is, the LOS is reflected from the *B1* interface and arrives at the antenna location from a positive elevation. The next signal is backscatter from below (R3), and is a reflection of the LOS from a combination of ground and interface (R3). At 5° elevation the *B2* interface is obstructed by the *B1* step. Close examination of Figure 13, shows some low level signal from the *B2* interface arrives at the antenna, but diffractive effects have reduced it's influence (diffracted R4).

The propagation mechanisms in this situation become evident at higher propagation angles. In Figure 14 the results for PETA at 12.5° satellite elevation are shown.

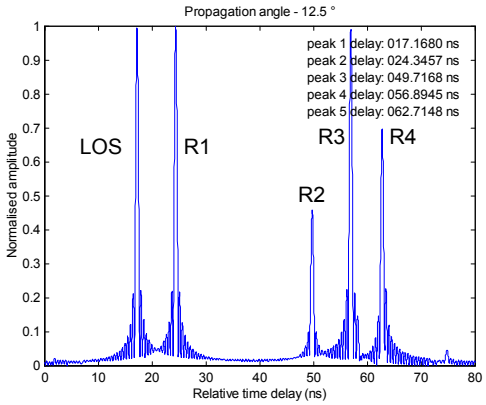


Figure 14 – Time Domain of Stepped Back-Scatter Problem (12.5°)

Figure 14 highlights the change in field due to a slightly higher satellite elevation. The multipath signal (R2) from *B1* (above the antenna) is no longer a direct reflection and is now a diffracted signal, hence its reduced signal level. The multipath signal from *B2* (R4) is becoming a direct reflection and thus it is becoming a stronger signal. At 15° (see Figure 15) the effect is more pronounced and the reflection from *B2* is essentially *in-the-clear*. It is noted that the reflection from *B1/ground* (R3) is unaffected, and that diffractive effects now dominate the reflection (R2) from the *B1* interface.

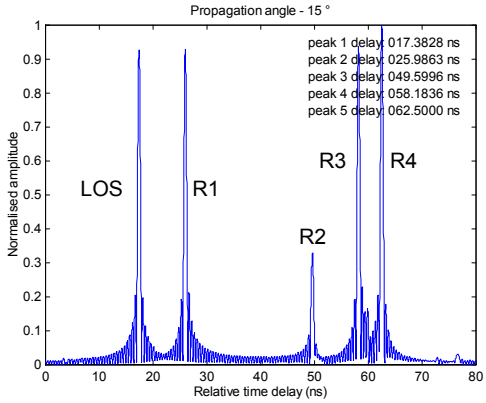


Figure 15 – Time Domain of Stepped Back-Scatter Problem (15°)

The total multipath propagation environment, as a function of elevation angle, is shown in Figure 16. Here we see the full influence of the diffractive effects for this situation.

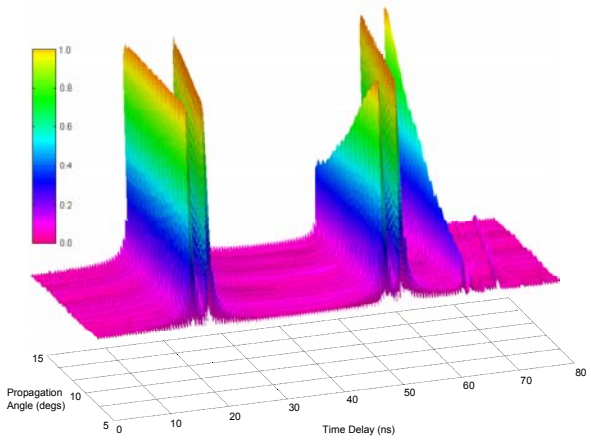


Figure 16 – Time delay profile as a function of satellite elevation

TERRESTRIAL PROPAGATION RESULTS

A common problem on many open cut mines is that of poor communications systems management. The open cut mine environment is an extreme environment for communications systems design due to the 'harshness' of the terrain. This 'harsh' terrain results in large shadow losses, requiring higher power transmissions and strategic placement of communications repeaters. The other factor affecting the communications problem, is that the shape of the mine is constantly changing. A mine communications system that was designed to provide 90% mine pit coverage in June 1997 for a strip mine, may not provide the required pit coverage in June 1998, since the next strip is now being processed. The increased distance from the communications transmitter now results in larger shadow losses, thus reducing the communications coverage in the pit.

Figure 17 shows the propagation over 4 pits on a typical open cut mine site, the 2D terrain profile was obtained from digital photogrammetry of a mine in Queensland. The simulation is for a 157Mhz Gaussian beam, mounted on a 20 m high tower, with the antenna main beam on the horizontal. If we wish to determine the change in performance by tilting the antenna down at 5º, we simply run the PE model again with a changed incident field condition. The results of this simulation appear in Figure 18. Comparison between these two figures shows that an improved coverage of the first pit can be made by tilting the antenna down at 5°. This does not affect the coverage in the second pit as can be seen in the simulations.

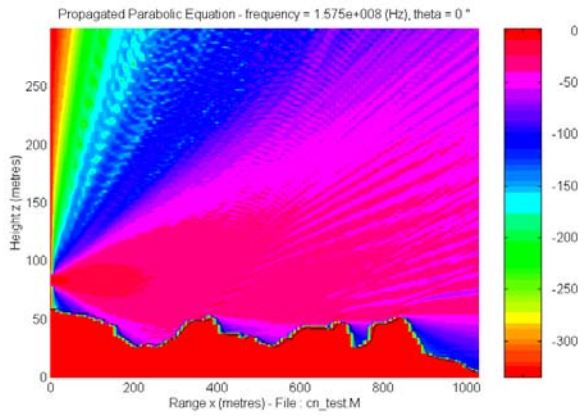


Figure 17 — Mine Repeater Propagation, 0° Antenna Tilt

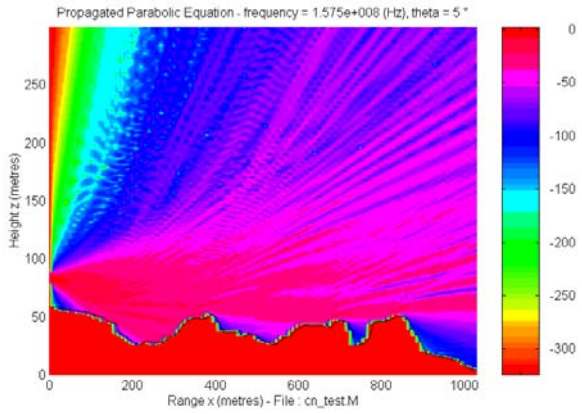


Figure 18 — Mine Repeater Propagation, 5° Antenna Tilt

CONCLUSIONS

It was seen in the stepped back-scatter example, how a relatively simple terrain environment can give rise to multiple delayed replicas of the GPS signal. It is hoped that these examples provide insight into the problems encountered when trying to overcome GPS multipath.

The advantage of numerical techniques, as discussed in this paper, is that the exact multipath nature of a complicated environment can be understood and decomposed. It is hoped that by combining a number of complex receiver models with the PE propagation models presented here, that a complete software-based satellite to user modelling system can be developed.

ACKNOWLEDGEMENTS

This work was carried out in the Cooperative Research Centre for Satellite Systems with financial support from the Commonwealth of Australia's Cooperative Research Centres Program.

B Hannah would like to thank the British Council for funding, through a postgraduate bursary award, a three month study visit to the Radio Communications Research Unit at Rutherford Appleton Laboratory (RAL) UK.

The authors would also like to thank Dr Mirielle Levy (RAL) for her continued support of this work.

REFERENCES

[1] R. A. Walker, "Numerical Modelling of GPS Signal Propagation," Presented at ION GPS-96, Kansas City, 1996.

[2] M. S. Braasch, "On the Characteristics of Multipath Errors in Satellite-Based Precision Approach and Landing Systems," PhD Thesis, *Department of Electrical and Computer Engineering*. Ohio University, 1992, pp. 203.

[3] B. Townsend and P. Fenton, "A Practical Approach to the Reduction of Pseudorange Multipath Errors in a L1 GPS Receiver," Presented at 7th International Technical Meeting of the Satellite Division of the Institute of Navigation., 1994.

[4] A. J. Van Dierendonck, P. Fenton, and T. Ford, "Theory and Performance of Narrow Correlator Spacing in a GPS Receiver," Presented at the Institute of Navigation National Technical Meeting, San Diego, CA, 1992.

[5] L. Garin, F. Van Diggelan, and J. Rousseau, "Strobe and Edge Correlator Multipath Mitigation for Code," Presented at the 9th International Technical Meeting of the Satellite Division of the Institute of Navigation., Kansas City, MI, 1996.

[6] R. D. J. van Nee, “Multipath and Multi-Transmitter Interference in Spread-Spectrum Communication and Navigation Systems,” PhD Thesis, Faculty of Electrical Engineering, Telecommuncation and Traffic Control Systems Group. Delft University of Technology, 1995, pp. 205.

[7] L. R. Weill, “Achieving Theoretical Accuracy Limits for Pseudoranging in the Presence of Multipath,” Presented at the 8th International Technical Meeting of the Satellite Division of the Institute of Navigation., Palm Springs, CA, 1995.

[8] K. H. Craig, “Propagation Modelling in the Troposphere: Parabolic Equation Method,” *Electronics Letters*, Vol. 24, pp. 1136-1139, 1988.

[9] R. A. Walker, “Operation and Modelling of GPS Sensors in Harsh Environments,” PhD Thesis, *School of Electrical and Electronic Systems Engineering*: Queensland University Of Technology, 1997.

[10] R. H. Hardin and F. D. Tappert, “Applications of the Split-Step Fourier Method to the Numerical Solution of Nonlinear and Variable Coefficient Wave Equations,” *Siam Review*, Vol. 15, pp. 423, 1973.

[11] A. E. Barrios, “Terrain and Refractivity Effects on Non-Optical Paths,” presented at AGARD Electromagnetic Wave Propagation Panel Symposium, Rotterdam, The Netherlands, 1993.

[12] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, “Broadband Modelling,” in *Computational Ocean Acoustics*, R. T. Beyer, Ed. New York: AIP Press, 1994.

A Short-Cut Method for Computing Positive Variance Component Estimates

Joachim Hartung

Summary

In a general variance component model with positive variance components a short-cut method is presented that yields almost everywhere for these components positive estimators that are invariant with respect to mean value translation and stay near the unbiasedness.

Key Words: Variance components, Minque, positive Minque, Invariance, approximate unbiased positive variance estimates.

Zusammenfassung

In einem allgemeinen Varianz-Komponenten-Modell mit positiven Varianzkomponenten wird eine verk|rzte Methode vorgestellt, welche f|r diese Komponenten fast |berall positive Schdtzer ergibt, die invariant bzgl. Mittelwerttranslationen sind und nahe der Unverzerrtheit bleiben.

1 Introduction

In a general variance components model there is the problem that unbiased quadratic estimators, or also maximum likelihood estimators if a distributional assumption is made, of the variance components can take on with a positive probability negative values for nonnegative variance components. These estimators are put then in such cases equal to zero, which for a usually strictly positive variance component is an unsatisfactory procedure. Therefore in the following a short-cut procedure is derived that overcomes this deficiency by yielding almost everywhere positive variance component estimators staying near the unbiasedness.

2 The Method

Let us consider the linear variance component model

$$z \sim \left(X\beta,\ \sum_{i=1}^{m} \alpha_i \cdot U_i \right) ,$$

that consists of an n-dimensional random variable z with mean value

$$\mathrm{E}\, z \quad = \quad X\beta$$

and variance-covariance matrix

$$\mathrm{Cov}(z) \quad = \quad \sum_{i=1}^{m} \alpha_i \cdot U_i \ ,$$

where the $(n \times k)$-design-matrix X and the m symmetric positive semi definite $(n \times n)$-matrices U_i, $i = 1, \ldots, m$, are known, while the parameter β varies in $\mathbb{R}^k$ and the parameter $\alpha = (\alpha_1, \ldots, \alpha_m)^T$ in $\mathbb{R}^m_+$, the positive orthant of $\mathbb{R}^m$, and we assume $\mathrm{rank}(X) < n$, and $\mathrm{Cov}(z)$ to be positive definite.

The problem considered here is to find quadratic estimates for the variance components $\alpha_1, \ldots, \alpha_m$, which should be positive almost everywhere, i.e. with probability one, and invariant with respect to the group Γ of mean value translations,

$$\Gamma \quad = \quad \left\{ z \longmapsto z + X\beta \mid \beta \in I\!R^k \right\} .$$

A maximal invariant linear statistic y with respect to Γ is given by

$$\begin{aligned} y &= \left(I - XX^+\right) z \\ &= \mathrm{Proj}_{|\mathrm{Range}(X)^\perp} z \ , \end{aligned}$$

where I is the $(n \times n)$-identity-matrix and X^+ is the pseudoinverse of X.

We get the reduced (by invariance) linear model

$$y \sim \left(0, \sum_{i=1}^{m} \alpha_i \cdot V_i \right) , \quad \alpha \in I\!R^m_+ \ ,$$

where $V_i = (I - XX^+)U_i(I - XX^+)$, i=1, ..., m.

Let now A be a symmetric $n \times n$-matrix, then a quadratic (invariant) estimator for a linear form $p^T\alpha > 0$, $p = (p_1, \ldots, p_m)^T$, $p \geq 0$, $p \in I\!R^m$, is given by y^TAy, and denoting "tr" the trace, its bias is given by

$$\begin{aligned} \mathrm{E}\, y^TAy - p^T\alpha &= \mathrm{E}\,\mathrm{tr}\, Ayy^T - p^T\alpha \\ &= \mathrm{tr}\, A\mathrm{E}\,(yy^T) - p^T\alpha \\ &= \mathrm{tr}\, A(\sum_{i=1}^{m} \alpha_i V_i) - p^T\alpha \\ &= \sum_{i=1}^{m} \alpha_i (\mathrm{tr}\, AV_i - p_i) \ , \end{aligned}$$

such that y^TAy is an unbiased estimator of $p^T\alpha$ if

$$\mathrm{tr}\, AV_i \quad = \quad p_i, \qquad \text{for all } i = 1, \ldots, m \ ,$$

and a solution A_0 of these equations, respectively the corresponding quadratic estimation function y^TA_0y is the (standard-) minimum norm invariant quadratic unbiased estimator (minque) of $p^T\alpha$, if A_0 has the minimum norm among all solutions.

Denote Sym the Hilbert space of all symmetric $(n \times n)$-matrices with the inner product of two matrices $A, B \in \mathrm{Sym}$ defined by $\mathrm{tr}\, AB$, which then induces the standard norm $\|A\| = \sqrt{\mathrm{tr}\, A^2}$. Furthermore let PSD denote the cone of positive semi definite matrices in Sym.

If the matrices $V_1, \ldots, V_m$ are linearly independent, which for simplicity may be assumed here, then the minque A_0 exists for all $p \in I\!R^m$. Since we only claim $A_0 \in \mathrm{Sym}$, of course we usually get

$$y^TAy \quad < \quad 0 \ \text{ with positive probability} \ ,$$

and only in rare cases $A_0 \in \mathrm{PSD}$.

Resticting in advance A to be in PSD has the consequence that the equations for unbiasedness are seldom fulfilled, so that these conditions had to be weakened, cf. Seely(1971), Rao(1972), Pukelsheim(1981), Lehmann and Casella(1998), and Hartung(1981), where in section 4 there is also a solution algorithm given, which however needs some numerical effort.

Therefore in the following a short-cut method is presented that yields an approximation in PSD to A_0 with a correction for bias.

Let us introduce the linear operator

$$\begin{aligned} g\colon \quad & \mathrm{Sym} \longrightarrow I\!R^m \\ & A \longmapsto gA = \begin{pmatrix} \mathrm{tr}\, AV_1 \\ \vdots \\ \mathrm{tr}\, AV_m \end{pmatrix} \ , \end{aligned}$$

then its adjoint g^* is given by

$$g^* : \begin{array}{l} I\!R^m \longrightarrow \mathrm{Sym} \\ a \longmapsto g^* a = \sum_{i=1}^m \alpha_i V_i, \quad a = (a_1, \ldots a_m)^T \end{array},$$

such that gg^* becomes the Gram-matrix G,

$$gg^* = G \quad = \quad \{\mathrm{tr}\, V_i V_j\}_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}} ,$$

of which the inverse G^{-1} exists because of the assumed linear independence of $V_1, \ldots, V_m$.

Denote g^+ the pseudoinverse operator of g, then the minque A_0 is given by

$$A_0 \quad = \quad g^+ p \ ,$$

which because of $g^+ = g^*(gg^*)^+$ permits the computational representation

$$A_0 \quad = \quad \sum_{i=1}^m a_i \cdot V_i, \text{ with } a = (a_1, \ldots, a_m)^T = G^{-1} p.$$

Let us define now the vectors $b = (b_1, \ldots, b_m)^T$ and $c = (c_1, \ldots, c_m)^T$ by

$$b_i = \begin{cases} a_i, & \text{if } a_i > 0 \\ 0, & \text{if } a_i \leq 0 \end{cases}, \quad \text{and} \quad c_i = \begin{cases} -a_i, & \text{if } a_i < 0 \\ 0, & \text{if } a_i \geq 0 \end{cases},$$

then

$$\begin{array}{rcl} A_0 & = & g^* a \\ & = & g^* b - g^* c \\ & =: & A_1 - A_2 \ , \end{array}$$

where $A_1 \in \mathrm{PSD}$, $A_2 \in \mathrm{PSD}$, and

$$gA_0 = p, \quad gA_1 = Gb =: q, \quad gA_2 = Gc =: r \ ,$$

such that

$$p^T \alpha \quad = \quad q^T \alpha - r^T \alpha \ ,$$

with the estimators

$$\begin{array}{rcl} \widehat{p^T \alpha} & = & \widehat{q^T \alpha} \ - \ \widehat{r^T \alpha} \\ & = & y^T A_1 y - y^T A_2 y. \end{array}$$

We assume $A_2 \neq 0$, otherwise $A_0 \in \mathrm{PSD}$.

Now A_1 is an approximation in PSD to A_0, with the estimate

$$\begin{array}{rcl} y^T A_1 y = \widehat{q^T \alpha} & = & \widehat{p^T \alpha} + \widehat{r^T \alpha} \\ & < & \widehat{r^T \alpha}, \text{ with positive probability} , \end{array}$$

although

$$q^T \alpha \ = \ p^T \alpha + r^T \alpha > r^T \alpha.$$

An additive correction of A_1, however, would lead again to possibly negative estimates. So the idea is now to work with a multiplicative correction term ψ, such that $\psi \cdot A_1$ replaces A_0 in $\widehat{p^T \alpha}$, and $\psi \cdot A_2$ replaces A_2 in $\widehat{r^T \alpha}$.

As determination equation for ψ we thus get

$$E\left\{\psi\cdot\widehat{q^T\alpha}+\psi\cdot\widehat{r^T\alpha}\right\} \stackrel{!}{=} q^T\alpha\,,$$

yielding

$$\psi = \frac{q^T\alpha}{q^T\alpha + r^T\alpha}\,,$$

and so we define our approximate solution in PSD as

$$A_{\text{PSD, appr.}(\psi)} := \psi\cdot A_1\,,$$

where ψ can be estimated by

$$\widehat{\psi} = \frac{y^TA_1y}{y^TA_1y + y^TA_2y}\,,$$

which gives for $p^T\alpha$ the desired, approximate estimate

$$\left(\widehat{p^T\alpha}\right)_{\text{PSD, appr.}(\widehat{\psi})} = \widehat{\psi}\cdot y^TA_1y.$$

We can remark that first simulation results show a good performance of this estimator.

Instead of our easily obtainable decomposition of A_0, we may also use the spectral decomposition of A_0 into $A_{1^*} - A_{2^*}$ as follows:

$$\begin{aligned} A_{1^*} &:= \sum_{\lambda\in\sigma(A_0)} \max\{0,\lambda\}\cdot P_\lambda\,, \\ A_{2^*} &:= A_{1^*} - A_0\,, \end{aligned}$$

where $\sigma(A_0)$ is the spectrum of A_0 and P_λ is the projection onto the eigenspace associated with λ, which needs a higher computational effort.

In more specified models also quite different and more detailed approximations may be derived, cf. e.g. Hartung(1999).

References

Hartung, J. (1981). Nonnegative minimum biased invariant estimation in variance component models. *The Annals of Statistics* **9**, 278–292.

Hartung, J. (1999). Ordungserhaltende positive Varianzschdtzer bei gepaarten Messungen ohne Wiederholungen. *Allgemeines Statistisches Archiv* **83**, erscheint.

Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation.* 2nd. ed., Springer, New York

Pukelsheim, F. (1981). On the existence of unbiased nonnegative estimates of variance covariance components. *The Annals of Statistics* **9**, 293–299.

Rao, C.R. (1972). Estimation of Variance and Covariance Components in Linear Models. *Journal of the American Statistical Association* **67**, 112–115.

Seely, J. (1971). Quadratic subspaces and completeness. *The Annals of Mathematical Statistics* **42**, 710–721.

Integral Equation Methods in Physical Geodesy

Bernhard Heck

1 Introduction

It is well-known since the days of G.G. Stokes (Stokes, 1849) that the main tasks of Geodesy, the determination of the geometry of the Earth's surface and its external gravity field, can be handled by solving geodetic boundary value problems. While Stokes's approach had been based on a reduction of observational data, related to the earth's surface, for gravitational effects induced by the topographical masses, M.S. Molodenskii provided a formulation in terms of an external boundary value problem associated to Laplace's differential equation with the topographical surface of the earth acting as boundary surface (Molodenskii et al., 1962). Further advances in the theory of the geodetic boundary value problem (GBVP) have been made in the past 30 years, especially by the work of H. Moritz, T. Krarup, P. Meissl, E. Grafarend and F. Sansò.

As a result, various formulations of the GBVP are discriminated today, depending on the type of boundary data given on the boundary surface and on the type and number of unknown functions to be solved for. A major criterion for the classification of the numerous types of the GBVP is the question whether the geometry of the boundary surface is known or to be determined from the boundary data itself as part of the GBVP. The concept and notion of "free" boundary value problems, involving a free boundary surface with unknown geometry, has first been introduced in Geodesy by E. Grafarend (Grafarend and Niemeier, 1971; Grafarend, 1972).

Since most of the original formulations of GBVPs are of non-linear type, the first step towards practically applicable solutions consists of a linearization of the primary, non-linear boundary conditions (observation equations) by introducing a reference ("normal") potential and – in the case of free BVPs – a reference surface ("telluroid") approximating the actual gravity potential and surface of the earth, respectively. In general, the linearized boundary conditions imply the derivative of the disturbing potential in a non-normal direction; thus the GBVPs at the level of the linearized problems are classified as fixed, oblique-derivative BVPs.

Further simplifications, e.g. the so-called spherical approximation and the planar approximation (Moritz, 1980, p. 349ff) are generally applied to the linear, oblique-derivative boundary operator in order to reduce the complexity of the GBVP. But still at this level of approximation the resulting BVPs cannot be solved in closed, analytical form due to the irregular boundary surface. Only at the level of the constant radius approximation, by replacing the topographic boundary surface by a sphere, closed solutions in the form of spherical integral formulae can be constructed by applying spherical harmonic expansions. For other geometrically simple substitutes of the boundary surface, e.g. a spheroid or ellipsoid of revolution, first-order solutions of the non-spherical GBVPs can be achieved by the procedure of ellipsoidal corrections (Heck, 1991, 1997; Seitz, 1997).

The more realistic case of an irregular, topographical boundary surface requires either direct discrete approaches such as finite element or finite difference methods (see the pioneering paper by Grafarend, 1975), or the integral equation approach, already applied by M.S. Molodenskii in combination with an analytical perturbation method. In the past decades the integral equation approach has been numerically adapted in the framework of the boundary Element Method (BEM); recent applications to the

GBVP proved the high flexibility and large potential of this promising approach (Klees, 1997; Lehmann, 1997).

The transformation of a BVP into an equivalent integral equation relies on the choice of a representation formula. For a BVP related to Laplace's differential equation admissible representation formulae are (generalized) Green's identities or the potentials of single or double layer mass distributions spread over the boundary surface. Taking advantage of the jump relations the representation formulae provide boundary integral equations which have to be solved for the unknown layer densitites or the potential on the boundary surface, Obviously any choice of representation formula yields a different boundary integral equation for one and the same boundary condition.

In the present paper several representation formulae are applied to the linearized, scalar free GBVP in spherical approximation ("Simple Molodensky Problem"). Section 2 gives a short review of the GBVP under consideration. Based on the representation of the disturbing potential by single and double layer potentials as well as by Brovar's generalized single layer and volume potentials, the transformation of the boundary condition is derived in section 3. For spherical boundary surfaces the solutions of the integral equations can be given in closed analytical form, which is the subject of section 4. Finally, section 5 summarizes some conclusions with respect to applications in Physical Geodesy.

2 The linearized, scalar free GBVP

In the formulation of the scalar free GBVP ("geodetic variant of Molodenskii's problem") it is presupposed that the "horizontal" coordinates of the point $P \in S$ situated on the closed boundary surface S – e.g. the geodetic coordinates with respect to an ellipsoid of revolution, fixed to the earth's rotating body – are known. As a consequence, this type of GBVP contains two unknown functions, identified by the ellipsoidal height H(P) of the boundary points and the gravity potential W(Q), fulfilling the extended Laplace equation

$$\text{Lap } W(Q) = 2T^2 \tag{2.1}$$

at any spatial point Q outside S; T denotes the angular velocity of the earth's rotation. Furthermore, the gravitational part V = W-Z (Z = ½ T^2p^2 centrifugal potential) of the gravity potential is regular at infinity,

$$V = 0\left(r^{-1}\right) \quad , \quad r = \left|\vec{X}(Q)\right|. \tag{2.2}$$

The information for the determination of the unknown functions H(P) and W(Q) has to be extracted from two types of boundary data, presupposed to be given in continuous form over the whole surface S. In the framework of the scalar free GBVP it common to use the observable modulus ɜ of the gravity vector and the geopotential number C with respect to a global fundamental point P_0 as boundary data. Assuming that the standard basic model of Physical Geodesy (Heck, 1997) holds, the relationship between the observables ɜ(P), C(P) at P,S and the unknown functions W, H(P) is provided by the non-linear observation equations

$$\Gamma(P) = \left|\text{grad}W(P)\right| \tag{2.3a}$$

$$C(P) = W\left(P_0\right) - W(P). \tag{2.3b}$$

Linearization of these equations can be achieved by introducing a reference potential w, e.g. a Somigliana-Pizzetti normal gravity field, fulfilling the relationships

$$\begin{aligned} &\text{Lap } w\,(Q) = 2T^2 \\ &w\text{-}Z \qquad = 0(r^{-1}) \quad , \; r = \left|\vec{X}(Q)\right| \quad , \end{aligned} \tag{2.4}$$

if the centrifugal parts in W and w are identical. A reference surface s э p suitable for linearization is constructed via Molodenskii's telluroid mapping (see Grafarend, 1978)

$$\varphi_g(p) = \varphi_g(P) \tag{2.5a}$$

$$\lambda(p) = \lambda(P) \tag{2.5b}$$

$$w(p) - w(p_0) = W(P) - W(P_0), \tag{2.5c}$$

where a one-to-one correspondence between the corresponding pairs of points p, P has been presupposed. The first and second mapping equation (2.5a,b) fix the telluroid point p,s on the ellipsoidal normal running through the surface point P,S; ν_g and 8 are the geodetic latitude and longitude, respectively, related to an ellipsoid of revolution with given size, form and orientation. The third equation (2.5c) provides the ellipsoidal height h(p)=h(ν_g,8) of the telluroid point p, which is numerically identical with the normal height of P.

Differencing the approximate quantities w, h from the original unknowns W, H yields the residual unknown ∗w (disturbing potential) and)h (height anomaly)

$$\delta w(Q) := W(Q) - w(Q) \tag{2.6a}$$

$$\Delta h \quad := H(P) - h(Q) \tag{2.6b}$$

where ∗w is assumed to be regular at infinity and harmonic in the space outside the telluroid s

$$\begin{aligned} &\mathrm{Lap}\, {*w} = 0 \\ &\delta w = 0(r^{-1}) \quad , \quad r = \left|\vec{X}(Q)\right|. \end{aligned} \tag{2.7}$$

After linearization of the boundary conditions (2.3a,b) with respect to the approximate information w, s and reducing for the unknown height anomaly)h the reduced linearized boundary condition

$$a \cdot \delta w + \langle \frac{\vec{\gamma}}{\gamma}, \mathrm{grad}\delta w \rangle = \Delta\gamma + a \cdot \Delta w_o \tag{2.8}$$

$$a = -\frac{\langle \vec{\gamma}, \mathrm{grad}\vec{\gamma} \cdot \vec{n}_e \rangle}{\gamma \cdot \langle \vec{\gamma}, \vec{n}_e \rangle} \tag{2.9}$$

is obtained, where $\vec{\gamma}$=grad w is the normal gravity vector with modulus $\gamma = |\vec{\gamma}|$, $\vec{n}_e$is the unit vector in the direction of the external ellipsoidal normal, $\Delta\gamma := \Gamma(P) - \gamma(p)$ the scalar gravity anomaly and $\Delta w_o := W(P_o) - w(p_o)$ an unknown potential constant. For the derivation of (2.8), (2.9) and the representation of this boundary condition in various curvilinear coordinates see e.g. Heck (1991, 1997).

The directional derivative in (2.8) is related to the direction of the normal gravity vector $\vec{\gamma}(p)$ which deviates from the radial direction by no more than 12 arcmin. globally. By approximating the direction of -$\vec{\gamma}$ by the direction of the radius vector $\vec{x}$ the boundary condition (2.8) simplifies considerably, resulting in the boundary condition of the "simple" Molodenskii problem

$$\left(-\frac{2}{r} \cdot \delta w - \frac{\partial \delta w}{\partial r} \right)_s = \Delta\gamma - \frac{2}{r} \Delta w_o . \qquad (2.10)$$

It should be noted that formally the same boundary condition in linear and spherical approximation is reproduced for the vectorial free GBVP. In the following, the unknown term proportional to $)w_o$ on the right hand side of (2.10) will be neglected, corresponding to a "proper" choice of the numerical value of the gravity potential at P_o.

3 Integral equations for the simple Molodensky problem

The transformation of partial differential equations, in particular Laplace's equation, into equivalent integral equations (considering the respective boundary conditions) can be achieved by applying either direct or indirect methods. The **direct method** is based on Green's identities: E.g. the standard BVPs of classical potential theory can be transformed by the aid of Green's 2nd or 3rd identities (Walter, 1971; Sigl, 1973), while the generalized Green's formula (Giraud, 1934) provides the transition for the oblique derivative BVP (Klees, 1992). A related procedure has been proposed by Molodensky (Molodenskii et al., 1962) and Moritz (Heiskanen and Moritz, 1967, p. 229) for the transformation of the simple Molodensky problem. A specific feature of the direct formulation is the fact that the potential function on the boundary surface can be solved for in a single step.

The **indirect methods** rely on the representation of the (harmonic) solution function by surface layer potentials, e.g. produced by single or double layer surface density functions defined over the boundary surface. Here the transformation into equivalent integral equations makes use of certain jump relations which occur when the computation point approaches the boundary surface in evaluating the surface layer potential or its derivatives. Indirect methods always provide two-step procedures: In the first step the integral equation for the unknown surface layer density, acting as an auxiliary unknown , is solved for; in the second step the representation formula has to be evaluated in order to calculate the potential function or its derivatives on or outside the boundary surface.

Originally, the integral equation method has been used in potential theory in order to prove the existence of solutions of various boundary value problems, this concept being strongly related to Fredholm's alternative (Martensen and Ritter, 1997). In the past two decades the integral equation approach has become the basis for numerical solutions, too, in the framework of the boundary element method (Hackbusch, 1989). Substantial numerical savings can be expected in many practical applications by reducing the dimension of the problem from 3 (dimension of the "spatial" Laplace operator) to 2 (dimension of the boundary surface on which the density function is defined).

Obviously, the transformation of a BVP into an equivalent integral equation is not unique, since any representation formula produces another type of integral equation for one and the same BVP. Since the analytical and numerical behaviour of these integral equations may be quite different, it is necessary to select, for a given BVP, those representations which possess optimal properties in this respect. In the following, several representation formulae related to the indirect approach will be applied to the simple Molodensky problem; for the special case of a spherical boundary surface the solution of the respective integral equation can be explicitly described.

3.1 Representation by a single layer potential

Since the potential of a single layer mass distribution on a closed surface, e.g. on the telluroid s, is harmonic in the external space and regular at infinity, a single layer potential can be used for representing the disturbing potential $*w$

$$\delta w(\vec{X}) = \frac{1}{4\pi} \int_s \frac{\mu(\vec{y})}{|\vec{X} - \vec{y}|} \cdot ds(\vec{y}) \tag{3.1}$$

$\vec{X}$ denotes the position vector of the point of evaluation in space, $\vec{y}$ of the variable point of integration on the boundary surface s where the density function takes the value $\mu(\vec{y})$. The single layer potential (3.1) is continuous throughout $|^3$, but in general not continuously differentiable with respect to each side of s. Considering the limiting relations of the normal derivative when the point $\vec{X}$ in space tends to the surface point $\vec{x}$, situated on the same surface normal (Martensen and Ritter, 1997), the gradient of the disturbing potential at the positive side of the surface s is given by the expression

$$\begin{aligned} (\text{grad}\,\delta w(\vec{x}))_+ &= -\frac{1}{2}\mu(\vec{x}) \cdot \vec{n}_\times - \\ &- \frac{1}{4\pi} \cdot \text{p.v.} \int_s \frac{\vec{x} - \vec{y}}{|\vec{x} - \vec{y}|^3} \cdot \mu(\vec{y}) \cdot ds(\vec{y}) \ , \end{aligned} \tag{3.2}$$

p.v. denoting Cauchy's principal value. Concerning the function spaces it should be presupposed that the surface is Hölder-continuously differentiable, $s \in C^{1+\alpha}, \alpha > 0$, and $\mu \in L^2(s)$ is quadratically integrable on the surface s.

Inserting the representation formula (3.1) and its gradient (3.2) into the reduced boundary condition (2.10) of the "simple" Molodenskii problem produces the following integral equation of second kind for the unknown auxiliary density function μ:

$$\frac{1}{2}\mu(\vec{x}) \cdot \cos(\vec{n}_x, \vec{x}) + \frac{1}{4\pi} \text{p.v.} \int_s \frac{|\vec{x}|^2 - |\vec{y}|^2 - 3|\vec{x} - \vec{y}|^2}{2|\vec{x}| \cdot |\vec{x} - y|^3} \cdot \mu(\vec{y}) . ds(\vec{y}) = \Delta\gamma(\vec{x}) \tag{3.3}$$

where $(\vec{n}_x, \vec{x})$ is the angle between the external surface normal and the position vector, which is roughly identical with the inclination angle β of the terrain. The integral equation (3.3) involves a pseudo-differential operator of order r=0 and contains a strongly singular integral kernel

$$k(\vec{x}, \vec{y} - \vec{x}) := \frac{|\vec{x}|^2 - |\vec{y}|^2 - 3|\vec{x} - \vec{y}|^2}{2|\vec{x}| \cdot |\vec{x} - \vec{y}|^3} \tag{3.4}$$

in conventional notation $(|\vec{x}| = r, |\vec{y}| = r', |\vec{x} - \vec{y}| = l)$

$$k(r, r', l) = \frac{r^2 - r'^2}{2r \cdot l^3} - \frac{3}{2r \cdot l} \tag{3.5}$$

The integral equation (3.3) was the starting point in M.S. Molodenskii's series expansion for the analytical solution of the GBVP (Moritz, 1980, p. 354 ff).

3.2 Representation by a double layer potential

Since the potential of a surface dipole distribution on the closed surface s is harmonicc outside the surface and regular at infinity, the double layer potential involving the density function ν can be used for representing the disturbing potential δw

$$\delta w(\vec{X}) = \frac{1}{4\pi}\int_s \frac{\partial}{\partial n_y}\frac{1}{|\vec{X}-\vec{y}|}\cdot\nu(\vec{y})\cdot ds(\vec{y}). \tag{3.6}$$

It is well-known (Martensen and Ritter, 1997) that the double layer potential is discontinuous when the point $\vec{X}$ in space tends to the surface point $\vec{x}$, fulfilling the limiting relations for the potential and its gradient

$$(\delta w(\vec{x}))_+ = \frac{1}{2}\nu(\vec{x}) + \frac{1}{4\pi}\int_s < \vec{n}_y, \frac{\vec{x}-\vec{y}}{|\vec{x}-\vec{y}|^3} > \nu(\vec{y})\cdot ds(\vec{y}) \tag{3.7}$$

$$\begin{aligned}(\mathrm{grad}\,\delta w(\vec{x}))_+ &= \frac{1}{2}\mathrm{Grad}\,\nu(\vec{x}) + \\ &+ \frac{1}{4\pi}\,\mathrm{p.f.}\int_s \left[\frac{\vec{n}_y}{|\vec{x}-\vec{y}|^3} - 3\frac{<\vec{n}_y,\vec{x}-\vec{y}>\cdot(\vec{x}-\vec{y})}{|\vec{x}-\vec{y}|^5}\right]\cdot\nu(\vec{y})\cdot ds(\vec{y})\end{aligned} \tag{3.8}$$

The integral (3.7) exists as an improper integral if it is assumed that s is piecewise Hölder-continuously differentiable, $s0C^{1+\alpha}$, $\alpha>0$ and ν is continuous, $\nu\in C^o$. In contrast, the integral in (3.8) has to be understood in the sense of Hadamard's part fini integral (Hackbusch, 1989, p. 284), presupposing $s\in C^{1+\alpha}$, $\nu\in C^{1+\alpha}(s)$, $\alpha>0$. Grad $\nu(\vec{x})$ denotes the surface gradient of the density function ν at $\vec{x}$.

Inserting (3.7) and (3.8) into the reduced boundary condition (2.10) of the simple Molodenskii poblem yields the following integro-differential equation for the unknown auxiliary density function ν:

$$\begin{aligned}&-\frac{1}{2}<\mathrm{Grad}\,\nu(\vec{x}),\frac{\vec{x}}{|\vec{x}|}> - \frac{\nu(\vec{x})}{|\vec{x}|} + \\ &+\frac{1}{4\pi}\,\mathrm{p.f.}\int_s\Big[-<\vec{n}_y,\vec{y}>\cdot\left[3\left(|\vec{x}|^2-|\vec{y}|^2\right)-|\vec{x}-\vec{y}|^2\right]+ \\ &+3<\vec{n}_y,\vec{x}>\cdot\left[|\vec{x}|^2-|\vec{y}|^2-|\vec{x}-\vec{y}|^2\right]\Big]\cdot\frac{\nu(\vec{y})\cdot ds(\vec{y})}{2|\vec{x}|\cdot|\vec{x}-\vec{y}|^5} = \Delta\gamma(\vec{x})\end{aligned} \tag{3.9}$$

This integro-differential equation involves a pseudo-differential operator of order r = 1 and contains a hypersingular integral kernel

$$\begin{aligned}k(\vec{x},\vec{y}-\vec{x}) := &\Big[-<\vec{n}_y,\vec{y}>\cdot\left[3\left(|\vec{x}|^2-|\vec{y}|^2\right)-|\vec{x}-\vec{y}|^2\right]+ \\ &+3<\vec{n}_y,\vec{x}>\cdot\left[|\vec{x}|^2-|\vec{y}|^2-|\vec{x}-\vec{y}|^2\right]\Big]\cdot\frac{1}{2|\vec{x}|\cdot|\vec{x}-\vec{y}|^5},\end{aligned} \tag{3.10}$$

in conventional notation

$$k(r,r',l)=\frac{3(r^2-r'^2)(r\cos\varepsilon-r'\cos\beta')}{2r\cdot l^5}+ \\ +\frac{r'\cos\beta'-3r\cos\varepsilon}{2r\cdot l^3} \tag{3.11}$$

where $\beta'=\sphericalangle(\quad)$ $(\vec{n}_y,\vec{x})$
the angle between the surface normal at $\vec{y}$ and the radius vector of the evaluation point $\vec{x}$.

3.3 Representation by Brovar's generalized single layer potential

Attempting to obtain simpler expressions for the solution of Molodenskii's problem, Brovar (1963, 1964) introduced two alternative representations of harmonic functions, regular at infinity, by generalized surface layer potentials. The first representation formula generalizes the single layer potential, extending the inverse distance kernel to the Stokes-Pizzetti kernel:

$$\delta w(\vec{X})=\frac{1}{4\pi}\int_s E_1(\vec{X},\vec{y})\cdot\lambda(\vec{y})\cdot ds(\vec{y}) \tag{3.12}$$

$$E_1(\vec{X},\vec{y}):=\frac{2}{|\vec{X}-\vec{y}|}-3\frac{|\vec{X}-\vec{y}|}{|\vec{X}|^2}-\frac{5}{|\vec{X}|^3}<\vec{X},\vec{y}>- \\ -\frac{3}{|\vec{X}|^2}\cdot<\vec{X},\vec{y}>\cdot\ln\frac{|\vec{X}|^2-<\vec{X},\vec{y}>+|\vec{X}|\cdot|\vec{X}-\vec{y}|}{2|\vec{X}|^2} \tag{3.13}$$

Despite of the extension of the kernel by a logarithmically singular term the integral (3.12) still exists as an improper integral if $s\in C^1$ (piecewise) and $\lambda\in L^2(s)$. Like in section 3.1 the generalized single layer potential (3.12) is continuous throughout ^{3}; its gradient is discontinuous, fulfilling the limiting relation

$$(\mathrm{grad}\,\delta w(\vec{x}))_+=-\lambda(\vec{x})\cdot\vec{n}_x+ \\ +\frac{1}{4\pi}\,\mathrm{p.v.}\int_s \mathrm{grad}_x E_1(\vec{x},\vec{y})\cdot\lambda(\vec{y})\cdot ds(\vec{y}) \tag{3.14}$$

where the integral is understood in the sense of Cauchy's principal value and $s\in C^{1+\alpha}$, $\lambda\in C^{\alpha}(s)$, $\alpha>0$.

Inserting the representation formula (3.12) and its gradient (3.14) into the reduced boundary condition (2.10) of the "simple" Molodenskii problem yields the following integral equation of second kind for the unknown auxiliary density function λ:

$$\lambda(\vec{x})\cdot\cos(\vec{n}_x,x)+\frac{1}{4\pi}\,\mathrm{p.v.}\int_s\left(\frac{|\vec{x}|^2-|\vec{y}|^2}{|\vec{x}|\cdot|\vec{x}-\vec{y}|^3}-3\frac{<\vec{x},\vec{y}>}{|\vec{x}|^4}\right)\cdot\lambda(\vec{y})\cdot ds(\vec{y})=\Delta\gamma(\vec{x}). \tag{3.15}$$

This integral equation again involves a pseudo-differential operator of order r=0 and contains a strongly singular integral kernel

$$k(\vec{x},\vec{y}-\vec{x}) := \frac{|\vec{x}|^2 - |\vec{y}|^2}{|\vec{x}|\cdot|\vec{x}-\vec{y}|^3} - 3\frac{<\vec{x},\vec{y}>}{|\vec{x}|^4} \quad , \tag{3.16}$$

in conventional notation

$$k(r,r',l) = \frac{r^2 - r'^2}{r\cdot l^3} - \frac{3r'\cos\psi}{r^3} \tag{3.17}$$

where $\psi = \angle(\vec{x},\vec{y})$ denotes the angle between the position vectors $\vec{x}$ (fixed point of evaluation) and $\vec{y}$ (variable point of integration). By comparing (3.17) with (3.5) it becomes obvious that the weakly singular term proportional to l^{-1} has disappeared; the additional term in (3.17) is essentially a spherical harmonic term of first degree.

3.4 Representation by Brovar's generalized "volume" potential

A second alternative surface layer representation of the disturbing potential, given by Brovar (1963, 1964) contains a kernel with an even weaker degree of singularity:

$$\delta w(\vec{X}) = \frac{1}{4\pi}\int_s E_2(\vec{X},\vec{y})\cdot\chi(\vec{y})\cdot ds(\vec{y}) \tag{3.18}$$

$$E_2(\vec{X},\vec{y}) := \frac{|\vec{X}-\vec{y}|}{|\vec{X}|^2} - \frac{<\vec{X},\vec{y}>}{|\vec{X}|^3}\left(1+\ln\frac{|\vec{X}|^2 - <\vec{X},\vec{y}> + |\vec{X}|\cdot|\vec{X}-\vec{y}|}{2|\vec{X}|^2}\right) \tag{3.19}$$

The spatial function (3.18) is harmonic in $\mathbb{R}^3$ besides on s, and regular at infinity. Due to the logarithmic (weak) singularity of $E_2(\vec{X},\vec{Y})$ the surface layer potential (3.18) is continuously differentiable in $\mathbb{R}^3$; since this property holds generally for volume potentials, the notion "generalized volume potential" has been chosen by Brovar. The gradient of this potential representation at the point $\vec{x}$ on the surface is given by the improper integral

$$\begin{aligned} \operatorname{grad}\delta w(\vec{x}) = \frac{1}{4\pi}\int_s \Bigg[& \frac{2|\vec{x}-\vec{y}|}{|\vec{x}|^3} - \frac{1}{|\vec{x}|\cdot|\vec{x}-\vec{y}|} + 3\frac{<\vec{x},\vec{y}>}{|\vec{x}|^4} + \\ & + \frac{2<\vec{x},\vec{y}>}{|\vec{x}|^4}\cdot\ln\frac{|\vec{x}|^2 - <\vec{x},\vec{y}> + |\vec{x}|\cdot|\vec{x},\vec{y}|}{2|\vec{x}|^2}\Bigg]\cdot\chi(\vec{y})\cdot ds(\vec{y}) \end{aligned} \tag{3.20}$$

Due to the continuity of $\operatorname{grad}\delta w(\vec{X})$, $\vec{X}\in\mathbb{R}^3$ there is no residual term outside the integral (3.20). For this reason an integral equation of first kind for the unknown density function χ is produced when the representation formula (3.18) and its gradient (3.20) are inserted into the reduced boundary condition (2.10) of the "simple" Molodenskii problem:

$$\frac{1}{4\pi}\int_s\left(\frac{1}{|\vec{x}|\cdot|\vec{x}-\vec{y}|}-\frac{<\vec{x},\vec{y}>}{|\vec{x}|^4}\right)\cdot\chi(\vec{y})\cdot ds(\vec{y})=\Delta\gamma(\vec{x}) \tag{3.21}$$

This integral equation involves a pseudo-differential operator of order r = -1 and contains a weakly singular integral kernel

$$k(\vec{x},\vec{y}-\vec{x}):=\frac{1}{|\vec{x}|\cdot|\vec{x}-\vec{y}|}-\frac{<\vec{x},\vec{y}>}{|\vec{x}|^4} \tag{3.22}$$

in conventional notation

$$k(r,r',l)=\frac{1}{r\cdot l}-\frac{3r'\cdot\cos\psi}{r^3} \tag{3.23}$$

Again the second term in (3.23) is essentially a spherical harmonic of first degree.

4 The special case of a spherical boundary surface

It is well-known that the formulae of Physical Geodesy become rather simple as soon as the boundary surface is a sphere. If any relationship in spherical approximation is applied, the respective problem in addition becomes a "normal" problem in the sense of potential theory, since the radial derivative is automatically a normal derivative on the spherical surface. Spherical BVPs play a dominant role in Physical Geodesy since on a global scale the earth can be approximated rather well by a sphere, the approximation error having the order of 0.3%. For this reason reduction methods, aiming at the creation of a "spherical" situation, have become very familiar; instead of calculating those reductions from prior information more rigorous approaches can be constructed on the basis of iterative schemes. These reduction procedures form the background of e.g. the so-called "ellipsoidal corrections" (Heck, 1997; Seitz, 1997).

In the following the integral equations derived in section 3 will be specified for a sphere of radius R acting as boundary surface s with surface element $ds=R^2Ad\Phi$. It is shown that the solutions of the integral equations for the various representations can easily be expressed in the frequency domain; in space domain the respective relationships are represented by spherical integrals.

4.1 Representation by a single layer potential

From equation (3.1) follows the representation of the disturbing potential by the potential of a single layer spread over the sphere with radius R

$$\delta w(\vec{X})=\frac{R^2}{4\pi}\int_\sigma\frac{1}{l}\cdot\mu(\vec{y})\cdot d\sigma(\vec{y}) \quad . \tag{4.1}$$

The Euclidean distance l between the points $\vec{X}$ in space and $\vec{y}$ on the sphere can be expressed by the angle P between the position vectors $\vec{X}$ and $\vec{y}$

$$l=\sqrt{r^2+R^2-2rR\cos\psi} \quad ; \tag{4.2}$$

for a computation point on the sphere $\left(\vec{X} \to \vec{x}, r = R\right)$ this relationship is simply

$$l_o = 2 \cdot R \cdot \sin \frac{\psi}{2} \quad , \tag{4.3}$$

hence

$$\delta w(\vec{x}) = \frac{R}{4\pi} \int_\sigma \frac{\mu(\vec{y})}{2 \cdot \sin \frac{\psi}{2}} \cdot d\sigma(\vec{y}) \quad . \tag{4.4}$$

In a similiar way the integral equation (3.3) reduces to

$$\frac{1}{2}\mu(\vec{x}) + \frac{1}{4\pi} \int_\sigma \left(\frac{-3}{4 \sin \frac{\psi}{2}} \right) \cdot \mu(\vec{y}) \cdot d\sigma(\vec{y}) = \Delta\gamma(\vec{x}) \quad . \tag{4.5}$$

Obviously the strongly singular integral kernel in (3.5) has now been transformed into a weakly singular kernel; conversely expressed this means that the strongly singular kernel in (3.5) is produced by the topography and ellipticity of the boundary surface.

Expanding the disturbing potential outside the boundary sphere into solid spherical harmonics

$$\delta w(\vec{X}) = \sum_{n=0}^{\infty} \left(\frac{R}{r} \right)^{n+1} \cdot \delta w_n(\vec{x}) \quad , \quad \vec{X} = \frac{r}{R} \cdot \vec{x} \tag{4.6}$$

and the functions $\Delta\gamma(\vec{x})$ and $\mu(\vec{x})$ in surface spherical harmonics

$$\Delta\gamma(\vec{x}) = \sum_{n=0}^{\infty} \Delta\gamma_n(\vec{x}) \quad , \quad \mu(\vec{x}) = \sum_{n=0}^{\infty} \mu_n(\vec{x}) \quad , \tag{4.7}$$

and inserting these series in (4.4) and (4.5) yields the following frequency-domain relations

$$\mu_n(\vec{x}) = \frac{2n+1}{n-1} \cdot \Delta\gamma_n(\vec{x}) \quad , \quad n \neq 1 \tag{4.8}$$

$$\delta w_n(\vec{x}) = \frac{R}{2n+1} \cdot \mu_n(\vec{x}) \quad . \tag{4.9}$$

These spectral relationships show that the single layer density μ as a function on the spherical boundary is about as rough as the gravity anomaly data; on the other hand the disturbing potential δw on the sphere is smoother than the density function since the high degree, short wavelength constituents are damped by the factor 1/(2n+1). Combining formulae (4.8) and (4.9) results in the well-known spectral Stokes formula (Heiskanen and Moritz, 1967)

$$\delta w_n(\vec{x}) = \frac{R}{n-1} \cdot \Delta\gamma_n(\vec{x}) \quad , \quad n \neq 1 \tag{4.10}$$

It should be noted that the first degree (n=1) terms are forbidden in (4.8) and (4.10), expressing the fact that $\mu_1(\vec{x})$ and $\delta w_1(\vec{x})$ cannot be determined from gravity anomaly data. On the other hand it must be postulated that the boundary data $\Delta\gamma$ fulfill the consistency condition

$$\Delta\gamma_1(\vec{x}) := \frac{3}{4\pi}\int_\sigma \Delta\gamma(\vec{y})\cdot\cos\psi\cdot d\sigma(\vec{y}) = 0 \quad . \tag{4.11}$$

Using the spherical harmonic expansion of the function 1/l the spectral relationship (4.8) can easily be transformed into the space domain, resulting in the spherical integral

$$\mu(\vec{x}) = 2\Delta\gamma(\vec{x}) + \frac{3}{4\pi}\int_\sigma \Delta\gamma(\vec{y})\cdot(S(\psi)-1)\cdot d\sigma(\vec{y}) + \mu_1(\vec{x}) \tag{4.12}$$

where S(ψ) denotes Stokes's function. By the combination of formulae (4.12) and (4.1), respectively (4.4) the solution of the simple Molodenskii problem in constant radius approximation is provided in two steps $(\Delta\gamma \to \mu \to \delta w)$, while a one-step procedure is based on a direct application of Stokes's integral formula equivalent to (4.10)

$$\delta w(\vec{x}) = \frac{R}{4\pi}\int_\sigma \Delta\gamma(\vec{y})\cdot(S(\psi)-1)\cdot d\sigma(\vec{y}) + \delta w_1(\vec{x}) \quad . \tag{4.13}$$

4.2 Representation by a double layer potential

Considering the fact that the normal derivative

$$\frac{\partial}{\partial n_y}\frac{1}{\left|\vec{X}-\vec{y}\right|} = \lim_{r'\to R}\frac{\partial}{\partial r'}\left(r^2 + r'^2 - 2rr'\cos\psi\right)^{-1/2}$$

$$= -\frac{1}{2Rl} + \frac{r^2 - R^2}{2Rl^3} \tag{4.14}$$

contains a part which acts as a spherical Dirac pulse for r→R, the representation formula (3.6) can be specified for a computation point situated on the spherical boundary

$$\delta w(\vec{x}) = \frac{1}{2}\nu(\vec{x}) - \frac{1}{4\pi}\int_\sigma \frac{\nu(\vec{y})}{4\cdot\sin\frac{\psi}{2}}\cdot d\sigma(\vec{y}) \quad . \tag{4.15}$$

In a similar way the integral equation (3.9) reduces to

$$-\frac{\nu(\vec{x})}{R} + \frac{1}{4\pi}\int_\sigma \frac{3\nu(\vec{y})}{8R\cdot\sin\frac{\psi}{2}}\cdot d\sigma(\vec{y}) - \frac{1}{4\pi}\,\text{p.f.}\int_\sigma \frac{\nu(\vec{y})}{8R\cdot\sin^3\frac{\psi}{2}} d\sigma(\vec{y}) = \Delta\gamma(\vec{x}) \quad . \tag{4.16}$$

The hypersingular part fini integral can be regularized by shifting the constant value $\nu(\vec{x})$ under the integral. This procedure results in the integral equation for the unknown double layer density ν:

$$\frac{3}{4\pi}\int_{\sigma}\frac{\nu(\vec{y})-\nu(\vec{x})}{8R\cdot\sin\frac{\psi}{2}}\cdot d\sigma(\vec{y})-\frac{1}{4\pi}\,p.v.\int_{\sigma}\frac{\nu(\vec{y})-\nu(\vec{x})}{8R\cdot\sin^3\frac{\psi}{2}}\cdot d\sigma(\vec{y})=\Delta\gamma(\vec{x}) \quad . \tag{4.17}$$

Obviously the part fini hypersingular integral degenerates into a simple Cauchy principal value integral containing a strongly singular kernel. Furthermore it can be recognized that in the spherical case the differential part of the integro-differential equation (3.9) disappears.

By the aid of the expansions (4.6) and (4.7) the following spectral domain relationships are obtained

$$\nu_n(\vec{x})=R\cdot\frac{2n+1}{n(n-1)}\cdot\Delta\gamma_n(\vec{x}) \quad , \quad n\neq 1 \tag{4.18}$$

$$\delta w_n(\vec{x})=\frac{n}{2n+1}\cdot\nu_n(\vec{x})=\frac{1}{2}\left(1-\frac{1}{2n+1}\right)\cdot\nu_n(\vec{x}) \quad , \tag{4.19}$$

proving that the double layer density ν as a function on the spherical boundary is smootehr than the gravity anomaly data; on the other hand the density function ν has the same degree of smoothness as the disturbing potential δw on the sphere. Again, the first degree terms $\nu_1(\vec{x})$ and $\delta w_1(\vec{x})$ cannot be determined from the gravity anomaly data, and the consistency condition (4.11) must be fulfilled. By combining equations (4.18) and (4.19) again the spectral Stokes's formula (4.10) is reproduced.

4.3 Representation by Brovar's generalized single layer potential

Brovar's first representation formula (3.12) can be easily specified for a computation point situated on the spherical boundary

$$\delta w(\vec{x})=\frac{R}{4\pi}\int_{\sigma}\lambda(\vec{y})\cdot(S(\psi)-1)\cdot d\sigma(\vec{y})+\delta w_1(\vec{x}) \tag{4.20}$$

where the kernel function is now Stokes's function; the strongly singular integral kernel $E_1(\vec{x},\vec{y})$ (3.13) has degenerated into a weakly singular one. The last term in (4.20) reflects the fact that the first degree terms of $\delta w(\vec{x})$ are indefinite.

In a similar way the integral equation (3.15) reduces to

$$\lambda(\vec{x})-\frac{3}{4\pi}\int_{\sigma}\cos\psi\cdot\lambda(\vec{y})\cdot d\sigma(\vec{y})=\Delta\gamma(\vec{x}) \quad . \tag{4.21}$$

The second term on the right hand side corresponds to the first degree harmonic term $\lambda_1(\vec{x})$ in the expansion of $\lambda(\vec{x})$. On the other hand, due to (4.11) the first-degree term in $\Delta\gamma$ is forced to be zero, thus it follows that $\lambda_1(\vec{x})=0$, too. As a consequence, equation (4.21) reduces to the "integral" equation

$$\lambda(\vec{x})=\Delta\gamma(\vec{x}) \quad , \tag{4.22}$$

i.e. the density function λ is identical with the boundary data $\Delta\gamma$.

A transformation of (4.20) and (4.22) into the spectral domain yields

$$\lambda_n(\vec{x}) = \Delta\gamma_n(\vec{x}) \tag{4.23}$$

$$\delta w_n(\vec{x}) = \frac{R}{n-1} \cdot \lambda_n(\vec{x}) \quad ,n \neq 1 \tag{4.24}$$

the combination of both resulting again in Stokes's formulae (4.10) and (4.13) in the spectral and in the space domain, respectively.

4.4 Representation by Brovar's generalized "volume" potential

Brovar's second representation formula (3.18) can be specified for a computation point situated on the spherical boundary

$$\begin{aligned}\delta w(\vec{x}) = \frac{R}{4\pi}\int_\sigma \left[-2\sin\frac{\psi}{2} - \cos\psi \cdot \left(1 + \ln\left(\sin\frac{\psi}{2} + \sin^2\frac{\psi}{2}\right)\right)\right] \cdot \\ \cdot \chi(\vec{y}) \cdot d\sigma(\vec{y}) + \delta w_1(\vec{x}) \quad .\end{aligned} \tag{4.25}$$

Again the first degree term $\delta w_1(\vec{x})$ is indefinite since the first degree term

$$\int_\sigma \cos\psi \cdot \chi(\vec{y}) \cdot d\sigma(\vec{y})$$

is subtracted on the right hand side of (4.25).

In a similar way the integral equation (3.21) reduces to

$$\frac{1}{4\pi}\int_\sigma \left(\frac{1}{2\sin\frac{\psi}{2}} - \cos\psi\right) \cdot \chi(\vec{y}) \cdot d\sigma(\vec{y}) = \Delta\gamma(\vec{x}) \quad . \tag{4.26}$$

Due to the fact that the boundary data $\Delta\gamma$ have to fulfill the consistency condition (4.11) the first degree term in the auxiliary density function χ vanishes too, i.e. $\chi_1(\vec{x}) = 0$, as the analysis of (4.26) proves. Consequently (4.26) reduces to the simple integral equation of first kind

$$\frac{1}{4\pi}\int_\sigma \frac{\chi(\vec{y})}{2\sin\frac{\psi}{2}} \cdot d\sigma(\vec{y}) = \Delta\gamma(\vec{x}) \tag{4.27}$$

A transformation of (4.25) and (4.27) into the spectral domain yields

$$\chi_n(\vec{x}) = (2n+1) \cdot \Delta\gamma_n(\vec{x}) \tag{4.28}$$

$$\delta w_n(\vec{x}) = \frac{R}{(2n+1)(n-1)} \cdot \chi_n(\vec{x}) \quad , \quad n \neq 1 \tag{4.29}$$

It can be recognized from (4.28) that the density function χ as a function on the boundary is rougher than the boundary data $\Delta\gamma$ since the short wavelength components in $\Delta\gamma$ are amplified by the factor (2n+1). This behaviour could be expected from (3.21), because the inverse of the operator $K:\chi\rightarrow\Delta\gamma$, being a pseudo-differential operator of order r = -1, naturally has a de-smoothing property and is unstable. On the other hand, the operator $I:\chi\rightarrow\delta w$ is strongly smoothing. As a consequence, a two-step approach for the solution of the GBVP, which is based on Brovar's second representation formula, will be senseless for numerical reasons, since the procedure used in the first step will not be stable. This behaviour is also visible when (4.28) is transformed into the space domain

$$\chi(\vec{x}) = \frac{1}{4\pi}\,\text{p.v.}\int_{\sigma}\left(\Delta\gamma(\vec{y}) - \Delta\gamma(\vec{x})\right)\left(\frac{1}{2\sin^3\frac{\psi}{2}} - 9\cos\psi\right)d\sigma(\vec{y}) + \chi_1(\vec{x}) \tag{4.30}$$

where the hypersingular integral has been regularized, leaving an integral in the sense of Cauchy's principal value.

5 Closing remarks

The preceding derivations have shown that there exist numerous alternative and competitive representations of the disturbing potential, providing as many integral equations for the solution of one and the same formulation of the GBVP. The two-step approach described above arrives at the solution after having solved the integral equation for the auxiliary density function which is inserted into the representation formula. For an arbitrary density function κ, $\kappa \in \{\mu, \nu, \lambda, \chi\}$ this is indicated by the sequence of mappings

$$\Delta\gamma \rightarrow \kappa \rightarrow \delta w \ .$$

In numerical solutions of the GBVP via the integral equation method (BEM) the properties of the respective operators play a dominant role (Klees, 1992, 1997; Lehmann, 1997). For numerical reasons it it advantageous to apply only non-desmoothing operators in this process. The variants described in sections 3.1, 3.2, 3.3 and 4.1, 4.2, 4.3 respectively are characterized by a sequence of two transformations, one of which retaining the same degree of roughness and the other one being of smooting type. An exception is provided in sections 3.4 and 4.4 where by the use of Brovar's second alternative of representation a desmoothing mapping $\Delta\gamma\rightarrow\chi$ has been applied which has to be counterbalanced in the second step $\chi\rightarrow\delta w$ by a much stronger smoothing. Since the degree of smooting of the composed mapping $\Delta\gamma\rightarrow\delta w$ is fixed, a smoothing gain in one step will be lost in the other step of the indirect BEM approach. For the same reason the use of surface layer representations involving higher order derivatives of the inverse distance

$$\frac{\partial^k}{\partial n^k}\left(\frac{1}{l}\right) \quad , \quad k \geq 2$$

cannot be recommended, in general.

Finally it should be noted that the integral equation method is applicable to the linearized GBVP in the strict sense, too, without presupposing spherical and planar approximations. The integral equation method in its modern numerical version, the Boundary Element Method, is capable of taking care of very irregular boundary surfaces, making it a most excellent and efficient tool for solving the GBVP. The considerable numerical expenditure can be managed today by the use of modern supercomputers (vector and parallel computers), as the results by Klees (1992, 1997) and Lehmann (1997) have confirmed.

Literature:

Brovar, V.V. (1963): Solutions of the Molodenskij Boundary Problem. Geodesy and Aerophotography (1963), 237-240.

Brovar, V.V. (1964): Fundamental Harmonic Functions With a Singularity on a Segment and Solution of Outer Boundary Problems. Geodesy and Aerophotography (1964), 150-155.

Giraud, G. (1934): Equations integrales principales. Annales Scientifiques de L'École Normale Supérieure, Troisième Série, 51 (1934), 251-372.

Grafarend, E. (1972); Die freie geodätische Randwertaufgabe und das Problem der Integrationsfläche innerhalb der Integralgleichungsmethode. In: E. Grafarend und N. Weck, Tagung Freie Randwertaufgaben. Mitt. Inst. f. Theoretische Geodäsie der Univ. Bonn, Nr. 4, 60-85.

Grafarend, E. (1975): The Geodetic Boundary Value Problem. In: B. Brosowski, E. Martensen (Hrsg.), Methoden und Verfahren der mathematischen Physik, Bd. 13, Part II. Bibliogr. Institut Mannheim/Wien/Zürich, 1-25.

Grafarend, E. (1978): The Definition of the Telluroid. Bull. Géod., 52, 25-37.

Grafarend, E.; Niemeier, W. (1971): The Free Nonlinear Boundary Value Problem of Physical Geodesy. Bull. Géod. No. 101, 1er Sept. 1971, 243-262.

Hackbusch, W. (1989): Integralgleichungen. Teubner-Verlag, Stuttgart 1989

Heck, B. (1991): On the Linearized Boundary Value Problems of Physical Geodesy. Rep. No. 407, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus/Ohio.

Heck, B. (1997): Formulation and Linearization of Boundary Value Problems: From Observables to a Mathematical Model. In: F. Sansò, R. Rummel (eds.): Geodetic Boundary Value Problems in View of the One Centimeter Geoid. Springer Lecture Notes in Earth Sciences, No. 65, 121-160.

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

Klees, R. (1992): Lösung des fixen geodätischen Randwertproblems mit Hilfe der Randelementmethode. Deutsche Geodätische Kommission, Reihe C, Heft Nr. 382, München 1992.

Klees, R. (1997): Topics on Boundary Element Methods. In: F. Sansò, R. Rummel (eds.): Geodetic Boundary Value Problems in View of the One Centimeter Geoid. Springer Lecture Notes in Earth Sciences, no. 65, 482-531.

Lehmann, R. (1997): Studies on the Use of the Boundary Element Method in Physical Geodesy. Deutsche Geodätische Kommission, Reihe A, Heft Nr. 113, München 1997.

Martensen, E.; Ritter, S. (1997): Potential Theory. In: Sansò, F.; Rummel R., (Eds.), Geodetic Boundary Value Problems in View of the One Centimeter Geoid. Lecture Notes in Earth Sciences, No. 65, Springer 1967, 19-66.

Molodenskii, M.S.; Eremeer, V.F.; Yurkina, M.I. (1962): Methods for Study of the External Gravitational Field and Figure of the Earth. Transl. from Russian (1960), Jerusalem, Israel Program for Scientific Translation.

Seitz, K. (1997): Ellipsoidische und topographische Effekte im geodätischen Randwertproblem. Deutsche Geodätische Kommission, Reihe C, Heft Nr. 483, München 1997.

Sigl, R. (1973): Einführung in die Potentialtheorie. H. Wichmann Verlag, Karlsruhe 1973.

Stokes, G.G. (1849): On the Variation of Gravity at the Surface of the Earth. Transactions of the Cambridge Philosophical Society, Vol. VIII, part V, 672-694.

Walter, W. (1971): Einführung in die Potentialtheorie. BI Hochschulskripten Bd. 765a, Bibliograph. Institut, Mannheim/Wien/Zürich, 1971 Deutsche Geodätische Kommission, Rh. C., Heft Nr. 382, München 1992.

Classical Electrodynamics: A Tutorial on its Foundations [1]

Friedrich W. Hehl, Yuri N. Obukhov and Guillermo F. Rubilar

Abstract:
We will display the fundamental structure of classical electrodynamics. Starting from the axioms of (1) electric charge conservation, (2) the existence of a Lorentz force density, and (3) magnetic flux conservation, we will derive Maxwell's equations. They are expressed in terms of the field strengths $(E, \mathcal{B})$, the excitations $(\mathcal{D}, H)$, and the sources (ρ, j). This fundamental set of four microphysical equations has to be supplemented by somewhat less general constitutive assumptions in order to make it a fully determined system with a well-posed initial value problem. It is only at this stage that a distance concept (metric) is required for space-time. We will discuss one set of possible constitutive assumptions, namely $\mathcal{D} \sim E$ and $H \sim \mathcal{B}$.

1 Introduction

Is it worthwhile to reinvent classical electrodynamics after it has been with us for more than a century? And after its quantized version, quantum electrodynamics (unified with the weak interaction) had turned out to be one of the most accurately tested successful theories? We believe that the answer should be affirmative. Moreover, we believe that this reformulation should be done such that it is also comprehensible and useful for experimental physicists and (electrical) engineers[2].

Let us collect some of the reasons in favor of such a reformulation. First of all an "axiomatics" of electrodynamics should allow us to make the fundamental structure of electrodynamics transparent, see, e.g., Sommerfeld [**?**] or [**?**, **?**, **?**]. We will follow the tradition of Kottler-Cartan-van Dantzig, see Truesdell & Toupin [**?**] and Post [**?**], and base our theory on two experimentally well established axioms expressed in terms of integrals, conservation of electric charge and magnetic flux, and a local axiom, the existence of the Lorentz force. All three axioms can be formulated in a 4-dimensional (spacetime) continuum without using the distance concept (i.e. without the use of a metric), see Schrödinger [**?**]. Only the fourth axiom, a suitable constitutive law, is specific for the "material" under consideration which is interacting with the electromagnetic field. The vacuum is a particular example of such a material. In the fourth axiom, the distance concept eventually shows up and gives the 4-dimensional continuum an additional structure.

Some of the questions one can answer with the help of such a general framework are: Is the electric excitation $\mathcal{D}$ a *microscopic* quantity like the field strength E? Is it justified to give $\mathcal{D}$

[1] *Dedication to Erik W. Grafarend on the occasion of his 60th birthday:* Wir wissen, dass heutzutage auch die Geodäten relativistische Effekte bei ihren "Triangulationen" berücksichtigen müssen. Wohl auch aus diesem Grunde hat Herr Grafarend immer ein offenes Ohr für entsprechende Theorien gehabt. Die einfachste relativistische klassische Feldtheorie, die wir kennen, ist die Elektrodynamik. Wir widmen diese Ausarbeitung Herrn Grafarend zu seinem 60.Geburtstage in der Hoffnung, dass er sich über die schönen Seiten dieser Darstellung genauso freut, wie wir es tun. Und dies umso mehr, als dass Herr Grafarend gleich am Anfang seiner Karriere sich intensiv mit Geometrie und Cartan-Formalismus auseinandergesetzt hat.

[2] For this reason, we apply in this article the more widespread formalism of tensor analysis ("Ricci calculus", see Schouten [**?**]) rather than that of exterior differential forms ("Cartan calculus", see Frankel [**?**]) which we basically prefer.

another dimension than E? The analogous questions can be posed for the magnetic excitation H and the field strength $\mathcal{B}$. Should we expect a magnetic monopole and an explicit magnetic charge to arise in such an electrodynamic framework? Can we immediately pinpoint the (metric-independent) constitutive law for a 2-dimensional electron gas in the theory of the quantum Hall effect? Does the non-linear Born-Infeld electrodynamics fit into this general scheme? How do Maxwell's equations look in an accelerated reference frame or in a strong gravitational field as around a neutron star? How do they look in a possible non-Riemannian spacetime? Is a possible pseudoscalar axion field compatible with electrodynamics? And eventually, on a more formal level, is the calculus of exterior differential forms more appropriate for describing electrodynamics than the 3-dimensional Euclidean vector calculus and its 4-dimensional generalization? Can the metric of spacetime be derived from suitable assumptions about the constitutive law?
It is really the status of electrodynamics within the whole of physics which comes much clearer into focus if one follows up such an axiomatic approach.

2 Foliation of the 4-dimensional spacetime continuum

From a modern relativistic point of view, the formulation of electrodynamics has to take place in a 4-dimensional continuum (differentiable manifold) which eventually is to be identified with spacetime, i.e. with a continuum described by one "time" coordinate x^0 and three "space" coordinates x^1, x^2, x^3 or, in short, by coordinates x^i, with $i = 0, 1, 2, 3$. Let us suppress one space dimension in order to be able to depict the 4-dimensional as a 3-dimensional continuum, as shown in Fig.**??**.

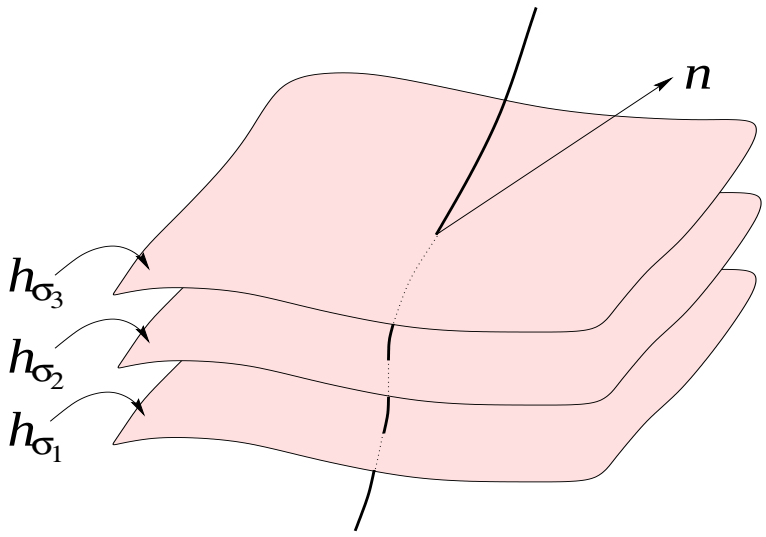


Figure 1: Foliation of spacetime: Each hypersurface h_σ represents, at a time σ, the 3-dimensional space of our perception one dimension of which is suppressed in the figure. The positive time direction runs upwards.

We assume that this continuum admits a foliation into a succession of different leaves or hypersurfaces h. Accordingly, spacetime looks like a pile of leaves which can be numbered by a monotonically increasing (time) parameter σ. A leaf h_σ is defined by $\sigma(x^i) = const.$ It represents, at a certain time σ, the ordinary 3-dimensional space surrounding us (in Fig.**??** it is 2-dimensional, since one dimension is suppressed).
At any given point in h_σ, we can introduce the covector $k_i := \partial_i \sigma$ and a 4-vector $n = (n^i) = (n^0, n^1, n^2, n^3) = (n^0, n^a)$ such that n is normalized according to

$$n^i \, k_i = n^i \partial_i \sigma = 1 \, . \tag{1}$$

Here $a, b \ldots = 1, 2, 3$ and $i, j \ldots = 0, 1, 2, 3$. Furthermore, summation over repeated indices is always understood. The vector n^i is "normal" to the leaf h_σ, whereas the covector k_i is tangential[3] to it.
With the pair (n, k) we can construct projectors which decompose all tensor quantities into *longitudinal* and *transversal* constituents with respect to the vector n, see Fig.**??**. Indeed, the matrices

$$L^i{}_j := n^i\, k_j \quad \text{and} \quad T^i{}_j := \delta^i_j - n^i\, k_j \quad \text{with} \quad L^i{}_j + T^i{}_j = \delta^i_j\,, \tag{2}$$

represent projection operators, i.e.

$$L^i{}_j\, L^j{}_k = L^i{}_k\,, \quad T^i{}_j\, T^j{}_k = T^i{}_k\,, \quad L^i{}_j\, T^j{}_k = T^i{}_j\, L^j{}_k = 0\,. \tag{3}$$

Taking an arbitrary covector U_i, we now can write it as

$$U_i = {}^{\perp}U_i + \underline{U}_i\,, \quad \text{where} \quad {}^{\perp}U_i := L^j{}_i\, U_j \quad \text{and} \quad \underline{U}_i := T^j{}_i\, U_j\,. \tag{4}$$

Obviously ${}^{\perp}U_i$ describes the longitudinal component of the covector and $\underline{U}_i$ its transversal component, with $n^i\, \underline{U}_i = 0$. Analogously, for an arbitrary vector V^i, we can write

$$V^i = {}^{\perp}V^i + \underline{V}^i\,, \quad \text{where} \quad {}^{\perp}V^i := L^i{}_j\, V^j \quad \text{and} \quad \underline{V}^i := T^i{}_j\, V^j\,. \tag{5}$$

Its transversal component $\underline{V}^i$ fulfills $\underline{V}^i\, k_i = 0$. This pattern can be straightforwardly generalized to all tensorial quantities of spacetime.
For simplicity, we confine our attention to the particular case when "adapted" coordinates $x^i = (\sigma, x^a)$ are used and when the "spatial" components of n vanish, i.e., $n^i = (1, 0, 0, 0)$. In that case, we simply have $k_i = \partial_i \sigma = (1, 0, 0, 0)$ and hence σ can be treated as a formal "time" coordinate.

3 Conservation of electric charge (axiom 1)

The conservation of electric charge was already recognized as fundamental law during the time of Franklin (around 1750) well before Coulomb discovered the force law in 1785. Nowadays, at a time, at which one can catch single electrons and single protons in traps and can *count* them individually, we are more sure than ever that electric charge conservation is a valid fundamental law of nature. Therefore matter carries as a *primary quality* something called electric charge which only occurs in positive or negative units of an elementary charge e (or, in the case of quarks, in 1/3th of it) and which can be counted in principle. Thus it is justified to introduce the physical dimension of charge q as a new and independent concept. Ideally one should measure a charge in units of $e/3$. However, for practical reasons, the SI-unit C (Coulomb) is used in laboratory physics.
Two remarks are in order: Charge is an additive (or extensive) quantity that characterizes the source of the electromagnetic field. It is prior to the notion of the electric field strength. Therefore it is *not* reasonable to measure, as is done in the CGS-system of units, the additive quantity charge in terms of the unit of force by applying Coulomb's law. Coulomb's law has no direct relation to charge conservation. Secondly, in the SI-system, for reasons of better realization, the Ampere A as current is chosen as the new fundamental unit rather than the Coulomb. We have $C = A\, s$ ($s =$ second).
As a preliminary step, let us remind ourselves that, in a 4-dimensional picture, the motion of a point particle is described, as in Fig.**??**, by a curve in spacetime, by a so-called worldline. The tangent vectors of this worldline represent the 4-velocity of the particle.

[3]The term "tangential" is used here in the sense of exterior calculus in which a covector (or 1-form) is represented by two ordered parallel planes – and the first plane is tangential to h_σ.

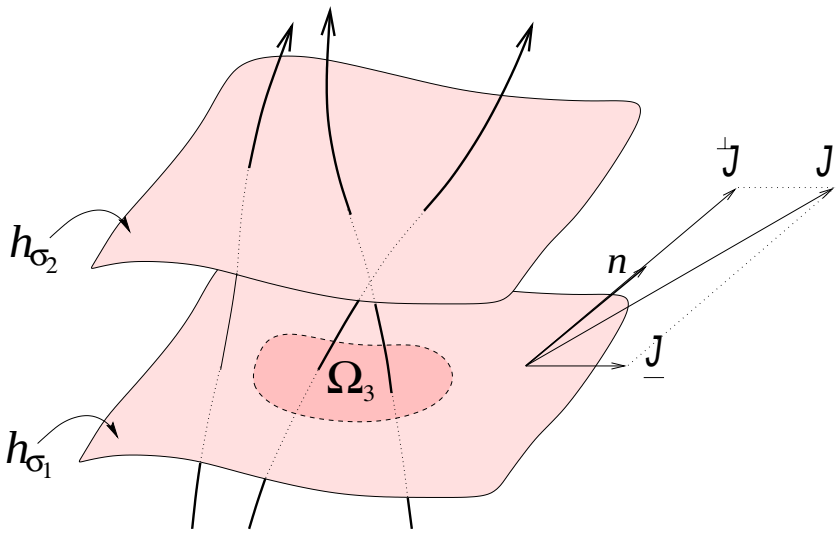


Figure 2: World lines, decomposition of the electric current into the piece ${}^{\perp}\mathcal{J}$ longitudinal to n and the transversal piece $\underline{\mathcal{J}}$, global conservation of charge.

If we mark a 3-dimensional volume Ω_3 which belong to a certain hypersurface h_σ, then the total electric charge inside Ω_3 is

$$Q = \int\limits_{\Omega_3 \subset h_\sigma} \rho \, dx^1 dx^2 dx^3 \,, \tag{6}$$

with ρ as the electric charge density. The total charge in space, which we find by integration over the whole of space, i.e., by letting $\Omega_3 \to h_\sigma$, is *globally* conserved. Therefore the integral in (**??**) over each hypersurface $h_{\sigma_1}, h_{\sigma_2}, \ldots$ keeps the same value.
The *local* conservation of charge, see Fig.**??**, translates into the following fact: If a number of worldlines of particles with one elementary charge enter a prescribed but arbitrary 4-dimensional volume Ω_4, then, in classical physics, the same number has to leave the volume. If we count the entering worldlines as negative and the leaving ones as positive (in conformity with the direction of their normal vectors), then the (3-dimensional) surface integral over the number of worldlines has to vanish.

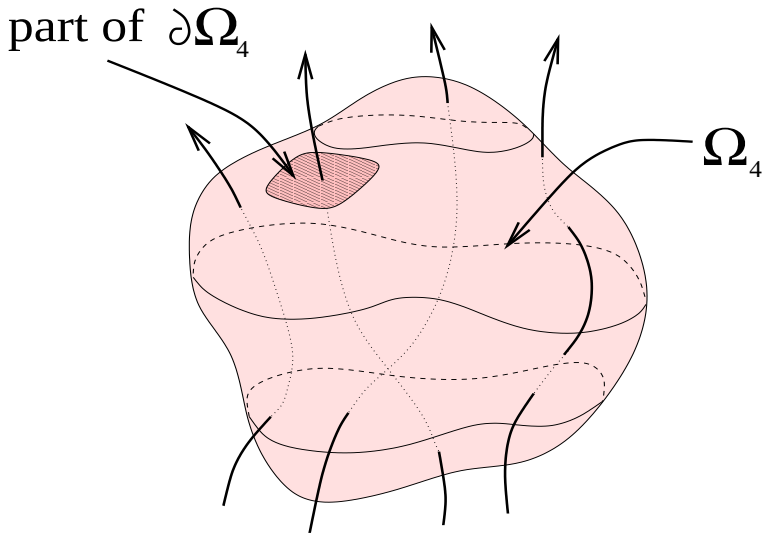


Figure 3: Local conservation of charge: Each worldline of a charged particle that enters the finite 4-volume Ω_4 via its boundary $\partial\Omega_4$ has also to leave Ω_4.

Now, the natural extensive quantities to be integrated over a 3-dimensional hypersurface are *vector densities,* see the appendix. Accordingly, in nature there should exist a 4-vector density $\mathcal{J}^i$ with 4 independent components which measures the charge piercing through an arbitrary 3-dimensional hypersurface. Therefore, it generalizes in a consistent 4-dimensional formalism the familiar concepts of charge density ρ and current density j^a. The axiom of local charge conservation then reads

$$\int\limits_{\partial\Omega_4} \mathcal{J}^i \, d^3S_i = 0\,, \tag{7}$$

where the integral is taken over the (3-dimensional) *boundary* of an arbitrary 4-dimensional volume of spacetime, with d^3S_i being the 3-surface element, as defined in the appendix.
If we apply Stokes' theorem, then we can transform the 3-surface integral in (**??**) into a 4-volume integral:

$$\int\limits_{\Omega_4} (\partial_i \mathcal{J}^i) \, d^4S = 0\,. \tag{8}$$

Since this is valid for an *arbitrary* 4-volume Ω_4, we find the local version of the charge conservation as

$$\partial_i \mathcal{J}^i = 0\,. \tag{9}$$

In this form, the law of conservation of charge is valid in arbitrary coordinates.
If one defines a particular foliation, then one can rewrite (**??**) in terms of decomposed quantities that are longitudinal and transversal to the corresponding normal vector n. The 4-vector density $\mathcal{J}^i$ decomposes as

$$\mathcal{J}^i = {}^{\perp}\mathcal{J}^i + \underline{\mathcal{J}}^i\,. \tag{10}$$

When adapted coordinates are used, the decomposition procedure simplifies and allows to define the 3-dimensional densities of charge ρ and of current j^a as

$$\rho := {}^{\perp}\mathcal{J}^0 = \mathcal{J}^0, \qquad j^a := \underline{\mathcal{J}}^a = \mathcal{J}^a. \tag{11}$$

With this, one can rewrite the definition of charge (**??**) in an explicitly coordinate invariant form

$$Q = \int\limits_{\Omega_3 \subset h_\sigma} \mathcal{J}^i \, d^3S_i\,, \tag{12}$$

since on h_σ we have $d^3S_0 = dx^1dx^2dx^3$ and $d^3S_a = 0$. Furthermore, Eq.(**??**) can be rewritten in (1+3)-form as the more familiar continuity equation

$$\partial_\sigma \rho + \partial_a j^a = 0\,. \tag{13}$$

The charge Q in (**??**) has the absolute dimension[4] q. The 4-current is a density in spacetime, and we have $[\mathcal{J}] = q/(t\,l^3)$. Thus the components carry the dimensions $[\rho] = [\mathcal{J}^0] = q/l^3 \stackrel{\text{SI}}{=} C/m^3$ and $[j^a] = [\mathcal{J}^a] = q/(t\,l^2) \stackrel{\text{SI}}{=} A/m^2$.

[4]A theory of dimensions, which we are using, can be found in Post [**?**], e.g.. A quantity has an *absolute* dimension, and if it is a density in spacetime we divide by $t\,l^3$. The *components* pick up a t (a t^{-1}) for an upper (a lower) temporal index and an l (an l^{-1}) for an upper (a lower) spatial index. A statement, see [**?**], that E and B must have the same dimension since they transform into each other is empty without specifying the underlying theory of dimensions.

4 The inhomogeneous Maxwell equations as consequence

Because of axiom 1 and according to a theorem of de Rham, see [**?**], the electric current density from (**??**) or (**??**) can be represented as a "divergence" of the *electromagnetic excitation*:

$$\mathcal{J}^i = \partial_j \mathcal{H}^{ij}\,, \qquad \mathcal{H}^{ij} = -\mathcal{H}^{ji}\,. \tag{14}$$

The excitation $\mathcal{H}^{ij}$ is a contravariant antisymmetric tensor density and has 6 independent components. One can verify that, due to the antisymmetry of $\mathcal{H}^{ij}$, the conservation law is automatically fulfilled, i.e., $\partial_i \mathcal{J}^i = \partial_i \partial_j \mathcal{H}^{ij} = 0$.
The 4-dimensional set (**??**) represents the inhomogeneous Maxwell equations. They surface here in a very natural way as a result of charge conservation. Charge conservation should not be looked at as a consequence of the inhomogeneous Maxwell equations, but rather the other way round, as shown in this tutorial. Of course, $\mathcal{H}^{ij}$ is not yet fully determined since

$$\tilde{\mathcal{H}}^{ij} = \mathcal{H}^{ij} + \epsilon^{ijkl}\,\partial_k \psi_l \tag{15}$$

also satisfies (**??**) for an arbitrary covector field ψ_i.
The $(1+3)$-decomposition of $\mathcal{H}^{ij}$ is obtained similarly to the decomposition of the current (**??**):

$$\mathcal{H}^{ij} = {}^{\perp}\mathcal{H}^{ij} + \underline{\mathcal{H}}^{ij}. \tag{16}$$

The nontrivial components of the longitudinal and transversal parts read

$$\mathcal{H}^{0a} = {}^{\perp}\mathcal{H}^{0a} = \mathcal{D}^a, \qquad \mathcal{H}^{ab} = \underline{\mathcal{H}}^{ab} = \epsilon^{abc}\, H_c\,, \tag{17}$$

with the electric excitation $\mathcal{D}^a$ (historical name: "dielectric displacement") and the magnetic excitation H_a ("magnetic field"). Here ϵ^{abc} is the totally antisymmetric 3-dimensional Levi-Civita tensor density with $\epsilon^{123} = 1$.
If we substitute the decompositions (**??**) and (**??**) into (**??**), we recover the 3-dimensional form of the inhomogeneous Maxwell equations,

$$\partial_a \mathcal{D}^a = \rho, \qquad \epsilon^{abc}\partial_b H_c - \partial_\sigma \mathcal{D}^a = j^a\,, \tag{18}$$

or, in symbolic notation,

$$\operatorname{div} \mathcal{D} = \rho, \qquad \operatorname{curl} H - \dot{\mathcal{D}} = j\,. \tag{19}$$

Since electric charge conservation is valid in microphysics, the corresponding Maxwell equations (**??**) or (**??**) are also *microphysical* equations and with them the excitations $\mathcal{D}^a$ and H_a are microphysical quantities likewise – in contrast to what is stated in most textbooks, see [**?**] and [**?**], compare also [**?**], e.g..
From (**??**) we can immediately read off $[\mathcal{D}^a] = [l\,\rho] = q/l^2 \overset{\text{SI}}{=} C/m^3$ and $[H_a] = [l\,j^a] = q/(t\,l) \overset{\text{SI}}{=} A/m$. Before we ever talked about *forces* on charges, charge conservation alone gave us the inhomogeneous Maxwell equations including the appropriate dimensions for the excitations $\mathcal{D}^a$ and H_a.
Under the assumption that $\mathcal{D}^a$ vanishes inside an ideal electric conductor, one can get rid of the indeterminacy of $\mathcal{D}^a$, as spelled out in (**??**), and we can measure $\mathcal{D}^a$ by means of two identical conducting plates ("Maxwellian double plates") which touch each other and which are *separated* in the $\mathcal{D}^a$-field to be measured. The charge on one plate is then measured. Analogous remarks apply to H_a. Accordingly, the excitations do have a direct *operational* significance.

5 Force and field strengths (axiom 2)

By now we have exhausted the information contained in the axiom 1 of charge conservation. We have to introduce new concepts in order to complete the fundamental structure of Maxwell's theory. Whereas the excitation $\mathcal{H} = (\mathcal{D}^a,\, H_a)$ is linked to the charge current $\mathcal{J} = (\rho,\, j^a)$, the electric and magnetic field strengths are usually introduced as forces acting on unit charges at rest or in motion, respectively. In the purely electric case with a test charge q, we have in terms of components

$$F_a \sim q\, E_a\,, \tag{20}$$

with F as force and E as electric field covector.
Let us take a (delta-function-like) test charge current $\mathcal{J} = (\rho,\, j^a)$ centered around a point with spatial coordinates x^a. Generalizing (**??**), the simplest relativistic ansatz for defining the electromagnetic field reads:

$$\text{force density} \quad \sim \quad \text{field strength} \times \text{charge current density}\,. \tag{21}$$

We know from Lagrangian mechanics that the *force* $\sim \partial L/\partial x^i$ is represented by a *covector* with the absolute dimension of action $\hbar$ (here $\hbar$ is *not* the Planck constant but rather only denotes its *dimension*). Accordingly, with the covectorial force density f_i, the ansatz (**??**) can be made more precise as axiom 2:

$$f_i = F_{ij}\, \mathcal{J}^j\,, \qquad F_{ij} = -F_{ji}\,. \tag{22}$$

The newly introduced covariant 2nd-rank 4-tensor F_{ij} is the electromagnetic field strength. The force density f_i was postulated to be normal to the current, $f_i\, \mathcal{J}^i = 0$. Thus the antisymmetry of the electromagnetic field strength is found, i.e., F_{ij} depends on 6 independent components. We know the notion of force from mechanics, the current density we know from axiom 1. Accordingly, axiom 2 is to be understood as an *operational* definition of the electromagnetic field strength F_{ij}.
With the decomposition

$$F_{ij} = {}^{\perp}F_{ij} + \underline{F}_{ij}\,, \tag{23}$$

we find the identifications for the electric field strength E_a and the magnetic field strength $\mathcal{B}^a$ (historical names: "magnetic induction" or "magnetic flux density"):

$$F_{a0} = {}^{\perp}F_{a0} = E_a, \qquad F_{ab} = \underline{F}_{ab} = \epsilon_{abc}\, \mathcal{B}^c. \tag{24}$$

These identifications are reasonable since for the spatial components of (**??**) we recover the Lorentz force density and, for the static case, Eq.(**??**):

$$f_a = F_{aj}\, \mathcal{J}^j = F_{a0}\, \mathcal{J}^0 + F_{ab}\, \mathcal{J}^b = \rho\, E_a + \epsilon_{abc}\, j^b\, \mathcal{B}^c\,. \tag{25}$$

Symbolically, we have

$$f = \rho\, E + j \times \mathcal{B}\,. \tag{26}$$

The time component of (**??**) represents the electromagnetic power density:

$$f_0 = F_{0a}\, \mathcal{J}^a = -E_a\, j^a\,. \tag{27}$$

6 Conservation of magnetic flux (axiom 3)

Axiom 2 on the Lorentz force gave us a new quantity, the electromagnetic field strength with the dimension $[F] = action/charge = \hbar/q =: \phi$, with $\phi = work \times time/charge = voltage \times time \stackrel{\text{SI}}{=} Vs = Wb$. Here Wb is the abbreviation for Weber. Thus its components carry the following dimensions: $[E_a] = [F_{a0}] = \phi/(t\, l) \stackrel{\text{SI}}{=} V/m$ and $[\mathcal{B}^c] = [F_{ab}] = \phi/l^2 \stackrel{\text{SI}}{=} Wb/m^2 = T$ (for Tesla).

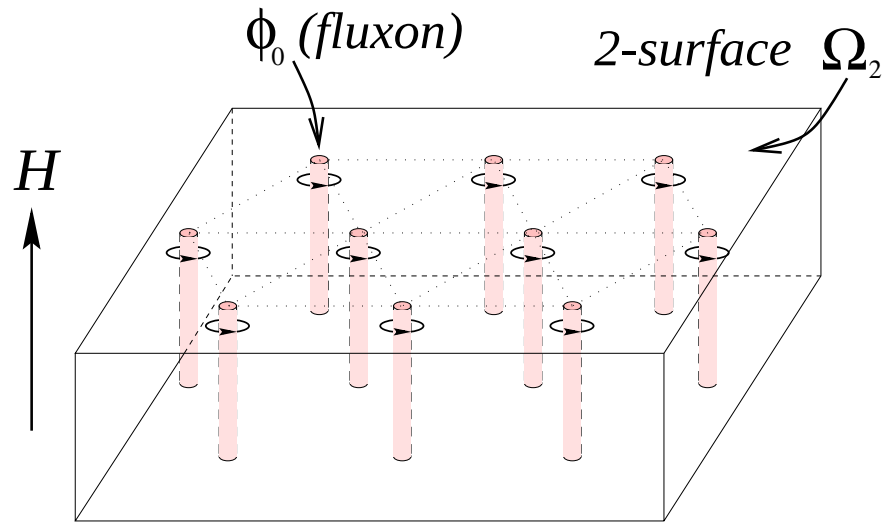


Figure 4: Sketch of an Abrikosov lattice in a type II superconductor in 3-dimensional space. In contrast to all the other figures, this is *not* an image of 4-dimensional spacetime.

We are in need of an experimentally established law that relates to F. And we would prefer, as in the case of the electric charge, to recur to a counting procedure. What else can we count in relation to the electromagnetic field? Certainly *magnetic flux lines* in the interior of a type II superconductor which is exposed to a sufficiently strong magnetic field. And these flux lines are quantized. In fact, they can order in a 2-dimensional triangular Abrikosov lattice, see Fig.**??**. These flux lines carry a unit of magnetic flux, a so-called flux quantum or *fluxon* with $\Phi_0 = h/(2e) = 2.07 \times 10^{-15}\, Wb$, see Tinkham [**?**]; here h *is* the Planck constant and e the elementary charge. These flux lines can move, via its surface, in or out of the superconductor, but they cannot vanish (unless two lines with different sign collide) or spontaneously come into existence. In other words, there is a strong experimental evidence that magnetic flux is a conserved quantity.
The number 2 in the relation $\Phi_0 = h/(2e)$ is due to the fact that the Cooper pairs in a superconductor consist of 2 electrons. Moreover, outside a superconductor the magnetic flux is *not* quantized, i.e., we cannot count the flux lines there with the same ease that we could use inside. Nevertheless, as we shall see, experiments clearly show that the magnetic flux is conserved also there.
As we can take from Fig.**??**, the magnetic flux should be defined as a 2-dimensional spatial integral. These flux lines are additive and we have

$$\Phi = \int\limits_{\Omega_2 \subset h_\sigma} \mathcal{B}^a \, d^2 S_a \,. \tag{28}$$

Here $\mathcal{B}^a$ is the magnetic field strength and $d^2 S_a$ the spatial 2-surface element. This definition of the magnetic flux should be compared with the definition (**??**) of the charge. Here, in (**??**), we integrate only over 2 dimensions rather than over 3 dimensions, as in the case of the charge in (**??**). Thus in a spacetime picture in which one space dimension is suppressed, see Fig.**??**, our magnetic flux integral looks like an integral over a finite interval $[A, B]$ embedded into the hypersurface h_{σ_1}.
Now we are going to argue again as in Sec.3. If $\Omega_2 \to \infty$, i.e., if we integrate over an infinite spatial 2-surface ($A \to +\infty$, $B \to -\infty$), then the total magnetic flux at time σ_1 is given by (**??**). If we propagate that interval into the (coordinate) future, see the interval on the hypersurface h_{σ_2}, then magnetic flux conservation requires the constancy of the integral Φ. In other words, if we orient the integration domain suitably, the loop integral, the domain of which is drawn in

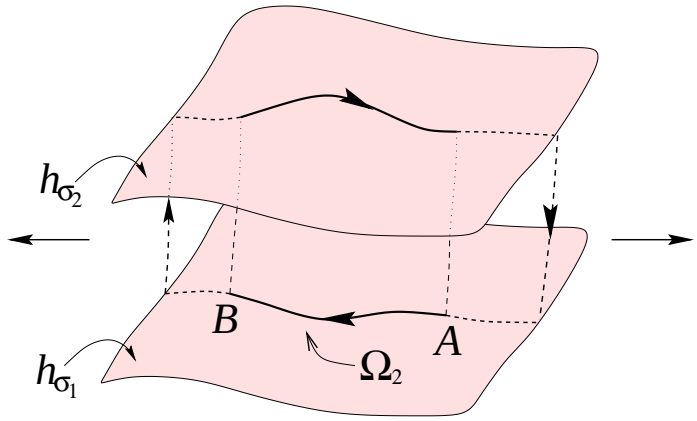


Figure 5: Magnetic flux in spacetime: The magnetic field $\mathcal{B}^a$, if integrated over the interval $[A, B]$, represents, at time σ_1, the magnetic flux piercing through this 2-dimensional integration domain.

Fig.**??**, has to vanish since no flux is supposed to leak out along the dotted "vertical" domains at spatial infinity.
Analogously as we did in the case of charge conservation, we want to formulate a corresponding local conservation law in an explicitly covariant way. We saw that the global conservation of magnetic flux is expressed as the vanishing of the integral of $\mathcal{B}$ over the particular 2-dimensional loop in Fig.**??**. In a 4-dimensional covariant formalism, the natural intensive objects to be integrated over a 2-dimensional region are second order antisymmetric covariant tensors, see the appendix. The magnetic field strength $\mathcal{B}$ is just a piece of the electromagnetic field strength F. Thus, it is clear that the natural local generalization of the magnetic flux conservation, our axiom 3, is

$$\int\limits_{\partial\Omega_3} \frac{1}{2} F_{ij}\, d^2S^{ij} = 0\,, \tag{29}$$

where the integral is taken over the boundary of an arbitrary 3-dimensional hypersurface of spacetime, as is sketched in Fig.**??**. We apply Stokes' theorem

$$\int\limits_{\Omega_3} \epsilon^{ijkl}\partial_{[j}\, F_{kl]}\, d^3S_i = 0\,, \tag{30}$$

and, since the volume is arbitrary, we have the local version of magnetic flux conservation as

$$\partial_{[i}F_{jk]} = 0. \tag{31}$$

We substitute the decomposition (**??**) into (**??**). Then we find the homogeneous Maxwell equations,

$$\partial_a\mathcal{B}^a = 0\,, \qquad \epsilon^{abc}\partial_b E_c + \partial_\sigma\mathcal{B}^a = 0 \tag{32}$$

or, symbolically,

$$\operatorname{div}\mathcal{B} = 0\,, \qquad \operatorname{curl} E + \dot{\mathcal{B}} = 0\,. \tag{33}$$

Thus both, the sourcelessness of $\mathcal{B}^a$ and the Faraday induction law follow from magnetic flux conservation. Both laws are experimentally very well verified and, in turn, strongly support the axiom of the conservation of the magnetic flux.
The recognition that Maxwell's theory, besides on charge conservation, is based on magnetic flux conservation, sheds new light on the possible existence of magnetic monopoles. First of all, careful search for them has not lead to any signature of their possible existence, see [**?**]. Furthermore, magnetic flux conservation would be violated if we postulated the existence of

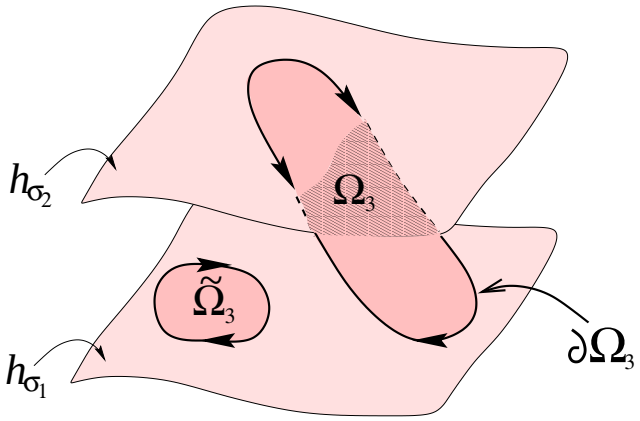


Figure 6: Conservation of magnetic flux in spacetime: Consider the arbitrary 3-dimensional integration domain Ω_3. The integral vanishes of the field strength F_{ij} over the 2-dimensional boundary $\partial\Omega_3$ of the 3-dimensional domain Ω_3. The analogous is true for the flux integral over $\partial\tilde{\Omega}_3$.

a current on the right hand side of (**??**). Now, Eq.(**??**) is the analog of (**??**), at least in our axiomatic set-up. Why should we believe in charge conservation any longer if we gave up magnetic flux conservation? Accordingly, we assume – in contrast to most elementary particle physicists, see Cheng & Li [**?**] – that in Maxwell's theory proper there is no place for a magnetic current[5] on the right hand side of (**??**).

7 Constitutive law (axiom 4)

The Maxwell equations (**??**) and (**??**) or, in the decomposed version, (**??**) and (**??**), respectively, encompass altogether 6 partial differential equations with a first order time derivative (the 2 remaining equations can be understood as constraints to the initial configuration). Since excitations and field strengths add up to $6 + 6 = 12$ independent components, certainly the Maxwellian set is underdetermined with respect to the time propagation of the electromagnetic field. What we clearly need is a relation between the excitations and the field strengths. As we will see, these so-called constitutive equations require additional knowledge about the properties of spacetime whereas the Maxwell equations, as derived so far, are of universal validity as long as classical physics is a valid approximation. In particular, in the Riemannian space of Einstein's gravitational theory the Maxwell equations look just the same as in (**??**) and (**??**). There is no adaptation needed of any kind, see [**?**].
If we investigate macroscopic matter, one has to derive from the microscopic Maxwell equations by statistical procedures the *macroscopic* Maxwell equations. They are expected to have the same structure as the microscopic ones. But let us stay, for the time being, on the microscopic level.
Then we can make an attempt with a *linear* constitutive relation between $\mathcal{H}^{ij}$ and F_{kl},

$$\mathcal{H}^{ij} = \frac{1}{2}\,\tilde{\chi}^{ijkl}\,F_{kl} = \frac{1}{2}\,f\chi^{ijkl}\,F_{kl}\,, \tag{34}$$

with the tensor density χ^{ijkl} that is characteristic for the spacetime under consideration. We require $[\chi] = 1$, i.e., for the dimensionfull scalar factor factor f we have $[f] = q/\phi = q^2/h \stackrel{\text{SI}}{=} C/(Vs) = A/V = 1/\Omega$. The dimensionless "modulus" χ^{ijkl}, because of the antisymmetries of $\mathcal{H}^{ij}$ and F_{kl}, obeys

$$\chi^{ijkl} = -\chi^{jikl} = -\chi^{ijlk}\,. \tag{35}$$

[5]This argument does not exclude that, for *topological* reasons, the integral in (**??**) could be non-vanishing, as in the case of a Dirac monopole with a string, see [**?**].

Moreover, if we assume the existence of a Lagrangian density for the electromagnetic field $\mathcal{L} \sim \mathcal{H}^{ij} F_{ij}$, then we have additionally the symmetries

$$\chi^{ijkl} = \chi^{klij}\,, \qquad \chi^{[ijkl]} = 0\,. \tag{36}$$

The vanishing of the totally antisymmetric part comes about since the corresponding Euler-Lagrange derivative of $\mathcal{L}$ with respect to the 4-potential A_i identically vanishes; here $F_{ij} = 2\partial_{[i}A_{j]}$. For χ^{ijkl}, this leaves 20 independent components[6]. One can take such moduli, if applied on a macrophysical scale, for describing the electromagnetic properties of anisotropic crystals, e.g.. Then also non-linear (for ferromagnetism) and spatially non-local constitutive laws are in use.
The simplest linear law is expected to be valid in vacuum. Classically, the vacuum of spacetime is described by its metric tensor $g^{ij} = g^{ji}$ that determines the temporal and spatial distances of neighboring events. Considering the symmetry properties of the density χ^{ijkl}, the only ansatz possible, up to an arbitrary constant, seems to be

$$\chi^{ijkl} = \sqrt{-\det g_{mn}}\,\left(g^{ik}g^{jl} - g^{jk}g^{il}\right)\,. \tag{37}$$

Note that χ^{ijkl} is invariant under a rescaling of the metric $g_{ij} \to \Omega^2 g_{ij}$, with an arbitrary function $\Omega(x^i)$. Using this freedom, we can always normalize the determinant of the metric to 1.
As an example, let us consider the *flat* spacetime metric of a Minkowski space in Minkowskian coordinates,

$$\eta^{ij} = \sqrt{c}\begin{pmatrix} c^{-2} & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{38}$$

If we substitute (**??**) into (**??**) and, in turn, Eq.(**??**) and $f = \sqrt{\varepsilon_0/\mu_0}$ into (**??**), then we eventually find the well-known vacuum ("Lorentz aether") relations,

$$\mathcal{H}^{ij} = \sqrt{\frac{\varepsilon_0}{\mu_0}}\eta^{ik}\eta^{jl}\,F_{kl} \qquad \text{or} \qquad \mathcal{D} = \varepsilon_0\,E\,, \quad H = (1/\mu_0)\,\mathcal{B}\,. \tag{39}$$

The law (**??**) converts Maxwell's equations, for vacuum, into a system of differential equations with a well-determined initial value problem.

Acknowledgments: We are grateful to Marc Toussaint for discussions on magnetic monopoles. G.F.R. would like to thank the German Academic Exchange Service DAAD for a graduate fellowship (Kennziffer A/98/00829).

A Four-dimensional calculus without metric and integrals

In a 4-dimensional space, in which arbitrary coordinates x^i are used, with $i = 0, 1, 2, 3$, one can define derivatives and integrals of suitable antisymmetric covariant *tensors* and antisymmetric contravariant *tensor densities* without the need of a metric. The tensors are used for representing intensive quantities (how strong?), the tensor densities for extensive (additive) quantities (how much?). The natural formalism for defining integrals in a coordinate invariant way is exterior calculus, see Frankel [**?**]. However, we will use here tensor calculus, see Schouten [**?**] and also Schrödinger [**?**], which is more widely known under physicists and engineers.

[6] With such a linear constitutive law it is even possible to *derive*, up to a conformal factor, a metric of spacetime, provided one makes one additional assumption, see [**?**, **?**].

Integration over 4-dimensional regions – scalar densities

Consider a certain 4-dimensional region Ω_4. Then a integral over Ω_4 is of the form

$$\int_{\Omega_4} \mathcal{A}\, d^4S\,, \tag{40}$$

where $d^4S := dx^0dx^1dx^2dx^3$ is the 4-volume element which is a scalar density of weight -1. We want this integral to be a scalar, i.e., that its value does not depend on the particular coordinates we use. Then the integrand $\mathcal{A}$ has to be a scalar density of weight $+1$. In other words, when using the tensor formalism, the natural quantity required to formulate an invariant integral over a 4-dimensional region is a scalar density of weight $+1$.

Integration over 3-dimensional regions – vector densities

Now we want to define invariant integrals over some 3-dimensional hypersurface Ω_3 in a four-dimensional space which can be defined by the parameterization $x^i = x^i(y^a)$, $a, b, c = 1, 2, 3$, where y^a are also arbitrary coordinates on Ω_3. Then we call

$$d^3S_i := \frac{1}{3!}\,\epsilon_{ijkl}\,\frac{\partial x^j}{\partial y^a}\frac{\partial x^k}{\partial y^b}\frac{\partial x^l}{\partial y^c}\,\epsilon^{abc}\,dy^1dy^2dy^3 \tag{41}$$

the *3-surface element on* Ω_3. This quantity is constructed by using only objects that can be defined in a general 4-dimensional space without metric or connection. It can be constructed as soon as we specify the parameterization of Ω_3. Here ϵ_{ijkl} is the 4-dimensional Levi-Civita tensor density of weight -1 and ϵ^{abc} the 3-dimensional Levi-Civita tensor density of weight $+1$ on Ω_3. Furthermore, this hypersurface element turns out to be a covector density of weight -1 with respect to 4-dimensional coordinate transformation. With this integration element to our disposal, the natural form of an invariant integral over Ω_3 is

$$\int_{\Omega_3} \mathcal{A}^i d^3S_i\,. \tag{42}$$

Therefore, the natural object to be integrated over Ω_3 in order to obtain an invariant result is a vector density of weight $+1$.

Integration over 2-dimensional regions – covariant tensors or contravariant tensor densities

Analogously, we can parameterize a 2-dimensional region Ω_2 by means of $x^i = x^i(z^\alpha)$, $\alpha, \beta = 1, 2$, where z^α are arbitrary coordinates on Ω_2. Then we can immediately construct the following 2-surface element

$$d^2S_{ij} := \frac{1}{2}\,\epsilon_{ijkl}\,\frac{\partial x^k}{\partial z^\alpha}\frac{\partial x^l}{\partial z^\beta}\,\epsilon^{\alpha\beta}\,dz^1dz^2, \tag{43}$$

where $\epsilon^{\alpha\beta}$ is the Levi-Civita density of weight $+1$ on Ω_2. This surface element is an antisymmetric second order covariant tensor density of weight -1. Then an invariant integral is naturally defined as

$$\int_{\Omega_2} \frac{1}{2}\mathcal{A}^{ij}\, d^2S_{ij}\,, \tag{44}$$

with $\mathcal{A}^{ij}$ being an antisymmetric second order contravariant tensor density of weight $+1$. Alternatively, one can write the same integral in terms on an antisymmetric second order covariant tensor $A_{ij} := \frac{1}{2}\epsilon_{ijkl}\mathcal{A}^{kl}$ and an antisymmetric second order contravariant surface element

$$d^2S^{ij} := \frac{\partial x^k}{\partial z^\alpha}\frac{\partial x^l}{\partial z^\beta}\epsilon^{\alpha\beta}\,dz^1dz^2\,, \tag{45}$$

such that

$$\int_{\Omega_2} \frac{1}{2}\mathcal{A}^{ij}\, d^2S_{ij} = \int_{\Omega_2} \frac{1}{2}A_{ij}\, d^2S^{ij}\,. \tag{46}$$

Since extensive quantities are represented by densities, we would take the first integral for them, whereas for intensive quantities the second integral should be used. Analogous considerations can be applied to (**??**) and (**??**).

Stokes' theorem

Stokes' theorem gives us as particular cases the following integral identities (see [**?**] p.67 et seq.):

$$\int_{\Omega_4} (\partial_i \mathcal{J}^i)\, d^4S = \int_{\partial\Omega_4} \mathcal{J}^i\, d^3S_i\,, \tag{47}$$

$$\int_{\Omega_3} (\partial_j \mathcal{H}^{ij})\, d^3S_i = \int_{\partial\Omega_3} \frac{1}{2}\,\mathcal{H}^{ij}\, d^2S_{ij}\,. \tag{48}$$

B Decomposition of totally antisymmetric tensors into longitudinal and transversal pieces

Here we provide the decomposition formulas for totally antisymmetric covariant and contravariant tensors, which are the natural generalization of the decomposition of vectors and covectors. We start by considering an antisymmetric covariant tensor of rank p, namely $U_{i_1\ldots i_p}$. Its longitudinal and transversal components are given by

$${}^{\perp}U_{i_1\ldots i_p} = p\; L^m_{\;[i_1|}U_{m|i_2\ldots i_p]}\,, \qquad \underline{U}_{i_1\ldots i_p} = (p+1)L^m_{\;[m}U_{i_1\ldots i_p]}\,, \tag{49}$$

where $k_i := \partial_i\sigma$, and $L^i_{\;j} := n^i k_j$. They fulfill the following properties:

$$n^{i_1}\underline{U}_{i_1\ldots i_p} = 0\,, \qquad n^{i_1}{}^{\perp}U_{i_1\ldots i_p} = n^{i_1}U_{i_1\ldots i_p}\,. \tag{50}$$

For $p = 1, 2, 3, 4$ we can explicitly write:

p	quantity	definition	explicitly
1	${}^{\perp}U_i$	$L^m_{\;i}U_m$	$n^m k_i U_m$
2	${}^{\perp}U_{ij}$	$2L^m_{\;[i|}U_{m|j]}$	$n^m\,(k_iU_{mj} - k_jU_{mi})$
3	${}^{\perp}U_{ijk}$	$3L^m_{\;[i|}U_{m|jk]}$	$n^m\,(k_iU_{mjk} + k_jU_{mki} + k_kU_{mij})$
4	${}^{\perp}U_{ijkl}$	$4L^m_{\;[i|}U_{m|jkl]}$	$n^m\,(k_iU_{mjkl} - k_jU_{mkli} + k_kU_{mlij} - k_lU_{mijk})$

Now we turn to $V^{i_1\ldots i_p}$, an antisymmetric contravariant tensors of rank p. We define the decomposition as

$${}^{\perp}V^{i_1\ldots i_p} = p\; L^{[i_1|}_{\;\;m}V^{m|i_2\ldots i_p]}\,, \qquad \underline{V}^{i_1\ldots i_p} = (p+1)L^{[m}_{\;\;m}V^{i_1\ldots i_p]}\,. \tag{51}$$

They fulfill

$$k_{i_1}\underline{V}^{i_1\ldots i_p} = 0\,, \qquad k_{i_1}{}^{\perp}V^{i_1\ldots i_p} = k_{i_1}V^{i_1\ldots i_p}\,. \tag{52}$$

For $p = 1, 2, 3, 4$ we have the following explicit expressions for the longitudinal components:

p	quantity	definition	explicitly
1	${}^{\perp}V^i$	$L^i_{\;m}V^m$	$n^i k_m V^m$
2	${}^{\perp}V^{ij}$	$2L^{[i|}_{\;\;m}V^{m|j]}$	$k_m\left(n^iV^{mj} - n^jV^{mi}\right)$
3	${}^{\perp}V^{ijk}$	$3L^{[i|}_{\;\;m}V^{m|jk]}$	$k_m\left(n^iV^{mjk} + n^jV^{mki} + n^kV^{mij}\right)$
4	${}^{\perp}V^{ijkl}$	$4L^{[i|}_{\;\;m}V^{m|jkl]}$	$k_m\left(n^iV^{mjkl} - n^jV^{mkli} + n^kV^{mlij} - n^lV^{mijk}\right)$

An analogous scheme is valid for the corresponding densities.

References

[1] F. Bopp, Prinzipien der Elektrodynamik. *Z. Physik* **169** (1962) 45-52.

[2] R.G. Chambers, Units – B, H, D, and E. *Am. J. Phys.* **67** (1999) 468-469.

[3] T. Cheng and L. Li, *Theory of Elementary Particle Physics*. Clarendon Press, Oxford (1984).

[4] V.L. Fitch, The far side of sanity. *Am. J. Phys.* **67** (1999) 467.

[5] T. Frankel, *The Geometry of Physics* – An Introduction. Cambridge University Press, Cambridge (1997).

[6] Y.D. He, Search for a Dirac magnetic monopole in high energy nucleus-nucleus collisions. *Phys. Rev. Lett.* **79** (1997) 3134-3137.

[7] F.W. Hehl, J. Lemke, and E.W. Mielke, Two Lectures on Fermions and Gravity, in *Geometry and Theoretical Physics,* Proceedings of Bad Honnef School, Feb. 1990, J. Debrus and A.C. Hirshfeld, eds.. Springer, Berlin (1991) pp. 56-140.

[8] J.D. Jackson, *Classical Electrodynamics*. 3rd ed.. Wiley, New York (1998).

[9] Y.N. Obukhov and F.W. Hehl, Space-time metric from linear electrodynamics. *Phys. Lett.* **B458** (1999) 466-470.

[10] E.J. Post, *Formal Structure of Electromagnetics* – General Covariance and Electromagnetics. North Holland, Amsterdam (1962) and Dover, Mineola, New York (1997).

[11] R.A. Puntigam, C. Lämmerzahl and F.W. Hehl, Maxwell's theory on a post-Riemannian spacetime and the equivalence principle. *Class. Quantum Grav.* **14** (1997) 1347-1356.

[12] W. Raith, *Bergmann–Schaefer, Lehrbuch der Experimentalphysik, Vol.2, Elektromagnetismus*, 8th ed.. de Gruyter, Berlin (1999).

[13] M. Schönberg, Electromagnetism and gravitation. *Rivista Brasileira de Fisica* **1** (1971) 91-122.

[14] J.A. Schouten, *Tensor Analysis for Physicists*. 2nd ed.. Dover, Mineola, New York (1989).

[15] E. Schrödinger, *Space-Time Structure*. Cambridge University Press, Cambridge (1954).

[16] A. Sommerfeld, *Elektrodynamik.* Vorlesungen über Theoretische Physik, Band 3. Dietrisch'sche Verlagsbuchhandlung, Wiesbaden (1948).

[17] M. Tinkham, *Introduction to Superconductivity*. 2nd ed.. McGraw-Hill, New York (1996).

[18] M. Toussaint, A gauge theoretical view of the charge concept in Einstein gravity. Gen. Relat. Grav. J., to appear 1999/2000. Los Alamos Eprint Archive gr-qc/9907024.

[19] C. Truesdell and R.A. Toupin, The Classical Field Theories. In *Handbuch der Physik,* Vol. III/1, S. Flügge ed.. Springer, Berlin (1960) pp. 226-793.

[20] M.R. Zirnbauer, *Elektrodynamik.* Tex-script of a course in Theoretical Physics II (in German), July 1998 (Springer, Berlin, to be published).

Grußwort des Vereins der "Freunde des Studiengangs Vermessungswesen der Universität Stuttgart e.V. (FVUS)"

Alfred Hils

Lieber, sehr geehrter Jubilar, meine sehr verehrten Damen und Herren, Wenn ich als Vertreter des Vereins der Freunde des Studiengangs Geodäsie und Geoinformatik der Universität Stuttgart zu Wort komme, darf ich darauf hinweisen, daß es vor 5 Jahren Herr Professor Grafarend war, der auf die Idee kam unseren Verein an der Universität Stuttgart ins Leben zu rufen.
Die Gründungsversammlung erfolgte dann 1995. Der Zweck des Vereins ist laut Satzung die Förderung der wissenschaftlichen Aus- und Weiterbildung sowie die fachliche Kontaktpflege mit dem Studiengang, Geodäsie und Geoinformatik. Der Satzungszweck wird verwirklicht indem durch Bereitstellung von Geldmitteln insbesondere Fachexkursionen der Studierenden und Vorträge im Rahmen der Geodätischen Kolloquien sowie Maßnahmen der beruflichen Fortbildung unterstützt werden.
Mitglieder unseres Vereins können alle jetzigen und ehemaligen Angehörige der Universität sowie Absolventen und Freunde unseres Studiengangs werden. Neben natürlichen Personen steht auch für juristische Personen die Mitgliedschaft offen.
Unser Verein besteht derzeit aus 88 natürlichen und 5 juristischen Personen.
Die Mitglieder stellen einen Querschnitt aus in der Lehre und in der Praxis tätigen Berufsträgern dar.
Professoren sowie Kollegen aus allen Sparten des Vermessungswesens, beamtete und angestellte Kollegen aus dem öffentlichen Dienst, aus der Landes- und Katastervermessung, der Flurneuordnung, von Sonderbehörden und Freiberufler öffentlich bestellte Vermessungsingenieure, Inhaber und Mitarbeiter von Ingenieurbüros finden sich bei uns.
Unser Verein ist sozusagen das Bindeglied zwischen Wissenschaft und Praxis und Sie, verehrter Herr Prof. Grafarend sind der geistige Vater unseres Zusammenschlusses.
Und nun zu Ihrem Geburtstag. Den Tag, der uns die Götter einmal nur im Leben gewähren können, feiere jeder hoch, meinte Goethe. Aber wir können auch fragen: Warum feiert man eigentlich seinen Geburtstag, mit welcher Berechtigung? Daß man geboren ist, dafür kann man nichts, und es gibt auch keinen Grund stolz darauf zu sein. Schließlich haben die Eltern diese Leistung zuwege gebracht. Noch weniger darf man sich die Tatsache als Verdienst anrechnen, daß man älter geworden ist. Das ist eine Leistung der Natur, für die man allenfalls Gott danken kann, an der man aber selber keinen Anteil hat.
So gesehen sind Geburtstagsfeiern unberechtigt. Betrachten wir die Angelegenheit einmal aus einem anderen Blickwinkel. Nehmen wir das Leben als ein wunderbares Geschenk. Jedes Jahr, das es länger währt, gibt uns neue Chancen, uns zu verwirklichen und Erfüllung zu finden. In der Feier unseres Geburtstags können wir die Freude darüber zum Ausdruck bringen. Ist es uns gelungen, nach unseren Vorstellungen zu leben, sind unsere Wünsche in Erfüllung gegangen, dann können wir unseren Geburtstag als eine Art Erntedankfest feiern.
Ich meine, so betrachtet haben wir allen Grund zum Feiern. Noch dazu geht es heute um einen sogenannten runden Geburtstag, der ja einen besonderen Stellenwert hat. Halten wir es also mit Goethe und geniessen wir diesen Augenblick. Herr Prof. Grafarend am 30. Oktober 1999, dem Tag Ihres Geburtstags, leben Sie 21915 Tage auf dieser schönen Erde. Sie sind geboren

im achten Zeichen des Tierkreises, als Skorpion. Die im Zeichen des Skorpion Geborenen haben eine tiefgründige und starke Persönlichkeit und große Reserven an emotionaler und physischer Energie. Der Skorpion ist zweifellos das verwirrendste und vielleicht auch am wenigsten verstandene aller Tierkreiszeichen. Die Skorpione selbst tragen wenig zur Lösung des Rätsels bei, weil es Ihnen gefällt, sich rätselhaft und geheimnisvoll zu geben. Er ist leicht verletzlich, aber sympathisch und mitfühlend, oft unendlich einsam.
Lieber Herr Grafarend, ich hoffe daß es mir gelungen ist Sie abseits Ihrer Stellung in der geodätischen Wissenschaft mit ein paar Federstrichen zu skizzieren.
Die Freunde des Studiengangs Geodäsie und Geoinformatik gratulieren Ihnen von Herzen zur Vollendung des 60. Geburtstags und wünschen Ihnen persönlich alles Gute und beruflich weiterhin viel Erfolg. Bleiben Sie so wie Sie sind, versuchen Sie nicht sich zu ändern. Aber ich bin überzeugt, daß Sie das sowieso nicht vorhaben. Und das ist gut so.

Erik W. Grafarend – Ist Größe messbar?

Bernhard Hofmann-Wellenhof

Vorgeschichte

Mit diesem Beitrag darf ich Erik W. Grafarend zum Geburtstag gratulieren. Ich nehme davon Abstand, in dieser Festschrift den Versuch zu starten, einen wissenschaftlichen Beitrag zu liefern, der einen Bezug zum universalen Arbeitsgebiet des nunmehr 60 Jahre jungen Geburtstagskinds herstellt. Ich möchte vielmehr einige Erinnerungen herausgreifen, die Grafarend in meine Lebenslinie eingeprägt hat. Und ich wähle die deutsche Sprache, weil mir die Erinnerungen in der Muttersprache besser aus der Feder fließen.

Die erste Begegnung

Märchen beginnen oft mit "Es war einmal...". Diese Geschichte beginnt auch mit "Es war einmal...", aber sie ist ein wahrgewordenes Märchen. Daher spielen Träume, Sehnsüchte und Phantasien die Hauptrolle, auch wenn es um die zentrale Frage "Ist Größe messbar?" geht, die durchaus geodätischen Ursprungs sein könnte.
Es war einmal vor ungefähr fünfundzwanzig Jahren. Als Student im höheren Semester hatte mich ein Vortragstitel oder der Name eines mir unbekannten Vortragenden in den Seminarraum gelockt, der zum Hochwissenschaftsbereich von Professor Helmut Moritz gehörte. Und herein trat ein junger Mann, der mich vom ersten Augenblick an faszinierte. Mich beeindruckten die Sicherheit, das Auftreten und die Fülle an Information, die in glänzender rhetorischer Akrobatik dem Publikum präsentiert wurde. Damals prägte ich mir den Namen Erik W. Grafarend ein! Ein kleines Detail von Grafarends "Erstauftritt" in meinem Leben möchte ich noch hinzufügen. Nach dem Vortrag kam es zu einer Diskussion mit dem Publikum. Auch hier kann ich mich weder an den Inhalt, dessen Sinn ich vermutlich kaum verstand, erinnern. Aber mir ist eine winzige Kleinigkeit unverrückbar ins Gedächtnis eingeschrieben.
Bevor ich sie preisgebe, sollte ich noch einen kleinen Abstecher in die Zeit vor fünfundzwanzig Jahren machen. Damals war es für Abiturienten keinesfalls selbstverständlich, Englisch als Konversationssprache einzusetzen, da – je nach Schulzweig – den klassischen Sprachen Latein und Griechisch der Vorzug gegenüber den lebenden Sprachen gegeben wurde. Mit meinen Englisch-Kenntnissen nach vier Jahren Mittelschule konnte ich zwar das parlamentarische System und die Historie des Commonwealth beschreiben, der Umgangssprache war ich aber bei weitem nicht mächtig.
Zurück zur Diskussion nach dem Vortrag. Ein Zuhörer, der als Gaststudent am Institut arbeitete, fragte, ob er eine Frage auch auf Englisch stellen dürfe, da seine Sprachkenntnisse aus Deutsch sehr gering seien. Erik W. Grafarend schaute den Fragenden an, nickte fast unmerklich mehrmals bejahend mit dem Kopf und sagte knapp: "Go ahead!" Dieses so sichere Auftreten, das eine völlige Beherrschung der Materie verriet (die ja schon damals gegeben war), bleibt mit unvergessen; mehr noch: ich versuchte, diesen Eindruck in die Leitmotive meines Lebens zu übernehmen. Erik W. Grafarend – ein Name hatte für mich Gestalt angenommen.

Varianzen, Kovarianzen und Varianz-Kovarianz-Matrizen

Nach dieser ersten Begegnung dauerte es eine Weile, ehe ich Erik W. Grafarend wiedersehen sollte, da ich noch einige Jahre benötigte, bevor ich mich selbst auf dem internationalen wissenschaftlichen Parkett zu bewegen versuchte. Und wieder möchte ich eine Kleinigkeit herausgreifen, die aber, wie ich glaube, charakteristisch für Grafarend ist. Wir schreiben irgendein Jahr, sind bei irgendeiner geodätischen Tagung und erleben gerade ein Grafarendsches Feuerwerk von Vortrag. Damals standen die Varianz-Kovarianz-Matrizen in der geodätischen Hitparade unangefochten auf Platz Eins. Die meisten Vortragenden kürzten den langwierigen Namen ab und sprachen – mathematisch sicherlich nicht ganz sauber – von Varianz oder Kovarianz; nicht jedoch Erik W. Grafarend. Penibel führte er in seinem Vortrag jedesmal die ganze Länge des Wortes an: Varianz-Kovarianz-Matrix. Der Begriff kam oft vor, aber Grafarend hielt die wissenschaftliche Akkuratesse mühelos durch.
Erinnerungen als kleine Mosaiksteinchen, die Kleinigkeiten im Leben Grafarends darstellen, die aber die Frage aufwerfen: Ist Größe messbar?

Erice, wo Träume niemals enden

Eine meiner schönsten Erinnerungen an Erik W. Grafarend führt mich nach Erice zu einer geodätischen Sommerschule. Ich bin leider nur durch Erzählungen anwesend, aber auch diese Anwesenheit lohnt sich. Nach einem lehrreichen Tag schlendere ich durch die pittoreske Altstadt. Ein heißer Tag neigt sich dem Ende zu. Die abendliche sizilianische Sonne wirft ein wunderbares Licht auf den verzauberten Ort, der Gestalt gewordene Sage ist. Die Elymerstadt hoch über dem Meer und dem Strand leitet ihren Namen von Eryx, Sohn eines Argonauten und der Aphrodite, ab. Der Elymerkönig ehrte seine göttliche Mutter mit einem Heiligtum, dem er seinen Namen gab. Im Wappen der 2000-Seelenstadt sieht man zwei Bergkuppen, über die eine Taube mit Frauenkopf fliegt, nämlich Aphrodite, die aus den Lüften ihr Heiligtum ansteuert.
Erice: ich schreite durch die engen Gassen, vorbei an alten, rauhen Mauern, stillen Häusern und halb arabischen, halb klösterlichen Höfen, wo die restlichen Bewohner Blumen in Töpfen züchten und Frauen bunte Teppiche nach alten Mustern weben. In dieses zeitlose Bild dringen plötzlich ferne Gitarrenklänge. Ich folge den Klängen und fühle mich in die Zeit des Minnesangs zurückversetzt. In einem Fensterrahmen sitzt eine dunkel-gekleidete Gestalt. Es ist Erik W. Grafarend. Sein Blick ist in die Ferne gerichtet, dorthin, wo sich an sizilianischen Gestaden die nimmermüden Wellen brechen. Ein Knie hat er wie spielerisch angezogen, die Gitarre, die den Klang in die engen Gassen von Erice zauberte, bildet mit dem Körper eine Einheit. Die Seele verschmilzt mit dem Klang zu einem Traum – Erice, wo Träume niemals enden. Und doch: Der Traum zerfließt – es war alles nur eine phantasievolle Vorstellung, die sich auf eine Erzählung aufbaute.
Einige Jahre später bin ich wirklich in Erice. Die Vormittagssonne steht schon hoch am Himmel und taucht die Mauern und Häuser in ein gleißendes Licht. Erinnerungen an ferne Klänge steigen hoch. Ich bin auf der Suche nach dem verlorenen Klang. Aber kein Gitarrenspiel füllt die Stadt. Bilder kommen und gehen. Und irgendwann tauchen sie wieder auf, die Bilder der Vergangenheit mit einer Gestalt am Fenster, den Blick in die Ferne gerichtet und dem Spiel der Gitarre – Erice, wo Träume niemals enden!

Will der Herr Graf ein Tänzlein wohl wagen?

Bei der Generalversammlung der IUGG in Wien, die Hans Sünkel organisierte, war auch ich an der Programmarbeit beteiligt und deshalb in regem Briefverkehr mit allen Präsidenten, Generalsekretären und Symposiumsverantwortlichen. Infolge meines doch langen Doppelnamens habe ich es mir zur Gewohnheit gemacht, bei der Unterschrift stets den Vornamen abzukürzen, also

lautet die Signatur B. Hofmann-Wellenhof. Gerade dieses "B." gab besonders einem Symposiumsverantwortlichen den Anstoß, sich für den Hintergrund zu interessieren. Als wir uns in Wien das erste Mal begegneten, war seine dringendste Frage, welchen Namen denn dieses B repräsentiere.
Ähnlich ging es mir mit dem Grafarendschen W, über dessen Hintergrund ich mir lange den Kopf zerbrach, da ich ohne Nachfragen auf die Lösung kommen wollte, die zu Erik ein passendes Pendant darstellte. Bei diesen gedanklichen Nachforschungen stufte ich einmal Grafarends Eltern als Anhänger der Musik von Richard Wagner ein und vermutete in Erik eine Anlehnung an den Fliegenden Holländer. Aus anderen Wagner-Opern hätte sich mühelos das W komponieren lassen: Wolfram von Eschenbach und Walther von der Vogelweide (Tannhäuser), Walther von Stolzing (Die Meistersinger von Nürnberg), Wotan (Der Ring des Nibelungen). Aber ich verwarf diese Varianten schließlich, da ich den Jäger Erik, der sich vergeblich um die romantisch schwärmende Senta bemüht, doch eher als Randfigur ohne jede Hoffnung auf Erfolg einstufe. Der Kontrast zum geodätischen Erik schien mir einfach zu groß.
Irgendwann wurde mir das W-Rätsel entschleiert, aber der geneigte Leser möge sich, falls ihm mein ehemaliges Rätsel immer noch Rätsel ist, selbst Gedankengebilde konstruieren, die nicht notwendigerweise in die Welt der Oper führen müssen.
Zum Namen Grafarend gibt es noch eine kleine Anekdote. Im Kreise einer Gesprächsrunde wurde irgendwann die Frage aufgeworfen, ob man die erste oder die zweite Silbe betone, also Gráfarend oder Grafárend zu sagen habe. Eine humoristische Stimme meinte, es wäre ihm wohl am liebsten, wenn man beide Silben betone, das klinge dann wie Graf Arend. Wer Erik W. Grafarend kennt, könnte sich durchaus auch Graf Arend vorstellen. Die Geburtstagsfeier könnte musikalisch leicht mit einer Arie aus Wolfgang Amadeus Mozarts Hochzeit des Figaro untermalt werden: Will der Herr Graf ein Tänzlein wohl wagen? Mag er's mir sagen, ich spiele ihm auf!

Lehr- und Wanderjahre

Unsere Begegnungen häuften sich. Bei zahlreichen Tagungen gab es immer wieder Gespräche, die für mich stets lehrreich waren. Ich erinnere mich an ein Symposium in Peking. An der äußersten Grenze meines Leistungsniveaus präsentierte ich eine Arbeit über Relativitätstheorie. Diskussionsbeiträge zu derartigen Themen bleiben eher eine Seltenheit, aber Erik W. Grafarend hatte eine Idee, einen komplizierten Formelapparat einfacher und doch allgemeiner darzustellen. In der Pause erklärte er mir seine Idee. Ich weiß heute noch nicht genau, ob ich alles verstanden habe, aber es freute mich, mit einem großen Geodäten wissenschaftlich diskutieren zu dürfen.
Über eine weitere Steigerung unserer wissenschaftlichen Beziehung kann ich noch berichten. Wir standen uns gegenüber und diskutierten über die Stuttgarter geodätische Schule. Ich führte an, wie wichtig es wäre, zum besseren Verständnis die Stuttgarter Denkweise und Nomenklatur einmal im Detail zu studieren. Daraufhin lud mich Erik W. Grafarend ein, nach Stuttgart zu kommen. In meiner Phantasie öffnete sich ein Tor zum geodätischen Olymp. Aber man braucht wohl auch für ein glückliches Leben Sehnsüchte und Wünsche, die sich nicht erfüllen. Die Zeit meiner Wanderjahre war schon im Ausklingen – ich nahm das Angebot nicht an, denn ich wollte nicht die Jugend meiner Kinder versäumen.

Ausklang – Geodesia, quo vadis?

Vermutlich kommen meine doch eher lyrischen Geburtstagswünsche in eine Festschrift, in der es rundum von komplizierten Formeln wimmelt. Manche Wissenschafter behaupten sogar, eine Publikation, die nicht zumindest ein Integral enthalte, könne keine wissenschaftliche Tiefe aufweisen. Um diesem Dilemma zu entrinnen, möchte ich meine Gedanken mit einer Formel

ausklingen lassen, die der Forderung nach einem Integral entspricht:

$$\int \text{Geodäsie} = \text{Erik W. Grafarend}\,.$$

Die Grenzen des Integrals lasse ich offen, denn wenn ich sie von minus Unendlich bis plus Unendlich anführte, dann gäbe es für den Jubilar keinen Anlass mehr, geodätisch weiterzuwirken. Und dann bekäme das Motto der Festveranstaltung eine düstere Klangfarbe: Geodesia, quo vadis?
Aber ein Ausklang zu einer Geburtstagsfeier darf nicht düster sein. "O Freunde, nicht diese Töne! Sondern lasst uns angenehmere anstimmen und freudenvollere!" (Aus "Ode an die Freude" von Friedrich von Schiller). Werfen wir daher noch einmal die Frage nach den integralen Grenzen auf. Sind diese Grenzen überhaupt definierbar, wenn die Gleichheit als Ergebnis Erik W. Grafarend ergeben muss? Ist Größe messbar?

Energiebetrachtungen für die Bewegung zweier Satelliten im Gravitationsfeld der Erde

Karl Heinz Ilk

Einleitung

Physikalische Grundlage der Ausmessung des Gravitationsfeldes der Erde sind die Feld- und Bilanzgleichungen der am Meßprozeß beteiligten physikalischen Systeme. Die Bilanzgleichungen beschreiben die Gesetzmäßigkeiten bei der Wechselwirkung der physikalischen Systeme. Unter Wechselwirkung wird dabei der Austausch dynamischer Größen, wie Energie, Impuls, Drehimpuls, usw. verstanden. Jeder Meßapparat stellt ein physikalisches System dar, in dem der Austausch dynamischer Größen in kontrollierter, d.h. bekannter und geeichter Form abläuft. Sind die Träger des Transports von Energie und Impuls durch materielle Körper modellierbar, dann sind sie i.a. einer kinematischen Beobachtung zugänglich. Die Wechselwirkung kann in diesem Fall durch kinematische Variable beschrieben werden und es können Bestimmungsgleichungen für die Parameter des Gravitationsfeldes formuliert werden. Künstliche Satelliten im Gravitationsfeld der Erde repräsentieren einen solchen "Meßapparat". Die klassische Methode der Satellitengeodäsie beruht auf der Analyse kinematischer Effekte akkumulierter Bahnstörungen zahlreicher Satelliten unterschiedlicher Bahnneigung und Flughöhe. Das hat zur Folge, daß sich der Meßprozeß über ein gewisses Zeitintervall erstreckt. Sowohl Feldparameter als auch zuzuordnende Zeitpunkte lassen sich nicht in lokalisierter Form angeben. Zukünftige satellitengestützte Meßmethoden beruhen dagegen auf einer lokalen Ausmessung des Gravitationsfeldes und zwar hinsichtlich des Orts- als auch des Zeitbereiches. Zur lokalen Ausmessung bieten sich insbesondere Energieaustauschbeziehungen des Sensors mit dem Gravitationsfeld an. Das Ziel der vorliegenden Arbeit ist, die Bilanzgleichungen für den Energieaustausch zwischen den am Meßprozeß beteiligten Körpern und dem Gravitationsfeld sowie Energie- und Bewegungsintegrale zu formulieren. Die abgeleiteten Beziehungen lassen sich sowohl zur Bestimmung von Parametern des Gravitationsfeldes verwenden, als auch zur Analyse und Validierung der Lösungen (z.B. Jekeli, 1998).
Für die vorliegende Problemstellung wird eine Beschreibung der Mechanik zugrunde gelegt, die auf einen Vorschlag von Falk (1966), bzw. Falk und Ruppel (1973, 1976) zurückgeht. Sie ist in das Gebäude der Thermodynamik integriert und unterscheidet sich in gewissen Details von der gewohnten Betrachtungsweise. Natürlich handelt es sich hierbei um keine neue Formulierung der klassischen Mechanik, Falk zeigt vielmehr, daß beispielsweise die gewohnte Hamilton-Theorie als thermodynamische Beschreibung der Mechanik angesehen werden kann (Falk, 1990). Diese Beschreibungsweise kann im Falle der Betrachtung von Energie- und Energieaustauschbeziehungen zwischen den Teilen eines mechanischen Systems vorteilhaft angewendet werden.

1 Die dynamische Formulierung des Problems

Ein ***physikalisches System***, bestehend aus Körpern und Feldern, kann durch den Austausch von dynamischen Größen wie Energie, linearer Impuls, Drehimpuls, etc. mit anderen physikalischen Systemen beschrieben werden. Sind alle austauschbaren Größen bekannt, dann ist das dynamische Gesamtsystem vollständig beschrieben, ebenso wie die Prozesse, an denen das physikalische System teilnimmt. Eine wichtige Austauschgröße ist die Energie, die in verschiedenen Formen ausgetauscht wird. Jede Energieform definiert ein Paar zueinander konjugierter Variabler, die „intensiven“ und die „extensiven“ Größen. Der Energieaustausch geschieht dabei durch die Änderung der extensiven Variab-

len. Das Produkt aus einer intensiven Größe und dem Differential einer extensiven Größe ergibt die Energieform.
Physikalische Prozesse können durch Angabe aller unabhängiger Energieformen $dE_j = \xi_j \, dX_j$ beschrieben werden, in denen das System Energie austauschen kann. Die Änderung dE der Energie E des Gesamtsystems wird durch die folgende Pfaffsche Form, die sog. ***Gibbssche Fundamentalform*** des Systems, dargestellt:

$$dE = \sum_{j=1}^{n} \xi_j \, dX_j \tag{1.1}$$

Die physikalischen Systeme selbst sind durch die funktionale Abhängigkeit der Energie E von allen extensiven Größen X_j beschrieben. Die Funktion $E(X_1, X_2, ..., X_n)$, die die Abhängigkeiten von den extensiven Größen beschreibt, wird als eine ***Gibbssche Funktion*** des Systems bezeichnet. Die intensiven Variablen sind durch die partiellen Ableitungen der Gibbsschen Funktion nach den konjugierten extensiven Variablen bestimmt:

$$\xi_j = \frac{\partial E(X_1, X_2, ..., X_n)}{\partial X_j} \tag{1.2}$$

Kinematische Bewegung ist Ortsveränderung geometrischer Punktkonfigurationen. Sie kann durch die zeitlichen Veränderungen der Ortsvektoren $\mathbf{r}(t)$ der geometrischen Punkte beschrieben werden. Die Ableitungen der Ortsvektoren nach der Zeit definieren die kinematischen Geschwindigkeiten $d\mathbf{r} / dt$. Im Falle einer starren Punktkonfiguration kann die kinematische Bewegung auch durch den Ortsvektor des Massenzentrums und den Orientierungsvektor $\varphi(t)$ der Punktkonfiguration beschrieben werden. Damit kann die kinematische Bewegung der einzelnen Punkte des starren Körpers beschrieben werden. Kinematische Bewegung ist also mit dem Begriff der Bahn eines geometrischen Punktes eng verknüpft.
Von der kinematischen Bewegung muß die ***dynamische Bewegung*** unterschieden werden. Unter dynamischer Bewegung versteht man den Transport von Energie und Impuls. Während die kinematische Geschwindigkeit die Bahngeschwindigkeit eines materiellen Punktes angibt, gibt die dynamische Geschwindigkeit $\mathbf{v}$ die Geschwindigkeit des Energie- und Impulstransportes an. Besteht der Energie-Impuls-Transport in der Bewegung eines geometrisch lokalisierbaren Körpers, beispielsweise eines Massenpunktes, so sind dynamische und kinematische Geschwindigkeiten gleich: $\mathbf{v} = d\mathbf{r} / dt$. Man beachte aber, daß es kinematische Bewegungen gibt, die keine dynamischen sind und umgekehrt.
Zur kinematischen Beschreibung der Bewegung des vorliegenden physikalischen Systems sind geeignete Modelle einzuführen. Während zur mathematischen Beschreibung des Systems „Gravitationsfeld“ eine Feldfunktion verwendet wird, eignet sich zur Beschreibung des starren Körpers „Erde“ das mathematische Bild einer starren Punktkonfiguration und für die Satelliten die Bilder von geometrischen Punkten. Kinematische Bewegung wird damit als kontinuierliche Aufeinanderfolge von räumlich geometrischen Punktkonfigurationen beschrieben. Beschrieben wird diese Aufeinanderfolge durch die Ortsvektoren zum Massenzentrum bzw. zu den die übrigen Körper repräsentierenden Punkten durch die kinematischen Geschwindigkeiten $d\mathbf{r}_j / dt$ bzw. durch die auf das Massenzentrum bezogene kinematische Winkelgeschwindigkeit $d\varphi_j / dt$.

In der vorliegenden Darstellung wird der Auffassung von Falk und Ruppel (1973, 1976) gefolgt und angenommen, daß das ***Gravitationsfeld*** ein eigenes physikalisches Gebilde ist, das nicht von den gravitierenden Körpern erzeugt wird, sondern von ihnen lediglich in seinem Zustand geändert werden kann. Diese Zustandsänderungen werden durch translatorische und rotatorische Verschiebungen der Körper im Feld herbeigeführt. Die Zustandsänderungen des Feldes beschränken sich dabei nur auf den Energieinhalt des Gravitationsfeldes (in der hier angenommenen statischen Näherung). Das Gravitationsfeld vermittelt zwar den Impuls- und Drehimpulsaustausch sowie den Energieaustausch zwischen den Körpern, behält aber von den dabei aufgenommenen und transportierten Impulsen und Drehimpulsen nichts zurück, wogegen es von der Energie einen Teil zurückbehalten kann (Falk und Ruppel, 1973, S.194). Sind Gesamtimpuls, Gesamtdrehimpuls und Gesamtenergie eines beliebigen (abgeschlossenen) Systems von wechselwirkenden Körpern und Gravitationsfeld konstant, so ist das zugrunde liegende Bezugssystem ein Inertialsystem. Ist das nicht der Fall, so ist ein weiteres Gebilde

beteiligt, mit dem das System wechselwirkt, d.h. mit dem es Impuls, Drehimpuls und Energie austauscht. Dieses weitere System ist das ***Trägheitsfeld***. Das Trägheitsfeld wechselwirkt mit jedem Gebilde, das Energie und Impuls besitzt.

2 Das System „Erde+Satellit-1+Satellit-2+Gravitationsfeld“

Die Gibbssche Funktion für das abgeschlossene physikalische System „Erde+Satellit-1+Satellit-2+ Gravitationsfeld“ kann in Abhängigkeit von den extensiven Variablen folgendermaßen angeschrieben werden (Ilk, 1983a,b; Abb. 3.1):

$$E(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes,\mathbf{r}_\otimes,\mathbf{r}_1,\mathbf{r}_2,\varphi_\otimes) = T(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes) + \hat{V}(\mathbf{r}_j, j=\otimes,1,2;\varphi_\otimes) + E_0 \ . \tag{2.1}$$

E_0 ist die innere Energie des Gesamtsystems und $T(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes)$ die kinetische Energie der beteiligten Körper

$$T(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes) = \frac{1}{2}\frac{\mathbf{p}_\otimes^2}{M_\otimes} + \frac{1}{2}\frac{\mathbf{p}_1^2}{M_1} + \frac{1}{2}\frac{\mathbf{p}_2^2}{M_2} + \frac{1}{2}\mathbf{L}_\otimes\cdot\mathbf{T}_\otimes^{-1}\cdot\mathbf{L}_\otimes \quad , \tag{2.2}$$

mit den Massen M_j der Körper und dem Trägheitstensor $\mathbf{T}_\otimes$ der Erde. Die kinetische Energie setzt sich aus translatorischen und rotatorischen Anteilen zusammen. Das System „Gravitationsfeld“ kann als räumlich ausgedehntes Gebilde betrachtet werden, das sich dadurch äußert, daß in jedem Raumpunkt Kräfte auf Massenpunkte wirken. Als geeignetes mathematisches Darstellungsmittel wird eine Funktion $\hat{V}(\mathbf{r}_j, j=\otimes,1,2;\varphi_\otimes)$ gewählt, die die potentielle Energie der Gravitationswechselwirkung aller beteiligten Körper beschreibt. Sie lautet für den vorliegenden Fall (Abb. 3.1):

$$\hat{V}(\mathbf{r}_j, j=\otimes,1,2;\varphi_\otimes) = \hat{V}_{\otimes 1}(\mathbf{r}_\otimes,\mathbf{r}_1,\varphi_\otimes) + \hat{V}_{\otimes 2}(\mathbf{r}_\otimes,\mathbf{r}_2,\varphi_\otimes) + \hat{V}_{12}(\mathbf{r}_1,\mathbf{r}_2) \ . \tag{2.3}$$

Dem betrachteten physikalischen System „Erde+Satellit-1+Satellit-2+ Gravitationsfeld“ kann Energie durch Änderung der Bewegungsenergie über die Änderung der linearen Impulse $\mathbf{v}_j\cdot d\mathbf{p}$, durch Änderung der Rotationsenergie über die Änderung der Drehimpulse $\mathbf{d}_j\cdot d\mathbf{L}_j$, sowie durch Änderung der Verschiebungsenergie über eine Translation $-\mathbf{K}_j\cdot d\mathbf{r}_j$ bzw. Orientierungsänderung $-\mathbf{M}_j\cdot d\varphi_j$ der beteiligten Körper zugeführt oder entzogen werden. Die in den Energieformen auftretenden Größen sind die Geschwindigkeiten der Massenzentren $\mathbf{v}_j$, die Änderungen der linearen Impulse $d\mathbf{p}_j$, die Winkelgeschwindigkeit der Erde $\mathbf{d}_\otimes$, die Änderung des Drehimpulses der Erde $d\mathbf{L}_\otimes$, die Gravitationswechselwirkungskräfte $\mathbf{K}_j$, die Lageänderungen $d\mathbf{r}_j$, das Gravitationswechselwirkungsdrehmoment $\mathbf{M}_\otimes$ und die Orientierungsänderung $d\varphi_\otimes$.
Die intensiven Variablen ergeben sich durch partielle Ableitungen der Gibbsschen Funktion nach den konjugierten extensiven Variablen. Für die dynamischen Geschwindigkeiten erhält man:

$$\mathbf{v}_j = \frac{\partial E(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes,\varphi_\otimes,\mathbf{r}_\otimes,\mathbf{r}_1,\mathbf{r}_2)}{\partial\mathbf{p}_j} = \frac{\mathbf{p}_j}{M_j}, \quad j=\otimes,1,2 \ , \tag{2.4}$$

entsprechend für die dynamische Winkelgeschwindigkeit der Erde

$$\mathbf{d}_\otimes = \frac{\partial E(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes,\varphi_\otimes,\mathbf{r}_\otimes,\mathbf{r}_1,\mathbf{r}_2)}{\partial\mathbf{L}_\otimes} = \mathbf{T}_\otimes^{-1}\cdot\mathbf{L}_\otimes \ , \tag{2.5}$$

für die Kräfte zufolge Gravitationswechselwirkung:

$$-\mathbf{K}_j = \frac{\partial E(\mathbf{p}_\otimes,\mathbf{p}_1,\mathbf{p}_2,\mathbf{L}_\otimes,\varphi_\otimes,\mathbf{r}_\otimes,\mathbf{r}_1,\mathbf{r}_2)}{\partial\mathbf{r}_j} = \nabla_{\mathbf{r}_j}\hat{V} = \sum_{\substack{k\\k\neq j}}\nabla_{\mathbf{r}_j}\hat{V}_{jk} = -\sum_{\substack{k\\k\neq j}}\mathbf{K}_{jk}, \quad j,k=\otimes,1,2, \tag{2.6}$$

sowie für das (in diesem Fall zu vernachlässigende) Drehmoment zufolge Gravitationswechselwirkung mit den Satelliten:

$$-\mathbf{M}_{\otimes} = \frac{\partial E(\mathbf{p}_{\otimes}, \mathbf{p}_1, \mathbf{p}_2, \mathbf{L}_{\otimes}, \varphi_{\otimes}, \mathbf{r}_{\otimes}, \mathbf{r}_1, \mathbf{r}_2)}{\partial \varphi_{\otimes}} = \mathbf{x} \times \nabla_{\mathbf{x}} \hat{V} = \mathbf{x} \times \nabla_{\mathbf{x}} \sum_{k=1}^{2} \hat{V}_{\otimes k} = -\sum_{k=1}^{2} \mathbf{M}_{\otimes k} \ . \tag{2.7}$$

Der Verschiebungsoperator $\nabla_{\mathbf{r}_j}$ und der auf das Geozentrum bezogene Drehoperator $\mathbf{x} \times \nabla_{\mathbf{x}}$, angewendet auf die potentielle Energie der Gravitationswechselwirkung, ergeben die Änderungsraten der Verschiebungsenergie bei infinitesimaler Translation des Körpers *j* als Ganzes bzw. bei infinitesimaler Drehung um sein Massenzentrum in folgendem Sinne:

$$\nabla_{\mathbf{r}_j} \hat{V}_{jk} = \lim_{|d\varphi| \to 0} \frac{\hat{V}_{jk}(\mathbf{r}_j + d\mathbf{r}_j) - \hat{V}_{jk}(\mathbf{r}_j)}{|d\mathbf{r}_j|}, \quad \mathbf{x} \times \nabla_{\mathbf{x}} \hat{V}_{jk} = \lim_{|d\varphi| \to 0} \frac{\hat{V}_{jk}(\varphi + d\varphi) - \hat{V}_{jk}(\varphi)}{|d\varphi|} \ . \tag{2.8}$$

Die Gibbssche Fundamentalform beschreibt nun die Änderung der Energie *E* des Gesamtsystems, dargestellt durch die verschiedenen Energieformen, in denen Energie ausgetauscht werden kann,

$$dE = \mathbf{d}_{\otimes} \cdot d\mathbf{L}_{\otimes} + \mathbf{v}_{\otimes} \cdot d\mathbf{p}_{\otimes} + \mathbf{v}_1 \cdot d\mathbf{p}_1 + \mathbf{v}_2 \cdot d\mathbf{p}_2 - \mathbf{K}_{\otimes} \cdot d\mathbf{r}_{\otimes} - \mathbf{M}_{\otimes} \cdot d\varphi_{\otimes} - \mathbf{K}_1 \cdot d\mathbf{r}_1 - \mathbf{K}_2 \cdot d\mathbf{r}_2 \ . \tag{2.9}$$

Da es sich im betrachteten Fall um ein abgeschlossenes System handelt gilt: $dE = 0$. Die Zerlegung der Gesamtenergie *E* in Summanden hat zur Folge, daß in der zugehörigen Gibbsschen Fundamentalform (2.9) die Größen dT_j und $d\hat{V}_j$ (in hoher Näherung) totale Differentiale der kinetischen Energie der Körper $\otimes,1,2$ und der potentiellen Energie des Feldes sind:

$$dT_{\otimes} = \mathbf{d}_{\otimes} \cdot d\mathbf{L}_{\otimes}, \quad dT_j = \mathbf{v}_j \cdot d\mathbf{p}_j \ , \quad d\hat{V}_{\otimes} = -\sum_{k=1}^{2} \mathbf{M}_{\otimes k} \cdot \mathbf{d}_{\otimes}, \quad d\hat{V}_j = -\sum_{\substack{k \\ k \neq j}} \mathbf{K}_{jk} \cdot d\mathbf{r}_j, \quad j = \otimes,1,2 \ . \tag{2.10}$$

Ziel ist, Bilanzgleichungen für den Energieaustausch zwischen den am Meßprozeß beteiligten Körpern und dem Gravitationsfeld zu formulieren. Im Falle eines ausgedehnten starren Körpers mit beliebiger Massenverteilung wird Bewegungsenergie und (translatorische) Verschiebungsenergie bzw. Rotationsenergie und (rotatorische) Verschiebungsenergie des Feldes ausgetauscht (Falk, 1973, Ilk, 1983b). Die Erhaltungssätze für den Energieaustausch gelten für jeden Zeitpunkt während des Bewegungsablaufes; sie beziehen sich in dieser Form zunächst auf ein raumfestes Bezugssystem. Dasselbe gilt für die Gibbssche Fundamentalform des Systems, die ja nichts anderes als die Summe der einzelnen Erhaltungssätze ist. Die Gesamtenergie des Systems bzw. die Gibbsschen Funktion bleibt während des Bewegungsablaufes ebenfalls konstant.
Führt man kinematische Geschwindigkeiten und Winkelgeschwindigkeiten ein, die sich für das vorliegende System von materiellen Körpern gleich den entsprechenden dynamischen Größen erweisen (Falk und Ruppel, 1973),

$$\mathbf{v}_j = \frac{d\mathbf{r}_j}{dt} =: \dot{\mathbf{r}}_j, \quad \mathbf{d}_{\otimes} = \frac{d\varphi_{\otimes}}{dt} =: \dot{\varphi}_{\otimes} \ , \tag{2.11}$$

so erhält man die Bilanzen für den Energieaustausch. Es ergeben sich

- für den Energieaustausch bei translatorischer Bewegung der einzelnen Körper (translatorische Einzelbewegung) mit dem Gravitationsfeld:

$$\mathbf{v}_j \cdot d\mathbf{p}_j - \mathbf{K}_j \cdot d\mathbf{r}_j = 0 \ . \tag{2.12}$$

  bzw. durch entsprechende Umformung:

$$M_j \, \dot{\mathbf{r}}_j \cdot d\dot{\mathbf{r}}_j - \mathbf{K}_j \cdot d\mathbf{r}_j = 0 \ , \quad \text{mit} \quad \mathbf{K}_j = \sum_{\substack{k \\ k \neq j}} \mathbf{K}_{jk} = \sum_{\substack{k \\ k \neq j}} \nabla_{\mathbf{r}_j} \hat{V}_{jk}, \quad j,k = \otimes,1,2 \ , \tag{2.13}$$

- und für den Energieaustausch bei rotatorischer Bewegung der Erde mit dem Gravitationsfeld:

$$\mathbf{d}_{\otimes} \cdot d\mathbf{L}_{\otimes} - \mathbf{M}_{\otimes} \cdot d\mathbf{r}_{\otimes} = 0 \ . \tag{2.14}$$

  bzw. durch entsprechende Umformung:

$$\dot{\varphi}_\otimes \cdot \mathbf{T}_\otimes \cdot d\dot{\varphi}_\otimes + \dot{\varphi}_\otimes \cdot \left(d\varphi_\otimes \times \mathbf{T}_\otimes \cdot \dot{\varphi}_\otimes\right) - \mathbf{M}_\otimes \cdot d\varphi_\otimes = 0,$$
$$\text{mit} \quad \mathbf{M}_\otimes = \sum_k \mathbf{M}_{\otimes k} = \sum_k \mathbf{x} \times \nabla_{\mathbf{x}} \hat{V}_{\otimes k}, \; k = 1,2\,. \tag{2.15}$$

Selbstverständlich sind die Kraft- bzw. Drehmomentanteile, die aus dem Energieaustausch der Satellitenbewegung bzgl. der Erde resultieren, für die Erde zu vernachlässigen. Für die Erde folgt damit sowohl für die Translation als auch für die Rotation Trägheitsbewegung. Der Energieaustausch durch diese Bewegungen soll im weiteren unberücksichtigt bleiben.

3 Energieaustauschbeziehungen im Quasi-Inertialsystem

Zur Ableitung der Energieaustauschbilanzen führt man in der Gibbsschen Funktion des Gesamtsystems anstelle der linearen Impulse $\mathbf{p}_\otimes, \mathbf{p}_1, \mathbf{p}_2$ die relativen Impulse $\mathbf{p}, \mathbf{P}, \mathbf{P}_{12}$ ein und anstelle der Koordinaten $\mathbf{r}_\otimes, \mathbf{r}_1, \mathbf{r}_2$ die Relativkoordinaten (Jacobi-Koordinaten) $\mathbf{r}, \mathbf{R}, \mathbf{R}_{12}$ (Abb. 3.1). Sie ergeben sich aus den Transformationen (Ilk, 1983a):

$$\begin{pmatrix} \mathbf{p} \\ \mathbf{P} \\ \mathbf{P}_{12} \end{pmatrix} = \mathbf{A}^{-T} \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_\otimes \end{pmatrix}, \quad \begin{pmatrix} \mathbf{r} \\ \mathbf{R} \\ \mathbf{R}_{12} \end{pmatrix} = \mathbf{A} \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_\otimes \end{pmatrix}, \tag{3.1}$$

mit den Transformationsmatrizen

$$\mathbf{A} = \begin{pmatrix} \frac{M_1}{m} & \frac{M_2}{m} & \frac{M_\otimes}{m} \\ \frac{M_1}{M_1 + M_2} & \frac{M_2}{M_1 + M_2} & -1 \\ -1 & 1 & 0 \end{pmatrix}, \quad \mathbf{A}^{-T} = \begin{pmatrix} 1 & 1 & 1 \\ \frac{M_\otimes}{m} & \frac{M_\otimes}{m} & -\frac{M_1}{M_1 + M_2} \\ -\frac{M_2}{M_1 + M_2} & \frac{M_1 + M_2}{m} & 0 \end{pmatrix}, \tag{3.2}$$

wobei für die Gesamtmasse gilt $m := M_1 + M_2 + M_\otimes$. Die Gibbssche Funktion des Gesamtsystems lautet in den neuen Variablen mit der inneren Energie E_0,

$$E(\mathbf{p}, \mathbf{P}, \mathbf{P}_{12}, \mathbf{L}_\otimes, \mathbf{R}, \mathbf{R}_{12}, \varphi_\otimes) = T(\mathbf{p}, \mathbf{P}, \mathbf{P}_{12}, \mathbf{L}_\otimes) + \hat{V}(\mathbf{R}, \mathbf{R}_{12}, \varphi_\otimes) + E_0 \quad , \tag{3.3}$$

und der kinetischen Energie,

$$T(\mathbf{p}, \mathbf{P}, \mathbf{P}_{12}, \mathbf{L}_\otimes) = \frac{1}{2}\frac{\mathbf{p}^2}{m} + \frac{1}{2}\frac{\mathbf{P}^2}{M} + \frac{1}{2}\frac{\mathbf{P}_{12}^2}{\mu_{12}} + \frac{1}{2}\mathbf{L}_\otimes \cdot \mathbf{T}_\otimes^{-1} \cdot \mathbf{L}_\otimes \quad , \tag{3.4}$$

mit dem Trägheitstensor $\mathbf{T}_\otimes$ der Erde sowie den Abkürzungen M und μ_{12} für die reduzierten Massen

$$M = \frac{M_\otimes (M_1 + M_2)}{M_1 + M_2 + M_\otimes}, \quad \mu_{12} = \frac{M_1 M_2}{M_1 + M_2}. \tag{3.5}$$

Die potentielle Energie der Gravitationswechselwirkung ist nun in den neuen Variablen zu formulieren:

$$\hat{V}(\mathbf{r}_j, j = \otimes,1,2; \varphi_\otimes) = \hat{V}(\mathbf{R}, \mathbf{R}_{12}, \varphi_\otimes)\,. \tag{3.6}$$

Eine Möglichkeit der mathematischen Formulierung der potentiellen Energie ist im Abschnitt 6 in Form von Reihenentwicklungen nach Kugelfunktionen gegeben.
Der Energieaustausch des betrachteten Systems ist mit den gewählten Variablen in den folgenden Energieformen möglich:

- Änderung der Bewegungsenergie durch Änderung der linearen Impulse:

$$\mathbf{v} \cdot d\mathbf{p}, \quad \mathbf{V} \cdot d\mathbf{P} \quad , \mathbf{V}_{12} \cdot d\mathbf{P}_{12} \tag{3.7}$$

- Änderung der Verschiebungsenergie durch Translation und Orientierungsänderung der beteiligten Körper:

$$-\mathbf{K}\cdot d\mathbf{r},\quad -\mathbf{K}_{(\otimes,1,2)}\cdot d\mathbf{R},\quad -\mathbf{K}_{(12)}\cdot d\mathbf{R}_{12}, \tag{3.8}$$

Die intensiven Variablen erhält man durch partielle Ableitung der Gibbsschen Funktion nach den konjugierten extensiven Variablen. Für die dynamischen Geschwindigkeiten $\mathbf{v},\mathbf{V},\mathbf{V}_{12}$ erhält man

$$\mathbf{v}=\frac{\partial E}{\partial \mathbf{p}}=\frac{\mathbf{p}}{m},\quad \mathbf{V}=\frac{\partial E}{\partial \mathbf{P}}=\frac{\mathbf{P}}{M},\quad \mathbf{V}_{12}=\frac{\partial E}{\partial \mathbf{P}_{12}}=\frac{\mathbf{P}_{12}}{\mu_{12}}, \tag{3.9}$$

und entsprechend für die Kräfte zufolge Gravitationswechselwirkung:

$$\begin{aligned}
-\mathbf{K}&=\frac{\partial E}{\partial \mathbf{r}}=\nabla_{\mathbf{r}}\hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes})=\mathbf{0}\,,\\
-\mathbf{K}_{(\otimes,1,2)}&=\frac{\partial E}{\partial \mathbf{R}}=\nabla_{\mathbf{R}}\hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes})=-\mathbf{K}_{1\otimes}-\mathbf{K}_{2\otimes}\,,\\
-\mathbf{K}_{(12)}&=\frac{\partial E}{\partial \mathbf{R}_{12}}=\nabla_{\mathbf{R}_{12}}\hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes})=-\mathbf{K}_{21}-\mathbf{G}_{(21)\otimes}\,,\\
&\text{mit } \mathbf{G}_{(21)\otimes}=\frac{M_1}{M_1+M_2}\mathbf{K}_{2\otimes}-\frac{M_2}{M_1+M_2}\mathbf{K}_{1\otimes}
\end{aligned} \tag{3.10}$$

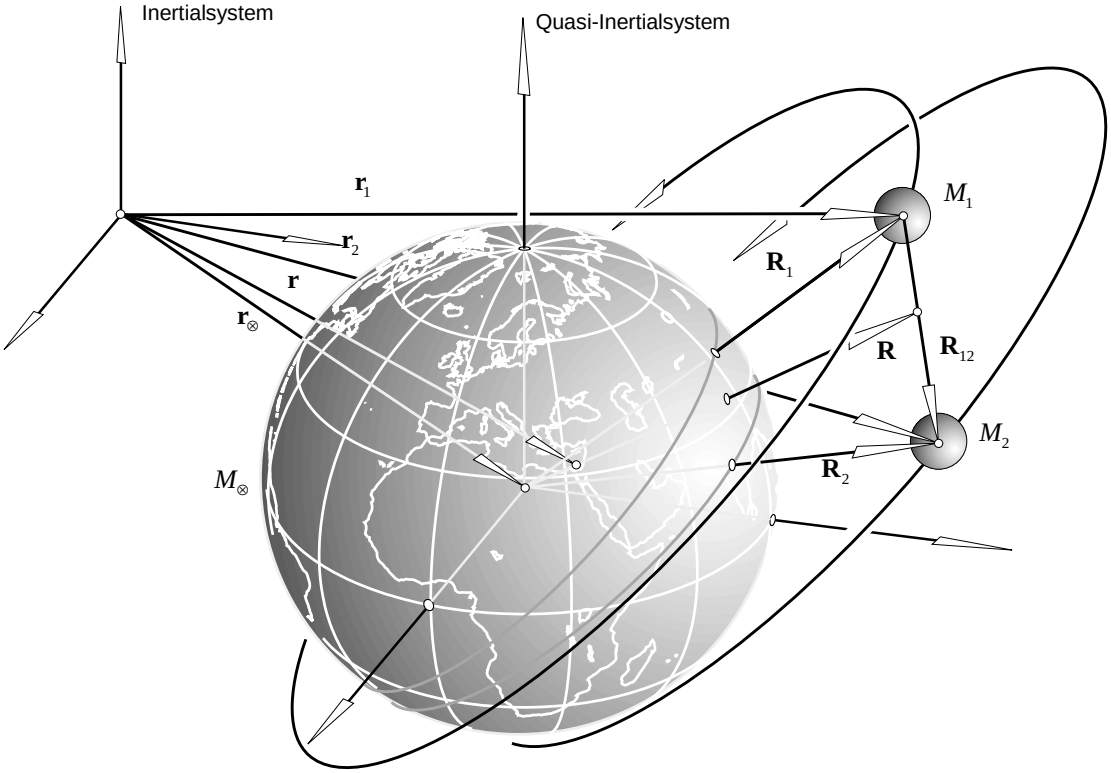


Abb. 3.1: Jacobi-Koordinaten

Die Bilanzgleichungen für den Energieaustausch des Systems $(\otimes,1,2)$ mit dem Gravitationsfeld ergeben sich,

- für die Bewegung des Gesamtsystems $(\otimes,1,2)$ bzgl. des Inertialsystems. Dabei findet kein Energieaustausch mit dem Gravitationsfeld statt. Zunächst gilt:

$$\mathbf{v}\cdot d\mathbf{p}-\mathbf{K}\cdot d\mathbf{r}=0\ , \tag{3.11}$$

bzw. nach Einführung der kinematischen Geschwindigkeit,

$$M\,\dot{\mathbf{r}}\cdot d\dot{\mathbf{r}}-\mathbf{K}\cdot d\mathbf{r}=0,\quad \text{wobei}\quad \mathbf{K}=\nabla_{\mathbf{r}}\hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes})=\mathbf{0}\quad\Rightarrow\quad M\,\dot{\mathbf{r}}\cdot d\dot{\mathbf{r}}=0\ , \tag{3.12}$$

- für die Relativbewegung des Teilsystem $(\ ,2)$ bzgl. der Erde $\otimes$:

$$\mathbf{V}\cdot d\mathbf{P}-\mathbf{K}_{(\otimes,1,2)}\cdot d\mathbf{R}=0\ , \tag{3.13}$$

bzw. nach entsprechender Umformung,

$$\left(M_1+M_2\right)\dot{\mathbf{R}}\cdot d\dot{\mathbf{R}}-\mathbf{K}_{(\otimes,1,2)}\cdot d\mathbf{R}=0\quad\text{wobei}\quad \mathbf{K}_{(\otimes,1,2)}=\nabla_{\mathbf{R}}\hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes})=\mathbf{K}_{1\otimes}+\mathbf{K}_{2\otimes}\ , \tag{3.14}$$

- für die Relativbewegung des Satellien 2 bzgl. des Satelliten 1:

$$\mathbf{V}_{12}\cdot d\mathbf{P}_{12}-\mathbf{K}_{(12)}\cdot d\mathbf{R}_{12}=0\ , \tag{3.15}$$

und nach Umformung,

$$\mu_{12}\,\dot{\mathbf{R}}_{12}\cdot d\dot{\mathbf{R}}_{12}-\mathbf{K}_{(12)}\cdot d\mathbf{R}_{12}=0\,,\quad\text{wobei}\quad \mathbf{K}_{(12)}=\nabla_{\mathbf{R}_{12}}\hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes})=\mathbf{K}_{21}+\mathbf{G}_{(21)\otimes}\ . \tag{3.16}$$

Die Relativbewegung des Satelliten 2 bzgl. des Satelliten 1 läßt sich auch als Abstandsänderung der beiden Satelliten und als Rotation der Verbindungslinie auffassen. Als Bezugssystem liege ein Quasi-Inertialsystem mit dem Ursprung im Satelliten 1 zugrunde. Zur Beschreibung der Relativbewegung von Satellit 2 bzgl. Satellit 1 werden als extensive Variable der Bahndrehimpuls $\mathbf{L}_{(12)}:=\mathbf{R}_{12}\times\mathbf{P}_{12}$, der Radialimpuls $\overline{P}_{12}:=\mathbf{e}\cdot\mathbf{P}_{12}$, der Orientierungsvektor $\varphi_{(12)}$ und der Radialabstand R_{12} eingeführt. Beachtet man, daß gilt

$$\mathbf{P}_{12}^2=\frac{\mathbf{L}_{(12)}^2}{R_{12}^2}+\overline{P}_{12}^2\ , \tag{3.17}$$

so läßt sich der Anteil "kinetische Energie" der Gibbsschen Funktion des Gesamtsystems (3.3) entsprechend (3.4) folgendermaßen angeben:

$$\begin{aligned}&T(\mathbf{p},\mathbf{P},\mathbf{P}_{12},\mathbf{L}_{\otimes})=\\ &=T(\mathbf{p},\mathbf{P},\mathbf{L}_{(12)},P_{12},\mathbf{L}_{\otimes})=\frac{1}{2}\frac{\mathbf{p}^2}{m}+\frac{1}{2}\frac{\mathbf{P}^2}{M}+\frac{1}{2}\frac{\mathbf{L}_{(12)}^2}{\mu_{12}R_{12}^2}+\frac{1}{2}\frac{\overline{P}_{12}^2}{\mu_{12}}+\frac{1}{2}\mathbf{L}_{\otimes}\cdot\mathbf{T}_{\otimes}^{-1}\cdot\mathbf{L}_{\otimes}\ ,\end{aligned} \tag{3.18}$$

Die anderen Terme der Gibbsschen Funktion in (3.3) bleiben gleich. Der Austausch von Bewegungsenergie bei der translatorischen Relativbewegung von 2 bzgl. 1 kann somit auch als Austausch von Rotationsenergie und Bewegungsenergie in Radialrichtung interpretiert werden,

$$\mathbf{V}_{12}\cdot d\mathbf{P}_{12}=\mathbf{d}_{(12)}\cdot d\mathbf{L}_{(12)}+\overline{V}_{12}d\overline{P}_{12}\ . \tag{3.19}$$

mit der dynamischen Winkelgeschwindigkeit der Verbindungslinie $\mathbf{d}_{(12)}$ und der dynamischen Geschwindigkeit in Radialrichtung $\overline{V}_{12}$. Entsprechend wird die Verschiebungsenergie des Teilsystems "Gravitationsfeld" durch die Bewegung der beiden Satelliten 1 und 2 in einen rotatorischen und einen translatorischen Anteil zerlegt:

$$\mathbf{K}_{(12)}\cdot d\mathbf{R}_{12}=\mathbf{M}_{(12)}\cdot d\varphi_{(12)}+\overline{K}_{(12)}d\overline{R}_{12}\ . \tag{3.20}$$

Man beachte, daß der Energieaustausch sowohl mit dem Gravitationsfeld als auch zwischen den beiden Komponenten der Relativbewegung stattfindet. Diese Kopplung läßt sich nicht beseitigen. Die intensiven Variablen erhält man durch partielle Ableitung der Gibbsschen Funktion mit der kinetischen Energie (3.18) nach den konjugierten extensiven Variablen $\mathbf{L}_{12}$,

$$\mathbf{d}_{(12)} = \frac{\partial E}{\partial \mathbf{L}_{(12)}} = \frac{\partial T(\mathbf{p},\mathbf{P},\mathbf{L}_{(12)},\overline{P}_{12},\mathbf{L}_{\otimes})}{\partial \mathbf{L}_{(12)}} = \frac{1}{\mu_{12} R_{12}^2}\mathbf{L}_{(12)} \ , \tag{3.21}$$

und dem Radialimpuls $\overline{P}_{12}$,

$$\overline{V}_{12} = \frac{\partial E}{\partial \overline{P}_{12}} = \frac{\partial T(\mathbf{p},\mathbf{P},\mathbf{L}_{(12)},\overline{P}_{12},\mathbf{L}_{\otimes})}{\partial \overline{P}_{12}} = \frac{\overline{P}_{12}}{\mu_{12}} \ . \tag{3.22}$$

Das Drehmoment $\mathbf{M}_{(12)}$ erhält man mit Hilfe des Drehoperators $\mathbf{x}\times\nabla_{\mathbf{x}}$, (siehe Formel (2.8)), angewendet auf die potentielle Energie der Gravitationswechselwirkung nach Formel (3.6),

$$-\mathbf{M}_{(12)} = \frac{\partial E}{\partial \varphi_{(12)}} = \mathbf{x}\times\nabla_{\mathbf{x}}\hat{V}(\mathbf{R},\mathbf{R}_{12}(\varphi_{(12)},R_{12}),\varphi_{\otimes}) = -\mathbf{R}_{12}\times\mathbf{G}_{(21)\otimes} \ , \tag{3.23}$$

und entsprechend die in radialer Richtung wirkende Kraft $\overline{K}_{(12)}$ mit Hilfe des Verschiebungsoperators $\nabla_{R_{12}}$ (siehe Formel (2.8)),

$$-\overline{K}_{(12)} = \frac{\partial E}{\partial R_{12}} = \nabla_{R_{12}}\hat{V}(\mathbf{R},\mathbf{R_{12}}(\varphi_{(12)},R_{12}),\varphi_{\otimes})\cdot\mathbf{e} = -\mathbf{K}_{21}\cdot\mathbf{e} - \mathbf{G}_{(21)\otimes}\cdot\mathbf{e} \ . \tag{3.24}$$

Die Bilanzgleichungen für den Energieaustausch bei Relativbewegung des Satelliten 2 bzgl. des Satelliten 1 lauten nun anstelle von (3.15) und (3.16),

- für die Rotation der Verbindungslinie $\overline{\ ,2}$:

$$\mathbf{d}_{(12)}\cdot d\mathbf{L}_{(12)} - \mathbf{M}_{(12)}\cdot d\varphi_{(12)} = 0 \ , \tag{3.25}$$

und mit

$$d\mathbf{L}_{(12)} = \mu_{12}\ R_{12}^2\ d\mathbf{d}_{(12)} = \mu_{12}\ R_{12}^2\ d\dot{\varphi}_{(12)} \ , \tag{3.26}$$

zunächst,

$$\mu_{12}\ R_{12}^2\ \dot{\varphi}_{(12)}\cdot d\dot{\varphi}_{(12)} - \mathbf{R_{12}}\times\mathbf{G}_{(21)\otimes}\cdot d\varphi_{(12)} = 0 \ , \tag{3.27}$$

bzw., nach Umformung, wobei die Größen $\dot{\overline{R}}_{12}$ und $d\overline{R}_{12}$ die Projektionen der vektoriellen Größen auf die Richtung $\mathbf{e}$ bezeichnen, alternativ zu:

$$\mu_{12}\left(\dot{\mathbf{R}}_{12}\cdot d\dot{\mathbf{R}}_{12} - \dot{\overline{R}}_{12}\ d\dot{\overline{R}}_{12} - \dot{\overline{R}}_{12}\left(\dot{\mathbf{R}}_{12}\cdot d\mathbf{e}\right)\right) - \mathbf{R}_{12}\times\mathbf{G}_{(21)\otimes}\cdot d\varphi_{(12)} = 0 \ , \tag{3.28}$$

- für die Abstandsänderung $\overline{\ ,2}$:

$$\overline{V}_{12}\ d\overline{P}_{12} - \overline{K}_{(12)}\ d\overline{R}_{12} = 0 \ , \tag{3.29}$$

und mit

$$\overline{V}_{12} = \mathbf{e}\cdot\mathbf{V}_{12} = \mathbf{e}\cdot\dot{\mathbf{R}}_{12} = \dot{\overline{R}}_{12}, \quad d\overline{P}_{12} = \mu_{12}\left(d\dot{\overline{R}}_{12} + d\mathbf{e}\cdot\dot{\mathbf{R}}_{12}\right) \ , \tag{3.30}$$

schließlich

$$\mu_{12}\ \dot{\overline{R}}_{12}\ d\dot{\overline{R}}_{12} + \mu_{12}\ \dot{\overline{R}}_{12}\left(d\mathbf{e}\cdot\dot{\mathbf{R}}_{12}\right) - \overline{K}_{(12)}\ d\overline{R}_{12} = 0 \ , \tag{3.31}$$

wobei

$$\overline{K}_{(12)}\ d\overline{R}_{12} = \left(\mathbf{e}\cdot\left(\mathbf{K}_{21} + \mathbf{G}_{(21)\otimes}\right)\right)\left(\mathbf{e}\cdot d\mathbf{R}_{12}\right) \ . \tag{3.32}$$

4 Energieaustauschbeziehungen im erdfesten Bezugssystem

Wird ein nichtinertiales Bezugssystem zugrunde gelegt, so sind die Wechselwirkungen des Systems mit dem Trägheitsfeld zu berücksichtigen (Abschnitt 1). Das bedeutet, daß die Gibbssche Funktion zu modifizieren ist. Im folgenden soll die Bewegung des Massenmittelpunktes der Satelliten 1 und 2 auf ein konstant rotierendes geozentrisches Bezugssystem bezogen werden. Die Rotation wird durch den Drehvektor $\Omega = (0,0,\omega)$ beschrieben. Entsprechend bezieht sich die Relativbewegung der beiden Satelliten auf dieses rotierende System. Die Trägheitsbewegung des Massenmittelpunktes des betrachteten abgeschlossenen Systems soll im folgenden nicht weiter betrachtet werden. Die gestrichenen Größen beziehen sich im folgenden auf das rotierende Bezugssystem. Die Gibbssche Funktion lautet für diesen Fall:

$$E' = \frac{1}{2}\frac{\mathbf{P}'^2}{M} + \frac{1}{2}\frac{\mathbf{P}_{12}'^2}{\mu_{12}} - \mathbf{P}' \cdot (\Omega \times \mathbf{R}') - \mathbf{P}_{12}' \cdot (\Omega \times \mathbf{R}_{12}') + \hat{V}'(\mathbf{R}', \mathbf{R}_{12}', \varphi_{\otimes}) + E_0 \ . \tag{4.1}$$

Die dynamischen Geschwindigkeiten $\mathbf{V}', \mathbf{V}_{12}'$ erhält man bzgl. des rotierenden Bezugssystems wiederum durch partielle Ableitung der Gibbsschen Funktion nach den konjugierten extensiven Variablen:

$$\mathbf{V}' = \frac{\partial E'}{\partial \mathbf{P}'} = \frac{\mathbf{P}'}{M} - \Omega \times \mathbf{R}', \quad \mathbf{V}_{12}' = \frac{\partial E'}{\partial \mathbf{P}_{12}'} = \frac{\mathbf{P}_{12}'}{\mu_{12}} - \Omega \times \mathbf{R}_{12}', \tag{4.2}$$

Man beachte, daß die linearen Impulse bzgl. des Inertialsystems und bzgl. des konstant rotierenden Bezugssystems identisch sind, während sich die Geschwindigkeiten unterscheiden:

$$\mathbf{P}' = \mathbf{P}, \quad \mathbf{P}_{12}' = \mathbf{P}_{12} \tag{4.3}$$

Für die Kräfte zufolge Wechselwirkung mit dem Gravitationsfeld und dem Trägheitsfeld erhält man:

$$\begin{aligned} -\mathbf{K}_{(\otimes,1,2)}' &= \frac{\partial E'}{\partial \mathbf{R}'} = -\mathbf{K}_{1\otimes}' - \mathbf{K}_{2\otimes}' + \Omega \times \mathbf{P}', \\ -\mathbf{K}_{(12)}' &= \frac{\partial E'}{\partial \mathbf{R}_{12}'} = -\mathbf{K}_{21}' - \mathbf{G}_{(21)\otimes}' + \Omega \times \mathbf{P}_{12}' \ , \end{aligned} \tag{4.4}$$

Die Bilanzgleichungen für den Energieaustausch des Systems $(\otimes,1,2)$ mit Gravitations- und Trägheitsfeldern erhält man:

- für die gemeinsame Bewegung der Satelliten 1 und 2, bezogen auf ein konstant rotierendes Bezugssystem im Geozentrum:

$$\mathbf{V}' \cdot d\mathbf{P}' - \mathbf{K}_{(\otimes,1,2)}' \cdot d\mathbf{R}' = 0 \ , \tag{4.5}$$

  bzw. nach entsprechender Umformung:

$$(M_1 + M_2)\, \dot{\mathbf{R}}' \cdot d\dot{\mathbf{R}}' - \mathbf{K}_{(\otimes,1,2)}' \cdot d\mathbf{R}' = 0 \ , \tag{4.6}$$

  wobei

$$\mathbf{K}_{(\otimes,1,2)}' = \mathbf{K}_{1\otimes}' + \mathbf{K}_{2\otimes}' - (M_1 + M_2)\, \Omega \times (\Omega \times \mathbf{R}') \ , \tag{4.7}$$

- für die Relativbewegung des Satelliten 2 bzgl. des Satelliten 1, bezogen auf ein konstant rotierendes Bezugssystem im Satelliten 1:

$$\mathbf{V}_{12}' \cdot d\mathbf{P}_{12}' - \mathbf{K}_{(12)}' \cdot d\mathbf{R}_{12}' = 0 \ , \tag{4.8}$$

  bzw. nach entsprechender Umformung:

$$\mu_{12}\, \dot{\mathbf{R}}_{12}' \cdot d\dot{\mathbf{R}}_{12}' - \mathbf{K}_{(12)}' \cdot d\mathbf{R}_{12}' = 0 \ , \tag{4.9}$$

  wobei

$$\mathbf{K}_{(12)}' = \mathbf{K}_{21}' + \mathbf{G}_{(21)\otimes}' - \mu_{12}\, \Omega \times (\Omega \times \mathbf{R}_{12}') \ . \tag{4.10}$$

Zerlegt man die Relativbewegung wieder in eine Dreh- und Radialbewegung, so erhält man folgende den Formeln (4.6) und (4.9) entsprechenden Bilanzgleichungen für den Energieaustausch:

- für die Rotation der Verbindungslinie $\overline{,2}$:

$$\mathbf{d}'_{(12)} \cdot d\mathbf{L}'_{(12)} - \mathbf{M}'_{(12)} \cdot d\varphi'_{(12)} = 0 \ , \tag{4.11}$$

bzw. mit

$$d\mathbf{L}'_{(12)} = \mu_{12}\ R_{12}^2\ d\mathbf{d}'_{(12)} = \mu_{12}\ R_{12}^2\ d\dot{\varphi}'_{(12)} \quad , \tag{4.12}$$

schließlich:

$$\mu_{12}\ R_{12}^2\ \dot{\varphi}'_{(12)} \cdot d\dot{\varphi}'_{(12)} - \mathbf{M}'_{(12)} \cdot d\varphi'_{(12)} = 0 \, , \tag{4.13}$$

wobei

$$\mathbf{M}'_{(12)} = \mathbf{R}'_{12} \times \mathbf{G}'_{(21)\otimes} - \mu_{12}\mathbf{R}'_{12} \times \left(\Omega \times \dot{\mathbf{R}}'_{12}\right) - \mu_{12}\mathbf{R}'_{12} \times \left(\Omega \times \left(\Omega \times \mathbf{R}'_{12}\right)\right) \ . \tag{4.14}$$

Alternativ kann der erste Term in (4.13) auch entsprechend wie in (3.28) angeschrieben werden.

- für die Abstandsänderung $\overline{,2}$:

$$\overline{V}'_{12}\ d\overline{P}'_{12} - \overline{K}'_{(12)}\ d\overline{R}'_{12} = 0 \ , \tag{4.15}$$

und mit

$$\overline{V}'_{12} = \mathbf{e}' \cdot \dot{\mathbf{R}}'_{12} =: \dot{\overline{R}}'_{12} \, , \quad d\overline{V}'_{12} = d\dot{\overline{R}}'_{12} \ , \tag{4.16}$$

schließlich

$$\mu_{12}\ \dot{\overline{R}}'_{12}\ d\dot{\overline{R}}'_{12} + \mu_{12}\ \dot{\overline{R}}'_{12}\left(d\mathbf{e}' \cdot \dot{\mathbf{R}}'_{12}\right) - \overline{K}'_{(12)}\ d\overline{R}'_{12} = 0 \ , \tag{4.17}$$

wobei

$$\overline{K}'_{(12)} = \mathbf{e}' \cdot \left(\mathbf{K}'_{21} + \mathbf{G}'_{(21)\otimes} - \mu_{12}\Omega \times \dot{\mathbf{R}}'_{12} - \mu_{12}\Omega \times \left(\Omega \times \mathbf{R}'_{12}\right)\right), \quad d\overline{R}'_{12} = \mathbf{e}' \cdot d\mathbf{R}'_{12} \ . \tag{4.18}$$

5 Energie- und Bewegungsintegrale

Die Einführung von Jacobi-Koordinaten ermöglichte die Abspaltung der Bewegung des Massenmittelpunktes des Gesamtsystems von den Einzelbewegungen. Damit ist der Gesamtimpuls **p** in Strenge konstant und der Massenmittelpunkt führt eine Trägheitsbewegung aus. Dasselbe gilt in hinreichender Näherung auch für den Drehimpuls der Erde, da die Gravitationswechselwirkung mit den Satelliten kein Drehmoment ausübt. Eine weitere Aufspaltung der Bilanzgleichungen ist nicht möglich, da die potentielle Energie von beiden Jacobi-Koordinaten abhängt. Allerdings gelingt es, Bilanzbeziehungen anzugeben, die bei gewissen Konfigurationen der beiden Satelliten nützlich sein können. Hierzu werden in der ***Gibbsschen Funktion*** des Gesamtsystems

$$E(\mathbf{p},\mathbf{P},\mathbf{P}_{12},\mathbf{L}_{\otimes},\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes}) = \frac{1}{2}\frac{\mathbf{p}^2}{m} + \frac{1}{2}\frac{\mathbf{P}^2}{M} + \frac{1}{2}\frac{\mathbf{P}_{12}^2}{\mu_{12}} + \frac{1}{2}\mathbf{L}_{\otimes} \cdot \mathbf{T}_{\otimes}^{-1} \cdot \mathbf{L}_{\otimes} + \hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes}) + E_0 \quad , \tag{5.1}$$

die aus den genannten Gründen konstanten Anteile zur inneren Energie addiert und man erhält das Energieintegral, bezogen auf ein ***Quasi-Inertialsystem*** in der folgenden Form, wenn noch die (dynamischen) Geschwindigkeiten statt der Impulse nach Formel (3.9) eingeführt werden:

$$E(\mathbf{V},\mathbf{V}_{12},\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes}) = \frac{M}{2}\mathbf{V}^2 + \frac{\mu_{12}}{2}\mathbf{V}_{12}^2 + \hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes}) = \text{const} \quad . \tag{5.2}$$

Da sich dieses ***Energieintegral*** auf ein Quasi-Inertialsystem bezieht, müssen die Potentialkoeffizienten nach Formel (6.4) vom erdfesten Bezugssystem in das Quasi-Inertialsystem transformiert werden. Durch Einführung der kinematischen Geschwindigkeiten wird das Energieintegral zu einem ***Bewegungsintegral***:

$$E(\dot{\mathbf{R}},\dot{\mathbf{R}}_{12},\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes}) = \frac{M}{2}\dot{\mathbf{R}}^2 + \frac{\mu_{12}}{2}\dot{\mathbf{R}}_{12}^2 + \hat{V}(\mathbf{R},\mathbf{R}_{12},\varphi_{\otimes}) = \text{const} \quad . \tag{5.3}$$

Bezieht man die Relativbewegungen $\mathbf{R}(t)$ und $\mathbf{R}_{12}(t)$ auf ein (beispielsweise konstant) ***rotierendes Bezugssystem***, so kann mit Hilfe der ***Gibbsschen Funktion*** nach Formel (4.1),

$$E'(\mathbf{P}',\mathbf{P}_{12}'^2,\mathbf{R}',\mathbf{R}_{12}') = \frac{1}{2}\frac{\mathbf{P}'^2}{M} + \frac{1}{2}\frac{\mathbf{P}_{12}'^2}{\mu_{12}} - \mathbf{P}'\cdot(\Omega\times\mathbf{R}') - \mathbf{P}_{12}'\cdot(\Omega\times\mathbf{R}_{12}') + \hat{V}'(\mathbf{R}',\mathbf{R}_{12}') + E_0 \quad , \tag{5.4}$$

und der Formeln für die dynamischen Geschwindigkeiten $\mathbf{V}',\mathbf{V}_{12}'$ bzgl. des rotierenden Bezugssystems,

$$\mathbf{V}' = \frac{\partial E'}{\partial \mathbf{P}'} = \frac{\mathbf{P}'}{M} - \Omega\times\mathbf{R}' \,, \quad \mathbf{V}_{12}' = \frac{\partial E'}{\partial \mathbf{P}_{12}'} = \frac{\mathbf{P}_{12}'}{\mu_{12}} - \Omega\times\mathbf{R}_{12}' \,, \tag{5.5}$$

das folgende ***Energieintegral*** erhalten werden:

$$E'(\mathbf{V}',\mathbf{V}_{12}',\mathbf{R}',\mathbf{R}_{12}') = \frac{M}{2}\mathbf{V}'^2 - \frac{M}{2}(\Omega\times\mathbf{R}')^2 + \frac{\mu_{12}}{2}\mathbf{V}_{12}'^2 - \frac{\mu_{12}}{2}(\Omega\times\mathbf{R}_{12}')^2 + \hat{V}'(\mathbf{R}',\mathbf{R}_{12}') = \text{const} \,, \tag{5.6}$$

Man beachte, daß bzgl. des konstant rotierenden Bezugssystems die Potentialkoeffizienten des Gravitationsfeldes der Erde konstant sind. Durch Einführung der kinematischen Geschwindigkeiten wird das Energieintegral zu einem ***Bewegungsintegral***:

$$E'(\dot{\mathbf{R}}',\dot{\mathbf{R}}_{12}',\mathbf{R}',\mathbf{R}_{12}') = \frac{M}{2}\dot{\mathbf{R}}'^2 - \frac{M}{2}(\Omega\times\mathbf{R}')^2 + \frac{\mu_{12}}{2}\dot{\mathbf{R}}_{12}'^2 - \frac{\mu_{12}}{2}(\Omega\times\mathbf{R}_{12}')^2 + \hat{V}'(\mathbf{R}',\mathbf{R}_{12}') = \text{const} \,. \tag{5.7}$$

Eine weitere Separierung des Bewegungsintegrals in zwei Anteile, die jeweils lediglich von einer der beiden Jacobi-Koordinaten abhängen, ist wegen der Kopplung über die potentielle Energie der Gravitationswechselwirkung $\hat{V}'(\mathbf{R}',\mathbf{R}_{12}')$ nicht möglich. Dies gelingt nur unter der speziellen Annahme, daß beide Bewegungsanteile keinen Einfluß aufeinander ausüben, daß also jeweils eine Jacobi-Koordinate während des Bewegungsablaufes konstant bleibt.
Man erhält bei einem bzgl. des erdfesten Bezugssystems konstanten Vektor $\mathbf{R}_{12}'$ den Erhaltungssatz:

$$E'(\dot{\mathbf{R}}',\mathbf{R}') = \frac{M}{2}\dot{\mathbf{R}}'^2 - \frac{M}{2}(\Omega\times\mathbf{R}')^2 - \tilde{V}'_{(\otimes,1,2)}(\mathbf{R}') = \text{const} \,, \tag{5.8}$$

Beachtet man, daß in diesem Fall der Term

$$W_{(\otimes,1,2)}(\mathbf{R}') = \frac{1}{2}(\Omega\times\mathbf{R}')^2 + \frac{1}{M}\tilde{V}'_{(\otimes,1,2)}(\mathbf{R}') \,, \tag{5.9}$$

das Schwerepotential der Erde ist, mit dem Gravitationspotential $\tilde{V}'_{(\otimes,1,2)}(\mathbf{R}')$ / M, ausgewertet nach Formel (6.21) mit $|\mathbf{R}_{12}'| = 0$, so ergibt sich ein Erhaltungssatz, der dem ***Jacobi-Integral*** entspricht:

$$E'(\dot{\mathbf{R}}',\mathbf{R}') = \dot{\mathbf{R}}'^2 - 2W_{(\otimes,1,2)}(\mathbf{R}') = \text{const} \,, \tag{5.10}$$

Für diesen speziellen Fall kann das Energieintegral (5.4) unter Beachtung von $\mathbf{P}' = \mathbf{P}$ und $\mathbf{P} = M\mathbf{V}$ und einem Rotationsvektor $\Omega = \omega\,\mathbf{e}_z$ umgeformt werden:

$$E'(\mathbf{V},\mathbf{R}') = \frac{M}{2}\mathbf{V}^2 - M\,\omega\,\mathbf{e}_z\cdot(\mathbf{R}'\times\mathbf{V}) + \hat{V}'(\mathbf{R}') = \text{const} \,, \tag{5.11}$$

Der Term $\mathbf{e}_z\cdot(\mathbf{R}'\times\mathbf{V})$ ist die Koordinate des Bahndrehimpulses des Massenzentrums in *z*-Richtung. Sie ist für axialsymmetrische Felder eine Erhaltungsgröße (siehe z.B. Schneider, 1992) und kann mit der Konstanten auf der rechten Seite zusammengefaßt werden. Im Falle eines axialsymmetrischen Gravitationsfeldes nimmt das Energieintegral also die folgende Form an:

$$E'(\mathbf{V},\mathbf{R}') = \frac{M}{2}\mathbf{V}^2 + \hat{V}(\mathbf{R}') = \text{const} \,, \tag{5.12}$$

Dies ist die totale Bahnenergie des Massenzentrums der Satelliten 1 und 2. In einem axialsymmetrischen Gravitationsfeld, wobei die Symmetrieachse mit dem Rotationsvektor des rotierenden Bezugssystems übereinstimmt, findet somit kein Energieaustausch mit dem Gravitationsfeld statt.
Für einen bzgl. dem erdfesten Bezugssystems konstanten Vektor $\mathbf{R}'$, also beispielsweise im Falle einer Kreisbahn folgt der Erhaltungssatz:

$$E'(\dot{\mathbf{R}}'_{12},\mathbf{R}'_{12}) = \frac{\mu_{12}}{2}\dot{\mathbf{R}}'^{2}_{12} - \frac{\mu_{12}}{2}\left(\Omega\times\mathbf{R}'_{12}\right)^2 + \hat{V}'(\mathbf{R}',\mathbf{R}'_{12}) = \text{const}\ , \tag{5.13}$$

Die potentielle Energie der Gravitationswechselwirkung $\hat{V}'(\mathbf{R}',\mathbf{R}'_{12})$ kann auch durch die Potentialfunktion (6.12) ersetzt werden, so daß man als Energieintegral erhält:

$$E'(\dot{\mathbf{R}}'_{12},\mathbf{R}'_{12}) = \frac{\mu_{12}}{2}\dot{\mathbf{R}}'^{2}_{12} - \frac{\mu_{12}}{2}\left(\Omega\times\mathbf{R}'_{12}\right)^2 - \tilde{V}'_{(12)}(\mathbf{R}',\mathbf{R}'_{12}) = \text{const}\ , \tag{5.14}$$

Das Jacobi-Integral (5.10) gilt für die Bahnbewegung jedes der beiden Satelliten. Bildet man die Differenz, so ergibt sich ein weiteres Energieintegral:

$$E'(\dot{\mathbf{R}}'_1,\mathbf{R}'_1,\dot{\mathbf{R}}'_2,\mathbf{R}'_2,) = \dot{\mathbf{R}}'_{12}\cdot\left(\dot{\mathbf{R}}'_2+\dot{\mathbf{R}}'_1\right) - 2\left(W_{(\otimes,1,2)}(\mathbf{R}'_2)\ - W_{(\otimes,1,2)}(\mathbf{R}'_1)\ \right) = \text{const}\ . \tag{5.15}$$

6 Modellierung der potentiellen Energie

Gelingt es, die potentielle Energie der Gravitationswechselwirkung durch die Jacobi-Koordinaten auszudrücken, so können die Kräfte (bzw. Drehmomente) als intensive Variable durch partielle Ableitungen der Gibbsschen Funktion nach den konjugierten extensiven Variablen abgeleitet werden. In der Geodäsie verwendet man vorzugsweise Reihenentwicklungen nach Kugelfunktionen. Geht man im vorliegenden Fall von der Formel für die potentielle Energie der Gravitationswechselwirkung aus, so können die einzelnen Summanden auf der rechten Seite

$$\hat{V}(\mathbf{r}_j, j=\otimes,1,2;\varphi_\otimes) = \hat{V}_{\otimes 1}(\mathbf{r}_\otimes,\mathbf{r}_1,\varphi_\otimes) + \hat{V}_{\otimes 2}(\mathbf{r}_\otimes,\mathbf{r}_2,\varphi_\otimes) + \hat{V}_{12}(\mathbf{r}_1,\mathbf{r}_2)\ . \tag{6.1}$$

beispielsweise die potentielle Energie der Erde und des Satelliten 1 aus der folgenden Formel (mit $\mathbf{R}_1(R_1,\vartheta_1,\lambda_1) = \mathbf{r}_1 - \mathbf{r}_\otimes$) erhalten werden:

$$\hat{V}_{\otimes 1}(\mathbf{r}_\otimes,\mathbf{r}_1,\varphi_\otimes) = -\frac{GM_1M_\otimes}{R_1}\sum_{n=0}^{\infty}\left(\frac{a_\otimes}{R_1}\right)^n\sum_{m=-n}^{n}\kappa_{nm}(\varphi_\otimes)\,Y_{nm}(\vartheta_1,\lambda_1) \tag{6.2}$$

Die Funktionen $Y_{nm}(\vartheta_1,\lambda_1)$ sind die komplexen Kugelflächenfunktionen des Grades n und der Ordnung m, für $n\geq 0$ und $-n\leq m\leq n$,

$$Y_{nm}(\vartheta,\lambda) = P_n^m(\cos\vartheta)e^{im\lambda},\qquad Y_{n,-m}(\vartheta,\lambda) = P_n^{-m}(\cos\vartheta)e^{-im\lambda} = (-1)^m\frac{(n-m)!}{(n+m)!}P_n^m(\cos\vartheta)e^{-im\lambda}\ , \tag{6.3}$$

und $\kappa_{nm}(\varphi_\otimes)$, $n\geq 0, -n\leq m\leq n$ die komplexen Potentialkoeffizienten. Wird ein raumfestes Koordinatensystem zugrunde gelegt, so hängen die Potentialkoeffizienten von der jeweiligen Orientierung $\varphi_\otimes$ der Erde bzgl. des Quasi-Inertialsystems ab und sind damit zeitabhängig. Geht man (ohne Einschränkung der Allgemeinheit) von einer konstanten Drehung der Erde um die z-Achse aus (Winkelgeschwindigkeit ω), so transformieren sich die Potentialkoeffizienten κ'_{nm}, bezogen auf ein erdfestes Koordinatensystem wie folgt:

$$\kappa_{nm} = e^{-im\omega t}\kappa'_{nm} \tag{6.4}$$

Eine entsprechende Formel wie (6.2) gilt für $\hat{V}_{\otimes 2}(\mathbf{r}_\otimes,\mathbf{r}_2,\varphi_\otimes)$. Die potentielle Energie der Gravitationswechselwirkung der beiden Satelliten ergibt sich zu (mit $\mathbf{R}_{12}(R_{12},\vartheta_{12},\lambda_{12}) = \mathbf{r}_2 - \mathbf{r}_1$)

$$\hat{V}_{12}(\mathbf{R}_{12}) = -\frac{GM_1M_2}{R_{12}} \ . \tag{6.5}$$

Die Kugelfunktionen können mittels der Transformationsformel (siehe z.B. Giacaglia, 1980)

$$\frac{1}{R_{\otimes 2}^{n+1}} Y_{nm}(\vartheta_{\otimes 2}, \lambda_{\otimes 2}) = \sum_{p=0}^{\infty} \sum_{q=-p}^{p} (-1)^{p+q} \frac{(n-m+p+q)!}{(n-m)!(p+q)!} R_{12}^p Y_{pq}(\vartheta_{12}, \lambda_{12}) \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1}, \lambda_{\otimes 1}) \tag{6.6}$$

zunächst auf den Ort des Satelliten 1 bezogen werden:

$$\begin{aligned} \hat{V}(\mathbf{R}, \mathbf{R}_{12}, \varphi_{\otimes}) = -GM_{\otimes} \sum_{n=0}^{\infty} a_{\otimes}^n \sum_{m=-n}^{n} \kappa_{nm}(\varphi_{\otimes}) \cdot \\ \cdot \Bigg[M_2 \Bigg(\sum_{p=1}^{\infty} \sum_{q=-p}^{p} (-1)^{p+q} \frac{(n-m+p+q)!}{(n-m)!(p+q)!} \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1}, \lambda_{\otimes 1}) R_{12}^p Y_{pq}(\vartheta_{12}, \lambda_{12}) \Bigg) + \\ + (M_1 + M_2) \frac{1}{R_{\otimes 1}^{n+1}} Y_{nm}(\vartheta_{\otimes 1}, \lambda_{\otimes 1}) \Bigg] - \frac{GM_1M_2}{R_{12}} . \end{aligned} \tag{6.7}$$

und wenn beachtet wird, daß für die Relativkoordinaten die folgende Beziehung gilt,

$$\mathbf{R}_{\otimes 1}(R_{\otimes 1}, \vartheta_{\otimes 1}, \lambda_{\otimes 1}) = \mathbf{R}(R, \vartheta, \lambda) - \frac{M_2}{M_1 + M_2} \mathbf{R}_{12}(R_{12}, \vartheta_{12}, \lambda_{12}) . \tag{6.8}$$

mit Hilfe der Transformationsformel

$$\begin{aligned} \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1}, \lambda_{\otimes 1}) = \\ = \sum_{l=0}^{\infty} \sum_{k=-l}^{l} (-1)^k \frac{(n+p-m+q+l+k)!}{(n+p-m+q)!(l+k)!} \left(\frac{M_2}{M_1 + M_2} \right)^l R_{12}^l Y_{lk}(\vartheta_{12}, \lambda_{12}) \frac{1}{R^{n+p+l+1}} Y_{n+p+l,m-q-k}(\vartheta, \lambda) \end{aligned} \tag{6.9}$$

auf den Massenmittelpunkt der beiden Satelliten 1 und 2. Allerdings ist die so erhaltene Formel verhältnismäßig kompliziert auszuwerten.
Für die praktische Anwendung scheint insbesondere für die Relativbewegung der beiden Satelliten eine alternative Potentialfunktion der Kräftefunktion

$$\mathbf{K}_{(12)} = \nabla_{\mathbf{R}_{12}} \tilde{V}_{(1,2)}(\mathbf{R}, \mathbf{R}_{12}, \varphi_{\otimes}) , \tag{6.10}$$

zweckmäßiger zu sein. Die Potentialfunktion setzt sich aus zwei Anteilen zusammen,

$$\tilde{V}_{(1,2)}(\mathbf{R}, \mathbf{R}_{12}, \varphi_{\otimes}) = \tilde{V}_{(12)\otimes}(\mathbf{R}, \mathbf{R}_{12}, \varphi_{\otimes}) + \tilde{V}_{12}(\mathbf{R}_{12}) \ , \tag{6.11}$$

der Potentialfunktion der Gezeitenkraft

$$\tilde{V}_{(12)\otimes}(\mathbf{R}, \mathbf{R}_{12}, \varphi_{\otimes}) = \frac{M_2}{M_1 + M_2} \hat{V}_{\otimes 1} - \frac{M_1}{M_1 + M_2} \hat{V}_{\otimes 2} \ , \tag{6.12}$$

und dem Term $\tilde{V}_{12}(\mathbf{R}_{12})$, der die potentielle Energie der Gravitationswechselwirkung der beiden Körper und 2 beschreibt. Die Potentialfunktion der Gezeitenkraft kann im Falle der beiden Punktmassen und 2 durch eine Entwicklung nach Kugelfunktionen dargestellt werden (vergleiche z.B. Ilk, 1983c; dort wurde sie mit negativem Vorzeichen als potentielle Energie der Gezeitenkraft eingeführt):

$$\begin{aligned} \tilde{V}_{(12)\otimes}(\mathbf{R}, \mathbf{R}_{12}, \varphi_{\otimes}) = GM_{\otimes} \mu_{12} \sum_{n=0}^{\infty} a_{\otimes}^n \sum_{m=-n}^{n} \kappa_{nm}(\varphi_{\otimes}) \cdot \\ \cdot \sum_{p=1}^{\infty} \sum_{q=-p}^{p} (-1)^{p+q} \frac{(n-m+p+q)!}{(n-m)!(p+q)!} \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1}, \lambda_{\otimes 1}) R_{12}^p Y_{pq}(\vartheta_{12}, \lambda_{12}) . \end{aligned} \tag{6.13}$$

Man erhält die Gezeitenkraft durch Bildung des Gradienten der Potentialfunktion:

$$\mathbf{G}_{(21)\otimes} = \nabla_{\mathbf{R}_{12}} \tilde{V}_{(12)\otimes}(\mathbf{R},\mathbf{R}_{12},\varphi_\otimes) = GM_\otimes \mu_{12} \sum_{n=0}^{\infty} a_\otimes^n \sum_{m=-n}^{n} \kappa_{nm}(\varphi_\otimes)\cdot$$
$$\cdot \sum_{p=1}^{\infty} \sum_{q=-p}^{p} (-1)^{p+q} \frac{(n-m+p+q)!}{(n-m)!(p+q)!} \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1},\lambda_{\otimes 1}) \nabla_{\mathbf{R}_{12}} \left(R_{12}^p Y_{pq}(\vartheta_{12},\lambda_{12}) \right). \tag{6.14}$$

mit dem Gradienten

$$\nabla_{\mathbf{R}_{12}} \left(R_{12}^p Y_{pq}(\vartheta_{12},\lambda_{12}) \right) = \frac{R_{12}^{p-1}}{2} \begin{pmatrix} (p+q)(p+q-1) Y_{p-1,q-1}(\vartheta_{12},\lambda_{12}) - Y_{p-1,q+1}(\vartheta_{12},\lambda_{12}) \\ i \left((p+q)(p+q-1) Y_{p-1,q-1}(\vartheta_{12},\lambda_{12}) + Y_{p-1,q+1}(\vartheta_{12},\lambda_{12}) \right) \\ 2(p+q) Y_{p-1,q}(\vartheta_{12},\lambda_{12}) \end{pmatrix} \tag{6.15}$$

Für das Drehmoment der Gezeitenkraft folgt entsprechend:

$$\mathbf{M}_{(12)} = \mathbf{x} \times \nabla_{\mathbf{x}} \tilde{V}_{(12)\otimes}(\mathbf{R},\mathbf{R}_{12},\varphi_\otimes) = \mathbf{R}_{12} \times \mathbf{G}_{(21)\otimes} = GM_\otimes \mu_{12} \sum_{n=0}^{\infty} a_\otimes^n \sum_{m=-n}^{n} \kappa_{nm}(\varphi_\otimes)\cdot$$
$$\cdot \sum_{p=1}^{\infty} \sum_{q=-p}^{p} (-1)^{p+q} \frac{(n-m+p+q)!}{(n-m)!(p+q)!} \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1},\lambda_{\otimes 1}) (\mathbf{x} \times \nabla_{\mathbf{x}}) \left(R_{12}^p Y_{pq}(\vartheta_{12},\lambda_{12}) \right). \tag{6.16}$$

mit

$$(\mathbf{x} \times \nabla_{\mathbf{x}}) \left(R_{12}^p Y_{pq}(\vartheta_{12},\lambda_{12}) \right) = R_{12}^p \begin{pmatrix} -\frac{i}{2} \left(Y_{p,q+1}(\vartheta_{12},\lambda_{12}) + (p+q)(p-q+1)\, Y_{p,q-1}(\vartheta_{12},\lambda_{12}) \right) \\ -\frac{1}{2} \left(Y_{p,q+1}(\vartheta_{12},\lambda_{12}) - (p+q)(p-q+1)\, Y_{p,q-1}(\vartheta_{12},\lambda_{12}) \right) \\ i\, q\, Y_{pq}(\vartheta_{12},\lambda_{12}) \end{pmatrix} \tag{6.17}$$

Der Term $\tilde{V}_{12}(\mathbf{R}_{12})$ in Formel (6.11) beschreibt die Potentialfunktion der Gravitationswechselwirkung der beiden Satelliten, betrachtet als Punktmassen,

$$\tilde{V}_{12}(\mathbf{R}_{12}) = \frac{GM_1M_2}{R_{12}} \ . \tag{6.18}$$

Eine entsprechende alternative Potentialfunktionen für die Kräftefunktion, die für die Relativbewegung des Massenzentrums der beiden Satelliten 1 und 2 bzgl des Geozentrums benötigt wird, erhält man auf folgende Weise:

$$\mathbf{K}_{(\otimes,1,2)} = \nabla_{\mathbf{R}} \tilde{V}_{(\otimes,1,2)}(\mathbf{R},\mathbf{R}_{12},\varphi_\otimes), \tag{6.19}$$

mit der Potentialfunktion,

$$\tilde{V}_{(\otimes,1,2)}(\mathbf{R},\mathbf{R}_{12},\varphi_\otimes) = \hat{V}_{\otimes 1} + \hat{V}_{\otimes 2} \ , \tag{6.20}$$

bzw. in einer Entwicklung nach Kugelfunktionen:

$$\tilde{V}_{(\otimes,1,2)}(\mathbf{R},\mathbf{R}_{12},\varphi_\otimes) = GM_2M_\otimes \sum_{n=0}^{\infty} a_\otimes^n \sum_{m=-n}^{n} \kappa_{nm}(\varphi_\otimes)\cdot$$
$$\cdot \left[\left(\sum_{p=1}^{\infty} \sum_{q=-p}^{p} (-1)^{p+q} \frac{(n-m+p+q)!}{(n-m)!(p+q)!} \frac{1}{R_{\otimes 1}^{n+p+1}} Y_{n+p,m-q}(\vartheta_{\otimes 1},\lambda_{\otimes 1}) R_{12}^p Y_{pq}(\vartheta_{12},\lambda_{12}) \right) + \right.$$
$$\left. + \frac{(M_1+M_2)}{M_2} \frac{1}{R_{\otimes 1}^{n+1}} Y_{nm}(\vartheta_{\otimes 1},\lambda_{\otimes 1}) \right]. \tag{6.21}$$

7 Zusammenfassung

Die abgeleiteten Bilanzgleichungen für den Energieaustausch gelten für jeden Zeitpunkt während des Bewegungsablaufes. Dasselbe gilt für die Gibbsschen Funktionen bzw. für die Energie- und Bewegungsintegrale. Damit kann die Konsistenz von beobachteten kinematischen Bewegungsgrößen mit den Parametern überprüft werden, die zur Beschreibung des Gravitationsfeldes eingeführt wurden. Die Konsistenzprüfung kann dabei für jeden Zeitpunkt des Bewegungsablaufes vorgenommen werden. Dies ist nicht ohne weiteres möglich, wenn eine Überprüfung durch die Lösung der Bewegungsgleichungen, beispielsweise durch numerische Integration erfolgt, da sich Abweichungen zu einem gewissen Zeitpunkt auf den weiteren Bahnverlauf auswirken. Ein Vergleich verschiedener Feldparameterbestimmungen weist dagegen den Nachteil auf, daß immer auch Regularisierungseffekte des instabilen Fortsetzungsprozesses nach unten zu gewissen Abweichungen beitragen können. Mit den Energie- und Energieaustauschbeziehungen steht ein Instrument zur Verfügung, das zur Verifizierung von Gravitationsfeldparametern dienen kann. Eventuell auftretende Abweichungen weisen darauf hin, daß am Energieaustausch weitere physikalische Systeme beteiligt waren. Dies wird in der Realität auch immer der Fall sein, da die effektiven Kräftefunktionen außer den gravitativen Wechselwirkungen weitere Anteile enthalten. Bei Verwendung der Energie- und Bewegungsintegrale muß allerdings beachtet werden, daß sich die axialsymmetrischen Anteile des Gravitationsfeldes in gewissen Bewegungsgrößen nicht bemerkbar machen. Numerische Erfahrungen einer solchen Verfizierung stehen noch aus.

Literaturverzeichnis

Falk, G.: *Theoretische Physik auf der Grundlage einer allgemeinen Dynamik, Heidelberger Taschenbücher, Band 7: Punktmechanik, Band 8: Aufgaben, Band 27: Thermodynamik, Band 28: Aufgaben*, Springer-Verlag, Berlin, Heidelberg, New York, 1966, 1968

Falk, G.: *Physik, Zahl und Realität*, Birkhäuser Verlag, Basel, Boston, Berlin, 1990

Falk, G., Ruppel, W.: *Mechanik, Relativität, Gravitation,*, Springer-Verlag, Berlin, Heidelberg, New York, 1973

Falk, G., Ruppel, W.: *Energie und Entropie*, Springer-Verlag, Berlin, Heidelberg, New York, 1976

Giacaglia, G.E.O.: *Transformations of spherical harmonics and applications to geodesy and satellite theory*, Studia Geophys. et Geod., Vol 24, pp. 1-11, 1980

Ilk, K.H.: *Eine Anmerkung zur Anwendung des Energieintegrals im Rahmen des SST-Verfahrens*, Veröff. Bayer. Kommission Int. Erdmess., Astron.-geod. Arb., Nr. 42, S. 123ff., München, 1982

Ilk, K.H.: *Formulierung von Energieaustauschbeziehungen zur Ausmessung des Gravitationsfeldes*, Veröff. Bayer. Kommission Int. Erdmess., Astron.-geod. Arb., Nr. 43, S. 128-166., München, 1983a

Ilk, K.H.: *On the Dynamics of a System of Rigid Bodies*, Manuscripta Geodaetica, Vol. 8, pp. 139-198, Stuttgart, 1983b

Ilk, K.H.: *Ein Beitrag zur Dynamik ausgedehnter Körper - Gravitationswechselwirkung*, DGK, Reihe C, Heft Nr. 288, München, 1983c

Jekeli, Ch.: *The determination of gravitational potential differences from satellite-to-satellite tracking*, submitted to Celestial Mechanics and Dynamical Astronomy, 1998

Schneider, M.: *Himmelsmechanik I - IV*, Spektrum Akademischer Verlag, Heidelberg, Berlin, 1992 - 1999

Über die Analyse von Beobachtungen in der Ausgleichungsrechnung – Äußere und innere Restriktionen

Ronald Jurisch, Georg Kampmann und Janette Linke

Zusammenfassung: In der Ausgleichungsrechnung spielt die sogenannte Geometrie der Beobachtungen eine entscheidende Rolle zur Begutachtung der erhaltenen Ergebnisse. Nachstehend wird eine weitreichende Analyse der äußeren und inneren Strukturen dieser Geometrie dargelegt.

Summary: Our contribution deals with the analysis of observations within the adjustment of observations. Introducing the normal form of the coefficient matrix and Plücker-Graßmann-Coordinates so called latent conditions (hidden condition equations within observation equations) can be detected.

1 Vermittelnde Ausgleichungsrechnung und äußere Restriktionen

Beginnen wollen wir unsere Betrachtungen mit der vermittelnden Ausgleichungsrechnung. Dies begründet sich in dem Umstand, daß man zielfunktionsinvariant die anderen Modelle der Ausgleichungsrechnung in dieses Modell überführen kann und wir die hier vorgestellten Begriffe an das Modell der vermittelnden Ausgleichungsrechnung gebunden sehen wollen. Es gilt für die vermittelnde Ausgleichung

$$\mathbf{A}\mathbf{x} = \mathbf{l} + \mathbf{v}, \;\; \mathrm{D}(\mathbf{l}) = \mathbf{P}^{-1}\sigma^2, \tag{A}$$

worin $\mathbf{l}$ einen (n,1) Vektor von (reduzierten) Beobachtungen, $\mathbf{A}$ eine (n,u) Matrix fester Koeffizienten mit rg $\mathbf{A} = q \le u$, $\mathbf{x}$ einen (u,1) Vektor fester, unbekannter Parameter und $\mathbf{v}$ einen (n,1) Vektor zufälliger Fehler bezeichnet, wobei die v_i als normalverteilt angenommen werden sollen mit $\mathbf{v} \sim N(0, \sigma^2\mathbf{E})$. Die Matrix $\mathbf{E}$ bezeichnet die Einheitsmatrix der betreffenden Dimension, n die Anzahl der Beobachtungen und u die Anzahl der unbekannten, zu schätzenden Parameter $\mathbf{x}$; D($\mathbf{l}$) bezeichnet die Varianz-Kovarianz-Matrix der Beobachtungen a priori, wobei die (n,n) Matrix $\mathbf{P}$ die diagonale Gewichtsmatrix und der Faktor σ^2 die Varianz der Gewichtseinheit (Referenz-Varianz) bezeichnet, (inkonsistentes Gleichungssystem vermittelnder Beobachtungen) vgl. (GRAFAREND, SCHAFFRIN, 1993).
Werden an dieses Modell Bedingungsgleichungen für die Unbekannten der Gestalt

$$\mathbf{H}\mathbf{x} = \mathbf{w} \tag{B}$$

angefügt, so sollen diese als *„äußere Restriktionen“* bezeichnet werden. Hierin bedeutet die (r,u) Matrix $\mathbf{H}$ die Restriktionsmatrix mit rg $\mathbf{H}$ = r und $\mathbf{w}$ ist der (r,1) Widerspruchsvektor.
Die Restriktionsgleichungen (B) entstehen dabei aus a priori Kenntnissen über die Geometrie des zugrunde liegenden Ausgleichungsproblems (z. B. geodätische Netzausgleichung).
Wegen der hinlänglich bekannten „Datumsinvarianz der Beobachtungen“ wollen wir zunächst die Problematik des möglichen Rangdefektes in $\mathbf{A}$ durch Hinzunahme der Matrix $\mathbf{E}$ mit $\mathbf{E}\mathbf{A}^{\mathrm{T}} = \mathbf{0}$ im Sinne der Analyse der Geometrie der Beobachtungen interpretiert sehen. Dabei gilt für die (u-q,u) Matrix $\mathbf{E}$ die Verfügung rg $\mathbf{E} = u - q$ und weiterhin die „äußeren Restriktionen“

$$\mathbf{E}\mathbf{x} = \mathbf{o}. \qquad \text{(C)}$$

Für alle Betrachtungen der Geometrie der Beobachtungen ist es nun notwendig, diese „äußeren Restriktionen“ zu eliminieren, um dann die „innere Geometrie der Beobachtungen“ untersuchen zu können. Ein solcher Eliminationsvorgang soll beispielhaft dargelegt werden. Es gilt:

$$\mathbf{A}\mathbf{x} = \mathbf{l} + \mathbf{v} \qquad \text{mit} \qquad \mathbf{H}\mathbf{x} = \mathbf{w}.$$

Dimensionen: $\mathbf{A}$ (n,u), $\mathbf{x}$ (u,1), $\mathbf{l}$ (n,1), $\mathbf{v}$ (n,1), $\mathbf{H}$ (r,u), $\mathbf{w}$ (r,1). Dieses Gleichungssystem kann wie nachfolgend umgeschrieben werden:

$$\begin{aligned} \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 &= \mathbf{l} + \mathbf{v} \\ \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 &= \mathbf{w} \end{aligned}$$

Unter der Bedingung, daß diese Aufteilung so durchgeführt worden ist, daß sich die (r,r) Matrix $\mathbf{H}_1$ invertieren läßt (regulär), gilt dann die folgende Darstellung:

$$\mathbf{x}_1 = \mathbf{H}_1^{-1}(\mathbf{w} - \mathbf{H}_2\mathbf{x}_2).$$

Mit der Festlegung für die Matrix $\mathbf{H}_1$ gilt dann für die übrigen Dimensionen $\mathbf{A}_1$ (n,r), $\mathbf{x}_1$ (r,1), $\mathbf{A}_2$ (n,u−r), $\mathbf{x}_2$ (u−r,1), $\mathbf{H}_1$ (r,r), $\mathbf{H}_2$ (r,u−r).
Die Gleichung $\mathbf{x}_1 = \mathbf{H}_1^{-1}(\mathbf{w} - \mathbf{H}_2\mathbf{x}_2)$ eingesetzt in die Gleichung $\mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{l} + \mathbf{v}$ ergibt:

$$\begin{aligned} & \mathbf{A}_1(\mathbf{H}_1^{-1}(\mathbf{w} - \mathbf{H}_2\mathbf{x}_2)) + \mathbf{A}_2\mathbf{x}_2 = \mathbf{l} + \mathbf{v} \\ = \; & \mathbf{A}_1\mathbf{H}_1^{-1}\mathbf{w} - \mathbf{A}_1\mathbf{H}_1^{-1}\mathbf{H}_2\mathbf{x}_2 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{l} + \mathbf{v} \\ = \; & (\mathbf{A}_2 - \mathbf{A}_1\mathbf{H}_1^{-1}\mathbf{H}_2)\mathbf{x}_2 = (\mathbf{l} - \mathbf{A}_1\mathbf{H}_1^{-1}\mathbf{w}) + \mathbf{v}, \end{aligned}$$

mit: $\mathbf{A}_R = (\mathbf{A}_2 - \mathbf{A}_1\mathbf{H}_1^{-1}\mathbf{H}_2)$ und $\mathbf{l}_R = (\mathbf{l} - \mathbf{A}_1\mathbf{H}_1^{-1}\mathbf{w})$
gilt: $\mathbf{A}_R\mathbf{x}_2 = \mathbf{l}_R + \mathbf{v}$.

Dieses Modell entspricht der ***Ausgleichung nach vermittelnden Beobachtungen ohne äußere Restriktionen***. Die Matrix $\mathbf{A}_R$ hat hierin vollen Spaltenrang (keinen surjektiven Rangdefekt).
Nachdem nun sichergestellt ist, daß keine äußeren Restriktionen im zu analysierenden Modell vorhanden sind, kann mit der Analyse der Geometrie der Beobachtungen begonnen werden. Wie sich zeigen wird, sind auch hierbei noch Restriktionen „zu fürchten“, die sich allerdings aus der Wahl des Ausgleichungsmodells bzw. der Beobachtungsanordnung ergeben. Wir bezeichnen diese Sachverhalte als „innere Restriktionen“. Diese sind Gegenstand nachstehender Darlegungen.

2 Analyse der inneren Geometrie von Beobachtungen in der vermittelnden Ausgleichungsrechnung

Ausgangspunkt der Betrachtungen sei das bekannte Modell der vermittelnden Ausgleichungsrechnung

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{l} + \mathbf{v} \\ \mathbf{v}^T\mathbf{v} &\rightarrow \min \end{aligned} \qquad (1)$$

Hierin sei $\mathbf{A}$ eine (n,u) Matrix mit vollem Spaltenrang (die sogenannte Designmatrix), $\mathbf{x} \in R^u$ ein Vektor unbekannter Parameter, $\mathbf{l} \in R^n$ der gegebene Beobachtungsvektor und $\mathbf{v} \in R^n$ ein Vektor unbekannter Verbesserungen. Mit dieser Voraussetzung für $\mathbf{A}$ enthält das Modell (1) der vermittelnden Ausgleichung keine „äußeren Restriktionen“ wie $\mathbf{H}\mathbf{x} = \mathbf{w}$ bzw. rg $\mathbf{A}$ = q < u, womit ein Rangdefekt zu beseitigen ist.

Zielfunktion ist die Quadratsumme der Verbesserungen, die minimiert werden soll (Methode der kleinsten Quadrate). Bereits durch die Formulierung in (1) wird deutlich, daß bei den Untersuchungen stochastische Aspekte zunächst nicht einbezogen werden sollen. Vielmehr geht es in den nachfolgenden Ausführungen im Besonderen um eine geometrische Analyse des Modells, insbesondere um geometrische Besonderheiten, wobei die Wechselbeziehungen zwischen der vermittelnden und der bedingten Ausgleichung eine große Rolle spielen.

2.1 Die Geometrie eines Modells

Aus streng mathematischer Sicht stellt sich die Behandlung des Modells (1) folgendermaßen dar. Die Spalten der Designmatrix **A** sind linear unabhängig und spannen einen u-dimensionalen Unterraum U des Beobachtungsraumes R^n auf (sie bilden eine Basis in U). Durch die Zielfunktion wird auf R^n (und damit auch auf $U \subset R^n$) die euklidische Metrik induziert.
Die Lösung des Problems liegt nun in der Konstruktion von orthogonalen Projektoren von R^n auf U. Die geometrischen Eigenschaften dieser Projektion hängen bekanntlich wesentlich von der „Lage“ von U bezüglich des Beobachtungsraumes R^n ab (U ist in R^n eingebettet). Ziel der Untersuchungen ist es daher, die Geometrie von U in R^n zu charakterisieren und daraus entsprechende Folgerungen für das Modell (1) abzuleiten.
Bei Wahl einer Basis in U kann der (orthogonale) Projektor auf U durch eine symmetrische und idempotente Matrix **C** realisiert werden. Nimmt man die Spalten von **A** als Basis für U, so erhält man als Projektionsmatrix **C** die sogenannte Hat-Matrix

$$\mathbf{C} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T. \tag{2}$$

Es sei jedoch schon an dieser Stelle vermerkt, daß die Matrix **C** nicht von der Wahl der Basis in U abhängt, was später noch von Wichtigkeit sein wird. Im Rahmen der Sensitivitätsanalyse (siehe z. B. CHATTERJEE, HADI 1988) spielen die Hauptdiagonalelemente c_{ii} (i = 1, … ,n) eine zentrale Rolle für Analysezwecke. Für sie gilt

$$0 \le c_{ii} \le 1. \tag{3}$$

Im Sinne der Ausgleichungsrechnung überführt die Matrix **C** die ursprünglichen Beobachtungen $\mathbf{l} \in R^n$ in die nach der Methode der kleinsten Quadrate ausgeglichenen Beobachtungen $\hat{\mathbf{l}} \in R^n$:

$$\mathbf{C}\,\mathbf{l} = \hat{\mathbf{l}} \tag{4}$$

bzw. das inkonsistente System in (1) in das konsistente System

$$\mathbf{A}\,\mathbf{x} = \hat{\mathbf{l}}\,. \tag{5}$$

Die zu bestimmenden Parameter $\hat{\mathbf{x}} \in R^u$ erweisen sich somit als Koordinaten der ausgeglichenen Beobachtungen bezüglich der Basis von U, die durch die Spalten von **A** gebildet wird. Für die zu bestimmenden Verbesserungen $\hat{\mathbf{v}} \in R^n$ erhält man mit der n-dimensionalen Einheitsmatrix **E**

$$-\hat{\mathbf{v}} = (\mathbf{E} - \mathbf{C})\,\mathbf{l}\,. \tag{6}$$

Auch die Verbesserungen spielen bei der Sensitivitätsanalyse eine wichtige Rolle. Die Hauptdiagonalelemente der Matrix **E** – **C** (in der Ausgleichungsrechnung Teilredundanzen r_i genannt) widerspiegeln den Anteil der i-ten Beobachtung an der Gesamtredundanz n – u (Freiheitsgrad der Ausgleichung).
Aus mathematischer Sicht stellt sich dieser Kontext folgendermaßen dar. Zu jedem u-dimensionalen Unterraum $U \subset R^n$ gibt es ein eindeutig bestimmtes orthogonales Komplement $U^\perp$, das wiederum einen Unterraum des R^n mit der Kodimension n – u darstellt.
Die Vektoren aus $U^\perp$ stehen senkrecht auf allen Vektoren von U und umgekehrt.

Statt der Projektion auf U kann selbstverständlich auch die Projektion auf $U^{\perp}$ betrachtet werden. Man wählt dazu eine Basis in $U^{\perp}$ und bildet damit eine Matrix **B** vom Format $(n, n-u)$ (das orthogonale Komplement zu **A**). Die Orthogonalität der Unterräume U und $U^{\perp}$ drückt sich nun in folgender Matrizenrelation aus

$$\mathbf{B}^T\mathbf{A} = \mathbf{0} \quad \text{bzw.} \quad \mathbf{A}^T\mathbf{B} = \mathbf{0} \tag{7}$$

Damit geht das Modell (1) in das äquivalente Modell

$$\begin{aligned} \mathbf{B}^T(\mathbf{l} + \mathbf{v}) &= \mathbf{o} \\ \mathbf{v}^T\mathbf{v} &\rightarrow \min \end{aligned} \tag{8}$$

über, welches als Modell der bedingten Ausgleichung bekannt ist (WOLF, 1997).
Die Lösung des Modells (8) besteht nun in der Konstruktion von orthogonalen Projektionen von R^n auf $U^{\perp}$. Ganz analog zu (2) erhält man diese Projektionen in Matrizenform

$$\mathbf{E} - \mathbf{C} = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T \tag{9}$$

und die Lösung $\hat{\mathbf{v}} \in R^n$ aus (6). Aus der Orthogonalität der Unterräume ergeben sich unmittelbar die bekannten Relationen

$$\mathbf{A}^T\mathbf{v} = \mathbf{o} \;\Leftrightarrow\; \mathbf{B}^T(\mathbf{l} + \mathbf{v}) = \mathbf{o}, \tag{10}$$

die in der Ausgleichungsrechnung zur Verprobung von numerischen Ergebnissen benutzt werden. Andererseits stellen diese Relationen aber auch a-priori-Informationen dar, die über geometrische Besonderheiten bei den Beobachtungen schon vor der eigentlichen Ausgleichung wichtige Hinweise aufzeigen, wie aufgezeigt werden wird.
Im weiteren wird es nun darum gehen, geometrische Besonderheiten in den Unterräumen U bzw. $U^{\perp}$ aufzudecken bzw. zu erkennen. Dies wird zunächst durch transparente Beispiele dargelegt, die danach mathematisch verallgemeinert werden sollen.

Beispiel 1: Wir betrachten den Fall einer Geradenausgleichung mit n = 5 und u = 2 für das in (CHATTERJEE, HADI, 1988) angeführte Beispiel, wo diesbezüglich auch auf ältere Quellen (BEHNKEN, DRAPER, 1972, DRAPER, SMITH, 1981) hingewiesen wird.
Die Designmatrix **A** des Problems (1) sei dabei gegeben durch

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 4 \end{pmatrix}.$$

Die fünfte Beobachtung (fünfte Zeile in **A**) hat hierin einen erheblichen Einfluß auf das Ausgleichungsergebnis, denn nach Streichung dieser Zeile wird das Restdesign singulär (Spaltensingularität). Deutlich wird dies auch durch Betrachtung des orthogonalen Komplements $U^{\perp}$, dargestellt durch die Matrix **B** vom Format $(n, n-u)$

$$\mathbf{B} = \begin{pmatrix} -2 & -1 & -1 \\ -2 & 1 & -2 \\ -1 & 1 & -1 \\ 5 & -1 & 4 \\ 0 & 0 & 0 \end{pmatrix},$$

die $\mathbf{B}^T\mathbf{A} = \mathbf{0}$ bzw. $\mathbf{A}^T\mathbf{B} = \mathbf{0}$ erfüllt. Die Matrix **B** enthält als fünfte Zeile eine Nullzeile, wobei dieser Sachverhalt unabhängig von einer speziell gewählten Basis in $U^{\perp}$ ist. Aus Gleichung (9) ergibt sich unmittelbar

$$r_{5j} = (\mathbf{E} - \mathbf{C})_{5j} = 0.0 \quad \text{für } j = 1, \ldots, 5$$

d. h. die Teilredundanz r_{55} der fünften Beobachtung verschwindet. Somit wird diese Beobachtung von keiner der restlichen Beobachtungen kontrolliert. Aus Gleichung (6) ergibt sich weiterhin

$$\hat{v}_5 = 0.0$$

Das Verschwinden der fünften Verbesserung stellt sich unabhängig vom konkreten Beobachtungsvektor **l** ein, ist also rein geometrisch bedingt aus der Struktur des Unterraums U bzw. $U^{\perp}$. Ein Datenfehler in der fünften Beobachtung wird somit durch die anderen Beobachtungen nicht „aufgefangen" und verfälscht das Ergebnis in erheblicher Weise, die fünfte Beobachtung wird zur „Restriktion" = Bedingungsgleichung für die Unbekannten.
Die Ursache für diese geometrische Besonderheit liegt hierbei offensichtlich in der Tatsache, daß die ersten vier Beobachtungen (Zeilen in **A**) den Charakter von Mehrfachbeobachtungen aufweisen.
Bezeichnet man im Modell $\mathbf{Ax} = \mathbf{l} + \mathbf{v}$ mit $\mathbf{Hx} = \mathbf{w}$ die Bedingungsgleichungen zwischen den Unbekannten $\mathbf{Hx} = \mathbf{w}$ als „äußere Restriktionen", so können „innere Restriktionen" im Modell der vermittelnden Beobachtungen folgendermaßen definiert werden.

Definition 1: Eine Beobachtung (i-te Zeile in **A**) heiße ***innere Restriktion***, falls ihre Verbesserung v_i unabhängig vom Beobachtungsvektor $\mathbf{l} \in R^n$ stets Null ist.

Folgerung 1: Folgende Aussagen sind äquivalent:
1) Die i-te Zeile in **A** ist eine innere Restriktion.
2) In der Projektionsmatrix **C** gilt: $c_{ii} = 1$ und $c_{ij} = 0$ für $j \neq i$.
3) Der i-te Einheitsvektor $\mathbf{e}_i \in R^n$ gehört zum Unterraum U, d. h. eine Beobachtungsrichtung ist im Projektionsraum enthalten: $\mathbf{e}_i \in U$.

Beweis: Es ist $-v_i = ((\mathbf{E} - \mathbf{C})\,\mathbf{l})_i = 0 \;\forall\, \mathbf{l} \in R^n$ genau dann, wenn $(\mathbf{E} - \mathbf{C})_{ij} = 0$ für $j = 1, \ldots, n$ und damit $c_{ii} = 1$ sowie $c_{ij} = 0$ für $j \neq i$. Dies ist äquivalent zu $(\mathbf{E} - \mathbf{C})\,\mathbf{e}_i = 0$, also $\mathbf{C}\,\mathbf{e}_i = \mathbf{e}_i$, was genau dann gilt, wenn $\mathbf{e}_i \in U$.

Folgerung 2: Eine Beobachtung im orthogonalen Komplement (j-te Zeile in **B**) ist eine innere Restriktion, falls gilt:

$$r_{ji} = 0 \text{ und } r_{jj} = 1 \text{ für } i = 1, \ldots, n \text{ mit } i \neq j.$$

Dies ist äquivalent zu:
1) $v_i + l_i = 0$,
2) $c_{ji} = 0$ für $i = 1, \ldots, n$,
3) $a_{ji} = 0$ für $i = 1, \ldots, n$ und
4) $\mathbf{e}_j \in U^{\perp}$.

Anmerkung: Eine innere Restriktion im orthogonalen Komplement **B** entspricht in der Ausgleichungsrechnung einer vollredundanten Beobachtung in der Designmatrix **A** (Nullzeile in **A**).
Aus den bislang dargelegten Sachverhalten der inneren Restriktionen entstehen folgende weiterführende Überlegungen. Teilredundanzen von Beobachtungen, die sehr klein im Vergleich zu den Übrigen sind (nahe bei Null), weisen darauf hin, daß die zugehörige Beobachtung nur wenig durch die anderen Beobachtungen kontrolliert wird (z. B. High-Leverage-Points). Damit ergibt sich Frage nach einer möglichen Umkehrung dieses Umstandes.
Werden Beobachtungen mit großen Teilredundanzen im allgemeinen gut kontrolliert?
Es wird nachstehend aufgezeigt, daß dies nicht generell der Fall sein muß. Insbesondere bei der geodätischen Netzausgleichung spielen solche Aspekte eine zentrale Rolle.

Einen weiteren Zugang erhält man durch folgende Überlegungen: In Beispiel 1 wurde deutlich, daß durch Streichung einer inneren Restriktion die verbleibende Design-Matrix singulär wird. Damit erhebt sich folgerichtig die Frage, ob dieser Effekt sich auch durch die gemeinsame Streichung von mehr als nur einer Beobachtung erreichen läßt, obwohl keine dieser betreffenden Beobachtungen eine innere Restriktion an sich bildet. Dieser Umstand würde dann auch dazu führen, daß beim Verbleib einer einzigen der zu streichenden Beobachtungen diese dann zu einer inneren Restriktion werden würde.
Als Schlußfolgerung erhält man dann die Aussage, daß diese Beobachtung nur durch die schon gestrichenen (herausgelassenen) Beobachtungen kontrolliert wurde, jedoch nicht von den Übrigen. Bekannt ist auch, daß durch die Weglassung von gerade $n-u$ Beobachtungen die Verbleibenden zu inneren Restriktionen werden (vorausgesetzt, daß das verbleibende Design regulär ist), da der Freiheitsgrad verschwindet und daher keine Überbestimmung mehr vorliegt.
Der bisher dargelegte Umstand steht in engem Zusammenhang mit folgender bekannter Tatsache (siehe z. B. CHATTERJEE, HADI 1988): die Teilredundanzen sind in Bezug auf die Anzahl der Beobachtungen nicht abnehmende Funktionen. Durch Hinzufügung entsprechender Beobachtungen können sie also lediglich größere numerische Werte annehmen. Durch Streichung von Beobachtungen tritt demzufolge ein gegenläufiger Effekt ein. Damit ergibt sich wiederum die Frage, ob durch Streichung einer oder mehrerer (wenige) Beobachtungen eine innere Restriktion entstehen kann.
Vergleichbare Überlegungen lassen sich auch für das orthogonale Komplement $U^{\perp}$ bzw. **B** anstellen. Letztendlich führen sie auf die Fragestellung, ob sich durch Streichung von Beobachtungen im orthogonalen Komplement innere Restriktionen erzeugen lassen, oder gleichbedeutend damit vollredundante Beobachtungen in U bzw. **A** entstehen.
Interessant ist hierbei der Umstand, daß sich Streichungen im orthogonalen Komplement (bedingte Ausgleichung) äquivalent als Streichungen von Unbekannten in der vermittelnden Ausgleichung interpretieren lassen, was im weiteren noch dargelegt wird.

Beispiel 2: Gegeben sei eine Koeffizientenmatrix **A** mit 6 Beobachtungen und 3 Unbekannten und ein dazugehöriges orthogonales Komplement **B**.

$$\mathbf{A} = \begin{pmatrix} 12 & 5 & 3 \\ 8 & 1 & 2 \\ 4 & 1 & 0 \\ 6 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 1 & \frac{3}{2} & -\frac{3}{4} \\ -1 & \frac{1}{2} & \frac{1}{4} \\ -1 & -2 & 0 \\ -4 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix}$$

	1	2	3	4	5	6
c_{ii}	0.85000	0.51667	0.65000	0.31667	0.26667	0.40000
r_{ii}	0.15000	0.48333	0.35000	0.68333	0.73333	0.60000
r_{ii} (1)	0	0.33333	0.33333	0.66667	0.66667	
r_{ii} (2)		0.33333	0.33333	0.66667	0.66667	0
r_{ii} (3)	0.14634	0.36585	0.17073		0.73171	0.58536

Tabelle 1: Teilredundanzen

Tabelle 1 zeigt folgende Sachverhalte auf. Bei Streichung der 1. bzw. 6. Beobachtung (Zeilen in **A**) wird die verbleibende (korrespondierende) Beobachtung zur inneren Restriktion (Teilredundanz Null) und somit von den restlichen Beobachtungen nicht mehr kontrolliert. Der verbleibende Freiheitsgrad $(n-u) = 2$ wird auf die Beobachtungen 2 bis 5 aufgeteilt, und zwar unabhängig davon, welche der Beobachtungen 1 oder 6 gestrichen wird. Streicht man andererseits die 4. Beobachtung, so verringern sich die Teilredundanzen der 1. und 6. Beobachtung nur unwesentlich zu den Restlichen.

Streicht man dagegen beide Beobachtungen Nr. 1 und Nr. 6, so wird das Restdesign (verbleibende 4 Zeilen in **A** weisen nur noch den Rang 2 auf) singulär. Die Beobachtungen 1 und 6 bilden also eine Gruppe, die von den restlichen Beobachtungen nur ungenügend kontrolliert wird.
Man beachte jedoch, daß dieser Effekt sich weder in der Design-Matrix **A**, noch an den Teilredundanzen unmittelbar aufzeigen läßt. Die Teilredundanz der 6. Beobachtung beträgt immerhin 0.60 und scheint somit gut kontrolliert.
Im orthogonalen Komplement **B** erkennt man, daß dort die Zeilen 1 und 6 linear abhängig sind.

Beispiel 3: Gegeben seien für 6 Beobachtungen und 3 Unbekannte

$$\mathbf{A} = \begin{pmatrix} 15 & 4 & 3 \\ 1 & 2 & 0 \\ 5 & 2 & 2 \\ 6 & 1 & 1 \\ 2 & 0 & 1 \\ 4 & 1 & 0 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ -\frac{7}{4} & -1 & \frac{1}{4} \\ \frac{1}{2} & 0 & -\frac{3}{4} \\ -1 & -2 & 0 \\ -4 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix}$$

	1	2	3	4	5	6
c_{ii}	0.73585	0.77358	0.62264	0.22642	0.26415	0.37736
r_{ii}	0.26415	0.22642	0.37736	0.77358	0.73585	0.62264

Tabelle 2: Teilredundanzen und Hauptdiagonale des orthogonalen Projektionsoperators zum Beispiel 3

In der Matrix **A** treten folgende Effekte auf: Die Streichung von zwei der Beobachtungen aus den Zeilen 1, 2 und 3 führt dazu, daß die verbleibende Beobachtung zur inneren Restriktion wird. Wie aus der Tabelle 2 ersichtlich ist, sind dies die Beobachtungen mit den größten Hauptdiagonalelementen des orthogonalen Projektionsoperators. Ebenso bilden aber auch die Beobachtungen 1, 4 und 6 eine Gruppe. Werden zwei beliebige Beobachtungen davon gestrichen, so wird die verbleibende Beobachtung zur inneren Restriktion.
Wiederum sind diese Effekte weder in der Design-Matrix **A** noch im orthogonalen Komplement **B** bzw. anhand der Teilredundanzen verifizierbar.

Beispiel 4: Gegeben sei eine Koeffizientenmatrix **A** mit 6 Beobachtungen und 4 Unbekannten, ein dazugehöriges orthogonales Komplement **B** und die Hauptdiagonalelemente des orthogonalen Projektionsoperators **C**.

$$\mathbf{A} = \begin{pmatrix} 5 & 3 & -1 & 2 \\ 9 & 4 & -1 & 3 \\ 13 & 5 & -1 & 4 \\ 2 & 2 & 5 & 2 \\ 4 & 3 & 8 & 3 \\ 6 & 4 & 11 & 4 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 2 & \frac{1}{2} \\ -4 & -1 \\ 2 & \frac{1}{2} \\ \frac{1}{2} & 1 \\ -1 & -2 \\ \frac{1}{2} & 1 \end{pmatrix}$$

	1	2	3	4	5	6
c_{ii}	0.83333	0.33333	0.83333	0.83333	0.33333	0.83333
r_{ii}	0.16667	0.66667	0.16667	0.16667	0.66667	0.16667

Tabelle 3: Teilredundanzen und Hauptdiagonale des orthogonalen Projektionsoperators zum Beispiel 4

Hier läßt sich nun Folgendes feststellen: die Streichung einer beliebigen Beobachtung in **A** führt zu einer Restriktion, z. B. bei Streichung der ersten Beobachtung wird die zweite Beobachtung zur inne-

ren Restriktion. Im orthogonalen Komplement **B** treten die Beobachtungen 1, 2 und 3 sowie 4, 5 und 6 jeweils als Mehrfachbeobachtungen auf. Anhand des Projektionsoperators **C** erkennt man eine weitere Besonderheit. **C** weist in diesem Fall eine Blockstruktur auf

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^1 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^2 \end{pmatrix}$$

mit $c_{11}^1 = c_{11}^2 = \frac{5}{6}$, $c_{22}^1 = c_{22}^2 = \frac{2}{6}$, und $c_{33}^1 = c_{33}^2 = \frac{5}{6}$.
Eine Verallgemeinerung der bisher aufgezeigten Effekte führt zum Begriff der latenten inneren Restriktion.

Definition 2: Eine Gruppe von k Beobachtungen (Zeilen in **A**) ($i \le k \le n-u$) erzeugt eine *latente innere Restriktion* k-ter Ordnung, falls nach Streichung von beliebigen $k-1$ Beobachtungen dieser Gruppe die verbleibende zur inneren Restriktion wird.

Diese Definition gilt analog für das orthogonale Komplement (Zeilen in **B**), wobei hier $1 \le k \le u$ ist. Der bisher verwendete Begriff der Restriktion ist in der Definition für k = 1 mit enthalten.
Es stellt sich nun die Frage, wie solche latente inneren Restriktionen zu erkennen sind.
Dazu sei zunächst der Spezialfall gerade einer Überbestimmung mit n = 3 und u = 2 betrachtet. Die Designmatrix kann durch zwei linear unabhängige Spaltenvektoren $\mathbf{a}^i \in R^3$ dargestellt werden, und zwar

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \left(\mathbf{a}^1 \; \mathbf{a}^2 \right).$$

Damit ist der Unterraum U eine Ebene im R^3, die von den Spaltenvektoren $\mathbf{a}^1$ und $\mathbf{a}^2$ aufgespannt wird, d. h. $U = \text{span}(\mathbf{a}^1, \mathbf{a}^2) = \{\mathbf{z} \in R^3 \mid \mathbf{z} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2, \; x_{1,2} \in \Re\}$. Wie hinlänglich bekannt ist, kann dieselbe Ebene U auch in der sogenannten Hesse-Form (parameterfreie Form) mit Hilfe eines sogenannten Normalenvektors $\mathbf{n} \in R^3$ dargestellt werden und zwar durch $\mathbf{n} \in U^\perp$ mit $U = \{\mathbf{z} \in R^3 \mid \mathbf{n}^T\mathbf{z} = 0\}$. Es gilt:

$$\mathbf{n} = \mathbf{a}^1 \times \mathbf{a}^2 = \begin{pmatrix} a_{21}a_{32} - a_{22}a_{31} \\ a_{31}a_{12} - a_{32}a_{11} \\ a_{11}a_{22} - a_{21}a_{12} \end{pmatrix} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$$

und damit $U^\perp = \text{span}(\mathbf{n})$.
Wie ersichtlich ist, sind die Komponenten von **n** Determinanten von 2-reihigen Untermatrizen von **A**. Es gilt der i-te Einheitsvektor $\mathbf{e}_i \in U$ genau dann, wenn $\mathbf{n}^T\mathbf{e}_i = 0 \;\Leftrightarrow\; n_i = 0$ (siehe Folgerung 1). Dies ist gleichbedeutend damit, daß die i-te Zeile in **A** eine Restriktion ist.
Von großem Interesse ist nun, wie dies auf den allgemeinen Fall $n > u \ge 1$ mit $n, u \in N$ übertragen werden kann. Die Lösung findet man in einem Hilfsmittel der algebraischen Geometrie, den sogenannten Plücker-Koordinaten, die in der Literatur auch als Plücker-Graßmann-Koordinaten (VAN DER WAERDEN, 1973) bekannt sind. Diese werden nachfolgend beschrieben.

2.2 Über Plücker-Koordinaten

Im linearen Modell der vermittelnden Ausgleichung (1) wird von den Spalten der Design-Matrix **A** ein u-dimensionaler Unterraum U des R^n aufgespannt. Man kann deshalb die u Spaltenvektoren der Länge n als Basis in U wählen (rg **A** = u)

$$\mathbf{A} = (\mathbf{a}^1, \ldots, \mathbf{a}^u) \;\text{ mit }\; \mathbf{a}^1, \ldots, \mathbf{a}^u \in R^n, \;\; U = \text{span}(\mathbf{a}^1, \ldots, \mathbf{a}^u).$$

Ebenso läßt sich die Matrix **A** durch ihre Zeilenvektoren der Länge u darstellen

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{pmatrix}$$

mit $\mathbf{a}_1, \ldots, \mathbf{a}_n \in R^u$. Wählt man nun u Zeilen aus **A**, sei also $\{i_1, \ldots, i_u\} \subset \{1, 2, \ldots, n\}$ eine Auswahl paarweise verschiedener Indizes, so bildet

$$\mathbf{A}(i_1, \ldots, i_u) = \begin{pmatrix} \mathbf{a}_{i_1}^T \\ \mathbf{a}_{i_2}^T \\ \vdots \\ \mathbf{a}_{i_u}^T \end{pmatrix}$$

eine quadratische Teilmatrix von **A**. Die zugehörige u-reihige Unterdeterminante d von **A**

$$d(i_1, \ldots, i_u) = \det \mathbf{A}(i_1, \ldots, i_u)$$

entspricht dann einer Plücker-Koordinate von U. Es gilt die folgende Aussage (VAN DER WAERDEN, 1973).

Satz 1: Die $\binom{n}{u}$ Plücker-Koordinaten von U sind homogene Koordinaten, die den Unterraum U eindeutig bestimmen.

Homogene Koordinaten sind nur bis auf einen Faktor c bestimmt und können nicht alle gleich Null sein. Die Eigenschaft der Homogenität der Plücker-Koordinaten wird deutlich, wenn man im Unterraum U einen Basiswechsel vornimmt. Dieser kann mittels einer regulären (u,u) Matrix **W** durch $\overline{\mathbf{A}} = \mathbf{A} \cdot \mathbf{W}$ realisiert werden. Für die neue Basis $\overline{\mathbf{A}}$ in U ergeben sich die Plücker-Koordinaten zu $\det \overline{\mathbf{A}}(i_1, \ldots, i_u) = \det \mathbf{W} \cdot \det \mathbf{A}(i_1, \ldots, i_u)$, sie unterscheiden sich damit von den Plücker-Koordinaten der Basis **A** nur um den konstanten Faktor det **W**.
Durch die Plücker-Koordinaten von U sind gleichzeitig die Plücker-Koordinaten von $U^\perp$ bestimmt und umgekehrt. Sei dazu $\Pi = (j_1, \ldots, j_n)$ eine Permutation von $\{1, \ldots, n\}$. Dann gilt folgender Zusammenhang zwischen den Plücker-Koordinaten von U, gebildet aus u Zeilen von **A** und den Plücker-Koordinaten von $U^\perp$, gebildet aus n – u Zeilen von **B**, Beweis siehe (VAN DER WAERDEN, 1973),

$$d(j_1, \ldots, j_u) = \operatorname{sgn} \Pi \, d^\perp(j_{u+1}, \ldots, j_n). \tag{11}$$

Im Beispiel 1 ist durch die Mehrfachbeobachtungen sofort ersichtlich, daß alle Plücker-Koordinaten von **A** verschwinden,

$$\det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0,$$

die die fünfte Beobachtung (die Restriktion im herkömmlichen Sinne) nicht enthalten. Im orthogonalen Komplement **B** werden alle Plücker-Koordinaten zu Null, die die fünfte Beobachtung (Nullzeile) enthalten, z. B.

$$\det \begin{pmatrix} -2 & -1 & -1 \\ -2 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} = 0.$$

Im Beispiel 2 werden alle Determinanten der (u,u)-Teilmatrizen von **A** zu Null, die drei der Beobachtungen 2, 3, 4 und 5 enthalten, d. h. die die beiden Beobachtungen nicht enthalten, die eine latente

innere Restriktion zweiter Ordnung erzeugen. Im orthogonalen Komplement **B** verschwinden alle Plücker-Koordinaten, die die Mehrfachbeobachtungen 1 und 6 enthalten.
Im dritten Beispiel werden die Plücker-Koordinaten der Teilmatrix von **A** zu Null, die aus den Beobachtungen 4, 5 und 6 bzw. aus den Beobachtungen 2, 3 und 5 gebildet wird. Im orthogonalen Komplement **B** verschwinden die beiden Plücker-Koordinaten, die die Beobachtungen 1, 2 und 3 oder 1, 4 und 6 enthalten, also jene, die eine latente Restriktion 3. Ordnung in **A** erzeugen.
Im Beispiel 4 werden alle Determinanten der (u,u)-Teilmatrizen von **A** zu Null, die gleichzeitig die Beobachtungen 1, 2, und 3 oder 4, 5 und 6 enthalten. Im orthogonalen Komplement **B** verschwinden alle Plücker-Koordinaten, die zwei der Mehrfachbeobachtungen von 1, 2 und 3 oder aus 4, 5 und 6 enthalten.
Anhand dieser Beispiele wird deutlich, daß verschwindende Plücker-Koordinaten die Ursache für latente innere Restriktionen sind. Ziel ist deshalb nun die systematische Aufdeckung von verschwindenden Plücker-Koordinaten.

Als Hilfsmittel können dazu ***Indexmengen*** benutzt werden, wie sie z. B. in (FINZEL, 1994) ausführlich beschrieben worden sind.

Es sei I_1 die Menge aller Indizes derjenigen Beobachtungen, die zum Verschwinden aller Plücker-Koordinaten führt, die eine dieser Beobachtungen enthalten:

$$I_1 = \{(i): i \in \{1, \dots, n\}, d(i,N) = 0 \ \forall N \subset \{1, \dots, n\}, \text{card } N = u-1\}.$$

Es sei I_2 die Menge aller Indizes derjenigen Paare von Beobachtungen, deren Indizes paarweise verschieden sind, jeweils nicht zu I_1 gehören und zum Verschwinden aller Plücker-Koordinaten führt, die ein Paar dieser Beobachtungen enthalten:

$$I_2 = \{(i_1,i_2): i_1,i_2 \in \{1, \dots, n\}, i_1 \neq i_2, (i_1), (i_2) \notin I_1, d(i_1,i_2,N) = 0 \ \forall N \subset \{1, \dots, n\}, \text{card } N = u-2\}.$$

Allgemein sei I_m die Menge aller Gruppen von m ($2 \leq m \leq u$) Beobachtungen, deren Indizes paarweise verschieden sind, jeweils nicht zu I_1 bis I_{m-1} gehören und zum Verschwinden aller Plücker-Koordinaten führt, die eine Gruppe von m dieser Beobachtungen enthalten:

$$I_m = \{(i_1, \dots, i_m): \ i_1, \dots, i_m \in \{1, \dots, n\}, \text{ paarweise verschieden}, \ \forall(i'_1, \dots, i'_\mu) \in I_\mu, \ 1 \leq \mu < m: \{i'_1, \dots, i'_\mu\} \not\subset \{i_1, \dots, i_m\}, d(i_1, \dots, i_m, N) = 0 \ \forall N \subset \{1, \dots, n\}, \text{card } N = u-m\}.$$

Nach (FINZEL, 1994) kann eine äquivalente Charakterisierung vorgenommen werden.

Satz 2: Folgende Aussagen sind äquivalent:
1) $(i_1, \dots, i_m) \in I_m$
2) Die Matrix (Teilmatrix von **A**)

$$\mathbf{A}(i_1, \dots, i_m) = \begin{pmatrix} \mathbf{a}_{i_1}^T \\ \mathbf{a}_{i_2}^T \\ \vdots \\ \mathbf{a}_{i_m}^T \end{pmatrix}$$

   ist vom Rang $m-1$ und $m-1$ beliebige Zeilen sind linear unabhängig.
3) Die Dimension des Durchschnitts gebildet aus dem Unterraum $U^\perp$ und dem Raum, der von den m Einheitsvektoren aufgespannt wird, ist $\dim(U^\perp \cap \text{span}\{\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_m}\}) = 1$,
   und falls der Vektor **w** in diesem Durchschnitt (außer dem trivialen Nullvektor) enthalten ist, also $\mathbf{w} \in U^\perp \cap \text{span}\{\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_m}\} \backslash \{\mathbf{o}\}$, dann gilt: $w_i \neq 0 \ \forall \ i \in \{i_1, \dots, i_m\}$.

Folgerung 3: Analoge Indexmengen können für $U^\perp$ ($\dim U^\perp = n-u$) aufgestellt werden:

$$J_m \text{ mit } 1 \leq m \leq n-u.$$

Satz 3:

1) Zur Indexmenge J_m gehören die Indizes von m Beobachtungen ($1 \le m \le n-u$) des orthogonalen Komplements genau dann, wenn die Zeilen $j_1, \dots, j_m$ eine latente Restriktion m-ter Ordnung in **A** bilden.
2) Ebenso gilt $(i_1, \dots, i_m) \in I_m$ ($1 \le m \le u$) genau dann, wenn die Zeilen $i_1, \dots, i_m$ eine latente Restriktion m-ter Ordnung in **B** bilden.

Beweis: 1) Aus 3) in Satz 2 folgt, daß es einen Vektor $\mathbf{w} \in U$ gibt mit $\mathbf{w} = \sum_{j=1}^{m} \alpha_{j_i} \mathbf{e}_{j_i}$, wobei alle $\alpha_{j_i} \neq 0$ sind. **w** kann dann dargestellt werden als

$$\mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix},$$

wobei $w_i = 0$ ist, genau dann, wenn $i \notin \{j_1, \dots, j_m\}$ ist. Durch Basistausch kann **w** in einer Matrix $\overline{\mathbf{A}}$ realisiert werden. Durch Streichen von beliebigen $m-1$ Zeilen der Gruppe $\{j_1, \dots, j_m\}$ entsteht stets eine Spalte mit genau einem Nichtnullelement, also eine innere Restriktion.
2) Für latente Restriktionen m-ter Ordnung im orthogonalen Komplement vollzieht sich der Beweis analog 1).

Für die oben beschriebenen Indexmengen ergibt sich nun im Einzelnen

- I_1 entspricht einer Nullzeile in **A** (= Vollredundanz in **A** und Restriktion in **B**),
- J_1 entspricht einer Nullzeile in **B** (= Vollredundanz in **B** und Restriktion in **A**),
- I_2 entspricht einer Mehrfachbeobachtung in **A** (= latente Restriktion 2. Ordnung in **B**),
- J_2 entspricht einer Mehrfachbeobachtung in **B** (= latente Restriktion 2. Ordnung in **A**),
- I_3 entspricht einer latenten Restriktion 3. Ordnung in **B**,
- J_3 entspricht einer latenten Restriktion 3. Ordnung in **A**,

usw.

Für das Beispiel 1 mit

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{und} \quad \mathbf{B} = \begin{pmatrix} -2 & -1 & -1 \\ -2 & 1 & -2 \\ -1 & 1 & -1 \\ 5 & -1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

sollen jetzt die Indexmengen angegeben werden. In **A** gibt es keine Beobachtung, die zum Verschwinden aller 2-reihigen Unterdeterminanten führt, die diese eine Beobachtung enthalten. Deshalb ist I_1 gleich der leeren Menge, also $I_1 = \varnothing$. In **B** gibt es genau eine Beobachtung, nämlich die fünfte Beobachtung, die zum Verschwinden aller 3-reihigen Unterdeterminanten führt, die diese Beobachtung enthalten. Somit enthält die Indexmenge J_1 als einziges Element die 5, also ist $J_1 = \{5\}$.
In der Matrix **A** verschwinden alle Plücker-Koordinaten, wenn man aus den ersten vier Zeilen zwei beliebige Beobachtungen auswählt. Die Indexmenge I_2 enthält die Indizes aller Paare dieser Beobachtungen, damit ist $I_2 = \{(1, 2, 3, 4)\}$. In **B** läßt sich kein Paar von Beobachtungen finden, so daß alle 3-reihigen Unterdeterminanten verschwinden, die dieses Paar von Beobachtungen enthalten. Es lassen sich auch keine drei Beobachtungen aus **B** auswählen, so daß alle Plücker-Koordinaten (auf Grund der Dimensionen ist dies jeweils nur eine) verschwinden, die diese drei Beobachtungen enthalten. Somit sind die Indexmengen J_2 und J_3 leer, $J_2 = J_3 = \varnothing$.

Die Indexmengen für die anderen Beispiele lauten wie folgt:

Bsp. 2: **A**: $I_1 = I_2 = \varnothing$, $I_3 = \{(2, 3, 4, 5)\}$
B: $J_1 = \varnothing$, $J_2 = \{(1, 6)\}$, $J_3 = \varnothing$
Bsp. 3: **A**: $I_1 = I_2 = \varnothing$, $I_3 = \{(4, 5, 6), (2, 3, 5)\}$
B: $J_1 = J_2 = \varnothing$, $J_3 = \{(1, 2, 3), (1, 4, 6)\}$
Bsp. 4: **A**: $I_1 = I_2 = \varnothing$, $I_3 = \{(4, 5, 6), (1, 2, 3)\}$, $I_4 = \varnothing$
B: $J_1 = \varnothing$, $J_2 = \{(1, 2, 3), (4, 5, 6)\}$

Aus den obigen Beispielen wird weiterhin ersichtlich, daß die Indexmengen von **A** und **B** nicht unabhängig sind, sondern sich sogar unmittelbar bedingen. Die Ursache hierfür liegt in der Beziehung (11) zwischen den Plücker-Koordinaten für U und $U^{\perp}$. Wir erläutern dies kurz anhand von Beispiel 2: Aus $J_2 = \{(1, 6)\}$ folgt, daß alle Plücker-Koordinaten von **A** verschwinden, die die Beobachtungen 1 und 6 nicht enthalten, also $I_3 = \{(2, 3, 4, 5)\}$ (beachte, I_3 besteht aus all den Dreierkombinationen von (2, 3, 4, 5)).
Der Nachteil der Verwendung von Plücker-Koordinaten ist der praktisch nicht vertretbare hohe Aufwand, um diese zu berechnen. Von großem Interesse ist es daher, ob es eine numerisch günstigere Möglichkeit gibt, die Indexmengen und damit die verschwindenden Plücker-Koordinaten zu bestimmen. Die Lösung bietet sich in der im folgenden Abschnitt dargestellten Normalform.

2.3 Erkennung innerer latenter Restriktionen anhand der Normalform

In (JURISCH, KAMPMANN, 1998) wurde der Begriff der Normalform eingeführt und wichtige Eigenschaften dieser Normalform aufgezeigt. Im weiteren wird dargelegt, wie Zusammenhänge zwischen der Normalform und den Plücker-Koordinaten bzw. den Indexmengen, die den inneren latenten Restriktionen entsprechen, hergestellt werden können.

Satz 4: In jedem u-dimensionalen Unterraum U des R^n gibt es eine Basis der Form:

$$\overline{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{a}}_1^T \\ \vdots \\ \tilde{\mathbf{a}}_n^T \end{pmatrix} \quad \text{mit} \quad \tilde{\mathbf{a}}_{i_j} = \mathbf{e}_{i_j} \quad j = 1, \ldots, u\,. \tag{12}$$

Die Zeilen von $\overline{\mathbf{A}}$ beinhalten die Zeilen einer u-dimensionalen Einheitsmatrix.

Beweis: Es sei U durch

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{pmatrix}$$

dargestellt. Da $\operatorname{rg}\mathbf{A} = u$ vorausgesetzt wurde, gibt es in **A** u linear unabhängige Zeilen. Diese seien $\mathbf{a}_{i_j}^T$, $j = 1, \ldots, u$. Faßt man diese Zeilen zu einer (u, u) Matrix **W** zusammen, d. h.

$$\mathbf{W} = \begin{pmatrix} \mathbf{a}_{i_1}^T \\ \vdots \\ \mathbf{a}_{i_u}^T \end{pmatrix} \tag{13}$$

und führt einen Basiswechsel in U durch Rechtsmultiplikation mit $\mathbf{W}^{-1}$ durch, so ergibt sich: $\overline{\mathbf{A}} = \mathbf{A} \cdot \mathbf{W}^{-1}$, wobei $\overline{\mathbf{A}}$ die durch (12) vorgegebene Eigenschaft besitzt.

Folgerung 4: Durch Zeilenvertauschung in **A** kann stets erreicht werden, daß gilt:

$$\overline{\mathbf{A}} = \begin{pmatrix} \mathbf{E} \\ \tilde{\mathbf{A}} \end{pmatrix} \tag{14}$$

mit der (u,u) Einheitsmatrix $\mathbf{E}$ und der $(n-u,u)$ Matrix $\tilde{\mathbf{A}}$. Die notwendigen Zeilenvertauschungen lassen sich dabei durch eine Permutation der Menge $\{1, \dots, n\}$ darstellen:

$$\{i_1, \dots, i_n\} = \Pi \{1, \dots, n\} \tag{15}$$

Die u Beobachtungen, die den ersten u Indizes entsprechen, also zur Erzeugung der Einheitsmatrix in (14) dienen, werden wir nachstehend als ***Beobachtungsbasis*** bezeichnen, mit anderen Worten: diese Beobachtungen bilden gleichfalls eine Basis im u-dimensionalen Zeilenraum von $\mathbf{A}$.
Neben dem Modell (1) kann nun in äquivalenter Weise das Modell

$$\overline{\mathbf{A}}\,\mathbf{y} = \bar{\mathbf{l}} + \mathbf{w} \qquad \mathbf{w}^T\mathbf{w} \rightarrow \min \tag{16}$$

(wobei $\bar{\mathbf{l}} = \Pi(\mathbf{l})$ den entsprechend der Permutation (15) umsortierten Beobachtungsvektor darstellt) betrachtet werden. Die Äquivalenz der Modelle (1) und (16) drückt sich dabei in folgendem Sachverhalt aus:

Folgerung 5: Für die Lösungen $\hat{\mathbf{x}}, \hat{\mathbf{v}}$ aus (1) und $\hat{\mathbf{y}}, \hat{\mathbf{w}}$ aus (16) gilt

$$\Pi(\hat{\mathbf{v}}) = \hat{\mathbf{w}}, \quad \hat{\mathbf{x}} = \mathbf{W}^{-1}\hat{\mathbf{y}} \tag{17}$$

Beweis: siehe (JURISCH, KAMPMANN, 1998).

Die Modelle (1) und (16) besitzen also dieselben Verbesserungen, die entsprechenden Parameter hängen über die Basistransformation zusammen. In (JURISCH, KAMPMANN, 1998) wurde auch der wichtige Umstand beschrieben, daß man mit der Normalform $\overline{\mathbf{A}}$ für U unmittelbar eine Basis für das orthogonale Komplement $U^{\perp}$ besitzt:

$$\overline{\mathbf{B}} = \begin{pmatrix} \tilde{\mathbf{A}}^T \\ -\mathbf{E} \end{pmatrix} \tag{18}$$

mit der $(n-u, n-u)$-dimensionalen Einheitsmatrix $\mathbf{E}$. Man beachte die offensichtliche Tatsache, daß

$$\overline{\mathbf{B}}^T\,\overline{\mathbf{A}} = \overline{\mathbf{A}}^T\,\overline{\mathbf{B}} = \mathbf{0}$$

gilt. Zusätzlich gilt der Umstand, daß die Indexvertauschungen in $\mathbf{A}$, die durch (15) beschrieben wurden, in gleicher Weise auf das ursprüngliche orthogonale Komplement $\mathbf{B}$ anzuwenden sind. Die Zeilen von $\mathbf{B}$, die der Matrix $\mathbf{E}$ in (18) entsprechen, bilden also eine Beobachtungsbasis im orthogonalen Komplement. Damit ergibt sich der nachstehende Sachverhalt:

Satz 5: Sei $\{i_1, \dots, i_n\} = \Pi \{1, \dots, n\}$ eine Permutation. Dann gilt: $\{i_1, \dots, i_u\}$ bildet eine Beobachtungsbasis in $U \Leftrightarrow \{i_{u+1}, \dots, i_n\}$ bildet eine Beobachtungsbasis in $U^{\perp}$.

Anders ausgedrückt bedeutet dies, daß mit u linear unabhängigen Zeilen aus $\mathbf{A}$ die entsprechenden komplementären Zeilen aus $\mathbf{B}$ auch linear unabhängig sind und umgekehrt. Aus (10) und der Tatsache, daß die Modelle (1) und (16) gleiche Verbesserungen aufweisen, ergibt sich unmittelbar folgende Beziehung:

$$\mathbf{w}^1 = -\tilde{\mathbf{A}}^T \mathbf{w}^2 \quad \Leftrightarrow \quad \mathbf{w}^2 + \bar{\mathbf{l}}^2 = \tilde{\mathbf{A}}(\mathbf{w}^1 + \bar{\mathbf{l}}^1) \tag{19}$$

Hierin sind $\mathbf{w}^1$ und $\bar{\mathbf{l}}^1$ die Komponenten von **v**, **l**, die der Beobachtungsbasis von **A** entsprechen sowie $\mathbf{w}^2$ und $\bar{\mathbf{l}}^2$ diejenigen in der komplementären Beobachtungsbasis in **B**. Zunächst gilt der nachstehende Umstand:

Satz 6: Die Zeilen $\{i_1, \dots, i_u\}$ aus **A** bilden genau dann eine Beobachtungsbasis, falls ihre Plücker-Koordinate nicht verschwindet,

$$d(i_1, \dots, i_u) \neq 0. \tag{20}$$

Wichtig in diesem Zusammenhang ist auch nachstehende Aussage.

Satz 7: Jedes Element $\tilde{a}_{ij}$ in der Matrix $\tilde{\mathbf{A}}$ aus (14) stellt eine Plücker-Koordinate dar.

Beweis: Wir indizieren die Matrix $\tilde{\mathbf{A}}$ durch $\tilde{\mathbf{A}} = (\tilde{a}_{ij})_{i=u+1,\dots,n}^{j=1,\dots,u}$. Wählt man nun ein Element $\tilde{a}_{ij}$ aus $\tilde{\mathbf{A}}$ aus, dann gilt:

$$d(1, \dots, j-1, j+1, \dots, u, i) = (-1)^{u+j} \tilde{a}_{ij} \quad i = u+1, \dots, n, \;\; j = 1, \dots, u\,. \tag{21}$$

d. h. die entsprechende Plücker-Koordinate ergibt sich, indem man die i-te Zeile aus $\tilde{\mathbf{A}}$ und alle Zeilen aus der Einheitsmatrix **E** außer der j-ten Zeile auswählt. Dies ist eine direkte Folge aus dem Entwicklungssatz von Laplace.

Folgerung 6: Jede Plücker-Koordinate, die von mehr als einer Zeile von $\tilde{\mathbf{A}}$ erzeugt wird, läßt sich als Funktion der $\tilde{a}_{ij}$ darstellen.

Auch dies ergibt sich durch sukzessive Anwendung des Entwicklungssatzes von Laplace. Die entstehenden Relationen zwischen den Plücker-Koordinaten werden in der Literatur auch als Plückerrelationen bezeichnet (VAN DER WAERDEN, 1973).
Die Spalten der Normalform in $\overline{\mathbf{A}}$ bzw. $\overline{\mathbf{B}}$ bilden eine spezielle Basis in U bzw. $U^{\perp}$, die in der Literatur auch als Plücker-Graßmann-Basis bezeichnet wird (VAN DER WAERDEN, 1973). Insbesondere erkennt man an Null-Elementen in $\tilde{\mathbf{A}}$ entsprechende verschwindende Plücker-Koordinaten, wodurch eine Verbindung zu den Indexmengen bzw. den latenten Restriktionen hergestellt werden kann. Genauer läßt sich dies folgendermaßen ausdrücken:

Satz 8: Die Zeilen $\{i_1, \dots, i_m\}$ von **A** bilden eine latente innere Restriktion der Ordnung m genau dann, wenn in jeder Beobachtungsbasis von **A** mindestens eine dieser Beobachtungen enthalten sein muß.

Beweis: Die Zeilen $\{i_1, \dots, i_m\}$ von **A** stellen genau dann eine latente innere Restriktion dar, falls die mit diesen Indizes gebildete Teilmatrix aus dem orthogonalen Komplement **B** den Rang $m-1$ besitzt und beliebige $(m-1)$ Zeilen linear unabhängig sind. Demzufolge können in jeder Beobachtungsbasis von **B** nur höchstens $(m-1)$ dieser Indizes auftreten. Die Aussage ergibt sich nun unmittelbar aus der Komplementarität der Beobachtungsbasen in **A** und **B**.

Durch Anwendung von Satz 8 auf die Normalform ergibt sich nun eine effektive Möglichkeit, latente Restriktionen in **A** aufzudecken.

Satz 9: Sei $\{i_1, \dots, i_m\}$ eine latente Restriktion in **A** und die Zeile i_1 in der Beobachtungsbasis zur Erzeugung von $\overline{\mathbf{A}}$ enthalten, die Zeilen $i_2, \dots, i_m$ nicht (sie erzeugen Zeilen in $\tilde{\mathbf{A}}$). Dann sind alle Elemente der i_1-ten Spalte in $\tilde{\mathbf{A}}$ Null, die nicht zu den Indizes $i_2, \dots, i_m$ gehören.

Beweis: Wie in Beweis von Satz 8 bilden wir die entsprechende Teilmatrix aus **B**. Diese enthält $m-1$ Zeilen der Einheitsmatrix **E** und eine Zeile aus $\tilde{\mathbf{A}}^T$. Die Aussage des Satzes ergibt sich jetzt direkt aus den Eigenschaften dieser Teilmatrix nach Satz 2.

Wir vermerken ausdrücklich, daß der Umkehrschluß, nämlich aus Nullelementen in den Spalten von $\tilde{\mathbf{A}}$ auf latente Restriktionen zu schließen nur in soweit richtig ist, daß die hiernach ermittelten latenten Restriktionen noch in latente Restriktionen niedrigerer Ordnung zerfallen können.
Selbstverständlich lassen sich die obigen Darlegungen auch auf Restriktionen im orthogonalen Komplement **B** übertragen. Diese zeigen sich dann anhand der Nullelemente der Spalten von $\tilde{\mathbf{A}}^T$, also den Zeilen von $\tilde{\mathbf{A}}$.
Zum Zwecke der transparenten Darlegung der beschriebenen Effekte werden hier die Beispiele untersucht.

Beispiel 1:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{und} \quad \mathbf{B} = \begin{pmatrix} -2 & -1 & -1 \\ -2 & 1 & -2 \\ -1 & 1 & -1 \\ 5 & -1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

Man erkennt unmittelbar, daß jede Beobachtungsbasis $\{i_1, i_2\}$ in **A** die fünfte Zeile (Beobachtung) enthalten muß. Umgekehrt enthält keine denkbare Beobachtungsbasis in **B** die fünfte Zeile. Wählt man etwa $\{1,5\}$ als Beobachtungsbasis aus, so ergibt sich:

$$\mathbf{W} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \quad \Rightarrow \quad \overline{\mathbf{A}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \\ 3 \\ 4 \\ 2 \end{pmatrix} \quad \text{und} \quad \overline{\mathbf{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

so daß anhand der Nullelemente in der zweiten Spalte von $\overline{\mathbf{A}}$ sofort die fünfte Beobachtung als Restriktion erkannt wird (die zweite Spalte von $\tilde{\mathbf{A}}$ enthält lediglich Nullelemente).

Beispiel 2:

$$\mathbf{A} = \begin{pmatrix} 12 & 5 & 3 \\ 8 & 1 & 2 \\ 4 & 1 & 0 \\ 6 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 1 & \frac{3}{2} & -\frac{3}{4} \\ -1 & \frac{1}{2} & \frac{1}{4} \\ -1 & -2 & 0 \\ -4 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix}$$

Wählt man als Beobachtungsbasis $\{1,2,3\}$, so erhält man

$$\overline{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} \qquad \overline{\mathbf{B}} = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{3}{4} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{4} \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

und erkennt die latente Restriktion $(1,6)$ der Ordnung 2 anhand der Nullelemente in der ersten Spalte von $\overline{\mathbf{A}}$. Würde man jedoch als Beobachtungsbasis $\{1,2,6\}$ wählen, so ergibt sich

$$\overline{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 2 & 2 \\ 1 & -1 & -2 \\ -2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 6 \\ 4 \\ 5 \\ 3 \end{pmatrix} \qquad \overline{\mathbf{B}} = \begin{pmatrix} -1 & 1 & 2 \\ 2 & -1 & 3 \\ 2 & -2 & 4 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

d. h. die latente Restriktion $(1,6)$ zeigt sich nicht mehr anhand von Nullelementen in $\tilde{\mathbf{A}}$, sondern durch die paarweise lineare Abhängigkeit der Spalte 1 und Spalte 3 in $\tilde{\mathbf{A}}$ bzw. der linearen Abhängigkeit der Zeilen 1 und 3 im orthogonalen Komplement $\overline{\mathbf{B}}$.

Beispiel 3:

$$\mathbf{A} = \begin{pmatrix} 15 & 4 & 3 \\ 1 & 2 & 0 \\ 5 & 2 & 2 \\ 6 & 1 & 1 \\ 2 & 0 & 1 \\ 4 & 1 & 0 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ -\frac{7}{4} & -1 & \frac{1}{4} \\ \frac{1}{2} & 0 & -\frac{3}{4} \\ -1 & -2 & 0 \\ -4 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix}$$

Anhand der Indexmengen erkennt man die latenten Restriktionen $(1,2,3)$ und $(1,4,6)$. Wählt man als Beobachtungsbasis z. B. $\{1,2,3\}$ so ergibt sich

$$\overline{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & -\frac{3}{4} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} \qquad \overline{\mathbf{B}} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{3}{4} \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Die latente Restriktion $(1,4,6)$ zeigt sich in der ersten Spalte von $\overline{\mathbf{A}}$, weil nur die Beobachtung Nr. 1 in der Beobachtungsbasis vorhanden ist. Die latente Restriktion $(1,2,3)$ zeigt sich allerdings nicht direkt. Sie äußert sich jedoch in Form von linearen Abhängigkeiten in $\tilde{\mathbf{A}}$ (die Summe der drei Spalten in $\tilde{\mathbf{A}}$ erzeugt den Nullvektor). Beide latenten Restriktionen würden sich nur dann direkt zeigen, wenn man als Beobachtungsbasis etwa $\{3,4,5\}$ wählen würde.

Folgerung 7: Enthält die Matrix $\mathbf{A}$ insgesamt u latente Restriktionen, so zeigt sich dies stets in der Normalform anhand von Nullelementen.

Beweis: Da jede Beobachtungsbasis mindestens eine der latenten Restriktionen enthalten muß, jedoch nur u Beobachtungen dazu gehören, muß von jeder latenten Restriktion genau eine in der Beobachtungsbasis enthalten sein. Die Aussage folgt damit aus Satz 9.

Bei den bisherigen Betrachtungen wurde davon ausgegangen, daß durch Streichung einer gewissen Anzahl von Beobachtungen aus dem Design **A** eine verbleibende Beobachtung zur Restriktion wird. Das folgende Beispiel zeigt jedoch, daß dabei auch mehr als eine Restriktion entstehen kann.

Beispiel 5:

$$\overline{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

Durch Streichung der letzten Beobachtung werden die erste und die zweite Beobachtung zur Restriktion. Anhand der Indexmengen erhält man folgende latente Restriktion in Beispiel 5: $(1,5)$ und $(2,5)$, die über die gemeinsame Beobachtung 5 miteinander verbunden (verkettet) sind. Dies führt nun zur folgenden Verallgemeinerung:

Definition 3: Mehrere latente innere Restriktionen der Ordnung m heißen verkettet, falls sie $(m-1)$ gemeinsame Indizes enthalten.

Folgerung 8: Es möge eine Gruppe von k latenten inneren Restriktionen der Ordnung m existieren. Durch entsprechende Zeilen- und Spaltenvertauschungen in der Designmatrix **A** kann dann stets erreicht werden, daß in einer Normalform zu **A** gilt:

$$\overline{\mathbf{A}} = \begin{pmatrix} \mathbf{E} \\ \tilde{\mathbf{A}} \end{pmatrix}, \quad \tilde{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \mathbf{0} & \tilde{\mathbf{A}}_{22} \end{pmatrix} \quad \text{mit} \quad \tilde{\mathbf{A}}_{11}(m-1,k) \tag{22}$$

Beweis: Wir wählen eine Beobachtungsbasis so, daß die gemeinsamen Indizes, die zur Verkettung führen, nicht darin enthalten sind. Demzufolge gehört zur Beobachtungsbasis von jeder dieser verketteten latenten Restriktionen genau eine zur Beobachtungsbasis. Die Struktur (22) ergibt sich damit aus Satz 9.

Folgerung 8 gestattet nun auch eine Erkennung latenter innerer Restriktionen schon anhand des ursprünglichen Designs von **A**.

Folgerung 9: Die Design-Matrix **A** weise nachstehende Blockstruktur auf (eventuell erst nach entsprechender Zeilen- und Spaltenvertauschung)

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix} \quad \text{mit} \quad \mathbf{A}_{11}(n_1,u_1),\ \operatorname{rg}\mathbf{A}_{22} = u_2 = u - u_1 \tag{23}$$

Dann gibt es u_1 verkettete, latente innere Restriktionen höchstens der Ordnung (n_1-u_1+1).

Beweis: Für die Teilmatrix $\mathbf{A}_{11}$ in (23) gilt: $\operatorname{rg}\mathbf{A}_{11} = u_1$, da stets $\operatorname{rg}\mathbf{A} = u$ vorausgesetzt wird.
Die Erzeugung der Normalform $\overline{\mathbf{A}}$ kann nun in zwei aufeinanderfolgenden Schritten durchgeführt werden. Zunächst werde durch die ersten u_1 Spalten in **A** eine (u_1,u_1) Einheitsmatrix $\mathbf{E}_1$ erzeugt und damit folgendes Teilergebnis erzielt:

$$\overline{\mathbf{A}}_1 = \begin{pmatrix} \mathbf{E}_1 & \mathbf{0} \\ \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \mathbf{0} & \tilde{\mathbf{A}}_{22} \end{pmatrix} \quad \text{mit} \quad \tilde{\mathbf{A}}_{11}(n_1 - u_1, u_1) \tag{24}$$

Aufgrund von rg $\mathbf{A}_{22} = u_2$ kann dieser Prozeß nun für die letzten u_2 Spalten von $\overline{\mathbf{A}}_1$ durchgeführt werden, wobei sich die ersten u_1 Spalten nicht mehr ändern. Durch entsprechende Zeilenvertauschungen erhält man:

$$\overline{\mathbf{A}}_1 = \begin{pmatrix} \mathbf{E}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_2 \\ \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \mathbf{0} & \tilde{\mathbf{A}}_{22} \end{pmatrix} \tag{25}$$

Die Aussage ergibt sich damit aus Folgerung 8.
Die Aussage von Folgerung 9 stellt dabei eine Verallgemeinerung eines Ergebnisses von (TUCHSCHERER, 1995) dar, worin der Spezialfall $n_1 = u_1$ dargestellt wurde. Man beachte, daß hierbei die Bedingung rg $\mathbf{A}_{22} = u_2$ automatisch wegen rg $\mathbf{A} = u$ erfüllt ist.
In konsequenter Fortführung der bisherigen Überlegungen betrachten wir das schon weiter oben angeführte Beispiel 4:

$$\mathbf{A} = \begin{pmatrix} 5 & 3 & -1 & 2 \\ 9 & 4 & -1 & 3 \\ 13 & 5 & -1 & 4 \\ 2 & 2 & 5 & 2 \\ 4 & 3 & 8 & 3 \\ 6 & 4 & 11 & 4 \end{pmatrix}$$

Wie schon weiter oben beschrieben wurde, führt hierbei die Streichung einer beliebigen Beobachtung dazu, daß eine andere verbleibende Beobachtung zur Restriktion wird. Die Ursache hierfür erkennt man sofort in der Normalform:

$$\overline{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 6 \\ 4 \\ 5 \\ 2 \end{pmatrix} \qquad \overline{\mathbf{B}} = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 6 \\ 4 \\ 5 \\ 2 \end{pmatrix}$$

Die Streichung der zweiten Beobachtung führt dazu, daß die Beobachtungen 1 und 3 zu Restriktionen werden (analog für die 5. Beobachtung, dann werden die 4. und die 6. zu Restriktionen). Im orthogonalen Komplement $\overline{\mathbf{B}}$ erweisen sich die Beobachtungen 1, 2 und 3 sowie 4, 5 und 6 als Mehrfachbeobachtungen. Offensichtlich gibt es zwei Gruppen verketteter, latenter innerer Restriktionen, nämlich

1) $(2,3), (2,1)$
2) $(5,4), (5,6)$

Die Matrix $\tilde{\mathbf{A}}$ weist eine Blockstruktur mit zwei komplementären Nullblöcken auf. Offenbar kann diese Blockstruktur durch Basistausch in der Beobachtungsbasis nicht zerstört werden, da ein solcher Tausch nur innerhalb der Beobachtungsgruppen $\{1,2,3\}$ bzw. $\{4,5,6\}$ erfolgen kann. Letztlich verweisen wir nochmals auf die Tatsache, daß der orthogonale Projektor **C** ebenfalls eine entsprechende Blockstruktur aufweist. Dies bedeutet jedoch, daß es sich in diesem Fall um zwei völlig unabhängige Ausgleichungsprozesse zwischen den Beobachtungen 1, 2, 3 bzw. 4, 5, 6 handelt.

Wir wollen nun diesen Sachverhalt verallgemeinern. Wir betrachten zu diesem Zweck den Spezialfall (22), der dem Zerfall der Design-Matrix **A** entspricht.

Folgerung 10: Die Design-Matrix **A** besitze (gegebenenfalls nach entsprechender Zeilen- und Spaltenvertauschung) folgende Struktur:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix} \tag{26}$$

mit spaltenregulären Blöcken $\mathbf{A}_{11}$ (n_1, u_1), $\mathbf{A}_{22}$ (n_2, u_2):

$$n_1 \geq u_1, \;\; n_2 \geq u_2, \;\; n_1 + n_2 = n, \;\; u_1 + u_2 = u \; .$$

Dann weisen sowohl die Normalgleichungen als auch der orthogonale Projektor eine entsprechende Blockstruktur auf:

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11}^T\mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^T\mathbf{A}_{22} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{22} \end{pmatrix}, \quad \mathbf{C}_{11}(n_1, n_1), \quad \mathbf{C}_{22}(n_2, n_2) \tag{27}$$

Die Aussage von Folgerung 10 stellt eine bekannte Tatsache dar und ergibt sich durch direkte Berechnung von $\mathbf{A}^T\mathbf{A}$ bzw. **C**.
Das Beispiel 4 zeigt jedoch, daß sich diese Blockstruktur der Form (26) nicht automatisch (als Folge der äußeren Geometrie der Ausgleichung) aufzeigt. Der nachstehende Satz zeigt, daß sich diese Eigenschaft anhand einer beliebig gewählten Normalform aufzeigen läßt.

Satz 10: Der orthogonale Projektor **C** weist genau dann eine Blockstruktur der Form (27) auf, wenn in einer Normalform eine analoge Blockstruktur in $\tilde{\mathbf{A}}$ auftritt:

$$\overline{\mathbf{A}} = \begin{pmatrix} \mathbf{E} \\ \tilde{\mathbf{A}} \end{pmatrix}, \quad \tilde{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{A}}_{22} \end{pmatrix} \quad \text{mit} \quad \tilde{\mathbf{A}}_{11}(n_1 - u_1, u_1), \quad \tilde{\mathbf{A}}_{22}(n_2 - u_2, u_2) \tag{28}$$

Beweis: Die erste Behauptung in Satz 10 (Blockstruktur des Projektors **C**) ergibt sich aus Folgerung 10 und der Tatsache, daß die Blockstruktur in (28) ein Spezialfall von (26) ist. Wir setzen nun die Gestalt (27) für **C** voraus. Aus der Symmetrie und der Idempotenz von **C** ergibt sich unmittelbar:

$$\mathbf{C}_{11}{}^T = \mathbf{C}_{11} \; , \;\; \mathbf{C}_{22}{}^T = \mathbf{C}_{22} \; , \;\; \mathbf{C}_{11}{}^2 = \mathbf{C}_{11} \; , \;\; \mathbf{C}_{22}{}^2 = \mathbf{C}_{22} \; . \tag{29}$$

Aufgrund von (29) läßt sich der Projektor **C** in die Summe zweier Projektoren zerlegen:

$$\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2 = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{22} \end{pmatrix}$$

Damit gilt weiterhin: sp $\mathbf{C}_1 = u_1$, sp $\mathbf{C}_2 = u_2$ mit $u_1 + u_2 = u$. Die Projektoren $\mathbf{C}_1$ und $\mathbf{C}_2$ projizieren den R^n auf einen u_1- bzw. u_2-dimensionalen Unterraum U_1 bzw. U_2, die sich als direkte, orthogonale Zerlegung von U interpretieren lassen. Aus der Invarianz der Unterräume bezüglich der Projektoren ergibt sich bei entsprechender Blockung der Matrix **A** gemäß:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \quad \text{mit} \quad \mathbf{A}_{11}(n_1, u_1), \quad \mathbf{A}_{22}(n_2, u_2)$$

$$\mathbf{C}_1\begin{pmatrix}\mathbf{A}_{11}\\ \mathbf{A}_{21}\end{pmatrix}=\begin{pmatrix}\mathbf{C}_{11}\mathbf{A}_{11}\\ \mathbf{0}\end{pmatrix}=\begin{pmatrix}\mathbf{A}_{11}\\ \mathbf{A}_{21}\end{pmatrix}\quad\Rightarrow\quad\mathbf{A}_{21}=\mathbf{0}$$

$$\mathbf{C}_2\begin{pmatrix}\mathbf{A}_{12}\\ \mathbf{A}_{22}\end{pmatrix}=\begin{pmatrix}\mathbf{0}\\ \mathbf{C}_{22}\mathbf{A}_{22}\end{pmatrix}=\begin{pmatrix}\mathbf{A}_{12}\\ \mathbf{A}_{22}\end{pmatrix}\quad\Rightarrow\quad\mathbf{A}_{12}=\mathbf{0}$$

und somit

$$\mathbf{A}=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_{22}\end{pmatrix}$$

Die entsprechende Blockstruktur (28) in der Normalform ergibt $\overline{\mathbf{A}}$ sich nun durch zweimalige Anwendung von Folgerung 9. Die Tatsache, daß dies unabhängig von der gewählten Beobachtungsbasis zur Erzeugung der Normalform ist, erkennt man sofort daran, daß bei einem möglichen Basistausch in (28) die Blockstruktur erhalten bleibt.
Zum Abschluß noch einige Ausführungen über den Zusammenhang zwischen latenten inneren Restriktionen und dem sogenannten Freiheitsgrad einer Ausgleichung. Dazu werden zunächst folgende bekannte Tatsachen verifiziert. Restriktionen tragen nicht zur Ausgleichung bei, so daß der Freiheitsgrad um die Anzahl der vorhandenen Restriktionen vermindert werden muß. Ebenso verringert sich der Freiheitsgrad beim Streichen einer gewissen Anzahl von Beobachtungen entsprechend. Streicht man nun (m−1) Beobachtungen einer latenten Restriktion m-ter Ordnung, so entsteht bekanntlich eine Restriktion, so daß der verbleibende Freiheitsgrad nicht (n−u) – (m−1), sondern (n−u) – m entsteht. Dieser Umstand eines verkleinerten Freiheitsgrades wird durch das Vorhandensein verketteter latenter Restriktionen noch verstärkt.
Analoge Betrachtungen können nun auch für das orthogonale Komplement aufgestellt werden. Wir betrachten dazu entsprechende Normalformen:

$$\overline{\mathbf{A}}=\begin{pmatrix}\mathbf{E}\\ \tilde{\mathbf{A}}\end{pmatrix}\quad\Leftrightarrow\quad\overline{\mathbf{B}}=\begin{pmatrix}\tilde{\mathbf{A}}^{\mathrm{T}}\\ -\mathbf{E}\end{pmatrix}$$

$\tilde{\mathbf{A}}_{\mathrm{R}}^{\mathrm{T}}$ sei die Matrix, die durch Streichung einer Zeile in $\tilde{\mathbf{A}}^{\mathrm{T}}$ entsteht (will man eine Beobachtung (Zeile) streichen, die der Einheitsmatrix $\mathbf{E}$ in entspricht, so ist diese zunächst durch Basistausch in $\tilde{\mathbf{A}}^{\mathrm{T}}$ einzutauschen, was stets möglich ist, falls diese Beobachtung keine Restriktion in $\overline{\mathbf{B}}$ darstellt). Damit ergibt sich nun durch Übergang zum orthogonalen Komplement:

$$\overline{\mathbf{B}}_{\mathrm{R}}=\begin{pmatrix}\tilde{\mathbf{A}}_{\mathrm{R}}^{\mathrm{T}}\\ -\mathbf{E}\end{pmatrix}\quad\Leftrightarrow\quad\overline{\mathbf{A}}_{\mathrm{R}}=\begin{pmatrix}\mathbf{E}\\ \tilde{\mathbf{A}}_{\mathrm{R}}\end{pmatrix}$$

Die Normalform $\tilde{\mathbf{A}}_{\mathrm{R}}$ entsteht jedoch aus $\overline{\mathbf{A}}$ durch Streichung einer Spalte in $\overline{\mathbf{A}}$ (Unbekannte im Modell (16)) und Streichung der entstehenden vollredundanten Beobachtung. Der Freiheitsgrad bleibt dabei erhalten. Analoges gilt für das Streichen mehrerer Beobachtungen im orthogonalen Komplement. Stehen jedoch die gestrichenen Beobachtungen im Zusammenhang mit latenten Restriktionen im orthogonalen Komplement, so verringert sich der Freiheitsgrad um mindestens 1.

2.4 Zusammenhänge zwischen multilinearer Graßmann-Algebra, Plücker-Koordinaten und Normalform

Wir verdanken Herrn Prof. Dr. mult. Erik W. Grafarend wertvolle Hinweise und Bemerkungen, die auf wichtige Zusammenhänge zwischen der von uns entwickelten Theorie und der multilinearen Graßmann-Algebra verweisen. Wir möchten uns auch an dieser Stelle nochmals für die wertvollen Anregungen bedanken. Im Folgenden sollen diese Zusammenhänge skizzenhaft aufgedeckt werden. Wir beginnen zunächst mit der Einführung der grundlegenden Begriffe dieser algebraischen Theorie.

Definition 4: Unter dem Graßmann-Bündel $G(n,u)$ über dem R^n verstehe man die Menge aller u-dimensionalen Unterräume des R^n:

$$G(n,u) = \{U \subseteq R^n \mid \dim U = u\} \quad 0 \le u \le n \tag{30}$$

Eine fundamentale Beziehung zwischen den Graßmann-Bündeln $G(n,u)$ und $G(n,n\text{-}u)$ ist durch die Dualisierung (auch Hodge-Dualisierung, Sternoperator genannt) gegeben. Sie beruht auf der Tatsache, daß sich jedem u-dimensionalen Unterraum $U \in G(n,u)$ ein sogenannter Dualraum $U^* \in G(n,n\text{-}u)$ in eineindeutiger Weise zuordnen läßt. Dieser Dualraum U^* ist jedoch nichts anderes als das orthogonale Komplement $U^\perp$ zu U. Wir vermerken, daß auch bei unseren Untersuchungen dieses orthogonale Komplement von entscheidender Bedeutung ist. Aus algebraischer Sicht vermittelt die Dualisierung eine Bijektion zwischen $G(n,u)$ und $G(n,n\text{-}u)$.

- $G^*(n,u) = G(n,n\text{-}u)$
- $U \in G(n,u) \Leftrightarrow U^* = U^\perp \in G(n,n\text{-}u)$

(31)

Eine weitere wichtige Begriffsbildung erzeugt einen Zusammenhang zwischen den Vektoren des Raumes R^n und den Elementen der Graßmann-Bündel. Dabei handelt es sich um das sogenannte Keilprodukt (auch äußeres, schiefes oder Graßmann-Produkt genannt). Aus Platzgründen verzichten wir hier auf eine tiefgreifende algebraische Einführung des Keilprodukts (siehe z. B. NEUTSCH, 1995), sondern beschränken uns auf einige wesentliche Zusammenhänge.
Sei dazu $U \in G(n,u)$ gegeben. Wir wählen nun eine Basis in U, die aus u linear unabhängigen Vektoren $\mathbf{a}^1, \ldots, \mathbf{a}^u \in U$ besteht. Diese spannen den Unterraum U auf, oder anders ausgedrückt, U ist die algebraische direkte Summe der durch $\mathbf{a}^i$, $i = 1, \ldots, u$ erzeugten eindimensionalen Unterräume. Diesen Zusammenhang zwischen den Vektoren $\mathbf{a}^i \in R^u$ und dem Unterraum U drücken wir nun durch das Keilprodukt aus:

$$U \mathrel{\hat{=}} \mathbf{a}^1 \wedge \mathbf{a}^2 \wedge \ldots \wedge \mathbf{a}^u \tag{32}$$

Umgekehrt wird hierdurch bei Vorgabe von u linear unabhängigen Vektoren aus R^n auch ein bestimmter Unterraum $U \in G(n,u)$ erzeugt. In der Literatur wird gezeigt, daß dieses Produkt multilinear und antisymmetrisch ist. Wir verdeutlichen dies anhand der 2-er Keilprodukte:

- $\mathbf{a}^1 \wedge \mathbf{a}^2 = -\mathbf{a}^2 \wedge \mathbf{a}^1$ (Umkehr der Orientierung)
- $(\alpha \mathbf{a}^1) \wedge \mathbf{a}^2 = \alpha \mathbf{a}^1 \wedge \mathbf{a}^2$
- $(\mathbf{a}^1 + \mathbf{a}^2) \wedge \mathbf{a}^3 = \mathbf{a}^1 \wedge \mathbf{a}^3 + \mathbf{a}^2 \wedge \mathbf{a}^3$

(33)

Im Fall der linearen Abhängigkeit kann man das Keilprodukt als nulldimensionalen Unterraum charakterisieren, so daß der Nullvektor im Keilprodukt genau wie die Zahl 0 bei gewöhnlichen Zahlenprodukten wirkt. Die Dualisierung eines Keilproduktes, bestehend aus u Faktoren, bewirkt ein Keilprodukt aus (n-u) Faktoren:

$$U \mathrel{\hat{=}} \mathbf{a}^1 \wedge \ldots \wedge \mathbf{a}^u \quad \Leftrightarrow \quad U^* \mathrel{\hat{=}} \mathbf{b}^1 \wedge \ldots \wedge \mathbf{b}^{n-u}, \tag{34}$$

wobei die $\mathbf{b}^i \in R^n$ eine Basis im $U^* = U^\perp$ bilden.
Im R^3 geht die Beziehung (34) über in:

$$^*(\mathbf{a}^1 \wedge \mathbf{a}^2) = \mathbf{a}^1 \times \mathbf{a}^2,$$

also das Kreuzprodukt zweier Vektoren des R^3.

Wir wenden uns nun der Frage der Parametrisierung der Elemente des Graßmann-Bündels $G(n,u)$ zu. Durch das Keilprodukt ist direkt eine mögliche Parametrisierung in Gestalt einer spaltenregulären Matrix **A** vom Format (n,u) induziert:

$$\mathbf{A} = \left(a_{ij}\right)_{i=1,\ldots,n}^{j=1,\ldots,u} = \left(\mathbf{a}^1 \ldots \mathbf{a}^u\right)$$

Jeder der Spaltenvektoren $\mathbf{a}^i \in R^n$ kann nun aber auch als Linearkombination bzgl. der kanonischen Basis des R^n dargestellt werden:

$$\mathbf{a}^i = \sum_{j=1}^{n} a_{ij}\mathbf{e}^j \tag{35}$$

Unter Benutzung der Eigenschaften des Keilprodukts erhält man:

$$\mathbf{a}^1 \wedge \ldots \wedge \mathbf{a}^u = \left(\sum_{j=1}^{n} a_{1j}\,\mathbf{e}^j\right) \wedge \ldots \wedge \left(\sum_{j=1}^{n} a_{uj}\,\mathbf{e}^j\right) = \sum_{i_1<\ldots<i_u} d_{i_1\ldots i_u}\,\mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_u} \tag{36}$$

In (36) stellen die Größen $d_{i1\ldots iu}$ jedoch nichts anderes als die Plücker-Koordinaten dar. In diesem Zusammenhang stellen sie eine Parametrisierung von U bzgl. der Keilprodukte der kanonischen Basisvektoren des R^n dar. Ein Vorteil dieser Parametrisierung von $U \in G(n,u)$ besteht in der sehr einfachen Möglichkeit der Dualisierung, da sich die Plücker-Koordinaten von $U^\perp$ direkt aus den Plücker-Koordinaten von U ergeben (siehe (21)). Die Darstellung von (36) impliziert auch die Aussage, daß die Dimension des Graßmann-Bündels $G(n,u)$ gleich $\binom{n}{u}$ ist. Diese Aussage ist jedoch falsch, wie wir im Weiteren darlegen werden. Wir vermerken dazu zunächst, daß die Plücker-Koordinaten nicht unabhängig voneinander gewählt werden können, da zwischen ihnen die schon erwähnten Plückerrelationen bestehen. Am deutlichsten wird dies anhand des Konzepts der Normalform:

$$\overline{\mathbf{A}} = \begin{pmatrix} \mathbf{E} \\ \tilde{\mathbf{A}} \end{pmatrix}$$

Wie dargelegt wurde, sind die Elemente von $\tilde{\mathbf{A}}$ selbst entsprechende Plücker-Koordinaten, alle anderen Plücker-Koordinaten lassen sich durch algebraische Funktionen aus diesen berechnen. Frei wählbar sind nur diejenigen Plücker-Koordinaten, die den Elementen von $\tilde{\mathbf{A}}$ entsprechen. Dies sind aber nur $u \cdot (n\text{-}u)$ Plücker-Koordinaten, hinzu kommt nur die Plücker-Koordinate, die der Beobachtungsbasis (Einheitsmatrix in $\overline{\mathbf{A}}$) entspricht. Durch die Normalform ist auch der Übergang von den Plücker-Koordinaten zur Matrixdarstellung bzw. Darstellung als Keilprodukt von U gegeben. Man wähle dazu eine beliebige nicht verschwindende Plücker-Koordinate und dividiere alle Plücker-Koordinaten durch diesen Wert. Die gewählte Plücker-Koordinate entspricht der Beobachtungsbasis (Einheitsmatrix in $\overline{\mathbf{A}}$). Die Elemente von $\tilde{\mathbf{A}}$ erhält man dann aus denjenigen Plücker-Koordinaten, die zur Beobachtungsbasis in dem Sinne assoziiert sind, daß sie genau u-1 Indizes aus der Beobachtungsbasis enthalten. Alle anderen Plücker-Koordinaten werden zur Konstruktion von $\overline{\mathbf{A}}$ bzw. U nicht benötigt. Damit erhält man abschließend folgendes Ergebnis.

Satz 11: Es gilt:

$$\dim G(n,u) = \dim G^*(n,u) = u \cdot (n\text{-}u) + 1 \tag{37}$$

Interessant ist in diesem Zusammenhang folgende Aussage.

Folgerung 11: Es gilt:

$$\binom{n}{u} = u \cdot (n-u) + 1 \quad \Leftrightarrow$$

- $n \leq 3,\ 0 \leq u \leq 3$
- $n \in N,\ u \in \{0,1\}$ und $u \in \{n\text{-}1, n\}$

Somit treten abweichende Effekte gegenüber der bisherigen Theorie erst ab n = 4 auf. Eine weitergehende algebraische Diskussion kann durchaus wichtige neue Erkenntnisse über die Struktur der Graßmann-Bündel erbringen.

3 Fazit und Ausblicke

Die innere Geometrie eines Ausgleichungsproblems stellt sich dar als Geometrie desjenigen Unterraumes U aus R^n (Beobachtungsraum), auf den projiziert werden soll. Diese Geometrie hat wesentlichen Einfluß auf die Ergebnisse der Ausgleichung (Projektion auf U entspricht den ausgeglichenen Beobachtungen, unbekannte Parameter, die Projektion auf $U^\perp$ entspricht den Verbesserungen). Insbesondere sind geometrische Besonderheiten aufzudecken, die durch eine spezielle Lage von U bzw. $U^\perp$ im R^n entstehen.
Diese Fälle werden durch den hier neu eingeführten Begriff der latenten Restriktionen charakterisiert, der eine Verallgemeinerung des bekannten Begriffes der Restriktion darstellt. Durch latente Restriktionen werden Beobachtungsgruppen (in U bzw. $U^\perp$) charakterisiert, die sich zwar gegenseitig kontrollieren, jedoch nicht durch die restlichen Beobachtungen kontrolliert werden können. Den theoretischen Rahmen für die latenten Restriktionen bilden die sogenannten Plücker-Koordinaten.
Latente Restriktionen entstehen durch verschwindende Plücker-Koordinaten, die wiederum durch entsprechende Indexmengen systematisch aufgedeckt werden können. Der hohe numerische Aufwand zur Berechnung der Plücker-Koordinaten kann durch einen Übergang zur sogenannten Normalform der Designmatrix (Plücker-Graßmann-Basis) vermindert werden.
Der Zusammenhang zwischen latenten Restriktionen und speziellen Beobachtungsbasen zur Erzeugung der Normalform wird charakterisiert. Damit können nun Begriffe wie Restriktionen, Mehrfachbeobachtungen, Kolinearitäten, zerfallende Ausgleichungsprobleme usw. in einen gemeinsamen Kontext gestellt und auch bearbeitet werden. Mit den hier dargestellten theoretischen Grundlagen können nun weiterführende Untersuchungen zur Geometrie von Beobachtungen angestellt werden. Wir erkennen hierfür zwei wesentliche Richtungen

1) Die Kenntnis latenter Restriktionen muß sich zwangsläufig auch in der Struktur und der Anwendung statistischer Testverfahren niederschlagen. Die Vorstellung, durch Streichung von Beobachtungen bzw. Unbekannten den Einfluß dieser Größen auf das Gesamtergebnis zu charakterisieren, ist in der Literatur mit dem Begriff „Omission Approach“ belegt (CHATTERJEE, HADI, 1988). Die Schwierigkeit bei der Anwendung liegt vor allem darin, gesichert zu erkennen, welche Größen (bzw. Gruppen von Größen) gestrichen werden sollen. Die praktische Durchrechnung aller möglichen Kombinationen ist von unvertretbar hohem numerischen Aufwand. Durch die Aufdekkung latenter Restriktionen werden somit wertvolle, ja unverzichtbare a priori Informationen für die statistischen Verfahren der Ausgleichungsrechnung geliefert.
2) Von großem praktischen Interesse sind auch diejenigen Fälle, die geometrisch bzw. numerisch sich „in der Nähe“ von latenten Restriktionen bewegen. Diese Fälle zeichnen sich durch vergleichsweise „kleine“ Plücker-Koordinaten bzw. betragsmäßig kleine Elemente in der Normalform aus, wobei wiederum das Problem der „richtig“ gewählten Beobachtungsbasis zur Erzeugung der Normalform entscheidend ist. Erste Untersuchungen zeigen einen wichtigen Zusammenhang mit den Balancierungsfaktoren (JURISCH, KAMPMANN, 1998). Diese Balancierungsfaktoren resultieren aus einem Vergleich zwischen Ist-Geometrie einer Ausgleichung und einer idealen Soll-Geometrie (Design im Beobachtungsraum) und geben bei entsprechender Normierung durch ihre Größenverhältnisse untereinander Informationen über besondere geometrische Verhältnisse der Ausgleichung.

Literatur

BEHNKEN, D. W., DRAPER, N. R.: Residuals and their Variance. Technometrics, 11, No. 1, p. 101-111, 1972.

CHATTERJEE, S., HADI, A. S.: Sensitivity Analysis in Linear Regression. John Wiley & Sons, New York, 1988.

DRAPER, N. R., SMITH, H.: Applied Regression Analysis. John Wiley & Sons, New York, 1981.

FINZEL, M.: Linear Approximation in l_n^∞. Journal of Approximation Theory, Vol. 76, p. 326-350, 1994.

GRAFAREND, E. W., SCHAFFRIN, B.: Ausgleichungsrechnung in linearen Modellen. BI Wissenschaftsverlag, Mannheim, 1993.

JURISCH, R., KAMPMANN, G.: Vermittelnde Ausgleichungsrechnung mit balancierten Beobachtungen – erste Schritte zu einem neuen Ansatz. Zeitschrift für Vermessungswesen, 123, S. 87-92, 1998.

NEUTSCH, W.: Koordinaten: Theorie und Anwendungen. Spektrum Akademischer Verlag GmbH, Heidelberg, Oxford, Berlin, 1995.

TUCHSCHERER, P.: Über den Sonderfall zweier Nicht-Null-Elemente in einer Spalte der Designmatrix für die Ausgleichung in linearen Modellen. Diplomarbeit an der Hochschule Anhalt (FH), 1995.

VAN DER WAERDEN, B. L.: Einführung in die algebraische Geometrie. Springer Verlag, Berlin, 1973.

WOLF, H.: Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. Ferdinand Dümmlers Verlag, Bonn, 1968.

WOLF, H.: Ausgleichungsrechnung I und II. 3. Auflage, Ferdinand Dümmlers Verlag, Bonn, 1997.

The Challenge of the Crustal Gravity Field

Juhani Kakkuri

1. Introduction

Two different methods have traditionally been used separately or together for determination of the geoid, namely 1) the *astrogeodetic levelling* method and 2) the *gravimetric* method. The former is based on the use of astrogeodetic deflections of the vertical as observables, which can be interpreted as horizontal gradients of the geoid undulation field, while the latter is based on the use mean gravity anomalies of surface blocks which should cover the whole surface of the Earth.
The advent of the artificial satellites has presented us with new methods to model the geoid. One of them is based on the use of ellipsoidal heights determined from GPS-observations. Namely, when confronting the ellipsoidal heigth H^* with the orthometric height H known from the precise levelling, the geoid undulation N is obtained simply by taking the difference $N = H^*-H$. This method leads to an extremely accurate determination of the geoid, provided naturally that sufficient number of accurate levelling points are available.
Details of the geoid can extensively be explored also by means of deep seismic sounding (DSS). This is possible because the data obtained from DSS can be used to construct a 3d-velocity structure model for the crust in the area to be studied. The velocity model can further be converted to a 3d-density model using the empirical relationship that holds between seismic velocities and crustal mass densities. Undulations of the geoid can then be estimated from the 3d-density model as shown by Wang, 1998 (also in Kakkuri and Wang, 1998).

2. Deep seismic sounding method

Deep seismic sounding and ocean drilling have revealed that the Earth's crust is not homogeneous but has a layered structure in the continental as well as in the oceanic areas. The vertical structure of thick continental crust is, however, more complicated than that of oceanic crust, and, in addition, in the continents the structure of ancient shield areas differ from that of younger basins. A three-layered crustal structure is observed in most parts of the shield areas, characterized by P-wave velocities of 6.0 - 6.5, 6.5 - 6.9 and 7.0 - 7.3 km/s, respectively. More complicated structures exist in quite a few places, mostly in the vicinity of the transition zones from continental crust to oceanic crust. The generalized structure of the basins is four-layered, a thick sediment cover being in the top and three igneous layers below.
Oceanic crust is only 5 - 10 km thick. Its top part consists of a layer of sediments that increases in thickness away from the oceanic ridges. The igneous oceanic basement consists of a thin (~0.5 km) upper layer of superposed basaltic lava flows underlain by a complex of basaltic intrusions, the sheeted dike complex. Below this the oceanic crust consists of gabbroic rocks (Lowrie, 1997).
The velocity at which compressional seismic P-waves travel through homogeneous materials can be expressed in the form

$$v_p = \sqrt{\frac{k + \frac{4}{3} n}{\rho}} \tag{1}$$

where ρ is the density, k is the bulk modulus and n is the shear modulus of the material. It can be seen that the velocity of P-waves depends on the elastic constants and the density of the material.

Thus, when the elastic parameters are known, the density can be calculated from the observed velocity. Unfortunately, as the elastic parameters are poorly known for materials inside the Earth, Eq. 1 is not applicable as such. For practical applications, it can be replaced by a linear relation known as Birch's law

$$v_p = a(\overline{m}) + b\rho \tag{2}$$

where a depends on the mean atomic weight $\overline{m}$ only, and b is a constant. For plutonic and metamorphic rocks, which are the main types of rocks in the shield areas, the mean atomic weight plays an insignificant role and can be safely neglected from the density-velocity relation (Gebrande, 1982). The following linear relations represent the shield areas (Chroston and Brooks 1989, Lebedev *et al.* 1977):

For upper crust	(v_p = 6.0, 6.5 km/s)	$v_p = 2.538\rho - 0.568 \pm 0.256$ km/s	(2a)
For mid-crust	(v_p = 6.5, 6.9 km/s)	$v_p = 3.184\rho - 2.580 \pm 0.122$ km/s	(2b)
For lower crust	(v_p = 6.8, 7.3 km/s)	$v_p = 2.717\rho - 1.250 \pm 0.120$ km/s	(2c)

Using the above relations we can estimate the velocity-density relations as follows:

Table 1. Density-velocity relations for plutonic and metamorphic rocks.

v_p (km/s)	ρ (g/cm³)
6.0	2.58 ± 0.11
6.4	2.80 ± 0.11
6.8	3.06 ± 0.05
7.3	3.15 ± 0.05

The velocities of seismic waves are generally found to be greater in igneous and crystalline rocks than in sedimentary ones (Parasnis, 1972). In the sedimentary rocks they tend to increase with depth of burial and geological age, and the application of Birch's law to sedimentary rocks is therefore questionable. Density data from drilling holes should be used instead of DSS-data in that case.

3. Mathematical modelling

Gravitational potential of a body can be written in the spherical coordinate system as follows (e.g. Heiskanen & Moritz 1967)

$$V(r,\theta,\lambda) = G\int \frac{\rho(r',\theta',\lambda')}{\sqrt{r^2 + r'^2 - 2rr'\cos\psi}}\, r'^2 \sin\theta'\, dr'\, d\theta'\, d\lambda' \tag{3}$$

where ψ is the angle between the vector $\overline{OQ}$ of the point $Q(r',\theta',\lambda')$ and the vector $\overline{OP}$ of the point $P(r,\theta,\lambda)$ as shown in Fig. 1, $\rho(r',\theta',\lambda')$ is the density of a mass element at point $Q(r',\theta',\lambda')$, and *G* is the Newtonian gravitational constant. In addition,

$$\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\lambda - \lambda') \tag{4}$$

The potential field of the crust can be constructed by slicing the crust into small spherical elements that take the form of a spherical prism and are filled with homogeneous masses, Fig. 2. The potential field of the whole crust is then the summation of potentials of the spherical prisms.

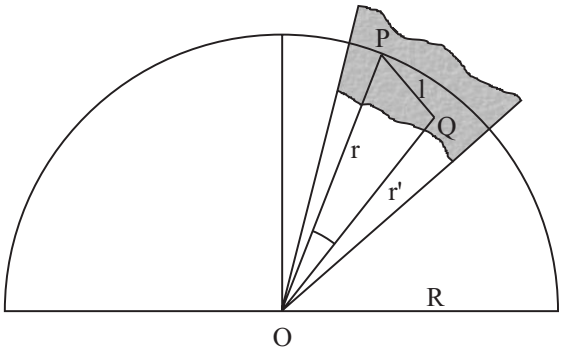


Fig. 1: The shaded area represents the whole crust from the surface down to The Moho.

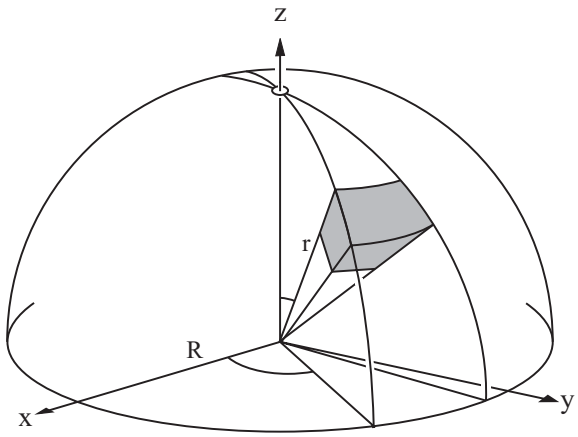


Fig. 2: A finite element of a mass body in a spherical prism form.

In order to evaluate Eq. 3 on the geoid, it is convenient to expand it into series of spherical harmonics. The expansion is to be performed separately for a case in which a mass element is above the reference sphere, i.e. for $\boldsymbol{r' > R = r}$, and separately for a mass element located below the reference sphere, i.e. for $\boldsymbol{r' < R = r}$.
The former, $r' > R = r$, is the case for most parts of the continental topographic masses. In this case Eq. 3 is given as follows:

$$V(r,\theta,\lambda) = G\rho \int d\lambda'\, d\theta'\, \sin\theta' \int dr'\, r' \sum_{l} \left(\frac{r}{r'}\right)^{l} P_l(\cos\psi) \tag{5}$$

where $P_l(\cos\psi)$ is the Legendre polynomial of degree *l*. Finally, according to Wang (1998), we have:

$$V = G\rho \sum_{l=0}^{\infty} H_l \sum_{m=-1}^{l} \overline{Y}_l^m(\theta,\lambda) \int d\lambda'\, d\theta'\, \sin\theta'\, \overline{Y}_l^m(\theta',\lambda') \tag{6}$$

where

$$\overline{Y}_l^m = \begin{cases} \overline{P}_l^m(\cos\theta)\cos m\lambda, m \geq 0 \\ \overline{P}^{|m|}\cos(\theta)\sin|m|\lambda, m < 0 \end{cases}$$

with $\overline{P}_l^m(\cos\theta)$ being the fully normalized associated Legendre function, and

$$H_l = \frac{R^2}{2l+1}\left(\frac{h_2 - h_1}{R} + \frac{1-l}{2}\frac{h_2^2 - h_1^2}{R^2} + \frac{(1-l)(-l)}{6}\frac{h_2^3 - h_1^3}{R^3} + \ldots\right)$$

where

$$\left.\begin{aligned} h_1 &= r_1 - R \\ h_2 &= r_2 - R \end{aligned}\right\} \qquad \text{with } r_1 < r_2 .$$

The latter, $r' < R = r$, is the case where masses are located below the geoid as in most parts of the Earth's crust. For derivation of the useful formulas, Eq. 3 is at first re-written as follows:

$$V(r,\theta,\lambda) = G\rho \int \frac{1}{\sqrt{1+\left(\frac{r'}{r}\right)^2 - 2\left(\frac{r'}{r}\right)\cos\psi}} \frac{r'^2}{r} \sin\theta'\, dr'\, d\theta'\, d\lambda' \tag{7}$$

and then developed into series as follows (Wang 1998):

$$V = G\rho \sum_{l=0}^{\infty} D_l \sum_{m=-l}^{l} \overline{Y}_l^m(\theta,\lambda) \int d\lambda'\, d\theta \sin\theta'\, \overline{Y}_l^m(\theta', \lambda'') \tag{8}$$

with

$$D_l = \frac{R^2}{2l+1}\left(\frac{D_2 - D_1}{R} - \frac{l+2}{2}\frac{D_2^2 - D_1^2}{R^2} + \frac{(l+2)(l+1)}{6}\frac{D_2^3 - D_1^3}{R^3} - \ldots\right)$$

where

$$\left.\begin{aligned} D_1 &= R - r_2 \\ D_2 &= R - r_1 \end{aligned}\right\} \qquad \text{with } D_1 < D_2 ;\ D \text{ being positive downwards.}$$

In order to investigate the contribution of the crust on the geoid, the geoidal undulation N caused by density anomalies in the crust is to be calculated. This is obtained from the well-known Bruns formula $N = T/\gamma$, where T is the disturbing potential on the geoid and γ is the normal gravity. The disturbing potential is the difference of the actual potential of the crust from the normal potential field. In order to calculate the normal potential field, the crust is to be divided into three homogeneous layers of equal thickness, Fig. 3. The depth of such a layer is the volume weighted mean depth of the corresponding layer of the actual crust, and its density is equal with the mean density of the actual layer.

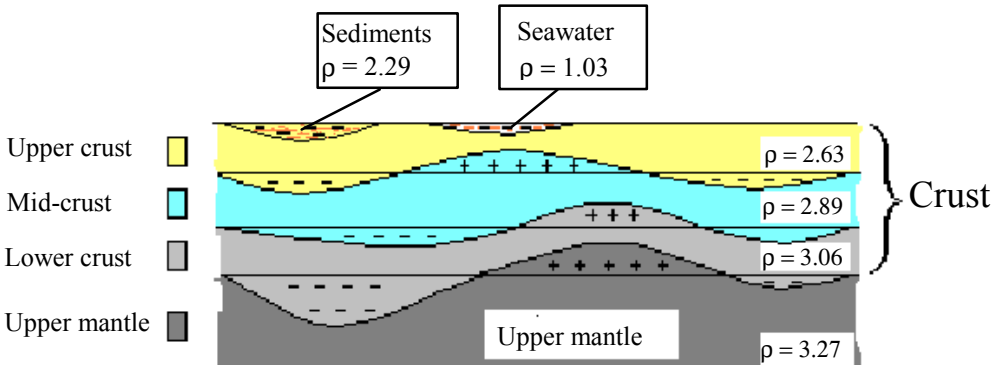


Fig.3: Mass models used for estimating the geoidal undulations from the crust. Straight lines show the boundaries of the normal (reference) mass model and curved lines those of the seismic (empirical) mass model. Positive and negative signs show the areas of mass surplus and mass deficiency, respectively.

4. Discussion

The deep seismic sounding method described was tested in Finland by Wang (1998) for estimating the contribution of the crust on the Fennoscandian gravimetric geoid. The work was the first contribution towards the solution of the problems related to this method. Influence of the layered structure of the crust on the geoid was found to be mainly due to the variation of the geometric shape of crustal layers. Variation of density inside the layers played a secondary role but was not insignificant. Accuracy obtained was found to be sufficient for the geophysical interpretation of the undulations of the Fennoscandian gravimetric geoid.
In the same way, the layered structure of the whole continental crust can be determined with the DSS for geophysical interpretation of the anomalies of the continental gravity field. To carry this out and to solve the problems related to the DSS method is a challenge to the geodesists and geophysicists in the next millenium.

References

Croston, P.N. and Brooks, S.G. (1989). Lower crustal seismic velocities from Lofoten-Vesterålen, North Norway. Tectonophysics, 157, 251-169.

Gebrande, H. (1982). Elasticity and inelasticity. In Landolt-Dörstein, Physical Properties of rocks, Vol. 16, pp. 1-99, ed. Angenmeister, G. Springer-Verlag, Berlin.

Kakkuri, J. and Z.T.Wang (1998). Structural efects of teh crust on the geoid modelled using deep seismic sounding interpretations. Geophys. J. Int 135, 495-504.

Lawrie, W. (1997). Fundamentals of Geophysics. Cambridge University Press.

Lebedev, T.S., Orovetskiy, Yu.P and Burtniy, P.A. (1977). A petrovelocity model of the Earth's crust based on the results of explosion seismology and high pressure experiments. Gerlands Beitr. Geophysik, Leipzig, 86, 303-312.

Parasnis, D.S. (1972). Principles of Applied Geophysics. Chapman and Hall.

Wang, Z.T. (1998). Geoid and Crustal Structure in Fennoscandia. Publ. Finn. Geod. Inst. 126.

Geodetic Pseudodifferential Operators and the Meissl Scheme

Wolfgang Keller

Abstract
The concept of pseudodifferential operators (PDO) is introduced as a generalization of the usual concepts of differential and integral operators. Based on the PDO concept in *Euclidean* spaces the concept of a PDO on a manifold is developed. It is demonstrated that for PDOs on a manifold the main part of the operator coincides with the usual planar approximation of the operator.
The so-called *Meissl* scheme is identified as the direct consequence of the homomorphy of the algebra of PDOs and the algebra of their symbols.

1 Introduction

Let $f : \mathcal{R}^n \to \mathcal{R}$ be a so-called function of moderate growth. The function $\hat{f}$, defined by

$$\hat{f}(\omega) := (2\pi)^{-\frac{n}{2}} \int_{\mathcal{R}^n} f(\mathbf{x}) e^{-\imath\omega^\top \mathbf{x}} d\mathbf{x} = \mathcal{F}\{f\}(\omega) \tag{1}$$

is called the *Fourier* transform of the function f. The function $\hat{f}$ is again a function of moderate growth and the so called inverse *Fourier* transform can be applied to it:

$$\mathcal{F}^{-1}\{\hat{f}\}(\mathbf{x}) := (2\pi)^{-\frac{n}{2}} \int_{\mathcal{R}^n} \hat{f}(\omega) e^{\imath\omega^\top \mathbf{x}} d\omega \tag{2}$$

The *Fourier* transform enjoys several useful properties:

- $$\mathcal{F}^{-1}\{\mathcal{F}\{f\}\} = f \tag{3}$$
- $$\mathcal{F}\{f * g\} = (2\pi)^{-\frac{n}{2}} \mathcal{F}\{f\}\mathcal{F}\{g\} \quad \text{convolution theorem} \tag{4}$$
- $$\mathcal{F}\{D^\alpha f\} = \mathcal{F}\{\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}\} = (-1)^{|\alpha|} \omega_1^{\alpha_1} \cdots \omega_n^{\alpha_n} \mathcal{F}\{f\} \tag{5}$$

  differentation theorem

The differentation theorem (5) of the *Fourier* transform is the starting point for the definition of the concept of pseudodifferential operators.
Let us consider the *Laplacian* in $\mathcal{R}^n$:

$$-\Delta u = -\sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_1^2}. \tag{6}$$

According to the differentiation theorem (5)

$$\mathcal{F}\{-\Delta u\} = \left(\sum_{i=1}^{n} \omega_i^2\right) \mathcal{F}\{u\} \tag{7}$$

holds. Applying the inverse *Fourier* transform to (7), one btains the following alternative representation of the *Laplacian*:

$$\begin{aligned} -\Delta u &= \mathcal{F}^{-1}\left\{\left(\sum_{i=1}^{n} \omega_i^2\right) \mathcal{F}\{u\}\right\} && (8) \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathcal{R}^n} \left(\sum_{i=1}^{n} \omega_i^2\right) \hat{u}(\omega) e^{\imath \omega^\top \mathbf{x}} d\omega && (9) \end{aligned}$$

This is the representation of the *Laplacian*, which is a differential operator, in the form of an integral. Hence, the name *pseudodifferential operator* is motivated for the following type of operators.

Definition 1 *The mapping*

$$pu := \mathcal{F}^{-1}\{a(\mathbf{x}, \omega)\mathcal{F}\{u\}\} \tag{10}$$

is called pseudodifferential operator and the function a is called its symbol.

Note that the concept of a pseudodifferential operator is much more general than the usual concept of a differential operator: If a is a polynomial in ω then the pseudodifferential operator coincides with a classical differential operator. If a is a suitable transcendental function, the corresponding PDO is a certain combination of a differential and a singular integral operator.

The symbol a also determines the order of the PDO.

Definition 2 *The PDO p is called a PDO of order r if*

$$|D_x^\beta D_\omega^\alpha a(\mathbf{x}, \omega)| \leq C_{\alpha\beta}(1 + |\omega|)^{r-|\alpha|} \tag{11}$$

holds.

Example 1 *For the Laplacian $-\Delta$ the symbol is*

$$symb\{-\Delta\} = |\omega|^2 \tag{12}$$

Hence, it holds

$$|D_x^\beta D_\omega^\alpha |\omega|^2| \leq |D_\omega^\alpha |\omega|^2| \leq |D_\omega^\alpha (1 + |\omega|)^2| \leq C(1 + |\omega|)^{2-|\alpha|} \tag{13}$$

This means that the Laplacian is a PDO of order 2.

Generally speaking: PDOs of negative order are smoothing operators and PDOs of positive order are de-smoothing operators. In most cases a PDO cannot be given by only one symbol but by a sequence of symbols with decreasing order.

Definition 3 *(extended)*
A mapping

$$pu := \sum_{k=0}^{\infty} \mathcal{F}^{-1}\{a_k(\mathbf{x}, \omega)\mathcal{F}\{u\}\} \tag{14}$$

with

$$|D_x^\beta D_\omega^\alpha a_k(\mathbf{x}, \omega)| \leq C_{k,\alpha\beta}(1 + |\omega|)^{r+k-|\alpha|} \tag{15}$$

is called a PDO of order r.

The part

$$p_0 u := \mathcal{F}^{-1}\{a_0(\mathbf{x}, \omega)\mathcal{F}\{u\}\} \tag{16}$$

is called the main part of p.

The main part represents the essential properties of p. In most cases the behaviour of p can be deduced from the behaviour of p_0.

2 PDOs on a manifold

The core of the definition of a PDO on a manifold is the fact that for a local patch the manifold has approximatively the same properties as an *Euklidean* space. Hence, an operator p is called a PDO on a manifold, if for every local coordinate patch it has the form (14).
Let us consider the concept in more detail. The manifold is denoted by Γ. Let $U_i \subset \Gamma, i = 1, 2, \ldots$ be a sequence of open subsets of Γ with the property

$$\bigcup_i U_i = \Gamma \tag{17}$$

These open subsets are called charts of Γ. For each chart U_i a mapping $\Phi_i : U_i \to \mathcal{R}^n$ is defined. For each $P \in U_i \subset \Gamma$ the real numbers $\Phi(P)$ are called local coordinates of P.

Definition 4 *A mapping $p : C^\infty(\Gamma) \to C^\infty(\Gamma)$ is called a PDO on the manifold Γ, if for every local coordinate patch U_i, the mapping*

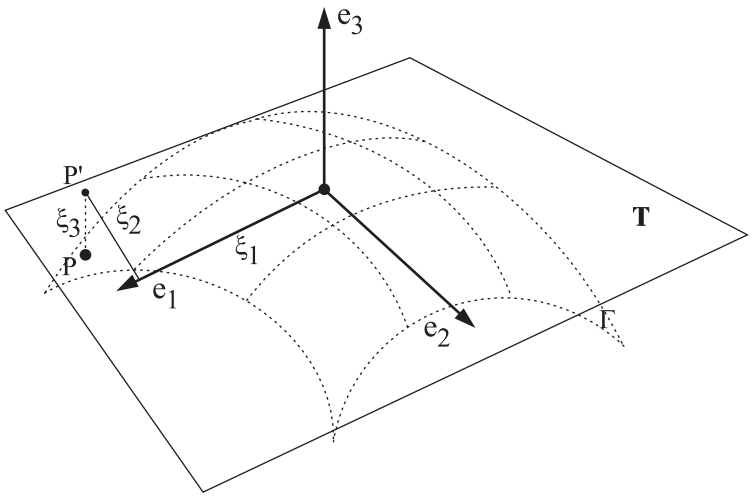
$$\Phi_i \circ p \circ \Phi_i^{-1} \tag{18}$$

is of the form (14).

Example 2 *Let Γ be a closed, orientable, smooth surface in $\mathcal{R}^3$. On Γ the following single-layer potential operator is defined:*

$$(pu)(\mathbf{x}) := \int_\Gamma \frac{u(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y} \tag{19}$$

In the neighbourhood of an arbitrary $\mathbf{x}_0 \in \Gamma$ local coordinates are introduced in the following way:



First a tangential plane T is attached to Γ in $\mathbf{x}_0$. Secondly, T is equipped with a cartesian coordinate system, having its origin in $\mathbf{x}_0$.

Let $P \in \Gamma$ be and $P' \in T$ its orthogonal projection onto the tangential plane. Let ξ_1, ξ_2 be the Cartesian coordinates of P' and ξ_3 the distance between P and P'. Then the local coordinates of $P \in \Gamma$ are defined by

$$\Phi(P) = (\xi_1, \xi_2, \xi_3) = \xi \tag{20}$$

Consequently, we have

$$(\Phi \circ p \circ \phi^{-1})u(\xi) = \int_{\mathcal{R}^3} \frac{u(\Phi^{-1}(\eta))}{|\Phi^{-1}(\xi) - \Phi^{-1}(\eta)|} |det(\Phi^{-1})'| d\eta \tag{21}$$

For Φ^{-1} the following Taylor expansion is valid

$$\begin{aligned} \Phi^{-1}(\eta) &= \Phi^{-1}(0) + (\Phi^{-1})'(0) && (22)\\ &= 0 + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \eta && (23)\\ &= (\eta_1, \eta_2, 0) && (24) \end{aligned}$$

Hence,

$$\begin{aligned} (\Phi \circ p \circ \Phi^{-1})u &\approx \int_{\mathcal{R}^2} \frac{u(\eta_1, \eta_2)}{\sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2}} d\eta && (25)\\ &= \frac{1}{|\xi|} * u && (26)\\ &= \frac{1}{2\pi} \mathcal{F}^{-1}\{\frac{1}{|\omega|} \mathcal{F}\{u\}\} && (27) \end{aligned}$$

This means that p is a PDO with the main part

$$p_0 u := \int_{\mathcal{R}^2} \frac{u(\eta)}{|\xi - \eta|} dy = \frac{1}{2\pi} \mathcal{F}^{-1}\{\frac{1}{|\omega|} \mathcal{F}\{u\}\} \tag{28}$$

3 Planar approximation

One typical technique in Physical Geodesy is the local approximation of globally defined integral operators. For this purpose the mean sphere S of the Earth is approximated by a tangential plane T. Consequently, the integral operator p defined on the sphere S has to be approximated by an integral operator p_0 on the tangential plane T. Usually, this is done by the following technique:

- In the point x_0 a *Cartesian* coordinate system is attached to the tangentialplane T, so that its ξ_3- axis coincides with the outer normal vector n of the sphere S in x_0
- A one-to-one relationship between S and T is established by orthogonal projection.
- The (ξ_1, ξ_2) coordinates of the projection are used as local coordinates on S.

It is easy to see that the mapping Φ^{-1} is given by

$$\Phi^{-1}(\xi_1, \xi_2) = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \sqrt{R^2 - \xi_1^2 - \xi_2^2} - R \end{pmatrix} \tag{29}$$

Let

$$(pu)(P) := \int_S K(\psi) u(Q) dS(Q) \tag{30}$$

be an invariant operator on S with ψ as the spherical distance between P and Q. Let the projections of the points P and Q be denoted by P' and Q', and let ξ and η denote their coordinates.
Obviously,

$$\psi = 2\arcsin(\frac{l}{2R}) \ , \quad l = \sqrt{|P' - Q'|^2 + (\xi_3 - \eta_3)^2} \tag{31}$$

holds and the representation of the invariant operator p in local coordinates is

$$pu = \int_{\mathcal{R}^2} K(2\arcsin(\sqrt{\frac{|P' - Q'|^2}{4R^2} + \frac{(\xi_3 - \eta_3)^2}{4R^2}}))u(Q')|det(\Phi^{-1})'|dQ' \tag{32}$$

Since $h := \frac{\xi_3 - \eta_3}{2R}$ is a small quantity, a *Taylor* expansion of K at the place $h = 0$ can be made. The replacement of K by the first term of its expansion is called *planar approximation* p_0 of p :

$$\begin{aligned} (p_0 u)(\xi) &= \int_{\mathcal{R}^2} K(2\arcsin(\frac{|\xi - \eta|}{2R})u(\eta)d\eta & (33)\\ &= K(2\arcsin(\frac{\bullet}{2R})) * u & (34)\\ &= \mathcal{F}^{-1}\{\hat{K} \cdot \hat{u}\} & (35) \end{aligned}$$

Now, the similarities between the main part of a PDO on a manifold and the planar approximation are obvious: The relation (29) defines the local coordinates, the representation $\Phi \circ p \circ \Phi^{-1}$ is given by (32) and the first term of the *Taylor* expansion gives the main part (33) of the corresponding PDO on the sphere S.
Usually, the planar approximation is understood intuitively. Its identification with the main part of the corresponding PDO gives an additional justification for this approximation: it already represents all essential properties of the original operator.

4 Meissl's Scheme

One of the most exiting things about PDOs is the homorphity of the algebra of PDOs with the algebra of its symbols. In detail this homomorphity is expressed by the following two relations

Theorem 1

$$\begin{aligned} symb(p + q) &= symb(p) + symb(q) & (36)\\ symb(p \circ q) &= symb(p) \cdot symb(q) & (37) \end{aligned}$$

In a maner of speaking, this means that one could work with the symbols instead of the operators themselves. Since the symbols are real function and the operators are mostly singular integral operators the handling of the former is much easier than the handling of the latter.

Example 3 *Let p be a PDO with the symbol $a(\omega)$*

$$pu = \mathcal{F}^{-1}\{a(\omega)\mathcal{F}\{u\}\} \tag{38}$$

and I the identity operator which also can be written as

$$Iu = \mathcal{F}^{-1}\{1 \cdot \mathcal{F}\{u\}\} \tag{39}$$

The determination of the inverse p^{-1} of p means that the following PDO-equation has to be solved:

$$p \circ p^{-1} = I \tag{40}$$

The corresponding symbol equation is

$$a(\omega) \cdot symb(p^{-1}) = 1 \tag{41}$$

which can be solved for $symb(p^{-1})$ *and giving the following representation of the inverse operator*

$$p^{-1}u = \mathcal{F}^{-1}\{\frac{1}{a(\omega)}\mathcal{F}\{u\}\} \tag{42}$$

The homomorphy means that a concatenation of several operators can be described by the multiplication of their symbols. For operators with geodetic relevance this relationship was already found earlier and independently of the context of PDO . It is called *Meissl Scheme* after its discoverer *P. Meissl*.

5 Construction of the Meissl scheme from the PDOs

The operators which are involved in the *Meissl* scheme are

- the upward continuation operator,
- the normal derivative operator,
- the gravity anomaly operator and the
- Stokes operator

For each of them the main part and its symbol has to be found. The upward continuation operator on the spehere is given by *Poisson*'s integral

$$Uu := u(r,\vartheta,\lambda) = \frac{R^2 - r^2}{4\pi}\int_\sigma \frac{u(\vartheta',\lambda')}{(R^2 - 2Rr\cos\psi + r^2)^{\frac{3}{2}}}\sin\vartheta' d\sigma(\vartheta',\lambda') \tag{43}$$

Its planar approximation, according to section 3, is the PDO

$$U_0 u = u(\mathbf{x},h) = \frac{1}{2\pi}\int_{\mathcal{R}^2} \frac{u(\mathbf{x}')}{(|\mathbf{x}-\mathbf{x}'|^2 + h^2)^{\frac{3}{2}}} d\mathbf{x}' \tag{44}$$

having the symbol $e^{-h\omega}$.
The normal derivative operator is derived from *Greens* representation theorem

Theorem 2 *(Greens representation theorem)*
Let u *be a harmonic function and* $\mathbf{n}$ *be the normal vector of* S*. For every* $\mathbf{x}$ *in* S *holds*

$$u(\mathbf{x}) = -\frac{1}{2\pi}\int_S \left(\frac{1}{|\mathbf{x}-\mathbf{y}|}\frac{\partial u}{\partial \mathbf{n}} - u\frac{\partial}{\partial \mathbf{n}}\frac{1}{|\mathbf{x}-\mathbf{y}|}\right) d\sigma(\mathbf{y}) \tag{45}$$

Denoting the single layer potential by s and the double layer potential by d

$$su = \frac{1}{2\pi}\int_S \frac{1}{|\mathbf{x}-\mathbf{y}|} u d\sigma \tag{46}$$

$$du = \frac{1}{2\pi}\int_S \frac{\partial}{\partial \mathbf{n}}\frac{1}{|\mathbf{x}-\mathbf{y}|} u d\sigma \tag{47}$$

the equation (45) can be rewritten as

$$Iu = -s(\frac{\partial u}{\partial \mathbf{n}}) + du \tag{48}$$

which can be solved for the normal derivative

$$nu := \frac{\partial}{\partial \mathbf{n}} u = -s^{-1}(I - d)u \tag{49}$$

The planar approximations of s and d are

$$s_0 u = \int_{\mathcal{R}^2} \frac{u}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y} \tag{50}$$

$$d_0 u = 0, \tag{51}$$

which leads to

$$n_0 u = -s_0^{-1} u. \tag{52}$$

Since the symbol of s_0 equals

$$symb(s_0) = 4\pi \frac{1}{|\omega|} \tag{53}$$

the main part of the normal derivative operator is given by

$$n_0 u = \frac{1}{4\pi} \mathcal{F}^{-1}\{|\omega| \mathcal{F}\{u\}\} \tag{54}$$

and its symbol is

$$symb(n_0) = \frac{1}{4\pi} |\omega| \tag{55}$$

In spherical approximation the gravity anomaly operator g is given by

$$gu := -\frac{\partial}{\partial \mathbf{n}} u - \frac{2}{R} u = -(n + \frac{2}{R} I)u \tag{56}$$

Obviously, its main part is

$$g_0 u = -n_0 u = s_0^{-1} u \tag{57}$$

The *Stokes* operator is given by

$$Stu := \frac{1}{4\pi\gamma R} \int_S S(\psi) u dS \tag{58}$$

with $S(\psi)$ being the *Stokes* function and γ being the normal gravity. The main part of St equals the planar approximation

$$St_0 u = \frac{1}{2\pi\gamma} \int_{\mathcal{R}^2} \frac{1}{|\mathbf{x} - \mathbf{y}|} u(\mathbf{y}) d\mathbf{y} = -\frac{1}{2\pi\gamma} s_0 u \tag{59}$$

having the symbol

$$symb(St_0) = \frac{1}{2\pi\gamma} \frac{1}{|\omega|} \tag{60}$$

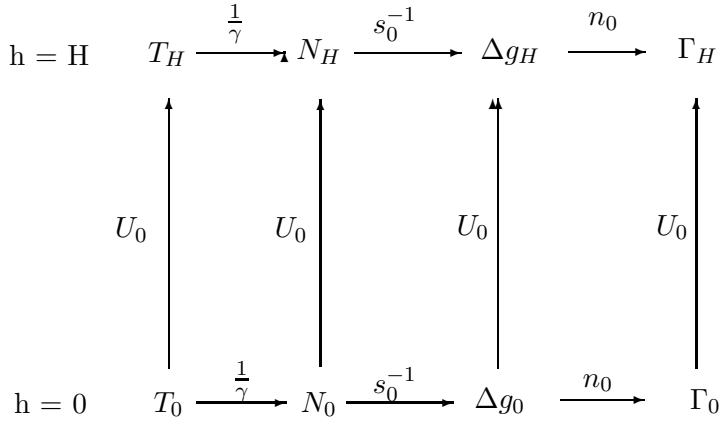
The following table summarizes the results

Name	main part	symbol
upward continuation U	$U_0 u = \frac{1}{2\pi} \int_{\mathcal{R}^2} \frac{u(\mathbf{x}')}{(\lvert\mathbf{x}-\mathbf{x}'\rvert^2+h^2)^{\frac{3}{2}}} d\mathbf{x}'$	$e^{-h\lvert\omega\rvert}$
normal derivation n	$n_0 u = \frac{1}{4\pi} \mathcal{F}^{-1}\{\lvert\omega\rvert \mathcal{F}\{u\}\}$	$\frac{1}{4\pi}\lvert\omega\rvert$
gravity anomaly g	$g_0 u = -n_0 u$	$-\frac{1}{4\pi}\lvert\omega\rvert$
Stokes St	$St_0 u = \frac{1}{2\pi\gamma} \int_{\mathcal{R}^2} \frac{1}{\lvert\mathbf{x}-\mathbf{y}\rvert} u(\mathbf{y}) d\mathbf{y}$	$\frac{1}{2\pi\gamma}\frac{1}{\lvert\omega\rvert}$

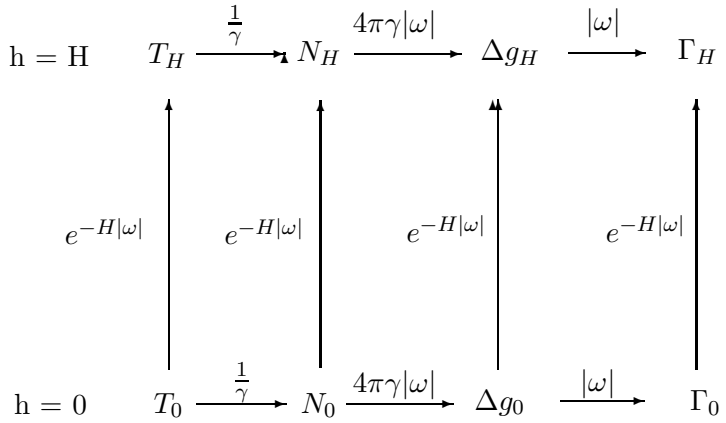
With the help of these four operators different geodetic quantities as

- disturbing potential T
- geoid undulations N
- gravity anomalies Δg
- vertical gravity gradients Γ

can be connected at ground level as well as at a certain height H. The following picture shows the commutative diagram of the previously mentioned quantities.



If this relationship is transformed into the frequency domain a relationship between the spectra of the used quantities is obtained.



This commutative diagram of the spectra is frequently called *Meissl* scheme.

6 Summary

The concept of a PDO is a usefull notion since it comprises both differential and integral operators under one term. The techniques, which were discussed here do not necessarily rely on PDOs, but the usage of the concept of PDOs simplifies the work much in the same way as matrix notation simplifies arithmetic calculations.
The use of singular integral operators in Physical Geodesy dates back to [4] and [2],[3] . In this papers the name PDO is never mentioned but the typical techniques are already used.
The introduction of PDOs into Geodesy was done by the famous article [6] and it is nowadays frequently used for the treatment of geodetic boundary value problems [5] and in connection with wavelets on the sphere [1].

References

[1] Freeden W. and Windheuser U. *Spherical wavelet transform and its discretization*, Advances in Computational Mathematics **11**(1995), pp 1-45

[2] Grafarend E.W. *The free nonlinear boundary value problem of physical geodesy*, Bull. Geodesique, **101**(1971), pp 243-262

[3] Grafarend E.W. *Die freie geodätische Randwertaufgabe und das Problem der Integrationsflächen innerhalb der Integralgleichungsmethode.* Mitt. Inst. f. Theoret. Geodäsie, No. 4, Bonn 1972

[4] Heiskanen W.A. and Moritz H. *Physical Geodesy*, W.H. Freeman, San Francisco, 1967

[5] Klees R. *Lösung des fixen geodätischen Randwertproblems mit Hilfe der Randelementmethode*, Reihe C, Heft 382, DGK, München, 1982

[6] Svensson S.L. *Pseudodifferential Operators – a New Approach to the Boundary Problems of Physical Geodesy*, Manuscr. Geodaetica,**8**(1983), pp 1-40

Analytical GPS Navigation Solution

Alfred Kleusberg

Abstract

The GPS navigation solution determines the coordinates $\boldsymbol{x} = (x, y, z)$ of the GPS receiver and the receiver clock offset cdT from measurements of at least four pseudo-ranges. We derive a direct solution of these observation equations without linearization and discuss the occurrence of unique solutions, double solutions, and infinitely many solutions, and the geometric conditions leading to these cases.

1. Introduction

The determination of the coordinates of a receiver position from measurement of pseudo-ranges to satellites is the standard mode of positioning for users of the Global Positioning System and similar systems; a minimum of four pseudo-ranges is necessary for three-dimensional positioning. Certain geometric constellations between the satellites and the receiver do not allow the determination of a unique position; we shall refer to these cases by using the term 'singularity'.

The equations linking the pseudo-ranges and the receiver coordinates are non-linear. The direct solution of these non-linear equations is possible, and several different solutions have been described in the literature. The widely used alternative is to linearize the pseudo-range equations and to use the tool of linear algebra in the position determination calculations.

When comparing results obtained from the solution of the non-linear and the linearized equation it was found, that differences occur for certain geometric constellations. In particular, the non-linear equations are solvable in some cases when the linearized equations lead to a singularity. To understand the reasons behind this behavior, we shall investigate the geometry leading to the above mentioned singularities.

2. The solution of the non-linear pseudo-range equations

Neglecting refraction effects, satellite clock offsets and measurement errors, the pseudo-range measured with a GPS receiver, p_i, is the sum of the satellite-to-receiver distance, s_i, and the receiver clock offset, dT, multiplied by the speed of light, c (Milliken and Zoller, 1980). The subscript i identifies the satellite.

The GPS navigation solution determines the coordinates (x, y, z) and the clock offset dT of a GPS receiver from pseudo-ranges p_i, i=0, 3 measured to four GPS satellites, and the coordinates (x_i, y_i, z_i), i=0, 3 of these satellites. These quantities are interrelated through the observation equations

$$p_i = [(x_i-x)^2 + (y_i-y)^2 + (z_i-z)^2]^{1/2} + c\cdot dT \qquad (1)$$

where we have used c as an abbreviation for the speed of light. The receiver clock offset can be eliminated from the observation equations (1) by subtracting p_0 from p_1, p_2, p_3. This yields equations

for three range differences $d_i = p_i - p_0$, i=1,3 represented in terms of satellite and receiver coordinates according to

$$d_i = [(x_i-x)^2 + (y_i-y)^2 + (z_i-z)^2]^{1/2} - [(x_0-x)^2 + (y_0-y)^2 + (z_0-z)^2]^{1/2} \quad (2)$$

From a geometric point of view, each of these three equations describes a hyperbolic surface of position. These surfaces intersect in the possible locations of the GPS receiver.

The non-linear eqns. (2) can also be solved directly without the process of linearization, thereby not requiring the availability of initial approximate values for the receiver position and being non-iterative as a consequence. Bancroft (1985) derived a rather elegant algebraic solution procedure for eqns. (2) and noted that his procedure "performs better than an iterative solution in regions of poor GDOP" (ibid.). His algorithm involves the inversion of a (4 x 4) matrix and the solution of a scalar equation of second order. Bancroft's method was further discusses and analyzed by Abel and Chaffee (1991) and by Chaffee and Abel (1994).

Krause (1987) published a two step algorithm for the direct solution of the eqns. (2). After the receiver clock offset $c \cdot dT$ is determined in a first step involving the inversion of a (2 x 2) matrix, the vectors from the satellites to the receiver can be evaluated and the receiver position is calculated through vector addition. Krause (ibid.) notes that "simulations under usual and extreme user and constellation situations showed absolute stability and precision for the algorithm".

The solution presented by Grafarend and Shan (1996) involves squaring eqn. (2) (to remove the square root), and then algebraically reducing the equations in order to provide the explicit solution for the receiver coordinates. This procedure includes the inversion of a (3 x 3) matrix.

The non-linear hyperbolic eqns. (3) were solved by Kleusberg (1994) using vector algebra. The algorithm is shown below in a modified and simplified version.

The distances b_i and the unit vectors $\boldsymbol{e}_i$ between the satellite S_0 and the satellites S_i are computed from the satellite coordinates (vectors are indicated by bold letters). These quantities completely describe the intrinsic geometry of the satellite configuration. The position of the receiver P is described by the (unknown) unit vector $\boldsymbol{e}$ pointing from S_0 to P, and the corresponding unknown distance s_0.

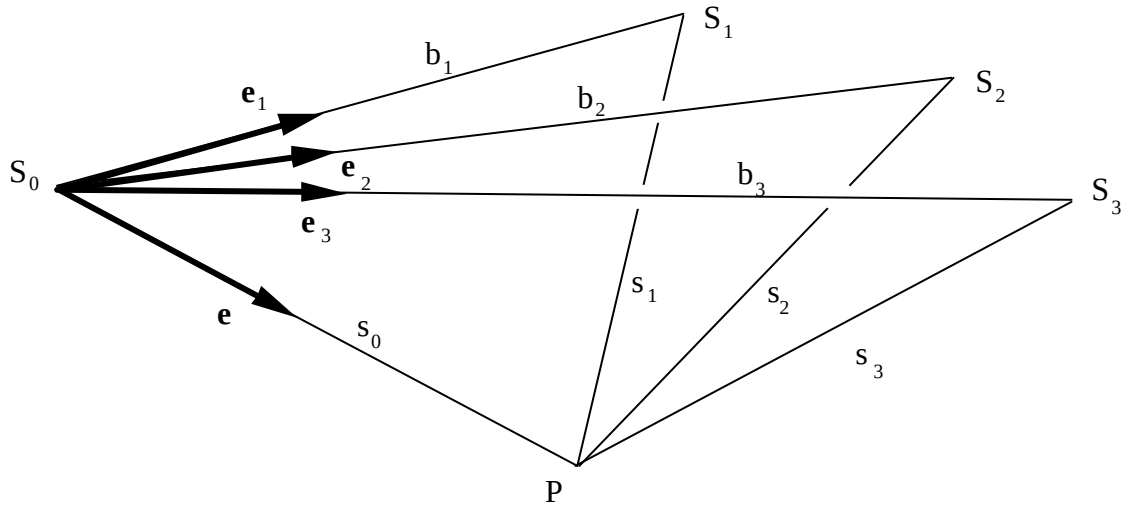


Figure 1: Geometry of the three-dimensional hyperbolic intersection

Noting that the cosine of the angle between the unit vectors $\boldsymbol{e}$ and $\boldsymbol{e}_i$ is equal to their scalar product $\boldsymbol{e} \cdot \boldsymbol{e}_i$, we can represent the geometry in each of the three triangles S_i - S_0 - P by the cosine rule according to

$$s_i^2 = b_i^2 + s_0^2 - 2 b_i s_0 \left(\mathbf{e} \cdot \mathbf{e}_i\right) \tag{3}$$

Further noting that the observation equation (2) can be rewritten as $s_i = d_i + s_0$, we also obtain in each of these three triangles a relation between the measurements of pseudo-range differences, and the distances between the receiver and the satellites

$$s_i^2 = d_i^2 + s_0^2 + 2 d_i s_0 \tag{4}$$

Equating equations (3) and (4) yields, after some basic algebraic manipulation

$$2 s_0 = \frac{b_i^2 - d_i^2}{d_i + b_i \left(\mathbf{e} \cdot \mathbf{e}_i\right)} \text{ , i = 1,3} \tag{5}$$

There are three unknowns in these three equations: the two independent components of the unit vector $\boldsymbol{e}$ and the distance s_0, all other terms are known.

In order to reduce the number of unknowns further, we equate the right hand sides of the first and the second of the eqns. (5), and similarly the second and the third, thereby eliminating the distance s_0.

$$\frac{b_i^2 - d_i^2}{d_i + b_i \left(\mathbf{e} \cdot \mathbf{e}_i\right)} = \frac{b_{i+1}^2 - d_{i+1}^2}{d_{i+1} + b_{i+1} \left(\mathbf{e} \cdot \mathbf{e}_{i+1}\right)} \text{ , i = 1,2} \tag{6}$$

These two equations can be rearranged by utilizing the distributive law of vector algebra to yield

$$\left[\frac{b_i}{b_i^2 - d_i^2} \mathbf{e}_i - \frac{b_{i+1}}{b_{i+1}^2 - d_{i+1}^2} \mathbf{e}_{i+1}\right] \cdot \mathbf{e} = \left[\frac{d_{i+1}}{b_{i+1}^2 - d_{i+1}^2} - \frac{d_i}{b_i^2 - d_i^2}\right] \text{ , i=1,2} \tag{7}$$

which reads in short form by using obvious abbreviations for the terms in square brackets

$$\begin{aligned} \mathbf{F}_1 \cdot \mathbf{e} &= U_1 \\ \mathbf{F}_2 \cdot \mathbf{e} &= U_2 \end{aligned} \tag{8}$$

These are two scalar equations for the components of the unit vector $\boldsymbol{e}$. In general, there will be two solutions for $\boldsymbol{e}$ satisfying eqn. (8). For the special case that $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are parallel, the solution is undefined.

The algebraic solution of equations (8) can be derived by applying the vector triple product identity to the product of $\boldsymbol{e}$, $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$

$$\boldsymbol{e} \times (\boldsymbol{F}_1 \times \boldsymbol{F}_2) = \boldsymbol{F}_1 \, \boldsymbol{e} \cdot \boldsymbol{F}_2 - \boldsymbol{F}_2 \, \boldsymbol{e} \cdot \boldsymbol{F}_1. \tag{9}$$

Replacing the scalar products on the right hand side of (9) with equation (8) we obtain

$$\boldsymbol{e} \times (\boldsymbol{F}_1 \times \boldsymbol{F}_2) = U_2 \boldsymbol{F}_1 - U_1 \boldsymbol{F}_2. \tag{10}$$

With the abbreviations

$$\boldsymbol{G} = \boldsymbol{F}_1 \times \boldsymbol{F}_2, \quad \boldsymbol{H} = U_2 \boldsymbol{F}_1 - U_1 \boldsymbol{F}_2 \tag{11}$$

we can rewrite equation (10) in a shorter form as

$$\boldsymbol{e} \times \boldsymbol{G} = \boldsymbol{H}. \tag{12}$$

Muliplying both sides of this equation by $\boldsymbol{G}$ from the left, and applying the triple vector product identity again to the left hand side of the resulting equation, we obtain

$$\boldsymbol{e}\, \boldsymbol{G} \cdot \boldsymbol{G} - \boldsymbol{G}\, \boldsymbol{G} \cdot \boldsymbol{e} = \boldsymbol{G} \times \boldsymbol{H}. \tag{13}$$

The scalar product in the second term of the left-hand side can be written in terms of the length of the vectors involved, and the angle β between them

$$\boldsymbol{G} \cdot \boldsymbol{e} = (\boldsymbol{G} \cdot \boldsymbol{G})^{1/2} \cos\beta . \tag{14}$$

The angle β also appears if we compute the length of the vector $\boldsymbol{H}$ from equation (12)

$$(\boldsymbol{H} \cdot \boldsymbol{H})^{1/2} = [(\boldsymbol{e} \times \boldsymbol{G}) \cdot (\boldsymbol{e} \times \boldsymbol{G})]^{1/2} = (\boldsymbol{G} \cdot \boldsymbol{G})^{1/2} \sin\beta . \tag{15}$$

Comparing eqns. (14) and (15) we get

$$\boldsymbol{G} \cdot \boldsymbol{e} = \pm (\boldsymbol{G} \cdot \boldsymbol{G})^{1/2} [1 - \boldsymbol{H} \cdot \boldsymbol{H} / \boldsymbol{G} \cdot \boldsymbol{G}]^{1/2} = \pm [\boldsymbol{G} \cdot \boldsymbol{G} - \boldsymbol{H} \cdot \boldsymbol{H}]^{1/2}. \tag{16}$$

Inserting this relation into equation (13) we obtain after some rearrangement the two solutions of equation (8)

$$\boldsymbol{e}^{1,2} = (\boldsymbol{G} \cdot \boldsymbol{G})^{-1} \{ \boldsymbol{G} \times \boldsymbol{H} \pm \boldsymbol{G} [(\boldsymbol{G} \cdot \boldsymbol{G}) - (\boldsymbol{H} \cdot \boldsymbol{H})]^{1/2} \}. \tag{17}$$

The distance s_0 from the satellite S_0 to the receiver P can now be determined from anyone of the three equations (5) according to

$$s_0^{1,2} = \frac{1}{2} \frac{b_i^2 - d_i^2}{d_i + b_i \left(\boldsymbol{e}^{1,2} \cdot \boldsymbol{e}_i \right)} \tag{18}$$

and the coordinates of the receiver are finally determined from (cf. Figure 1)

$$\boldsymbol{x}^{1,2} = \boldsymbol{x}_0 + s_0^{1,2} \boldsymbol{e}^{1,2} . \tag{19}$$

3. Discussion of results

Qualitatively, the results obtained from eqns. (18) and (19) can be classified in the following way:

a) The two unit vectors determined from eqn. (17) are different, and eqn. (18) yields two positive distances s_0. In this case, there are two intersections of the hyperbolic surfaces of position. Both

solutions satisfy the observation eqns. (2). The correct solution can be identified if *a priori* information about the approximate receiver location is available.

b) The two unit vectors determined from eqn. (17) are different, and eqn. (18) yields one positive and one negative distance s_0. In this case, only the solution belonging to the positive distance satisfies the observation eqns. (2).

c) The two unit vectors determined from eqn. (17) are the same. In this case, the two intersections of the hyperbolic surfaces of position coincide. It will be shown elsewhere that in this particular case the receiver-to-satellite unit vectors are on a conic surface. It is known that in this case the linearized pseudo-range cannot be solved uniquely.

d) The receiver is located on the extension of one of the baselines b_i. In this case, $d_i = \pm b_i$ and one of the denominators in eqn. (7) is zero. This critical geometric situation is know from 2-dimensional hyperbolic positioning (e.g. LORAN). In 3-dimensional satellite positioning it does not occur if the receiver location is significantly lower than the orbits of the satellites.

e) The two vectors $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are parallel. In this case, the denominator in eqn. (17) is zero, and there are infinitely many solutions $\boldsymbol{e}$ satisfying eqns. (8). Since $\boldsymbol{F}_1$ is in the plane defined by the satellites S_0, S_1 and S_2, and $\boldsymbol{F}_2$ is in the plane defined by satellites S_0, S_2 and S_3, the four satellite positions are coplanar in this particular case. Not all coplanar satellite positions will lead to this singularity; it will be shown elsewhere that only the arrangement of the satellite positions in a conic section allows infinitely many solutions of the observation equations (2).

References

Abel, J.S. and Chaffee, J.W. (1991): Existence and Uniqueness of GPS Solutions. IEEE Transactions on Aerospace and Electronic Systems, Vol. 27, No. 6, pp. 952-956

Bancroft, S. (1985): An Algebraic Solution of the GPS Equations. IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-21, No. 6, pp. 56-59

Chaffee, J. and Abel, J. (1991): On the Exact Solutions of Pseudorange Equations. IEEE Transactions on Aerospace and Electronic Systems, Vol. 30, No. 4, pp. 1021-1030

Grafarend, E.W. and Chan, J. (1996): A Closed-form Solution of the Nonlinear Pseudo-Ranging Equations (GPS), ARTIFICIAL SATELLITES Planetary Geodesy, No. 28, pp. 133-147

Kleusberg, A. (1994): Die direkte Lösung des räumlichen Hyperbelschnitts. Zeitschrift für Vermessungswesen, Vol. 119, No. 4, pp. 188-192

Krause, L.O. (1987): A Direct Solution to GPS-Type Navigation Equations. IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-23, No. 2, pp. 225-232

Milliken, R.J, Zoller, C.J. (1980): Principle of Operation of NAVSTAR and System Characteristics. in: Global Positioning System Papers published in NAVIGATION, Vol. I, pp. 3-14, The Institute of Navigation, Washington, D.C.

Grundprinzipien der Bayes-Statistik

Karl-Rudolf Koch

Zusammenfassung:
In drei wesentlichen Punkten unterscheidet sich die Bayes-Statistik von der traditionellen Statistik. Zunächst beruht die Bayes-Statistik auf dem Bayes-Theorem, mit dessen Hilfe unbekannte Parameter zu schätzen, Konfidenzregionen für die Parameter anzugeben und Hypothesen für die Parameter zu prüfen sind. Ferner nimmt die Bayes-Statistik eine Erweiterung des Wahrscheinlichkeitsbegriffs vor, indem die Wahrscheinlichkeit von Aussagen definiert wird, wobei die Wahrscheinlichkeit ein Maß für die Plausibilität der Aussage gibt. Schließlich sind die unbekannten Parameter der Bayes-Statistik Zufallsvariable, was aber nicht bedeutet, daß sie keine Konstanten repräsentieren dürfen. Auf vielfältige Anwendungen der Bayes-Statistik bei geodätischen Problemstellungen wird hingewiesen.

1 Einführung

Im Mittelpunkt der Bayes-Statistik steht das Bayes-Theorem. Mit seiner Hilfe lassen sich unbekannte Parameter schätzen, Konfidenzregionen für die unbekannten Parameter festlegen und die Prüfung von Hypothesen für die Parameter ableiten. Dieser einfache und anschauliche Weg der Herleitung ist der traditionellen Statistik versperrt, da sie sich nicht auf das Bayes-Theorem gründet. Insofern besitzt die Bayes-Statistik einen wesentlichen Vorteil gegenüber der traditionellen Statistik.
Für die Bayes-Statistik wird eine Erweiterung des Begriffs der Wahrscheinlichkeit vorgenommen, indem die Wahrscheinlichkeit für Aussagen definiert wird. Dagegen beschränkt sich die traditionelle Statistik auf die Wahrscheinlichkeit von zufälligen Ereignissen. In der Bayes-Statistik wird also die Wahrscheinlichkeit nicht nur für zufällige Ereignisse, sondern ganz allgemein für Aussagen eingeführt. Die Wahrscheinlichkeit gibt die Plausibilität von Aussagen an, durch sie wird der Zustand des Wissens über eine Aussage ausgedrückt. Für die Wahrscheinlichkeit von Aussagen lassen sich durch logisches und konsistentes Schließen drei Gesetze ableiten, aus denen die gesamte Wahrscheinlichkeitstheorie entwickelt werden kann.
Im Gegensatz zur traditionellen Statistik, bei der die unbekannten Parameter feste Größen repräsentieren, sind die unbekannten Parameter in der Bayes-Statistik Zufallsvariable. Das bedeutet aber nicht, daß die unbekannten Parameter nicht Konstanten repräsentieren dürfen, wie beispielsweise die Koordinaten eines festen Punktes. Mit dem Bayes-Theorem erhalten die unbekannten Parameter Wahrscheinlichkeitsverteilungen, aus denen die Wahrscheinlichkeit, also die Plausibilität von Werten der Parameter folgt. Die Parameter selbst können daher feste Größen darstellen und brauchen nicht aus Zufallsexperimenten zu resultieren.
Im folgenden soll auf die Gesetze der Wahrscheinlichkeit und auf die Überlegungen zur Parameterschätzung, zur Festlegung von Konfidenzbereichen und zum Test von Hypothesen eingegangen werden. Außerdem werden Anwendungen genannt. Nur einige Ergebnisse können präsentiert werden, Details sind bei Koch (1990;1999) nachzulesen.

2 Gesetze der Wahrscheinlichkeit

Die Bayes-Statistik arbeitet ausschließlich mit bedingten Wahrscheinlichkeiten, denn im allgemeinen hängt eine Aussage davon ab, ob eine weitere Aussage wahr ist. Man schreibt $A|B$, um auszudrücken, daß A wahr ist unter der Bedingung, daß B wahr ist. Die Wahrscheinlichkeit von $A|B$, auch bedingte Wahrscheinlichkeit genannt, wird mit

$$P(A|B) \tag{2.1}$$

bezeichnet. Sie gibt ein Maß für die Plausibilität der Aussage $A|B$ an. Bedingte Wahrscheinlichkeiten sind geeignet, empirisches Wissen auszudrücken, denn durch $P(A|B)$ wird die Wahrscheinlichkeit angegeben, mit der vorhandenes Wissen für weiteres Wissen relevant ist.
Wie von Cox (1946) und Jaynes (1995) gezeigt, lassen sich durch logisches und konsistentes Schließen das Produktgesetz der Wahrscheinlichkeit

$$P(AB|C) = P(A|C)P(B|AC) = P(B|C)P(A|BC) \tag{2.2}$$

mit

$$P(S|C) = 1 \tag{2.3}$$

und das Summengesetz

$$P(A|C) + P(\bar{A}|C) = 1 \tag{2.4}$$

ableiten. Hierin bedeuten A, B und C allgemeine Aussagen, S die sichere Aussage und $\bar{A}$ die Negation der Aussage A.
Aus diesen drei Gesetzen lassen sich alle Gesetze der Wahrscheinlichkeitstheorie ableiten, die für die Bayes-Statistik benötigt werden. Berücksichtigt man die Bedingung C im Produktgesetz (2.2) nicht und löst nach $P(A|B)$ auf, erhält man die Definition der bedingten Wahrscheinlichkeit der traditionellen Statistik. Sie wird dort durch die relative Häufigkeit erklärt, was im Gegensatz zu der Ableitung, die auf (2.2) führt, weniger einleuchtend ist.
Löst man das Produktgesetz (2.2) nach $P(A|BC)$ auf, erhält man das Bayes-Theorem

$$P(A|BC) = \frac{P(A|C)P(B|AC)}{P(B|C)} \; . \tag{2.5}$$

Bei den üblichen Anwendungen des Bayes-Theorems bedeutet A die Aussage über ein unbekanntes Phänomen, B die Aussage, die Information über das unbekannte Phänomen enthält, und C eine Aussage über zusätzliches Wissen. Man bezeichnet $P(A|C)$ als Priori-Wahrscheinlichkeit, $P(A|BC)$ als Posteriori-Wahrscheinlichkeit und $P(B|AC)$ als Likelihood. Die Priori-Wahrscheinlichkeit der Aussage über das unbekannte Phänomen wird also durch die Likelihood modifiziert, die Information über das Phänomen enthält, um die Posteriori-Wahrscheinlichkeit zu erhalten.
Im folgenden wird noch das verallgemeinerte Bayes-Theorem angegeben, das für Verteilungen gilt.

3 Verteilungen

Die Aussagen sollen sich im folgenden auf die numerischen Werte von Variablen in Form von reellen Zahlen beziehen. Diese Variablen werden in der traditionellen Statistik Zufallsvariablen genannt, da ihre Werte aus Zufallsexperimenten stammen. Diese Einschränkung besteht in der Bayes-Statistik nicht, die Aussagen können die Werte beliebiger Variablen beinhalten. Dennoch wird die Bezeichnung Zufallsvariable beibehalten.
Gegeben sei eine diskrete Zufallsvariable X mit den diskreten Werten $x_i \in \mathbb{R}$ für $i \in \{1, \ldots, m\}$. Die Wahrscheinlichkeit $P(X = x_i|C)$, daß X den Wert x_i unter der Bedingung der Aussage C annimmt, die zusätzliche Information enthält, wird mit

$$p(x_i|C) = P(X = x_i|C) \quad \text{für} \quad i \in \{1, \ldots, m\} \tag{3.1}$$

bezeichnet. Man nennt $p(x_i|C)$ die diskrete Wahrscheinlichkeitsdichte oder auch diskrete Verteilung für die diskrete Zufallsvariable X.
Die diskrete Dichte der n-dimensionalen Zufallsvariablen $X_1, \ldots, X_n$ ist (3.1) entsprechend gegeben durch

$$p(x_{1j_1}, \ldots, x_{nj_n}|C) = P(X_1 = x_{1j_1}, \ldots, X_n = x_{nj_n}|C) \quad \text{mit} \quad j_k \in \{1, \ldots, m_k\},\ k \in \{1, \ldots, n\}\ . \tag{3.2}$$

In abgekürzter Schreibweise erhält man

$$p(x_1, \ldots, x_n|C) = P(X_1 = x_1, \ldots, X_n = x_n|C) \tag{3.3}$$

oder für den $n \times 1$ Zufallsvektor $\boldsymbol{x}$, dessen Werte ebenfalls mit $\boldsymbol{x}$ bezeichnet werden,

$$\boldsymbol{x} = |x_1, \ldots, x_n|'$$

die Dichte

$$p(\boldsymbol{x}|C)\ . \tag{3.4}$$

Auf eine entsprechende Form läßt sich auch die Dichte einer n-dimensionalen stetigen Zufallsvariablen $X_1, \ldots, X_n$ mit den Werten $x_1, \ldots, x_n \in \mathbb{R}$ in den Intervallen $-\infty < x_k < \infty$ mit $k \in \{1, \ldots, n\}$ bringen, so daß (3.4) die Dichte eines diskreten oder stetigen Zufallsvektors $\boldsymbol{x}$ bezeichnet.
Aus dem Produktgesetz (2.2) folgt die bedingte diskrete oder stetige Dichte für den diskreten oder stetigen Zufallsvektor $\boldsymbol{x}_1$ unter der Bedingung gegebener Werte für den diskreten oder stetigen Zufallsvektor $\boldsymbol{x}_2$ und der zusätzlichen Bedingung C mit

$$p(\boldsymbol{x}_1|\boldsymbol{x}_2, C) = \frac{p(\boldsymbol{x}_1, \boldsymbol{x}_2|C)}{p(\boldsymbol{x}_2|C)}\ . \tag{3.5}$$

Über die bedingte Dichte wird die bedingte Unabhängigkeit von Zufallsvariablen eingeführt. Sind $\boldsymbol{x}_i, \boldsymbol{x}_j$ und $\boldsymbol{x}_k$ diskrete oder stetige Zufallsvektoren, dann sind $\boldsymbol{x}_i$ und $\boldsymbol{x}_j$ genau dann voneinander unabhängig, falls gilt

$$p(\boldsymbol{x}_i|\boldsymbol{x}_j, \boldsymbol{x}_k, C) = p(\boldsymbol{x}_i|\boldsymbol{x}_k, C)\ . \tag{3.6}$$

Sind $\boldsymbol{x}$ und $\boldsymbol{y}$ diskrete oder stetige Zufallsvektoren, erhält man mit (3.5)

$$p(\boldsymbol{x}|\boldsymbol{y}, C) = \frac{p(\boldsymbol{x}, \boldsymbol{y}|C)}{p(\boldsymbol{y}|C)}\ , \tag{3.7}$$

worin der Vektor $\boldsymbol{y}$ gegebene Werte enthält, oder entsprechend

$$p(\boldsymbol{y}|\boldsymbol{x}, C) = \frac{p(\boldsymbol{x}, \boldsymbol{y}|C)}{p(\boldsymbol{x}|C)}\ . \tag{3.8}$$

Löst man (3.7) und (3.8) nach $p(\boldsymbol{x}, \boldsymbol{y}|C)$ auf und setzt die sich ergebenden Ausdrücke gleich, erhält man das verallgemeinerte Bayes-Theorem

$$p(\boldsymbol{x}|\boldsymbol{y}, C) = \frac{p(\boldsymbol{x}|C)p(\boldsymbol{y}|\boldsymbol{x}, C)}{p(\boldsymbol{y}|C)}\ . \tag{3.9}$$

Da der Vektor $\boldsymbol{y}$ feste Werte enthält, ist $p(\boldsymbol{y}|C)$ konstant. Das Bayes-Theorem wird daher häufig in der Form angewendet

$$p(\boldsymbol{x}|\boldsymbol{y}, C) \propto p(\boldsymbol{x}|C)p(\boldsymbol{y}|\boldsymbol{x}, C)\ , \tag{3.10}$$

in der $\propto$ das Proportionalitätszeichen bedeutet.

Der Zufallsvektor $\boldsymbol{x}$ enthalte unbekannte Parameter. Die Werte, die $\boldsymbol{x}$ annehmen kann, werden, wie bereits erwähnt, ebenfalls mit $\boldsymbol{x}$ bezeichnet. Die Menge der Werte $\boldsymbol{x}$ bezeichnet man als Parameterraum $\mathcal{X}$, also $\boldsymbol{x} \in \mathcal{X}$. Der Zufallsvektor $\boldsymbol{y}$ repräsentiere Daten. Die Dichte $p(\boldsymbol{x}|C)$ enthält Information über die Parameter $\boldsymbol{x}$, bevor die Daten $\boldsymbol{y}$ erhoben wurden, also Vorinformation. Man nennt daher $p(\boldsymbol{x}|C)$ die Priori-Dichte. Mit Berücksichtigung der Beobachtungen $\boldsymbol{y}$ folgt die Dichte $p(\boldsymbol{x}|\boldsymbol{y}, C)$, die als Posteriori-Dichte für die Parameter $\boldsymbol{x}$ bezeichnet wird. Da die Daten $\boldsymbol{y}$ vorliegen, wird $p(\boldsymbol{y}|\boldsymbol{x}, C)$ nicht als Funktion der Daten $\boldsymbol{y}$, sondern als Funktion der Parameter $\boldsymbol{x}$ interpretiert. Die Dichte $p(\boldsymbol{y}|\boldsymbol{x}, C)$ wird daher als Likelihoodfunktion bezeichnet. Die Daten modifizieren also durch die Likelihoodfunktion die Priori-Dichte und führen auf die Posteriori-Dichte für die Parameter.
Im Bayes-Theorem wird der Vektor $\boldsymbol{x}$ der unbekannten Parameter als Zufallsvektor definiert, dem eine Priori- und Posteriori-Dichte zugeordnet wird. Das bedeutet jedoch nicht, daß der Parametervektor $\boldsymbol{x}$ keine Konstanten repräsentieren dürfe wie beispielsweise die Koordinaten eines festen Punktes. Durch die Priori- und die Posteriori-Dichte wird die Wahrscheinlichkeit bestimmt, daß die Werte der Parameter in gewissen Bereichen liegen. Die Wahrscheinlichkeit gibt die Plausibilität dieser Aussagen an, also die Plausibilität von Werten der Parameter. Die Parameter können daher konstante Größen repräsentieren, sie brauchen nicht als Ergebnisse von Zufallsexperimenten interpretiert zu werden.
Die Kenntnis der Posteriori-Dichte $p(\boldsymbol{x}|\boldsymbol{y}, C)$ genügt, um die unbekannten Parameter zu schätzen, um Hypothesen für die unbekannten Parameter zu testen und um Bereiche anzugeben, in denen die Parameter mit vorgegebener Wahrscheinlichkeit liegen. Hierauf wird im folgenden Kapitel eingegangen.

4 Parameterschätzung, Konfidenzregionen und Hypothesenprüfung

Um Parameter zu schätzen oder Hypothesen zu testen, sind verschiedene Wege möglich, und für einen muß man sich entscheiden. Die Entscheidung ist zu beurteilen, um zu wissen, ob eine gute Entscheidung getroffen wurde. Dies hängt von dem wahren Zustand des Systems ab, in dem die Entscheidung zu treffen ist. Das System werde durch den Zufallsvektor $\boldsymbol{x}$ der unbekannten Parameter repräsentiert. Daten mit Information über das System existieren, sie seien in dem Zufallsvektor $\boldsymbol{y}$ zusammengefaßt. Um eine Entscheidung zu fällen, wird die Entscheidungsregel $\delta(\boldsymbol{y})$ aufgestellt, die bestimmt, welche Aktion in Abhängigkeit von den Daten $\boldsymbol{y}$ gestartet wird. Mit den Kosten der durch $\delta(\boldsymbol{y})$ ausgelösten Aktion soll die Entscheidung beurteilt werden. In Abhängigkeit von $\boldsymbol{x}$ und $\delta(\boldsymbol{y})$ wird die Kostenfunktion $L(\boldsymbol{x}, \delta(\boldsymbol{y}))$ eingeführt. Betrachtet werden die a posteriori zu erwartenden Kosten, die mit der Posteriori-Dichte $p(\boldsymbol{x}|\boldsymbol{y}, C)$ berechnet werden,

$$E[L(\boldsymbol{x}, \delta(\boldsymbol{y}))] = \int_{\mathcal{X}} L(\boldsymbol{x}, \delta(\boldsymbol{y}))p(\boldsymbol{x}|\boldsymbol{y}, C)d\boldsymbol{x} \; . \tag{4.1}$$

Die Entscheidungsregel $\delta(\boldsymbol{y})$ wird nun derart festgelegt, daß die a posteriori zu erwartenden Kosten (4.1) minimal werden. Dies bezeichnet man als Bayes-Strategie.
Es sei $\hat{\boldsymbol{x}}$ die Schätzung des Vektors $\boldsymbol{x}$ der unbekannten Parameter, so daß $\delta(\boldsymbol{y}) = \hat{\boldsymbol{x}}$ gilt. Eine einfache Kostenfunktion ergibt sich mit der Quadratsumme $(\boldsymbol{x} - \hat{\boldsymbol{x}})'(\boldsymbol{x} - \hat{\boldsymbol{x}})$ der Fehler $\boldsymbol{x} - \hat{\boldsymbol{x}}$ der Schätzung, die noch durch die Inverse $\boldsymbol{\Sigma}^{-1}$ der positiv definiten Kovarianzmatrix $D(\boldsymbol{x}) = \boldsymbol{\Sigma}$ der Parameter $\boldsymbol{x}$ gewichtet wird, so daß die quadratische Kostenfunktion

$$L(\boldsymbol{x}, \hat{\boldsymbol{x}}) = (\boldsymbol{x} - \hat{\boldsymbol{x}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \hat{\boldsymbol{x}}) \tag{4.2}$$

erhalten wird. Die Bayes-Strategie führt dann auf die Bayes-Schätzung $\hat{\boldsymbol{x}}_B$ mit

$$\hat{\boldsymbol{x}}_B = E(\boldsymbol{x}|\boldsymbol{y}) \tag{4.3}$$

oder mit der Definition des Erwartungswertes auf

$$\hat{\boldsymbol{x}}_B = \int_{\mathcal{X}} \boldsymbol{x} p(\boldsymbol{x}|\boldsymbol{y}, C) d\boldsymbol{x} \; . \tag{4.4}$$

Mit der Kostenfunktion der absoluten Fehler erhält man die Median-Schätzung und mit Null-Eins-Kosten die MAP-Schätzung $\hat{\boldsymbol{x}}_M$, die Maximum-A-Posteriori-Schätzung

$$\hat{\boldsymbol{x}}_M = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{y}, C) \; . \tag{4.5}$$

Wegen ihrer einfachen Berechnung wird sie häufig angewendet. Sie entspricht der Maximum-Likelihood-Schätzung der traditionellen Statistik.
Mit der Posteriori-Dichte $p(\boldsymbol{x}|\boldsymbol{y}, C)$ für den Vektor $\boldsymbol{x}$ der unbekannten Parameter aus dem Bayes-Theorem läßt sich mit

$$P(\boldsymbol{x} \in \mathcal{X}_u|\boldsymbol{y}, C) = \int_{\mathcal{X}_u} p(\boldsymbol{x}|\boldsymbol{y}, C) d\boldsymbol{x} \tag{4.6}$$

die Wahrscheinlichkeit berechnen, daß der Vektor $\boldsymbol{x}$ im Unterraum $\mathcal{X}_u$ des Parameterraums $\mathcal{X}$ mit $\mathcal{X}_u \subset \mathcal{X}$ liegt. Um als Konfidenzregion zu dienen, ist der Unterraum derart festzulegen, daß er für maximale Dichten einen großen Teil der Wahrscheinlichkeit enthält, zum Beispiel 95%. Als Konfidenzregion $\mathcal{X}_B$ mit $\mathcal{X}_B \subset \mathcal{X}$ wird daher eine Region höchster Posteriori-Dichte, auch H.P.D.-Region genannt, definiert durch

$$P(\boldsymbol{x} \in \mathcal{X}_B|\boldsymbol{y}, C) = \int_{\mathcal{X}_B} p(\boldsymbol{x}|\boldsymbol{y}, C) d\boldsymbol{x} = 1 - \alpha$$

und

$$p(\boldsymbol{x}_1|\boldsymbol{y}, C) \geq p(\boldsymbol{x}_2|\boldsymbol{y}, C) \quad \text{für} \quad \boldsymbol{x}_1 \in \mathcal{X}_B \; , \; \boldsymbol{x}_2 \notin \mathcal{X}_B \; . \tag{4.7}$$

Den Wert für $1 - \alpha$ bezeichnet man als Konfidenzniveau und wählt in der Regel $\alpha = 0,05$. Es läßt sich zeigen, daß das Hypervolumen der Konfidenzregion (4.7) minimal im Vergleich zu den Hypervolumen beliebiger Konfidenzregionen mit dem Konfidenzniveau $1 - \alpha$ ist.
Es seien $\mathcal{X}_0 \subset \mathcal{X}$ und $\mathcal{X}_1 \subset \mathcal{X}$ Unterräume des Parameterraums $\mathcal{X}$, und $\mathcal{X}_0$ und $\mathcal{X}_1$ seien disjunkt, also $\mathcal{X}_0 \cap \mathcal{X}_1 = \emptyset$. Die Annahme $\boldsymbol{x} \in \mathcal{X}_0$ bezeichnet man als Nullhypothese und $\boldsymbol{x} \in \mathcal{X}_1$ als Alternativhypothese. Die Nullhypothese H_0 ist gegen die Alternativhypothese H_1 zu testen, folglich

$$H_0 : \boldsymbol{x} \in \mathcal{X}_0 \quad \text{gegen} \quad H_1 : \boldsymbol{x} \in \mathcal{X}_1 \; . \tag{4.8}$$

Null-Eins-Kosten werden eingeführt, indem der richtigen Entscheidung für eine korrekte Nullhypothese H_0 oder eine korrekte Alternativhypothese H_1 keine Kosten aufgebürdet werden. Die Bayes-Strategie führt dann auf die Entscheidungsregel, falls

$$\frac{\int_{\mathcal{X}_0} p(\boldsymbol{x}|\boldsymbol{y}, C) d\boldsymbol{x}}{\int_{\mathcal{X}_1} p(\boldsymbol{x}|\boldsymbol{y}, C) d\boldsymbol{x}} > 1 \; , \text{ akzeptiere} \quad H_0 \; . \tag{4.9}$$

Andernfalls ist H_1 anzunehmen. Ähnliche Entscheidungsregeln erhält man, falls spezielle Priori-Dichten für die Hypothesen eingeführt werden.

5 Anwendungen

Wendet man die Bayes-Schätzung (4.3) im linearen Modell an, ergeben sich Ergebnisse, die mit denen der traditionellen Statistik übereinstimmen, so daß die Bayes-Statistik die Resultate der traditionellen Statistik enthält. Die Ergebnisse sind daher allgemeiner, denn beispielsweise bei der Ableitung robuster Schätzverfahren mit Hilfe der Bayes-Statistik gewinnt man den Vorteil,

daß Konfidenzbereiche angegeben oder Hypothesentests vorgenommen werden können (Koch und Yang 1998a), was mit der traditionellen Statistik nicht möglich ist. Auch das robuste Kalman-Filter ist mit der Bayes-Statistik leicht angebbar (Koch und Yang 1998b). Statistische Inferenz von Varianzkomponenten, deren Schätzung von Grafarend (1978) und Grafarend und D'Hone (1978) erstmalig für geodätische Problemstellungen umfassend untersucht wurden, läßt sich ebenfalls mit der Bayes-Statistik lösen (Koch 1988; Ou und Koch 1994). Das Modell der Prädiktion und Filterung macht erst dann wirklich Sinn, wenn es im Sinne der Bayes-Statistik interpretiert wird (Koch 1994). Die automatische Interpretation digitaler Bilder mit Hilfe von Markoff-Zufallsfeldern benötigt die Bayes-Statistik (Klonowski 1998; Köster 1995; Koch 1995b). Schließlich beruhen die Bayes-Netze, die Entscheidungen in Systemen mit Unsicherheit ermöglichen, auf der Bayes-Statistik (Koch 1999). Bayes-Netze eignen sich ebenfalls zur automatischen Interpretation digitaler Bilder (Koch 1995a; Kulschewski 1999) oder für Entscheidungen im Zusammenhang mit Informationssystemen (Stassopoulou et al. 1998). Bei der Analyse geodätisch relevanter Daten ist also die Bayes-Statistik nicht mehr fortzudenken.

Literatur

Cox, R.T. (1946) Probability, frequency and reasonable expectation. *American Journal of Physics*, 14:1–13.

Grafarend, E. (1978) Schätzung von Varianz und Kovarianz der Beobachtungen in geodätischen Ausgleichungsmodellen. *Allgemeine Vermessungs-Nachrichten*, 85:41–49.

Grafarend, E. und A. d'Hone (1978) *Gewichtsschätzung in geodätischen Netzen.* Reihe A, 88. Deutsche Geodätische Kommission, München.

Jaynes, E.T. (1995) Probability theory: The logic of science. *http://bayes.wustl.edu/pub/Jaynes/book.probability.theory.*

Klonowski, J. (1998) *Segmentierung und Interpretation digitaler Bilder mit Markoff-Zufallsfeldern.* Reihe C. Deutsche Geodätische Kommission, München (im Druck).

Koch, K.R. (1988) *Parameter Estimation and Hypothesis Testing in Linear Models.* Springer, Berlin.

Koch, K.R. (1990) *Bayesian Inference with Geodetic Applications.* Springer, Berlin.

Koch, K.R. (1994) Bayessche Inferenz für die Prädiktion und Filterung. *Z Vermessungswesen*, 119:464–470.

Koch, K.R. (1995a) Bildinterpretation mit Hilfe eines Bayes-Netzes. *Z Vermessungswesen*, 120:277–285.

Koch, K.R. (1995b) Markov random fields for image interpretation. *Z Photogrammetrie und Fernerkundung*, 63:84–90, 147.

Koch, K.R. (1999) *Einführung in die Bayes-Statistik.* Manuskript. Institut für Theoretische Geodäsie der Universität, Bonn.

Koch, K.R. und Y. Yang (1998a) Konfidenzbereiche und Hypothesentests für robuste Parameterschätzungen. *Z Vermessungswesen*, 123:20–26.

Koch, K.R. und Y. Yang (1998b) Robust Kalman filter for rank deficient observation models. *J Geodesy*, 72:436–441.

Köster, M. (1995) *Kontextsensitive Bildinterpretation mit Markoff-Zufallsfeldern.* Reihe C, 444. Deutsche Geodätische Kommission, München.

Kulschewski, K. (1999) *Modellierung von Unsicherheiten mit Bayes-Netzen zur qualitativen Gebäudeerkennung und -rekonstruktion.* Institut für Theoretische Geodäsie der Universität Bonn, Bonn (in Vorbereitung).

Ou, Z. und K.R. Koch (1994) Analytical expressions for Bayes estimates of variance components. *Manuscripta geodaetica*, 19:284–293.

Stassopoulou, A., M. Petrou und J. Kittler (1998) Application of a Bayesian network in a GIS based decision making system. *Int J Geographical Information Science*, 12:23–45.

GLONASS Carrier Phases

Alfred Leick

ABSTRACT

Processing of GLONASS carrier phase observations differs from that of GPS. These differences are briefly reviewed. Presently GLONASS does not contain selective availability (SA). Simply graphing between-satellite differences reveals parts of SA that is implemented on GPS satellite signals.
The single difference and double difference carrier phase solutions are analyzed in terms of their suitability for baseline determination with GLONASS carrier phases. The single difference and double difference receiver bias terms for phases, labeled SDRB and DDRB respectively, are introduced. The DDRB is numerically verified from observations.
The double difference fixed solution depends on the initial receiver (rover) coordinates. The single difference solution does not have such a dependency. For a test data set the coordinates estimated from both solutions agree within one millimeter even though the initial coordinates were in error by 1.7 m. The double difference ambiguities were fixed using the LAMBDA technique. Using both GPS and GLONASS carrier phases, the ambiguities could be fixed correctly at all epochs, including the first one, with L1 phases only.

INTRODUCTION

There is a strong interest for including GLONASS satellites in any GPS positioning solution. It is well known that additional satellites and frequencies strengthen the solution. The benefits of additional GLONASS satellites are especially noticeable when attempting OTF (On-The-Fly) ambiguity resolution. The fact that GLONASS satellites transmit at different frequencies has attracted much attention, primarily by individuals interested in precise positioning, for example Raby and Dale (1993), Leick et al. (1995), Rossbach and Hein (1996), Hall et al. (1997), and Kozlov and Thachenk (1997).
We will revisit the topic of ambiguity fixing with GLONASS carrier phases and pay attention to frequency-dependent receiver errors. A well-known strength of double differencing GPS carrier phase observations is that the receiver channel bias cancels. This bias is the same for each satellite observed at the same receiver, but differs between receivers. In case of a hybrid GPS/GLONASS receiver the biases for GPS and GLONASS differ. They do not cancel when double differencing the phase observations from satellites of both systems. In fact, the GLONASS channel biases might even exhibit a small variation as a function of temperature and cable length (Dodsen et al. 1999). These variations are not discussed in this paper.
There are several other aspects of the GLONASS system that have been discussed widely in the literature. For example Bykhanov (1999) discusses the GLONASS time system. The differences between the PZ-90 and WGS-84 reference coordinate system have been studied for many years. Russian scientists reported some of their work in Bazlov et al. (1999). Many questions regarding the implications of the different timing and coordinate reference systems for GLONASS and GPS will be answered by the international IGEX campaign (Pascal, 1999). Finally the GLONASS broadcast ephemeris parameterization differs from that of GPS (Stewart and Tsakiri, 1998).

The data sets for this contribution were observed with R100 receivers manufactured by 3S Navigation of Irvine, California, in connection with a general study to asses GLONASS observations (Leick et al. 1998). The pseudoranges and carrier phases were recorded for GPS (L1 only) and GLONASS (L1 & L2). The receivers were located on the roof of the 3S Navigation offices at Irvine.

Data set A consists of several 1-2 week long observation series made with the same receiver at a recording interval of 5 minutes. The Data set was used primarily to compute UREs for GLONASS. The results are reported elsewhere.
Data set B was recorded on June 12, 1998 using a 10 s recording interval. Two receivers operated independently, i.e. they were not connected to an atomic clock.
We follow the RINEX conventions for naming the satellites. For example, G15 and R15 denote the GPS satellite PRN 15 and GLONASS satellite with almanac number 15 respectively.

SA FROM BETWEEN SATELLITE DIFFERENCES

Between satellite differences (BSD) do not depend on receiver clock errors. Their variation over time reveals, among other things, the satellite clock errors. Because there is no selective availability (SA) implemented on GLONASS satellites, the GLO-GPS differences will be affected by the SA dither on GPS. Figure 1 shows several L1 BSD carrier phases with respect to the GLONASS satellite R17. The dither of the GPS clocks is clearly visible from the dashed lines. The GLO-GLO pairs follow a more or less flat line around zero (solid lines).

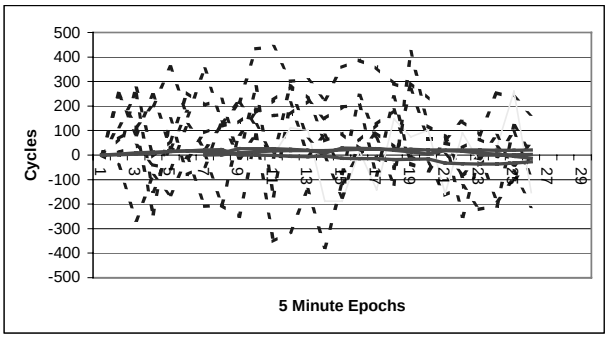


Figure 1: Between Satellite Differences (2 hours from Data set A on DOY 068)

ESTIMATING BASELINES FROM SINGLE DIFFERENCES

An advantage of the single difference formulation is that the signals from GPS and GLONASS satellites are not differenced explicitly. In the context of *single difference solutions*, the terminology *fixed solution* refers to the fact that GPS/GPS and GLO/GLO *double difference ambiguities* have been constrained to integers. For such fixed solutions the adjusted single difference ambiguities are still non-integers.
The mathematical model for carrier phase as applied to short baselines is written as

$$\varphi^{q}_{km,1,GPS} = \frac{f_1}{c}\rho^{q}_{km} + N^{q}_{km,1} + d_{km,1,GPS} - f_1\, dt_{km} \quad (1)$$

$$\varphi^{s}_{km,1,GLO} = \frac{f_1^{s}}{c}\rho^{s}_{km} + N^{s}_{km,1} + d_{km,1,GLO} - f_1^{s} dt_{km} \quad (2)$$

The superscripts $q = 1 \ldots S_{GPS}$ and $s = 1 \ldots S_{GLO}$ identify the satellites. The symbols $d_{km,1,GPS}$ and $d_{km,1,GLO}$ denote single difference receiver biases (SDRB) for the respective systems. The model assumes only one bias term per satellite system, i.e. it does not include frequency dependent terms for GLONASS that may result from temperature variation and other sources.

We combine the single difference ambiguity and the SDRB into a new parameter ξ as follows

$$\xi^{q}_{km,1,GPS} \equiv N^{q}_{km,1,GPS} + d_{km,1,GPS} \tag{3}$$

$$\xi^{s}_{km,1,GLO} \equiv N^{s}_{km,1,GLO} + d_{km,1,GLO} \tag{4}$$

The unknown receiver (rover) coordinates, the parameters ξ, and the receiver clock differences dt_{km} can now be estimated every epoch using Kalman filtering. The outcome of the i[th] epoch is the estimated parameter vector denoted by $X_i(+)$ and its covariance matrix $P_i(+)$. The parameter vector includes the epoch estimates $\xi^{q}_{km,1,GPS}(+)$ and $\xi^{s}_{km,1,GLO}(+)$.

Next, we transform the single difference estimates to double differences. Let p or r denote the GPS or GLONASS base satellite respectively, then the transformation is given by

$$\begin{bmatrix} \varphi^{pq}_{km,GPS}(+) \\ \varphi^{rs}_{km,GLO}(+) \end{bmatrix} = D \begin{bmatrix} \xi^{q}_{km,GPS}(+) \\ \xi^{s}_{km,GLO}(+) \end{bmatrix} \tag{5}$$

$$\Sigma_i = DC_iD^T \tag{6}$$

The matrix D has $(S_{GPS} + S_{GLO} - 2)$ rows and $(S_{GPS} + S_{GLO})$ columns. The matrix C_i is a submatrix of $P_i(+)$. Equation (6) follows from variance-covariance propagation. The symbol Σ_i denotes the covariance matrix of the double difference ambiguities at epoch i.

The transformation (5) generates only GPS/GPS or GLO/GLO pairs of double differences. These do not depend on the SDRB; the respective ambiguities are conceptually integers. It is now possible to attempt to determine the integer double differences ambiguities using a technique such as LAMBDA (Teunissen (1993). The input is the real-valued double difference ambiguities, $\left(\varphi^{pq}_{km,GPS}, \varphi^{rs}_{km,GLO}\right)$ of (5), and the covariance matrix, Σ, of (6). The outcome is a set of integers $\left(\Psi^{pq}_{km,GPS}, \Psi^{rs}_{km,GLO}\right)$. As a last step the epoch Kalman filter solution can be constrained to these integer values. The result is a single difference epoch solution with fixed double difference ambiguities.

The various steps discussed above are repeated for each epoch. Let's denote the updated ξ–parameters by $\hat{\xi}^{q}_{km,1,GPS}$ and $\hat{\xi}^{s}_{km,1,GLO}$. These are the values obtained after the double difference ambiguities have been constrained to integers. The fractional part for the GPS/GLO differences

$$\Delta\xi^{ps}_{km,1} = \hat{\xi}^{p}_{km,1,GPS} - \hat{\xi}^{s}_{km,1,GLO} \tag{7}$$

is the estimated double difference receiver bias (DDRB). This bias is expected to be constant with time and estimates the difference $d_{km,1,GPS} - d_{km,1,GLO}$.

For the sake of completeness let it be stated that the transformation (5) can also be directly implemented in (1) and (2).

Numerical Results: We used L1 pseudoranges and carrier phases of Data Sets B to investigate (7) as a function of time. All ambiguities could be correctly fixed for all epochs, including even the first one. Here we do not address the conditions under which it is possible to fix ambiguities at single epochs or for short intervals. Teunissen et al. (1998) provide an interesting contribution regarding the reliability of ambiguity resolution in such cases.

Figure 2 shows the DDRB differences (7) for the GPS-GPS and GPS-GLO the float solutions, i.e. the double difference ambiguity constraints are not yet imposed. The differences are taken with respect to satellite G5. It is readily seen that, after convergence of the Kalman filter, the GPS-GPS differences are located around zero. A variation of the order of a couple of hundredths of a cycle is seen, although the theoretical value is zero since all GPS satellites transmit on the same frequency.

The mixed GPS-GLO differences are offset by about 0.35 cycles and differ among each other by several hundredths of a cycle as well. Since this variation is of the same size as the one observed for

the GPS-GPS differences, it seems that this data set does not allow one to make the definitive statement about the dependency of the SDRB on the various GLONASS frequencies.

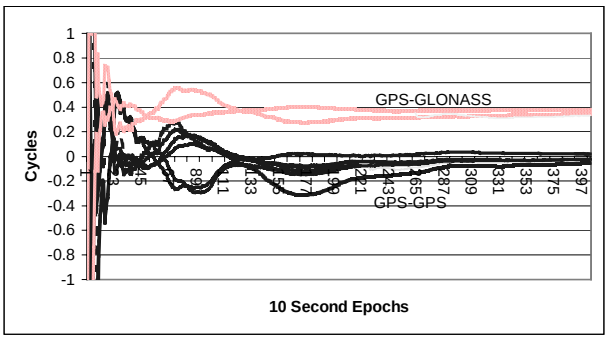


Figure 2: DDRB differences with Float Double Difference Ambiguities (Data set B)

Figure 3 shows the estimated DDRB differences (7) for the fixed solution, i.e. the double difference GPS-GPS and GLO-GLO integer ambiguities have been fixed. The initial variation prior to convergence of the Kalman filter is not present in this figure because the double difference ambiguities could be fixed at all epochs. Because all double difference ambiguities could be fixed, the figure shows identical graphs for each GLONASS satellite. The DDRB differences seem to vary by a couple of hundredths of a cycle over time. The cause for this variation must still be investigated. Figure 3 seems to suggest that is it permissible to constrain the DDRB differences (7) to a constant.

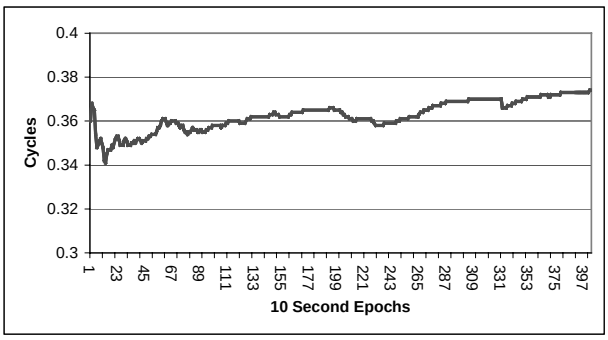


Figure 3: DDRB differences between GPS and GLONASS with fixed GPS/GPS and GLO/GLO Double Difference Integer Ambiguities.

Figure 4 shows the estimated length of the baseline for the float solutions, and the respective plus and minus standard deviations. The straight line at 1.751 m is the length estimated from the fixed solution. It is readily seen that the float and fixed solutions converge and that the fixed solution provides the correct position even at the first epoch. The standard deviations for the double difference ambiguities (not shown in the figure) are in the range of millimeter, whereas those for the single difference ambiguities (same fixed solution) and the receiver clock difference are about 1 cycle and 0.001 μs respectively. Successfully fixing the double difference ambiguities does not imply that the single difference ambiguities can be fixed as well (due to the correlation between ξ and dt_{km}).

The receiver clocks drifted about 440 μs. If we exclude the GLONASS observations, several epochs are needed to fix the ambiguities correctly.

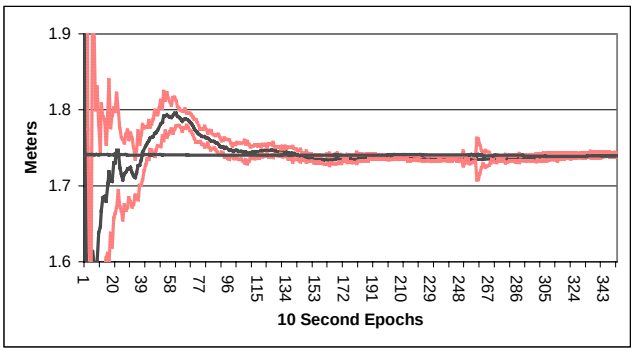


Figure 4: Epoch Position Solutions for Data set B

DRAWBACKS OF DOUBLE DIFFERENCING

Conventional double differencing for GLONASS observations gives

$$\begin{aligned}\varphi^{rs}_{km,1,GPS,GLO} &\equiv \varphi^{r}_{km,1} - \varphi^{s}_{km,1} \\ &= \frac{f_1^{\,r}}{c}\rho^{r}_{km} - \frac{f_1^{\,s}}{c}\rho^{s}_{km} + N^{rs}_{km,1} - \left(f_1^{\,r} - f_1^{\,s}\right)dt_{km}\end{aligned} \tag{8}$$

The double differences depend on the receiver clock error and the frequencies. Figure 5 displays this dependency; the O-C values were computed using known coordinates for the stations and then translated to zero at the first epoch. The dependency on the frequency can readily be seen from the figure; the reference satellite is G5 (1575.42 MHz). The GPS-GPS differences graphically coincide with the horizontal axis and are not visible in this figure.

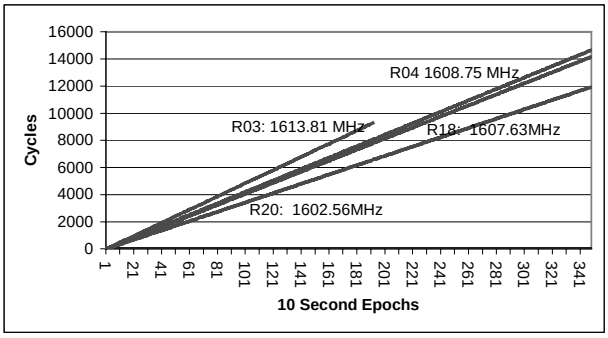


Figure 5: Double difference O-C values for known baseline (Data set B)

Scaling the carrier phases to distance, or to a mean GLONASS frequency, or to $f_1^{\,r}$ or $f_1^{\,s}$ for the (r,s)pair eliminates the receiver clock term but introduces a linear combination of single difference ambiguities whose coefficients are non-integer. The transformed double difference

$$\begin{aligned}\tilde{\varphi}^{rs}_{km,1,GLO} &\equiv \frac{f_1^{\,s}}{f_1^{\,r}}\varphi^{r}_{km,1,GLO} - \varphi^{s}_{km,1,GLO} \\ &= \frac{f_1^{\,s}}{c}\rho^{rs}_{km} + \tilde{N}^{rs}_{km} + \frac{f_1^{\,s}}{f_1^{\,r}}N^{r}_{km,1,0} + \eta^{rs}\end{aligned} \tag{9}$$

$$\eta^{rs} = \left(\frac{f_1^{\,s} - f_1^{\,r}}{f_1^{\,r}}\right)dN^{r}_{km,1} \le 0.01\,dN^{r}_{km,1} \tag{10}$$

contains an integer term $\tilde{N}_{km}^{rs}$ and a small term η^{rs}. The symbol $N_{km,1,0}^{r}$ represents an integer approximation of the single difference ambiguity $N_{km,1}^{r}$ which can be derived from pseudoranges. The size of the small η-term depends on the quality of the initial estimate of $N_{km,1,0}^{r}$, and constitutes a model error when neglected in the fixed solution. This limitation does not apply to the float solution.
Figure 6 shows the double difference residuals G5 - GLO for the batch least-squares implementation of (9) using Data set B. The double difference ambiguities GPS/GPS and GLO/GLO are fixed. The top set of lines is based on approximate coordinates which were in error by about 1.7 m, thus $dN_{km,1}^{r}$ is correspondingly large. Using approximate coordinates that are even less accurate, one would eventually recognize a frequency dependency within this band of lines. The accurate coordinates were used for the bottom set of lines, thus $dN_{km,1}^{r}$ is correspondingly small. The model error (10) causes the shift between both sets of lines. The model error falsifies the position estimate even when the ambiguities are formerly fixed. The bottom set of lines can be directly compared with Figure 3 for the single difference solution. Again, the DDRB differences could be modeled by a constant.

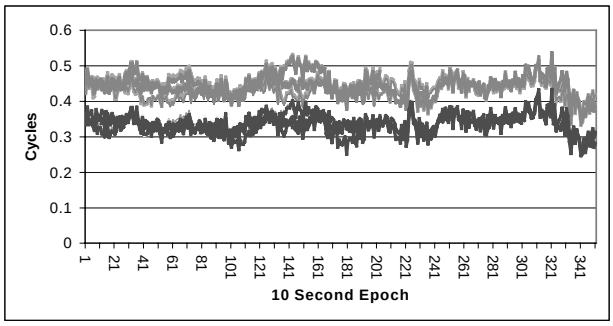


Figure 6: DDRB from Double difference Solution

SUMMARY

The model error that occurs in GLONASS double difference fixed solutions does conceptually not occur with GPS observations.
When processing GLONASS carrier phase observations, caution should be exercised. Ambiguity search might identify the wrong integers and, as such, introduce a bias in the fixed solution. For double differencing to work correctly one must have good a priori knowledge of single difference ambiguities which, in turn, are derived from pseudoranges. Since the accuracy of pseudoranges are potentially effected by multipath, one might be inclined to favor the single difference formulation and fix the propagated double difference ambiguities.

REFERENCES

Bazlov Y., Galazin V., Kaplan B., Maksimov V. & Rogozin V. (1999). GLONASS to GPS. A New Coordinate Transformation, *GPS World*, January, 54 – 58.

Bykhanov, E. (1999). Timing and positioning with GLONASS and GPS. *GPS Solutions*, 3(1).

Dodson, A. H., Moore, T., Baker, D. F. & Swann, J W. (1999). Hybrid GPS + GLONASS. *GPS Solutions*, 3(1).

GLONASS ICD-95. 1995. GLONASS Interface Control Document. International Civil Aviation Organization (ICAO, RTCA Paper No. 639-95)

Hall T., Burke B., Pratt M. & Misra O. 1997. Comparison of GPS and GPS+GLONASS Positioning Performance, *Proc. ION-GPS-97*, 1543-1550

Kozlov, D. & Tkachenko M.. 1997. Instant RTK cm with Low Cost GPS+GLONASS C/A Receivers, *ION-GPS-97*, 1559-1569.

Leick A. 1995. *GPS Satellite Surveying*, J. Wiley & Sons, New York.

Leick A., Li J., Beser J. & Mader G. 1995. Processing GLONASS Carrier Phase Observations - Theory and First Experience, *Proc. ION-GPS 95*, 1041-1047

Leick, A., Beser J., Rosenboom, P., & Wiley B. (1998) Assessing GLONASS Observation. *Proc. ION-GPS-98*, Nashville, pp. 1605-1612

Pascal W. (1999). IGEX-98: International GLONASS Experiment. *GPS Solutions* 3(2).

Raby P. & Daly P. 1993. Using GLONASS System for Geodetic Survey, *Proc. ION-GPS 93*, 1129-1138.

Rossbach U., & Hein G. 1996. Treatment of Integer Ambiguities in GDPS/DGLONASS Double Difference Carrier Phase Solutions, *Proc. ION-GPS 96*, 909-916

Stewart M. & Tsakiri M. (1998). GLONASS Broadcast Orbit Computations. *GPS Solutions*, 2(2)

Teunissen P. J. G. 1993. Least-squares estimation of the integer GPS ambiguities. Delft Geodetic Computing Centre *LGR Series* No. 6. Delft University of Technology.

Teunissen P. J. G., P. Joosten, & Odijk D. 1998. The Reliability of GPS Ambiguity Resolution. *GPS Solutions*, 2(3)

BIOGRAPHY

Dr. Alfred Leick is a professor at the University of Maine, Department of Spatial information Science and Engineering. He is author of the book *GPS Satellite Surveying*. He spent the academic year 1985-86 at the University of Stuttgart collaborating with Prof. Grafarend while being supported by the Alexander von Humboldt foundation. He revisited the Geodetic Institute in August 1999, being supported again by the Humboldt foundation.

Analytical versus Numerical Integration in Satellite Geodesy

Dieter Lelgemann and Chunfang Cui

„The purpose of computation is insight, not numbers“ (Hamming).

I Introduction and historical developments

In the near future a big step has again to be expected in satellite geodesy. Extremely precise measuring systems (accelerometer, low-low SST (relative accuracy 10^{-11}), gradiometer) in satellites orbiting as low as possible will allow not only the determination of the global gravity field in form of harmonic coefficients up to a limit somewhere below n = m = 180, but even the tracking of the extremely small signals in the gravity field due to mass redistributions in atmosphere, oceans and solid Earth down to the inner core.

An good understanding for all possible shortcomings in the data analysis process is required last but not least to avoid an interpretation of „geodetic observation errors“ as physical signals of mass redistributions.

Regarding the praxis of data analysis most professionals seem to be convinced today that only the use of numerical integration techniques will allow a reasonable analysis of these precise data.
We are now facing the danger that from two alternative and competitive procedures, namely

- the analysis in the time domain
- and the analysis in the spectral domain,

the latter will only insufficiently be supported. A meaningful analysis in the spectral domain, however, can only be performed on the basis of an analytical integration, since only an analytical solution connects the periodic effects in orbital data with the force parameters in a physically judicious mode.

The complete or general solution of N differential equations of first order contains a set of N arbitrary constants. Assigning particular numerical values to those constants one gets a so-called particular solution of the differential equation system.

Any numerical integration of an initial state problem corresponds to such a particular solution. By a suitable variation of the initial state vector (as well as the force parameters (variational equations)) one can generate a set of particular solutions.

From this set we usually pick up that particular solution which is in best agreement with observations, that is, which provides for the squares of the residuals the least sum (least squares adjustment).

However, a complete or general solution will be required for a clear and concrete understanding of the geometrical/physical nature of an energy process as well as for an understanding of the information at hand about this process, that are the gravity field parameters and the measurements. Those solutions

are well-known in celestial mechanics and in satellite geodesy as analytical solutions; the derivation of such kind of solutions may be called analytical integration.

Perturbation theory is the general concept underlying both approaches, analytical and numerical integration. In the latter case a numerically derived orbit is used as a reference and afterwards the so-called variational equations are used to determine the „perturbations“ or corrections.

Analytical integration is more complex and based on several tools such as

- Problemoriented choice of orbital variables (Hill-, Kepler-, Delauney-variables, etc.)
- Infinitesimal transformation of variables based on series expansions (Trig. series, power series)

How did it come to the present situation, that is to an overestimation of the numerical integration approach and an underestimation of the analytical integration approach?

At the early times of satellite geodesy until the seventies only relatively inaccurate measurements (Baker-Nunn camera, Laser 1. generation) have been available. Based on analytical solutions (Brouwer, King-Hele, Kaula, Kozai, Gaposchkin etc.) special attention was devoted in particular to resonance effects, because only in relation to those effects the relatively inaccurate data could provide a reasonable signal/noise ratio. First developments of the so-called „lumped coefficient concept“, that is the analysis in the spectral domain, have later been carried on and extended, but only off the main path.

Due to the big jump in accuracy of the observations (Laser 3. generation, altimeter, GPS-phase observations) in the seventies the accuracy of the analytical solutions of first order (relative accuracy: 10^{-6}) as on hand at that time was insufficient for the analysis of those high precision data. We all have been forced at that time to restrict ourselves to numerical integration and analysis in the time domain; only this approach delivered the accuracy for the analysis of those high-precision data such as Laser and allowed the inclusion of all kind of force fields, gravitational and non-gravitational, in a systematic manner.

Of course, the praxis could not wait at that time whether eventually an analytical solution of high-precision would be developed under an inclusion of all kind of force fields in a systematic manner. As a result, the use of spectral analysis as a tool was very often considered with uneasy feelings and finally often not be clearly understood anymore.

A really alternative method of data analysis in the spectral domain could only be expected if an analytical solution could be developed of equally high accuracy as the numerical ones.

The authors have worked after a stay of the first author in the USA in 1975/76 to develop such kind of analytical solution. (Cui 1997) presents an analytical solution of second order (relative accuracy: 10^{-9}), which will be extended in the next future to a solution of third order (relative accuracy 10^{-12} corresponding to the accuracy of upcoming SST-data of 0.5×10^{-11}). An outline of the strategy for its further development is shown in (Cui 1999).

Some basic criteria which should guide the development of any analytical solution designed for data analysis in the spectral domain are given in the sequel.

Of course, at this stage we have to ponder and discuss again the merits and drawbacks of both approaches, the numerical and analytical one. The following article should be considered as a first attempt in this respect, probably still biased in the moment from the point of view of the authors.

II Some basic aspects of the numerical integration approach

A very good and pleasantly short description of the approach can be found in (Beutler 1996). However, to discuss merits and drawbacks some comments may be opportune which may be structured into 4 sections:

- Reference orbit and its differential equations
- Generation of reference orbit data (model data)
- Variation of parameters (state vectors and force parameters, auxiliary parameters)
- Spectral analysis of strings of data (orbital variables, model data, residuals, observations etc.)

Reference orbit. The 6 differential equations for an orbit (primary equations) can be numerically integrated using a suitable and sufficiently accurate technique if and only if **numerical values** are given for

- the parameters P_f of models for all force vectors $\mathbf{k}(P_f)$
- 6 (in case of SST 12) orbital variables $P_{isv}(t_0)$ describing position vector $\mathbf{r}(P_{isv})$ and velocity vector $\dot{\mathbf{r}}(P_{isv})$ at the reference epoch t_0 (initial state vector)

The movement of the center of mass of the satellite will further be described either by 6 instantaneous orbital variables $P_t(t)$ or by the 6 Cartesian components of the instantaneous state vectors $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ at usually equidistant epochs $t = t_k$ (e.g. $t_{k+1} - t_k = 1\,\mathrm{min}$); at any other epoch *t* the instantaneous state vectors may be obtained by a suitably chosen interpolation procedure.

In case an observation is connected to a station at the Earth surface (or to a point at the ocean surface as in altimetry), position and velocity of this station or point, respectively, must also be expressed by a function depending on time and certain constants, but we will restrict ourselves here to the simple case of SST.

The form of the differential equations will of course vary with the 6 orbital variables applied. The equations using Cartesian components as variables are given e.g. in (Beutler 1996). Using as an other example Hill variables we get the 6 differential equations

$$\begin{pmatrix} d\dot{r}/dt \\ dG/dt \\ dH/dt \\ dr/dt \\ du/dt \\ d\Omega/dt \end{pmatrix} = \begin{pmatrix} G^2/r^3 \\ 0 \\ 0 \\ \dot{r} \\ G/r^2 \\ 0 \end{pmatrix} + \begin{pmatrix} \partial V/\partial r \\ \partial V/\partial u \\ \partial V/\partial \Omega \\ 0 \\ -\partial V/\partial G \\ -\partial V/\partial H \end{pmatrix} + \begin{pmatrix} F_1 \\ rF_2 \\ r(\cos i F_2 - \sin i F_3) \\ 0 \\ (r/G)\cot i \sin u F_3 \\ r(G\sin i)^{-1} \sin u F_3 \end{pmatrix}$$

Here *V* is the gravity potential (geopotential, disturbing potential due to the other celestial bodies, tide induced potential, etc.) and F_i are the components of the non-gravitational forces with respect to the Gaussian basis (for details see Cui and Mareyen 1992).

Model data. Any data can be modeled as a function of the instantaneous orbital variables by including additional parameters P_m describing properties of the measurement process (e.g. eccentricity vector between center of mass and phase center of the antenna, tropospheric/ionospheric reduction model etc.):

$$\tilde{l} = \tilde{l}\left(\overline{X}_I(P_t), \dot{\overline{X}}_I(P_t), \overline{X}_{II}(P_t), \dot{\overline{X}}_{II}(P_t), P_m \right)$$

Those equations are called observation equations. Forming the differences with measured values *l*,

$$\Delta l = l - \tilde{l}$$

we get information how good our model describes the real process of motion.

Variation of parameters; Perturbations. We may look at the total differentials of the 6 instantaneous state variables as linear functions of the differentials of the initial state variables and the force parameters; that may be expressed by

$$[\Delta P_t] = \left[\frac{\partial P_t}{\partial P_{isv}}\right][dP_{isv}] + \left[\frac{\partial P_t}{\partial P_f}\right][dP_f].$$

The matrices of partial derivatives can be computed by the solution of a differential equation system, the variational equations.

The number of the equations of this system is equal to the number of the parameters P_{isv} (6 or 12, respectively) and P_f. If a huge set of force parameters should be determined from the data, as will be the case in SST, one is confronted with a huge variational equation system.

However, „whereas highest accuracy is required in the integration of the primary equations, the requirements are less stringent for the variational equations“ (Beutler 1996, p. 78). The equations

$$d\tilde{l} = \left[\frac{\partial \tilde{l}}{\partial P_t}\right][dP_t] + \left[\frac{\partial \tilde{l}}{\partial P_m}\right][dP_m]$$

are called linearized observation equations (sometimes misleading also variational equations) in case the parameters $P = \{P_{isv}, P_f, P_m\}$ are considered as unknowns in an adjustment procedure.

Spectral analysis and mean orbit. This is by far the weakest point in the pure numerical integration approach. Of course, one can always use an empirical spectral model; the orbit length (short/long arc) provides then the smallest and the numerical integration step the highest frequency.

However, those empirical frequencies depend on an arbitrary chosen computation model and have nothing to do with the frequencies of the orbit perturbations generated by the physical forces. In fact, people often skip the spectral investigations therefore, presenting as a substitute illustrating pictures (see e.g. Beutler 1996, p. 59ff).

Since „the osculating elements are not well suited to study the long term evolution of the satellite systems“ so-called mean elements are often introduced. „The purpose is the same in all cases: one would like to remove the higher frequency part of the spectrum in the time series of the elements“.

„There are many different ways to define mean orbital elements starting from a series of osculation elements.“ In fact, there is, but only if an arbitrarily defined empirical spectrum is introduced and certainly **not** if the generating forces will define the spectrum as it corresponds to reality.

III Some basic aspects of the analytical integration approach

Regarding satellite data analysis one can hardly overestimate one big advantage of analytical orbit integration: spectral analysis or the „lumped coefficient concept“, respectively, may not only be used for efficient algorithms but over all for a much better insight into the information content of data.
Having this in mind an efficient analytical solution should be designed fulfilling some important criteria, among them at least

1) The global accuracy of analytical integration should meet all present and future accuracy requirements
2) A suitable technique for comparisons with results of numerical integration should be available
3) It must be possible to introduce all kind of force fields in a systematic and unified manner
4) The analytical solution should be designed to be as efficient as possible for applications of spectral analysis techniques
5) Regarding applications the basic structure of the analytical solution should be most simple and lucid even though details may remain fairly complex.

Those criteria have been developed in the course of the derivation of a second order solution which will be used here to illustrate those general comments in this section.

1) The global (and not just some local) accuracy of analytical solutions can be determined in powers of $\varepsilon = c_{2(0)} \approx 10^{-3}$. We have to distinguish

- solutions of first order: $\varepsilon^2 \approx 10^{-6}$ (e.g. Kaula's solution)
- solutions of second order: $\varepsilon^3 \approx 10^{-9}$ (e.g. Cui's solution)
- solutions of third order: $\varepsilon^4 \approx 10^{-12}$ (in development)

The accuracy requirements depend of course strongly on the data accuracy. Upcoming SST-data of the GRACE-mission will have a relative accuracy of 0.5×10^{-11}; therefore a third order solution would be necessary if numerical integration should entirely be avoided in the data analysis process. For the solution of the variational equations a second order solution only will be sufficient.

2) In view of a comparison of numerical and analytical integration one has to recognize the fact that the technique of numerical integration is extremely inflexible in contrast to analytical ones. As a consequence the analytical solution must be adopted to the numerical ones.

Principally, the analytical solutions are based on the parameters of a mean orbit, the numerical ones on the numerical values for the orbital variables describing the initial state vector.

The inverse analytical solution, that is, the computation of the elements of the mean orbit from given initial (or any instantaneous) state vector is of utmost importance in view of comparisons with numerical integration results.

Moreover, the inverse analytical solution will provide a very efficient tool for the definition of a physically meaningful mean orbit.

Last but not least the inverse analytical solution will provide an extremely efficient tool to use also in satellite data analysis the traditional geodetic concept of data reduction with the goal to simplify the functional model. As well-known this concept is very often and efficiently applied in other domains of geodesy, where it has a long tradition.

A technique to proceed from the initial state vector to elements of the mean orbit and vice versa is described in (Cui and Lelgemann 1995).

3) Regarding the force field it is of uppermost importance that physically two completely different sources of forces govern the orbit

- gravitational forces (Earth and Earth-tide, Moon, Sun etc.),
- non-gravitational or surface forces (air drag, radiation pressure etc.).

The instantaneous state variables may be separated into the sum

$$P_t(t) = \overline{P}_t(t) + \delta P_{fg}(t) + + \delta P_{fs}(t)$$

The perturbations of the orbital variables due to non-gravitational forces are always defined as being zero at the initial state epoch t_0. Those orbit perturbations are growing irregularly; their description using trig. functions may therefore not an efficient concept.

The best way to proceed may be the following. The instantaneous surface force can be expressed by its components with respect to the Gaussian basis in form of an empirical time series (numerical values at equidistant ($\Delta t = const$) epochs

- either using data of an accelerometer as foreseen for the GRACE-mission
- or otherwise using an empirical model such as developed in (Arfa-Karboodvand, 1997)

Using the observation equations for the Hill variables (see section 1) together with crude and approximate instantaneous orbital variables one can express the effect of the surface forces accurate enough by empirical, equidistant epoch data $F_i^{sf}(t_k)$.

4) It can be shown that furthermore **all** gravitational forces will result in a **secular** movement just only of the ascending node (right ascension of the ascending node Ω) and of the perigee (argument of perigee ω) of a quasi-secular rotating ellipse as a reference (mean) orbit. All gravitational induced perturbations of such a reference orbit are then purely periodic (inclusive constant terms), having the simple functional form

$$\delta P_t = \sum_q \sum_m \sum_k \left[a_{kmq}(i, G, e; P_{fg}) \cos(ku + mh + qf) + b_{kmq}(i, G, e; P_{fg}) \sin(ku + mh + qf) \right]$$

where u , $h = \Omega - \Theta$ and f are the argument of latitude, the geographical longitude of ascending node and the true anomaly, all of the reference orbit.

The amplitudes of the trig. functions depend on the constant orbital elements i , G and e of the reference orbit as well as on the gravitational force field parameters. They are often called „lumped coefficients“ in the literature.

We have already checked that a third order solution will have the same structure, that is, very small additional terms only for the periodic perturbations have to be added to a second order solution (Cui 1999).

Theoretically, the summations have to be extended to the limits $-\infty < k < \infty$, $0 \le m < \infty$ and $-\infty < q < \infty$, but for applications the smallest summation limits should be fixed according to the accuracy requirements.

Despite some details of the solution may be fairly complex (see e.g. Cui 1997) its basic structure is obviously very simple.

The solution can also approximately be expressed by reducing the numbers of angular variables.

In case of nearly circular orbits it can be shown that the true anomaly can be expressed with sufficient accuracy as a function of the argument of latitude

$$f = (1-\sigma)u + const$$

where $\sigma = \sigma(i, G; \tilde{P}_{fg}) = O(c_{2(0)})$ is a very small number and $\tilde{P}_{fg}$ is a subset of the gravitational force field parameters.

In case of geosynchronous (repeating) orbits there will be a fixed ratio between the revolutions of the satellite and of the Earth,

$$-\frac{\dot{u}}{\dot{h}} = \frac{\dot{u}}{\dot{\Theta} - \dot{\Omega}} = \frac{p}{q}, \qquad p, q = \text{integer numbers}$$

that is, during q revolutions of the Earth the satellite will perform p revolutions. In such cases the perturbations can be expressed as a function of just one angular variable.

In case of so-called „deep resonance“ (critical inclination $i = 63.4°$, exact polar orbits $i = 90°$, geosynchronous orbits) some lumped coefficients will become infinitely large whereas the corresponding frequency becomes infinitely small. With other words the perturbations become similar to secular effects. In this case a Taylor series expansion may be used together with Encke's technique. The same method may also be used to investigate possible coupling effects of gravitational with non-gravitational forces.

5) Regarding the application of analytical solutions for spectral analysis we may separate the (linearised) observation equation system into

$$[d\tilde{l}] = \left[\frac{\partial \tilde{l}}{\partial P_{isv}}\right][dP_{isv}] + \left[\frac{\partial \tilde{l}}{\partial P_{fg}}\right][dP_{fg}] + \left[\frac{\partial \tilde{l}}{\partial P_{fs}}\right][dP_{fs}] + \left[\frac{\partial \tilde{l}}{\partial P_{m}}\right][dP_{m}]$$

that is, into terms according to

- initial state variables P_{isv} (or mean orbit variables $\overline{P}$)
- gravity force field parameters P_{fg}
- surface force parameters P_{fs}
- measurement technique related parameters P_m

If as a goal the determination of the gravity field is intended the second term will be of major importance. Neglecting just for the moment the two last terms we may express even the **non-linear** observation equations in the form

$$\tilde{l} = \tilde{l}_0(\overline{P}) + \delta\tilde{l}$$

where $\delta\tilde{l}$ may be expressed by a formula similar to those for the perturbations of the orbital variables (see Cui 1997)

$$\delta\tilde{l} = \sum_q \sum_m \sum_k \left[a_{kmq}(i, G, e; P_{fg}) \cos(ku + mh + qf) + b_{kmq}(i, G, e; P_{fg}) \sin(ku + mh + qf) \right],$$

where a_{kmq} and b_{kmq} are now the lumped coefficients of the observable *l*.
Consequently, we may separate the corresponding matrix into the product of two matrices

$$\left[\frac{\partial \tilde{l}}{\partial P_{fg}} \right] = \left[\frac{\partial \tilde{l}}{\partial a_{kmq}} \;\middle|\; \frac{\partial \tilde{l}}{\partial b_{kmq}} \right] \begin{bmatrix} \frac{\partial a_{kmq}}{\partial P_{fg}} \\ \frac{\partial b_{kmq}}{\partial P_{fg}} \end{bmatrix} = \mathbf{BT}$$

This separation is fundamental for the application of the „lumped coefficient concept“, that is the spectral analysis technique.

The matrix $\mathbf{N} = \mathbf{B}^T\mathbf{B}$ will become for a sufficient length of a data set a **diagonal dominant** matrix, in the case of an unlimited length of the data set a diagonal one.

The matrix **T** is a **very sparse** matrix separating into small diagonal block matrices connecting the gravity force field parameters P_{fg}, among them in particular the harmonic coefficients c_{nm} and s_{nm}, with the lumped coefficients.

The effects of a variation of either the mean variables or the initial state variables must be carefully analyzed with respect to the question whether periodic effects will occur with analogue frequencies as due to gravitational force effects in order to avoid aliasing in the framework of a determination of force parameters P_{fg}.

The same must be done for effects of a variation of surface forces or of auxiliary parameters P_m on the data. If aliasing may occur the determination of the force parameters must be done with extreme due care.

The spectral analysis technique was already extremely helpful in the framework of altimeter data analysis. Using older GEOSAT-ephemeris provided by NOAA with a radial orbit error of about 5 m it turned out that the largest part of those orbital errors have been generated by the use of an inadequate Earth gravity model; the radial orbital error could be reduced to 0.30 m using crossover-differences as data (Cui and Lelgemann 1995). In contrast, as a study in progress has shown, the orbital error of ERS-ephemeris of about 10 cm **cannot** be explained by an insufficient Earth gravity field model.

In any case such kind of investigations can only be done in the spectral domain on the basis of a precise and suitably designed analytical solution.

IV Numerical versus analytical integration

Having in mind a comparison of the results of both approaches we have to clarify first for an unbiased judgement possible problematic sides of both techniques, since those may be the origin of imperfections of the results.

Since the authors had uneasy feelings with a pure numerical approach they have started two decades ago with the development of a precise analytical solution of higher order. Those uneasy feelings are based on the following arguments.

1) Insufficient cognition of the information content of a specific kind of observational data (like Laser, altimeter, SST etc.), that is, insufficient cognition of the unknowns which can unobjectively be determined from those observations.
2) Insufficient comprehension of the „correlation“ effects occurring in the determination of the unknowns in case of a completely filled up (not-sparse), bad conditioned normal equation matrix of huge size (more than 30,000 unknowns in case of a gravity field resolution of n=m<180). The computation of the correlation matrix will never be sufficient for a necessary insight.
3) Extremely high may be the also danger that near the absolute minima for the sum of pvv there are relative minima. In such a case it depends just on the approximate starting values for the unknowns in the Newton-Raphson iteration which minima will be reached with the final solution.
4) Insufficient knowledge about the effects of the definition of the orbital arc length (short arc, long arc). An unobjectionable determination of the force parameters will only be possible if aliasing of force effects into the initial state parameters is excluded.
5) Insufficient knowledge about the frequencies of a given data string, that is, about the forms in the variation of measurements generated by specific force field components, by a variation of initial state parameters etc.
6) Insufficient knowledge about the consequences of deep resonance effects which occur for geosynchronous orbits (repeating orbits) as often chosen for Earth observing or navigational satellites.

Moreover, the error estimate of analytical integration (e.g. 10^{-12} for a theory of 3. order) is always a global error estimate. In contrast, numerical integration provides in fact efficient local error estimates, but poorly global ones.

7) Insufficient **global** error estimates in case of very low flying satellites (large longperiodic superimposed by very small shortperiodic disturbations), that is, if very small step sizes are required.
8) Error estimation in case of huge systems of more than 30,000 variational equations as will occur in SST-data analysis.

All those possible problems must be carefully considered in case of e.g. SST-data analysis. For the case of analytical integration approaches the authors do not see similar problems, but of course our point of view may be biased and an „advocatus diabolus“ would be desirable.

One problem using analytical integration may be the correctness of the complex formulas connecting the lumped coefficients with the force parameters. Good theoreticians may check, however, the derivation of those formulas on the one hand and comparisons with simulation results using numerical integration may give hints about yet incorrect analytical terms.

Despite the fact that a lot of individual objectives remain to be investigated the basic concepts for at least one high-precision analytical solution (there may be other analytical methods providing an even more efficient or a simpler solution) has been developed providing an alternative method to the numerical integration approach for the analysis of the extremely precise tracking data as will come up in the near future.

V Final remarks

This article was not written to stir up war between partisans of the analytical and the numerical approach. But it is a fact that we have two alternative concepts to analyze the SST-data obtained in the next future from expensive missions and **we should use both concepts**.

The basic theoretical developments for spectral domain analysis have already been performed and the next steps (e.g. software developments, simulation studies, analysis of measurements) will require team work and with this financial support. The development of todays high-precision numerical approach software was very time consuming and expensive; the future development in the framework of a high-precision analytical approach will certainly be faster and cheaper.

It seems to be urgent now to discuss openly possible merits and shortcomings of both approaches and over all use both methods or moreover a combination of both, a semi-analytical one, for the data analysis.

The comparison of the results of both techniques may startle both sides but will certainly give, according to our opinion, an enormous progress not only for the theoretical foundation of our beloved science but also for the interpretation of the results of the new geodetic satellite missions.

Acknowledgment

Dear Erik. „**Nichts ist so praktisch wie eine gute Theorie**“ has, like all good proverbs, two meanings. It provides not only a definition for the conception „good theory“ but states also how important theory may be for practical work. You have always insisted that the theoretical background of geodesy has to be cleared and extended in our times of fast progressing observation techniques and computers. Thank you for this insistence, old buddy.

References:

Arfa-Karboodvand, K. (1997) Zur präzisen Berechnung der Oberflächenkräfte eines erdgebundenen Satelliten auf Basis der Hill-Variablen. DGK, Reihe C, Nr.476.

Beutler, G. (1996) GPS Satellite Orbits, in GPS for Geodesy, edited by Kleusberg, A and Teunissen, P.J.G., Springer, Berlin, Heidelberg, New York.

Cui, Ch. (1997) Satellite Orbit Integration based on Canonical Transformations with Special Regard to the Resonance and Coupling Effects. DGK, Reihe A, Nr.112.

Cui, Ch. & Mareyen, M. (1992) Gauss's equations of motion in terms of Hill variables and first application to numerical integration of satellite orbits. manuscripta geodaetica, 17, 155-163

Cui, Ch., Lelgemann, D. (1995) Analytical dynamic orbit improvement for the evaluation of geodetic-geophysic satellite data. Journal of Geodesy, 70: 83-97.

Cui, Ch. (1999) Satellite Orbital Theory: Strategy to Derive a Third Order Analytical Solution Using Lie-Series. In: M. Orhan Altan and Lothar Gründig (Eds.): Proc. of the Third Turkish-German Joint Geodetic Days. Towards a Digital Age. Istanbul/Turkey, June 1-4, 1999. pp. 699-709

About the generalised analysis of network-type entities

Klaus Linkwitz

We find network-type entities in many technical fields, either as abstracted physical realities or as theoretical models to describe a typical structure of the underlying problem. In traffic engineering we have transportation- and in telecommunication communication networks. The electrical engineer thinks in the context of electrical networks and sewage waters are canalised in hydromechanic networks. The structural engineers encounters prestressed cable-nets. The geodesist, finally, is involved in geodetic networks with many subdivisions: levelling- and triangulation nets, trilateration- or „combined“ nets, satellite- nets.

Each profession has developed its peculiar skills, techniques, and algorithms to „calculate“ and „analyse“ the networks appertaining to its own profession, as if it were a peculiar field belonging just to its own profession.

However, it will be shown, that the underlying theories can be generalised leading to a uniform, systematic approach to all network-type entities and their calculation and analysis. This is not only intellectually interesting but means also, that the huge number of existing numerical algorithms can be transferred - with only slight adaptations - from one field to the other. Moreover it means a definite step towards interdisciplinarity. When it comes to networks, the electrical- may understand the transportation engineer, and the geodesist may become a professional in certain fields of structural engineering.

1. Tools of graph theory for the description of nets

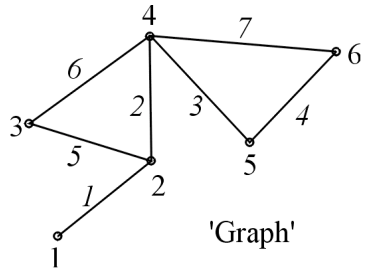


Fig. 1: Graph of 6 nodes and 7 branches

As a prerequisite for a generalised analysis of netlike entities, the *topological* and *semantical* properties of the net have to be separated. The adequate tool to describe the topology is furnished by graph theory and its matrix formulation. In the context of graph theory a „net“ consists of a set of „nodes“ and a set of „branches“, connecting the nodes. By no means the nodes must be always be points with coordinates in 2D- or 3D space. In the *physical realisation* of a net a „node“ may be realised by a certain time, the state of a system, an agglomeration of vehicles, or, naturally, a point in space. Thus physical realisations of nets could be geodetic nets, levelling nets, nets of transportation or traffic, electrical nets, netlike structures in architecture, etc.

When describing the topological relationships in a net, we have to discriminate between „incidence“ and adjacency relationships. The relationships are called „topological“ as they are invariant with respect to mappings. They may be represented in matrices.

The topological relations between nodes and branches are described in the „branch-node“ matrix $\mathbf{C}$, which easily can be constructed following its definition

$$c_{i,k} = +1 \; or \; -1 \text{ if branch } \textit{from} \text{ node } \textit{to} \text{ node k, else } c_{i,k,} = 0$$

Branch	*Nodes*	
1	*1*	*2*
2	*2*	*4*
3	*4*	*5*
4	*5*	*6*
5	*2*	*3*
6	*3*	*4*
7	*4*	*6*

Fig. 2: List of graph

As an example for a net we take Fig. 1 with the branch node matrix $\mathbf{C}$. We observe, that each row of the matrix has exactly one element +1 and one element -1; the rest is filled by zeros. As the sum of all columns is the 0-vector, we conclude, that the columns are mutually linearly dependent. In an interconnected net the rank deficiency is m -1. Generally speaking, the rank deficiency is equal to the number of independent partial nets. Also we have to emphasise, that $\mathbf{C}$ does not contain any metric information about the net.

Branches \ Nodes	1	2	3	4	5	6
1	1	-1				
2		1		-1		
3				1	-1	
4			C		1	-1
5		1	-1			
6			1	-1		
7						-1

Fig. 3 Branch-node matrix of graph

Two further important matrices.

The symmetrical node-node matrix $\mathbf{C}^{\mathbf{T}} \cdot \mathbf{C}$ has some important and interesting properties:
- Elements in the main diagonal = number of branches tied to the node,
- Non-zero element -1 outside the main diagonal: $\Leftrightarrow$ The two nodes are connected by a branch,
- Nodes in the net in chain-like sequences can be derived

The also symmetrical branch-branch matrix $\mathbf{C} \cdot \mathbf{C}^{\mathbf{T}}$ has the properties
- All elements in the main diagonal are „2“ $\Leftrightarrow$ every branch has two nodes at its ends
- Elements outside the main diagonal are 0,+1,-1:
  - 0 , if the branches do no intersect each other
  - +1 , if two branches with the same direction are connected to the node m,
  - -1 , if two branches with opposite directions are connected to the node m

This very elementary matrix description of the topological structure of a net can serve as the point of departure to discriminate in a net between its

- *topological* structure(characterised by the number of nodes, branches and their connections with the nodes), and the
- *physical realisation* (i.e. embodiment in a physical or mathematical context) of a net, which could be a geodetic-, elasto-mechanical-, hydro-mechanical-, electrical-, transportation-net, etc.

In this context and thus preparing arbitrary physical realisations we assign to

a.) each node (which may be from one to three-dimensional) a node variable x_i („Something is or appertains to the node“), and arrange them in the vector $\mathbf{x}$

$$\mathbf{x}^{\mathbf{T}} = (x_1, x_2, ..., x_n),$$

b.) to each branch a branch variable t_k („Something happens between the nodes“), arranged in the vector $\mathbf{t}$

$$\mathbf{t}^{\mathbf{T}} = (t_1, t_2, ..., t_m)$$

In the case that the variables $\mathbf{x}$ and $\mathbf{t}$ have more than one dimension the vectors of the node variables may be written in the form of

$$\mathbf{x}^{\mathbf{T}} = (x_1, x_2, ..., x_n)$$
$$\mathbf{y}^{\mathbf{T}} = (y_1, y_2, ..., y_n) \qquad \Rightarrow \ \bar{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \end{pmatrix}$$
$$\mathbf{z}^{\mathbf{T}} = (z_1, z_2, ..., z_n)$$

The same expansion may also be necessary for the branch variables $\mathbf{t}$;consequently the branch node matrix $\mathbf{C}$ has to be expanded to.

$$\mathbf{C} = \begin{pmatrix} \mathbf{I_k} & -\mathbf{I_k} & 0 \\ 0 & \mathbf{I_k} & -\mathbf{I_k} \\ -\mathbf{I_k} & 0 & \mathbf{I_k} \end{pmatrix}$$

where $\mathbf{I_k}$ = identity matrix of k dimensions

2. Elementary relationships in nets and numerical analysis and solutions

After this preparation we may formulate four elementary and essential linear relationships between the variables $\mathbf{x}$ and $\mathbf{t}$ and the matrix $\mathbf{C}$. In the case if higher dimensions of the branch and node variables we replace $\mathbf{x} \Rightarrow \bar{\mathbf{x}}$, $\mathbf{u} \Rightarrow \bar{\mathbf{u}}$, $\mathbf{t} \Rightarrow \bar{\mathbf{t}}$, etc.

1.) Regarding the nodes connected by branches we generate the vector $\mathbf{u}$ of the *differences* of the node variables $\mathbf{x}$. This is achieved by the multiplication

$$\mathbf{u} := \mathbf{C} * \mathbf{x} \qquad <=> \qquad \mathbf{C} * \mathbf{x} = \mathbf{u} \tag{1}$$

and differentiation with respect to $\mathbf{u}$ gives the Jacobian

$$\left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right) = \mathbf{C} \tag{2}$$

2.) In each node the sum of the branch variables connected to that node is

$$\mathbf{r} = \mathbf{C}^{\mathbf{T}} * \mathbf{t} \tag{3}$$

where $\mathbf{r}$ is a vector whose elements are the sums in the individual nodes.

3.) For the branch-mesh matrix $\mathbf{M}$ we have the relationship

$$\mathbf{M}^{\mathbf{T}} * \mathbf{u} = \mathbf{0} \tag{4}$$

for all $\mathbf{x}$ and $\mathbf{u} := \mathbf{C} * \mathbf{x}$ (since $\mathbf{M}^{\mathbf{T}} * \mathbf{C} = \mathbf{0}$) (5)

4.) It is always possible - depending in the individual case on the *physical realisation* of the net - to establish a (non)-linear relation between the branch variables $\mathbf{t}$ and the node variable-differences $\mathbf{u}$ in the general form of

$$\mathbf{t} = \mathbf{f}(\mathbf{u}) \tag{6}$$

The above equations (1) - (6) are the basis to generalise the „solutions“ of net-analysis problems, irrespective of their physical realisation.

Combining (6) and (3) we get

$$\mathbf{C}^{\mathrm{T}} * \mathbf{f}(\mathbf{u}) = \mathbf{r} \tag{7}$$

where the elements of $\mathbf{u}$ are the differences of the node variables $\mathbf{x}$. Equations (7) are generally non-linear, can be interpreted as equations of „*equilibrium*“ in each node, and minimise certain quadratic forms inherent in the problem.

In table 2 a few examples of the linear relationships of equations (1) - (7) are shown.

In analysis-problems mostly the node variables are the unknowns. They have to be determined from the (non)-linear equations (7). Introducing for them starting values $\mathbf{x}_0$, and also starting values $\mathbf{u}_0$, compatible with the $\mathbf{x}_0$

$$\mathbf{x} = \mathbf{x}_0 + \Delta\mathbf{x}_1 \quad \mathbf{u} = \mathbf{u}_0 + \Delta\mathbf{u}_1 \tag{8}$$

we can apply Newton's method to solve (7) in a process of iteration. The $\mathbf{x}_0$ may be estimated, „pre-calculated“, or taken from a model.
Using the *linear* term of a Taylor expansion of (7) expressed by the appropriate Jacobian

$$\mathbf{C}^{\mathrm{T}} * \mathbf{f}(\mathbf{u}) - \mathbf{r} = \mathbf{0} \;<=>\; \mathbf{C}^{\mathrm{T}} * \mathbf{f}(\mathbf{u}_0) + \mathbf{C}^{\mathrm{T}} * \frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{x}} * \Delta\mathbf{x} - \mathbf{r} = \mathbf{0}$$

we find

$$\mathbf{C}^{\mathrm{T}} * \frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{x}} * \Delta\mathbf{x} = \mathbf{r} - \mathbf{C}^{\mathrm{T}} * \mathbf{f}(\mathbf{u}_0)\,. \tag{9}$$

However, using the chain rule of calculus we can express the Jacobian as a matrix product, and keeping in mind (2) we get

$$\frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{x}}\right) = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) * \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right) = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) * \mathbf{C}\,. \tag{10}$$

Substituting (10) into (9) yields the first step of the consequent iteration process

$$(\mathbf{C}^{\mathrm{T}} * \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) * \mathbf{C}) * \Delta\mathbf{x} = \mathbf{r} - \mathbf{C}^{\mathrm{T}} * \mathbf{f}(\mathbf{u}_0) \tag{11}$$

which easily may be generalised and can be considered the basic iteration equation for the „numerical solution“ of network-type entities.

Here, three essential remarks are necessary

1.) When applying the chain rule (10) to the Jacobian's it is possible to *reintroduce* the branch node matrix $\mathbf{C}$. Thus also (11) reflects the discrimination between the topology of the net - contained in $\mathbf{C}$ - and its physical realisation contained in the Jacobian $\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right)$.

2.) To use the generalised equations (11) for the numerical solution of an actual physical realisation of a net, the Jacobian $\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right)$ *appertaining to this very realisation* has to be found, which may constitute the main „difficulty" for the analysis of an individual realisation of a net. The examples following below show also, how just this difficulty can be overcome.

3.) The degree of „exactness" in the interpretation of the Jacobian $\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right)$, consequently, is decisive for characteristic properties of the individual solution, yielding what may be called a 1^{st} *order solution* and a 2^{nd} *order solution*

3. Examples of application

When applying the general theory to an actual physical realisation, we always start by *identifying appropriately* the node variables $\mathbf{x}$ and the branch variables $\mathbf{t}$ for the individual case to be investigated.

3.1 Electrical network

1st step: identification

Node variables	$\mathbf{x}$	$\equiv$	electrical potential in the nodes
Branch variables	$\mathbf{t}$	$\equiv$	branch currencies
Differences	$\mathbf{u}$	$\equiv$	$\mathbf{C}*\mathbf{x}$ $\equiv$ branch tensions.

2nd step: relationships

The needed relationship between node-variable-differences $\mathbf{u}$ and branch variables $\mathbf{t}$ is expressed by Ohm's law

$$\mathbf{t} = \mathbf{Q} * \mathbf{u} \tag{12}$$

thus the needed Jacobian is $\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = \mathbf{Q}$ and the above general equations of „equilibrium" (3) now reflect Kirchhoff's law.

$$\mathbf{C}^{\mathbf{T}} * \mathbf{t} = \mathbf{r} \tag{13}$$

By substituting and appropriately using (11) we immediately find the standard method of electrical network-analysis

$$(\mathbf{C}^{\mathbf{T}} * \mathbf{Q} * \mathbf{C}) * \mathbf{x} = \mathbf{r} - \mathbf{C}^{\mathbf{T}} * \mathbf{Q} * \mathbf{u} \tag{14}$$

3.2 Network of spirit levelling (without regarding gravity)

1st step: identification

Node variables	$\mathbf{x}$	$\equiv$	elevation of points
Branch variables	$\mathbf{t}$	$\equiv$	$\mathbf{P} \cdot \mathbf{v} = \mathbf{P} \cdot \mathbf{u} - \mathbf{P} \cdot \mathbf{l}$ weighted residuals
Differences	$\mathbf{u}$	$\equiv$	$\mathbf{l} + \mathbf{v}$ adjusted height differences.

2nd step: relationships

By substituting we find $\mathbf{u} = \mathbf{l} + \mathbf{v} \Leftrightarrow \mathbf{v} = \mathbf{u} - \mathbf{l}$, and thus

$$\mathbf{C}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{v} = \mathbf{0} \Rightarrow \mathbf{C}^{\mathbf{T}} \cdot \mathbf{P} \cdot (\mathbf{u} - \mathbf{l}) = \mathbf{0}$$

as „equations of equilibrium“. Since

$$\mathbf{u} := \mathbf{C} * \mathbf{x}$$

we get immediately

$$\left(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{C}\right) \cdot \mathbf{x} = \mathbf{C}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{l} \tag{15}$$

corresponding exactly to the normal equations of least squares adjustment.

Here already the original problem is linear. Had we introduced approximate values $\mathbf{x_0}$ we could have used - because the needed Jacobian in our case is now

$$\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = \mathbf{P} \tag{16}$$

directly the general equations (11) (since: $\mathbf{f}(\mathbf{u}_0) = \mathbf{t}_0 = \mathbf{P} \cdot (\mathbf{u}_0 - \mathbf{l})$) and the result would have been

$$\left(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{C}\right) \cdot \Delta\mathbf{x} = \mathbf{C}^{\mathbf{T}} \cdot \mathbf{P} \cdot (\mathbf{l} - \mathbf{u}_0) \tag{17}$$

with

$$\mathbf{u}_0 := \mathbf{C} * \mathbf{x}_0 .$$

3.3 Network of watersupply

1st step: identification

Node variables	$\mathbf{x}$	≡	pressure in nodes
Branch variables	$\mathbf{t}$	≡	flow in branches
Differences	$\mathbf{u}$	≡	differences in pressure

2nd step: relationships

(Non)-linear relationships between differences in pressure and flow

$$\mathbf{T} \cdot \mathbf{t} = \mathbf{2} \cdot \mathbf{R}^{-1} \cdot \mathbf{u}$$

where $\mathbf{R}$ = coefficients of friction, and $\mathbf{T} = diag(|\mathbf{t}|)$.

The law of equilibrium, stating that incoming flows must be equal to outgoing flows in each node is now

$$\mathbf{C}^{\mathbf{T}} \cdot \mathbf{t} = \mathbf{r}$$

and the necessary Jacobian comes out as

$$\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = \mathbf{T}^{-1} \cdot \mathbf{R}^{-1}.$$

By substituting we find immediately the equations to analyse the system

$$\mathbf{C}^{\mathrm{T}} \cdot \left[\mathbf{T}^{-1} \cdot \mathbf{R}^{-1}\right]_0 \cdot \mathbf{C} \cdot \Delta\mathbf{x} = \mathbf{r} - \mathbf{C}^{\mathrm{T}} \cdot \mathbf{t}_0 \qquad (18)$$

3.4 Network of transport

1st step: identification

Node variables	$\mathbf{x}$	traffic potentials in the nodes
Branch variables	$\mathbf{t}$	flow of traffic on branch; $\mathbf{t} := \mathbf{P} \cdot \mathbf{u}$
Differences	$\mathbf{u}$	differences in potentials: $\mathbf{u} = \mathbf{C} \cdot \mathbf{x}$

2nd step: relationships

In this - very elementary ! - treatment of transportation networks we assign to every branch a cost-factor p_i, accomodated consequently in the diagonal matrix $\mathbf{P}$. Moreover the general equations (3) $\mathbf{C}^{\mathrm{T}} \cdot \mathbf{t} = \mathbf{r}$ describe now the relationships of equilibrium pertinent to each node. Finally, we find directly the always needed Jacobian of the relationship $\mathbf{t} = \mathbf{f}(\mathbf{u}) \Rightarrow \mathbf{t} := \mathbf{P} \cdot \mathbf{u}$ to be here

$$\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = \left(\frac{\partial [\mathbf{P} \cdot \mathbf{u}]}{\partial \mathbf{t}}\right) = \mathbf{P}$$

and our general equations (11) turn out to be now

$$(\mathbf{C}^{\mathrm{T}} \cdot \mathbf{P} \cdot \mathbf{C}) \cdot \Delta\mathbf{x} = \mathbf{r} - \mathbf{C}^{\mathrm{T}} \cdot \mathbf{P} \cdot \mathbf{u}_0$$

They can be solved, if at least one node is assigned a given potential.

3.5 Network of trilateration (in 2 or 3 dimensions)

1st step: identification

Node variables	$\mathbf{x}$	2- or 3-dimensional coordinates of nodes
Branch variables	$\mathbf{t}$	$\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) \cdot \mathbf{P} \cdot (\mathbf{g}(\mathbf{u}) - \mathbf{l})$ (!), vide (19) ff. below,
Differences	$\mathbf{u}$	differences in coordinates: $\mathbf{u} = \mathbf{C} \cdot \mathbf{x}$

2nd step: relationships

To accommodate this problem in our general treatment we have to take some deviations, because the identification, especially of the branch variables $\mathbf{t}$, is not so obvious.
First, instead of formulating - as we are accustomed to in geodesy - the „observation equations“ in the conventional form

adjusted observations = function of the unknown coordinates

$$\mathbf{l} + \mathbf{v} = \mathbf{f}(\mathbf{x})$$

we formulate now

(adjusted observations) = function unknown of coordinate *differences*

$$\mathbf{l} + \mathbf{v} = \mathbf{g}(\mathbf{u}) \text{ with } \mathbf{u} = \mathbf{C} \cdot \mathbf{x} \text{ as above.}$$

The minimum principle of adjustment theory is $\mathbf{v}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{v} \Rightarrow \min$ and differentiation with respect to the unknowns $\mathbf{x}$ and observing the chain rule yields after some calculation

$$\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{v} = \mathbf{0}.$$

To transcribe this to the above general form we use again the chain rule

$$\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right) = \left(\frac{\partial \mathbf{v}}{\partial \mathbf{u}}\right) \cdot \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right) = \left(\frac{\partial \mathbf{v}}{\partial \mathbf{u}}\right) \cdot \mathbf{C} = \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\right) \cdot \mathbf{C}$$

and, after transposing, get

$$\mathbf{C}^{\mathbf{T}} \cdot \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\right)^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{v} = \mathbf{0}$$

$$\mathbf{v} = \mathbf{g}(\mathbf{u}) - \mathbf{l}$$

$$\mathbf{C}^{\mathbf{T}} \cdot \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\right)^{\mathbf{T}} \cdot \mathbf{P} \cdot (\mathbf{g}(\mathbf{u}) - \mathbf{l}) = \mathbf{0}. \tag{19}$$

Comparing now (19) with the general form (3) and keeping in mind that $\mathbf{r} = \mathbf{0}$ we can identify and define

$$\mathbf{t} := \left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\right)^{\mathbf{T}} \cdot \mathbf{P} \cdot (\mathbf{g}(\mathbf{u}) - \mathbf{l}) \tag{20}$$

as already stated above, wherein

$\mathbf{g}(\mathbf{u}) := \mathbf{w} = (w_1, w_2, ..., w_n)$	lengths between nodes calculated from coordinate differences ≡ adjusted distances
$\mathbf{P}$	= diagonal matrix of weights,
$\mathbf{l}^{\mathbf{T}} = (l_1, l_2, ..., l_m)$	= observed lengths between nodes,
$\left(\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\right)^{\mathbf{T}}$	= Jacobian, the elements of which are the partial derivatives of the coordinate differences with respect to the lengths, i.e. which are $\mathbf{u} / \mathbf{w} = \cos\alpha$ corresponding to the direction cosines of the (adjusted) distances in space.

For the subsequent matrix - notation we assign to the vectors $\mathbf{w}, \mathbf{u}, \mathbf{l}$, etc. diagonal matrices, $\mathbf{W}, \mathbf{U}, \mathbf{L}$, allowing us to write (20) in the form

$$\mathbf{t} := \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P} \cdot (\mathbf{w} - \mathbf{l}) \tag{20}$$

From (20) we can drive the (conventional) „linear“ as the (non-conventional) „non-linear“ approach for the solution.

(a) Conventional „linear approach“.

Needing, as always, the Jacobian $\left(\frac{\partial [\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}}\right) = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right)$ we consider in (20) the part $(\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P}) = const.$ as constant, meaning that the geometry of the net calculated from *approximate* coordinates and the geometry calculated form *adjusted* coordinates are *identical*, i.e. we consider the residuals $\mathbf{v}$ to be *small*. Then the Jacobian is

$$\left(\frac{\partial [\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}}\right) = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P}) \cdot \left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right) = (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P}) \cdot \mathbf{W}^{-1} \cdot \mathbf{U}$$

and the general equations (11) become

$$\underbrace{[\mathbf{C}^{\mathbf{T}} \cdot (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}) \cdot \mathbf{C}]}_{\mathbf{A}^{\mathbf{T}} \quad \mathbf{A}} \cdot \mathbf{x} = \underbrace{\mathbf{C}^{\mathbf{T}} \cdot \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P} \cdot (\mathbf{l} - \mathbf{w}_0)}_{\mathbf{A}^{\mathbf{T}}} \tag{21}$$

in which we immediately recognise the well known „normal equations“ from geodetic adjustment theory by identifying $\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} = \mathbf{A}^{\mathbf{T}} \Leftrightarrow \mathbf{W}^{-1} \cdot \mathbf{U} = \mathbf{A}$.

(b) Non-conventional „non-linear approach“.

Again starting from (20) we now let drop the above assumption of $(\mathbf{D}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P}) = const.$ and the $\mathbf{v}$ to be small. To begin, we convert (20) to the form

$$\mathbf{t} = \mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot (\mathbf{e} - \mathbf{l} / \mathbf{w}) \tag{20}$$

and, after introducing the newly defined variable

$$\mathbf{q} := \mathbf{P} \cdot (\mathbf{e} - \mathbf{l} / \mathbf{w}), \tag{22}$$

we get, still shorter

$$\mathbf{t} = \mathbf{U}^{\mathbf{T}} \cdot \mathbf{q} \tag{23}$$

Now, finally, we have found in (23) an appropriate representation of $\mathbf{t}$ to calculate the „complete“ Jacobian in such a form that we can compare the result immediately with the conventional linear approach **(a).**

In manipulating (23) we first have to observe the product rule

$$\frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}} = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = \left(\frac{\partial[\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Big|_{\mathbf{U}=const.} + \left(\frac{\partial[\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Big|_{\mathbf{q}=const.}$$

Tackling the first term we find (with $\mathbf{e}^{\mathbf{T}} = (1,1,\ldots\ldots\ldots,1)$)

$$\left(\frac{\partial[\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Big|_{\mathbf{U}=const.} = \mathbf{U}^{\mathbf{T}} \cdot \left(\frac{\partial \mathbf{q}}{\partial \mathbf{u}}\right) = \mathbf{U}^{\mathbf{T}} \cdot \frac{\partial[\mathbf{P} \cdot (\mathbf{e} - \mathbf{l}/\mathbf{w})]}{\partial \mathbf{w}} \cdot \left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right),$$

where $\frac{\partial[\mathbf{P} \cdot (\mathbf{e} - \mathbf{l}/\mathbf{w})]}{\partial \mathbf{w}} = \mathbf{P} \cdot \mathbf{L} \cdot \mathbf{W}^{-2}$ and $\left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right) = \mathbf{W}^{-1} \cdot \mathbf{U}$.

Thus we find the first term

$$1st\ term = \mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{L} \cdot \mathbf{W}^{-2} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}$$

For the second term we find easily - using again conversions between vectors and appertaining diagonal matrices –

$$2^{nd}\ term = \left(\frac{\partial[\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Big|_{\mathbf{q}=const.} = \mathbf{e} \cdot \mathbf{q}^{\mathbf{T}} \Rightarrow \mathbf{Q}$$

Therefore the complete Jacobian is

$$\frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}} = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = [\mathbf{Q} + (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{L} \cdot \mathbf{W}^{-3} \cdot \mathbf{U})] \tag{24}$$

and our general equations of solutions (11) now become

$$\left\{\mathbf{C}^{\mathbf{T}} \cdot [\mathbf{Q} + (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{L} \cdot \mathbf{W}^{-3} \cdot \mathbf{U})] \cdot \mathbf{C}\right\} \cdot \Delta\mathbf{x} = \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P} \cdot (\mathbf{l} - \mathbf{w}_{\mathbf{0}}) \tag{25}$$

Equations (25), whose matrix of coefficients consists of the two terms

(I) $(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{C})$

(II) $\left\{\mathbf{C}^{\mathbf{T}} \cdot (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{L} \cdot \mathbf{W}^{-3} \cdot \mathbf{U}) \cdot \mathbf{C}\right\}$

represent the normal equations of trilateration in the non-linear approach. Term(I) is the *main additional term* for the non-linear solution. Term (II) corresponds nearly exactly to the conventional linear approach in geodesy, because

$$(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{W}^{-2} \cdot \mathbf{U} \cdot \mathbf{C}) = (\mathbf{C}^{\mathbf{T}} \cdot \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{P} \cdot \mathbf{W}^{-1} \cdot \mathbf{U} \cdot \mathbf{C}) = (\mathbf{A}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{A})$$
$$\mathbf{A}^{\mathbf{T}} \quad \cdot \mathbf{P} \cdot \quad \mathbf{A}$$

in conventional notation. Thus

$$(\text{II}) = \left\{\mathbf{C}^{\mathbf{T}} \cdot (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{L} \cdot \mathbf{W}^{-3} \cdot \mathbf{U}) \cdot \mathbf{C}\right\} = (\mathbf{A}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{A}) \cdot \mathbf{L} \cdot \mathbf{W}^{-1}$$

Now, it is true, that $\mathbf{L} \cdot \mathbf{W}^{-1} \approx \mathbf{E}$ since the quotient of observed and adjusted lengths is appr. equal to 1 if the $\mathbf{v}$ are small. Insofar the non-linear solution contains essentially the additional Term (I), i.e the matrix $\mathbf{Q}$ has to be added to the normal equations.

3.6 Network of pin-jointed trusses in space: prestressed cable nets (in 2 or 3 dimensions)

1st step: identification

Node variables	$\mathbf{x}$	$\equiv$	coordinates of nodes
Branch variables	$\mathbf{t}$	$\equiv$	$\mathbf{k} \cdot (\mathbf{u} / \mathbf{w})$ components of forces, vide (7) below
Differences	$\mathbf{u}$	$\equiv$	differences in coordinates; $\mathbf{u} = \mathbf{C} \cdot \mathbf{x}$
Sums	$\mathbf{s}$	$\equiv$	sums of forces in each node; $\mathbf{s} = \mathbf{C}^{\mathbf{T}} \cdot \mathbf{t}$

2nd step: relationships

The lengths in space between nodes are

$$\mathbf{w}_i = \sqrt{\Delta\mathbf{x}_i^{\,2} + \Delta\mathbf{y}_i^{\,2} + \Delta\mathbf{z}_i^{\,2}}$$

and the equations of equilibrium in each node

$$\mathbf{C}^{\mathbf{T}} \cdot \mathbf{t} = \mathbf{s}\,. \tag{26}$$

after defining

$$\mathbf{k} = \text{ absolute forces in bars}$$

we can also define

$$\mathbf{t} := \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{k} \;\Rightarrow \mathbf{t} = (\textit{absolute forces}) \cdot \frac{\textit{coordinate differences}}{\textit{lengths in space}} = \textit{force components}$$

$$\mathbf{t} := \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{k} \tag{27}$$

Comparing (27) with (20) and introducing Hooke's coefficients $\mathbf{H}$ of elasticity we may also write the forces in the form of

$$\mathbf{k} = \mathbf{H} \cdot \mathbf{L}^{-1} \cdot (\mathbf{w} - \mathbf{l}) = \mathbf{H} \cdot (\mathbf{w} / \mathbf{l} - \mathbf{e}) \tag{28}$$

Taking now equations (26) and (27) as starting points we have, as in the example of the trilateration above, the alternatives of the linear and non-linear approach and additionally we can derive the method of „force densities"

(a) „Linear approach" = 1st-order theory

As above we consider in (27) the part $(\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1}) = const.$ i.e. only small deformations under loads occur. Then the seeked Jacobian is

$$\frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}} = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1}) \cdot \left(\frac{\partial \mathbf{k}}{\partial \mathbf{u}}\right) = (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1}) \cdot \left(\frac{\partial \mathbf{k}}{\partial \mathbf{w}}\right) \cdot \left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right)$$

and since

$$\left(\frac{\partial \mathbf{k}}{\partial \mathbf{w}}\right) = \mathbf{H} \cdot \mathbf{L}^{-1}$$

and

$$\left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right) = \mathbf{W}^{-1} \cdot \mathbf{U},$$

the Jacobian is now

$$\left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-2} \cdot \mathbf{H} \cdot \mathbf{L}^{-1} \cdot \mathbf{U}).$$

Therefore the equations of solution corresponding to (11) are

$$[\mathbf{C}^{\mathbf{T}} \cdot (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-2} \cdot \mathbf{H} \cdot \mathbf{L}^{-1} \cdot \mathbf{U}) \cdot \mathbf{C}] \cdot \Delta\mathbf{x} = \mathbf{s} - \mathbf{C}^{\mathbf{T}} \cdot \mathbf{U}^{\mathbf{T}} \cdot \mathbf{W}^{-1} \cdot \mathbf{k}_{\mathbf{0}} \qquad (29)$$

(b) „Non-linear approach“ = 2nd-order theory

Using the same procedure as above, defining $\mathbf{q} := \mathbf{W}^{-1} \cdot \mathbf{k}$, or

$$\mathbf{q} := \mathbf{W}^{-1} \cdot \mathbf{k} = \mathbf{W}^{-1} \cdot \mathbf{H} \cdot (\mathbf{w} / \mathbf{l} - \mathbf{e}) = \mathbf{H} \cdot (\mathbf{1} / \mathbf{l} - \mathbf{1} / \mathbf{w})$$

we have in

$$\mathbf{t} := \mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}$$

the same starting equations as in (23), however with a slight difference in physical meaning. Exactly as above we formulate

$$\frac{\partial[\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}} = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}}\right) = \cdot \left(\frac{\partial [\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Bigg|_{\mathbf{U}=const.} + \cdot \left(\frac{\partial [\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Bigg|_{\mathbf{q}=const.}$$

Turning to the first term we find now

$$\left(\frac{\partial [\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}}\right)\Bigg|_{\mathbf{U}=const.} = \mathbf{U}^{\mathbf{T}} \cdot \left(\frac{\partial \mathbf{q}}{\partial \mathbf{u}}\right) = \mathbf{U}^{\mathbf{T}} \cdot \frac{\partial [\mathbf{H} \cdot (\mathbf{1} / \mathbf{l} - \mathbf{1} / \mathbf{w})]}{\partial \mathbf{w}} \cdot \left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right)$$

where

$$\frac{\partial [\mathbf{H} \cdot (\mathbf{1} / \mathbf{l} - \mathbf{1} / \mathbf{w})]}{\partial \mathbf{w}} = \mathbf{H} \cdot \mathbf{W}^{-2}$$

and

$$\left(\frac{\partial \mathbf{w}}{\partial \mathbf{u}}\right) = \mathbf{W}^{-1} \cdot \mathbf{U}.$$

Thus we find

$$\text{first term} = (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{H} \cdot \mathbf{W}^{-1} \cdot \mathbf{W}^{-2} \cdot \mathbf{U}).$$

For the second term we find easily - using again conversion between vectors and diagonal matrices –

$$\left(\frac{\partial [\mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}]}{\partial \mathbf{u}} \right)\Bigg|_{\mathbf{q}=const.} = \mathbf{e} \cdot \mathbf{q}^{\mathbf{T}} \Rightarrow \mathbf{Q}$$

Therefore the Jacobian is

$$\frac{\partial [\mathbf{f}(\mathbf{u})]}{\partial \mathbf{u}} = \left(\frac{\partial \mathbf{t}}{\partial \mathbf{u}} \right) = [\mathbf{Q} + (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{H} \cdot \mathbf{W}^{-3} \cdot \mathbf{U})] \tag{30}$$

and the general solution

$$\left\{ \mathbf{C}^{\mathbf{T}} \cdot [\mathbf{Q} + (\mathbf{U}^{\mathbf{T}} \cdot \mathbf{H} \cdot \mathbf{W}^{-3} \cdot \mathbf{U})] \cdot \mathbf{C} \right\} \cdot \Delta \mathbf{x} = \mathbf{s} - \mathbf{C}^{\mathbf{T}} \cdot \mathbf{t}_{\mathbf{0}} \tag{31}$$

represents the first step in a Newton iteration procedure. This solution can also accomodate large deformations of the net (and consequently large elastic elongations of the bars between the nodes) and thus be used for prestressed cable-net analysis and also for the analysis of discretised membranes.

Also the intrinsic relationship between the method of least squares and the elastomechanical analysis of pin-jointed trusses becomes evident very clearly when comparing and interpreting equations (25) and (31).

(c) Method of „Force-densities“

Starting directly from (26) and using our definition $\mathbf{t} := \mathbf{U}^{\mathbf{T}} \cdot \mathbf{q}$ we can create from the vector $\mathbf{q}$ a hyperdiagonal matrix $\mathbf{Q}$ and from $\mathbf{U}^{\mathbf{T}}$ the vector $\mathbf{u}$,arriving at the form

$$\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{u} = \mathbf{s}$$

which we can evaluate further by observing (1) $\mathbf{u} = \mathbf{C} \cdot \mathbf{x}$ to

$$(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{C}) \cdot \mathbf{x} = \mathbf{s} \tag{32}$$

Equations (32) represent a *linear system* of equations permitting - after the force densities contained in the elements of $\mathbf{Q}$ have been estimated - the direct solution for the unknowns $\mathbf{x}$ (= figure of equilibrium) in *one direct numerical* step by solving the linear system (32)! Naturally the vector $\mathbf{x}$ has to be composed of the *unknown, variable* coordinates *plus given, fixed* coordinates, resulting in a *regular system* of equations (32).

4. Final Remarks

Summarising, we may state

1. Using the generalised approach and identifying adequately the node and branch variables appertaining to the specific realisation of the net the consequent specific method of net-analysis reveals itself, which is always embedded in and consistent with the general solution and its mathematical structure.
2. Regarding the numerical computation, always a system of equations with a matrix of coefficients of the type $(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Z} \cdot \mathbf{C})$ or $(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{C})$ has to be solved in which the place and number of non-zero elements do not change if $\mathbf{Z}$ or $\mathbf{Q}$ are changed. However, the numbering of the nodes may be changed, so that computing time - using for example. sparse matrix-techniques, can be influenced and also minimised. Always a symbolic factorisation is possible by investigating only the matrices

   $(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{C})$

   $(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{C})$ is symmetrical if $\mathbf{Q}$ is symmetrical,

   $(\mathbf{C}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{C})$ is symmetrical if $\mathbf{Q}$ is positive definite.

3. A relative small number of basic equations - applied and interpreted individually and appropriately - constitute the „inner“, „natural“ laws of net-like entities.

5. Literature

[1] Linkwitz, K.: Fehlertheorie und Ausgleichung von Streckennetzen nach der Theorie elastischer Systeme.(Theory of Errors and Least Squares Analysis of Trilateration Networks by Using the Theory of Elastomechanics“); Doctor Thesis, Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften, Reihe C, Heft 46, München 1961.

[2] Linkwitz, K. und Schek, H.J.: Einige Bemerkungen zur Berechnung von vorgespannten Seilnetzkonstruktionen. („Remarks concerning the Analysis of Prestressed Cable- Structures“) Ingenieur - Archiv 1971, S.145-158.

[3] Linkwitz, K.: New Methods for the Determination of Cutting Pattern of Prestressed Cable Nets and their Application to the Olympic Roofs Munich; Proc. 1971 IASS Pacific Symposium, TOKYO and KYOTO , Architectural Institute of JAPAN

[4] Linkwitz, K., Schek, H.J.: A New Method of Analysis of Prestressed Cable Networks and its Use on the Roofs of the Olympic Game Facilities at Munich; IABSE, Amsterdam 1972

[5] Linkwitz, K.: Die Ermittlung des Zuschnitts für Dächer der Olympiasportstätten München. („The Determination of the Cutting Patterns of the Olympic Roofs Munich 1972“)ZfV 1972, Heft 9, 10.

[6] Schek, H.J.: The force densities method for form finding and computation of general networks. Computer Methods in Applied Mechanics and Engineering, 1974, S. 115-134.

[7] Gründig, L.: Die Berechnung vorgespannter Seil- und Hängenetzen unter Berücksichtigung ihrer topologischen und physikalischen Eigenschaften und der Ausgleichungsrechnung. („The Analysis of prestressed Cable- and Hanging-Nets with regard to their topological and physical properties and to Least Squares Theory“) Doctor- Thesis; Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften, Reihe C, Heft 246, München 1976.

[8] Linkwitz, K.: Die Berechnung vorgespannter Seilnetze als Aufgabe der Ausgleichungsrechnung.(„The Analysis of Prestressed Cable-Nets as an Application of Least Squares Theory“ Internationaler Kurs für Ingenieurvermessungen hoher Präzision. Darmstadt 1976.

[9] Schek, H.J.: Über Ansätze und numerische Methoden zur Berechnung großer netzartiger Strukturen. („About Theories and Numerical Methods for the Analysis of large, netlike Entities“) Thesis of Habilitation, Universität Stuttgart 1977.

[10] Linkwitz, K.: Über eine neue Methode der Gauß'schen Methode der kleinsten Quadrate. Die Formfindung und statische Analyse von räumlichen Seil- und Hängenetzen. („About a New Method within the Theory of Least Squares: Formfinding and Analysis of Spatial Prestressed and Hanging Nets") Braunschweigische Wissenschaftliche Gesellschaft, Sonderveröffentlichung zum 200. Geburtstag C.F. Gauß, Braunschweig 1977.

[11] Linkwitz, K.: Zur Systematik der Analyse von Netzen. („On a Systematic Representation of the Analysis of Nets") Special Lecture Intern. Summer School Ettore Majorana, Trapani/Sizilien 1978.

[12] Bahndorf, J.: Zur Systematisierung der Seilnetzberecnung und zur Optimierung von Seilnetzes. („About a Systematic Approach to the Analysis and Optimisation of Cable Nets") Doctor Thesis; Deutsche Geodätische Kommission, Reihe C, Heft 373, München 1991.

[13] Neureither, M.: Modellierung geometrisch-topologischer Daten zur Beschreibung und Berechnung netzartiger und flächenhafter Strukturen.(„Modeling of Geometrical-Topological Data for the Analysis of Netlike Entities") Doctor Thesis Deutsche Geodätische Kommission, Reihe C, Heft 387, München 1992

[14] Linkwitz, K.: Formfinding of lightweight surface structures by geodetic methods. Proceedings of the International Symposium on the application of Geodesy to Engineering, Stuttgart 1991, Springer Verlag, Heidelberg, New York 1992.

[15] Linkwitz, K., Bahndorf, J.: Separation of Topological and Physical Properties in the Non-linear Analysis of Surface Lightweight Structures, International Conference on Computational Methods in Structural and Geotechnical Engineering, Hong Kong 1994

[16] Singer,S.: Die Berechnung Minimalflächen, Seifenblasen, Membranen und Pneus aus geodätischer Sicht („Analysis and Computation of soap-films, membranes, and pneus from the view point of the geodetic sciences") Doctor Thesis, Deutsche Geodätische Kommission Reihe C, Heft 448, Muenchen 1995

[17] Linkwitz, K., Ströbel.D.,Singer P.: Die Analytische Formfindung („Analytical Formfinding"), in Prozess und Form Natuerlicher Konstruktionen, Ernst & Sohn, Berlin 1996, ISBN 3-433-02883-4

[18] Ströbel, D.: Die Anwendung der Ausgleichungsrechnung auf elastomechanische Systeme („Application of the Theory of Least Squares on the Analysis of Elastomechanical Systems") Doctor Thesis, Bayerische Akademie der Wissenschaften, Reihe C, Heft 478, München 1997

[19] Röder, R.: Zur Analyse und Optimierung von Transportnetzen („On the Analysis of Transportation Networks") Doctor Thesis, Bayerische Akademie der Wissenschaften, Reihe C, Heft 484, München 1997

[20] Linkwitz, K. : Students' Training in Formfinding and Analysis of Tension Structures, Newsletter December 1997 Structural Morphology Group, Working Group 15 of IASS, Delft University Printing Office 1997

[21] Linkwitz, K.: About Formfinding of Double-Curved Structures; Engineering Structures, Volume 21, Number 8, August 1999; Elsevier Science Ltd. England

[22] Linkwitz, K.: Formfinding by the 'Direct Approach' and Pertinent Strategies for the Conceptual Design of Prestressed and Hanging Structures; Volume 14, 1999; International Journal of Space Structures, Space Structures Research Centre, Surrey, Guildford, United Kingdom

Intrinsic Parameters and Satellite Orbital Elements

Evangelos Livieratos

ABSTRACT

The relations between the curvature and torsion of the satellite orbit with the orbital elements and the equipotential surface counterparts are revisited, using some angular quantities which define the geometry of the orbit and its relation to the equipotential surface and the line of force, i.e. the *slope* of the orbit (ζ), the *zenith distance* of the orbit (Z), the *separation* (θ) of the orbital plane from the equipotential surface and the *separation* (β) of the orbital plane from the Frenet osculating plane.

1. INTRODUCTION

The motion of a satellite along its orbit, a curve S in space, is governed by the well known differential equation of celestial mechanics $\ddot{\vec{x}} = \vec{g}(\vec{x})$, where $\ddot{\vec{x}}$ is the acceleration vector of the satellite and $\vec{g}$ the gravity vector, at the position $\vec{x}$, ignoring additional small forces due to drag, solar pressure and luni-solar attraction. Differences in acceleration at two indeed neighbouring points on S, e.g. $\vec{x}$ and $\vec{x} + \mathrm{d}\vec{x}$/dS, establish the linear relation of differential changes of acceleration with relevant changes of satellite position, namely

$$\frac{\mathrm{d}\ddot{\vec{x}}}{\mathrm{dS}} = \frac{\mathrm{d}\vec{g}(\vec{x})}{\mathrm{dS}} = w(\vec{x})\ \frac{\mathrm{d}\vec{x}}{\mathrm{dS}} \tag{1.1}$$

where $w(\vec{x})$ is a linear operator (homography) synthesising all the mechanical properties of the gravity field. In terms of matrix notation, $w(\vec{x})$ is represented by the gravity gradient tensor (or the Bruns tensor), **W**, in the equivalent linear transformation,

$$\frac{\mathrm{d}\ddot{\mathbf{x}}}{\mathrm{dS}} = \frac{\mathrm{d}\mathbf{g}}{\mathrm{dS}} = \mathbf{W}\ \frac{\mathrm{d}\mathbf{x}}{\mathrm{dS}} \tag{1.2}$$

where $\ddot{\mathbf{x}}$ is the acceleration components, **g** the geocentric components of the gravity vector and **x** the geocentric co-ordinates of the satellite which refer to the geocentric reference frame $\mathbf{e_X}$ represented by a triad of mutually orthogonal unit vectors ($\mathbf{e_X}$**:** $\vec{e}_x$, $\vec{e}_y$, $\vec{e}_z$), where $\vec{e}_x$ is directed to the vernal equinox and $\vec{e}_z$ to the pole. Gravity gradient tensor **W** is a basic topic of study in satellite gradiometry *(Rummel 1986)*. It describes fully (see, e.g., *Marussi 1985*) the intrinsic geometry of the gravity field *(Grafarend 1974-)* since it contains the curvatures and torsions of the equipotential surface as well as the curvatures of the line of force, at the satellite point. On the other hand, the differential changes of satellite acceleration, $\mathrm{d}\ddot{\mathbf{x}}$/dS, can be expressed in terms of curvature and torsion of the satellite orbit. The same holds for the differential change of position d**x**/dS, since it can be shown the relation between the variation of relevant Kepler elements with the curvature and torsion of the satellite orbit. This interrelation between the intrinsic properties of the gravity field with those of the satellite orbit has not been studied extensively in the geodetic literature. Some indeed isolated examples can only be mentioned treating the satellite orbit in terms of its intrinsic properties *(Hotine 1969)* and in relation with the intrinsic properties of the gravity field *(Marussi 1962)*. In this paper the relations between the curvature and torsion of the satellite orbit and the equipotential surface counterparts are revisited,

using some angular quantities which define the geometry of the orbit and its relation with the equipotential surface and the line of force, i.e. the slope of the orbit (ζ), the zenith distance of the orbit (Z), the separation (θ) of the orbital plane from the equipotential surface and the separation (β) of the orbital plane from the Frenet osculating plane.

2. FRAMES AND TRANSFORMATIONS

Traditionally, the geometry of the satellite orbit S is respectively associated with the geocentric and the perigee-related reference frames $\mathbf{e}_X$ and $\mathbf{e}_p$, the second represented here by the triad of mutually orthogonal unit vectors ($\mathbf{e}_p$: $\vec{e}_p$, $\vec{e}$, $\vec{e}_n$), where $\vec{e}_p$ is directed to the perigee and $\vec{e}_n$ is normal to the orbital plane. The rotational transformation of these frames are given by

$$\mathbf{e}_X = \mathbf{R}_3(-\Omega)\, \mathbf{R}_1(-i)\, \mathbf{R}_3(-\omega)\, \mathbf{e}_p \tag{2.1}$$

where ω the argument of perigee, i the inclination of the orbit and Ω the longitude of the ascending node, three quantities which define the space orientation of the orbit, with respect to the geocentric frame $\mathbf{e}_X$. One more triad of mutually orthogonal unit vectors, is also used, as a reference frame in satellite geodesy, namely the moving orbital triad ($\mathbf{e}$: $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$), where $\vec{e}_1$ is collinear with the radial vector from the geo-centre to the satellite, $\vec{e}_2$ is directed along the satellite orbit and $\vec{e}_3 (= \vec{e}_n)$ is normal to the orbital plane. This moving frame is related with $\mathbf{e}_p$ via true anomaly f

$$f = \cos^{-1} (\vec{e}_p \cdot \vec{e}_1) , \tag{2.2}$$

by the transformation

$$\mathbf{e} = \mathbf{R}_3(f)\, \mathbf{e}_p \ . \tag{2.3}$$

True anomaly f and the radial distance r, of the satellite from the geo-centre, define as polar co-ordinates, the position of the satellite with respect to the perigee-related frame $\mathbf{e}_p$. Considering the unit tangent vector of the orbit $\vec{t}$, we can define the "slope" of the orbit ζ, as the angle from the radial vector $\vec{e}_1$

$$\zeta = \cos^{-1} (\vec{e}_1 \cdot \vec{t}) . \tag{2.4}$$

If β is the small angle separating the orbital plane from the osculating plane of the orbit, in terms of the Frenet triad, ($\mathbf{e}_F$: $\vec{t}$, $\vec{n}$, $\vec{b}$), where $\vec{t}$ the unit tangent vector, $\vec{n}$ the unit normal and $\vec{b}$ the unit binormal, the relation between the $\mathbf{e}$ triad and the Frenet triad is given by the transformation

$$\mathbf{e}_F = \mathbf{R}_1(\beta)\, \mathbf{R}_3(\zeta)\, \mathbf{e} \tag{2.5}$$

which combined with (2.1) and (2.3) gives the relation between the Frenet and the geocentric triads

$$\boxed{\mathbf{e}_F = \mathbf{R}_1(\beta)\, \mathbf{R}_3(q)\, \mathbf{R}_1(i)\, \mathbf{R}_3(\Omega)\, \mathbf{e}_X} \tag{2.6}$$

where q,

$$q = \omega + f + \zeta \tag{2.7}$$

is the orientation of the orbit-tangent with respect to the equatorial plane. At each satellite point, the relevant equipotential surface (W=const.) intersects the orbital plane by an angle θ. The intersection of

the equipotential surface with the orbital plane defines the satellite trajectory on the equipotential surface, associated with the surface unit tangent vector $\vec{t}^*$. The angle ε, between the tangent to the orbit and its counterpart on the equipotential surface is, thus

$$\varepsilon = \cos^{-1}\ (\vec{t}\cdot\vec{t}^*) \tag{2.8}$$

where ε, when added to q, gives the orientation of the tangent of the satellite trajectory on the equipotential surface, with respect to the equatorial plane. The unit vector $\vec{N}$, normal to the equipotential surface along the vertical at the satellite point, is obviously orthogonal with $\vec{t}^*$, both belonging to a "natural" triad of mutually orthogonal unit vectors ($\mathbf{e}_N$: $\vec{t}^*$, $\vec{T}$, $\vec{N}$), where $\vec{t}^*$ defines the direction of the orbit on the equipotential surface, $\vec{T}$ the perpendicular direction, both vectors $\vec{t}^*$ and $\vec{T}$, on the horizontal plane, and $\vec{N}$ the opposite direction of the gravity vector $\vec{g}$,

$$\vec{N} = -\frac{1}{g}\vec{g} \tag{2.9}$$

where g the intensity of gravity at the satellite orbit. The zenith distance Z, of the orbit, is defined by

$$Z = \cos^{-1}\ (\vec{N}\cdot\vec{t}\,) \tag{2.10}$$

and due to (2.8) and (2.9), it is

$$Z = 90° - \varepsilon \tag{2.11}$$

$$\vec{g}\cdot\vec{t} = -\,g\cos Z = -\,g\sin\varepsilon \tag{2.12}$$

The definitions of the Frenet and the natural frames give their rotational transformation, as function of the angles β, ε, θ,

$$\boxed{\mathbf{e}_N = \mathbf{R}_1(\theta)\,\mathbf{R}_3(\varepsilon)\,\mathbf{R}_1(-\beta)\,\mathbf{e}_F} \tag{2.13}$$

from which, combining with (2.6), we obtain

$$\boxed{\mathbf{e}_N = \mathbf{R}_1(\theta)\,\mathbf{R}_3(\varepsilon+q)\,\mathbf{R}_1(i)\,\mathbf{R}_3(\Omega)\,\mathbf{e}_X} \quad . \tag{2.14}$$

For the model spherical field and for a polar circular orbit (eccentricity zero, inclination i = 90°), the above angular quantities ζ, θ, β, ε reduce to

$$\zeta = \theta = 90° \tag{2.15}$$
$$\beta = \varepsilon = 0° \tag{2.16}$$

and consequently, the frames $\mathbf{e}_F$ and $\mathbf{e}_N$ coincide with $\mathbf{e}$,

$$\begin{aligned} \vec{e}_1 &= -\vec{n} = \vec{N} \\ \vec{e}_2 &= \vec{t} = \vec{t}^* \\ \vec{e}_3 &= \vec{b} = \vec{T} \end{aligned} \tag{2.17}$$

In such approximation, $\vec{e}_1$ is in the vertical direction, $\vec{e}_2$ is the tangent and $\vec{e}_3$ normal to the orbital plane. The slope ζ is thus, the zenith distance Z of the orbit. It is clear that the slope of the orbit ζ, the

angular separations β and θ of the orbital plane from the Frenet osculating plane and from the equipotential surface respectively as well as the angular separation ε of the orbit-tangent from its counterpart on the equipotential surface, reflect the contribution of the anomalous field of forces which affect the satellite orbit.

3. ACCELERATION

The differential change of $\mathbf{e}_F$ along the orbit S, is given by the Frenet-Serret relation

$$\frac{d\mathbf{e}_F}{dS} = \mathbf{F}\mathbf{e}_F \tag{3.1}$$

$$\mathbf{F} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \tag{3.2}$$

where κ, τ are respectively the curvature and the torsion of the orbit. The time derivative of $\vec{x}$ is the velocity vector $\dot{\vec{x}}$ along the unit tangent vector $\vec{t}$, of the orbit

$$\dot{\vec{x}} = v\ \vec{t}\ . \tag{3.3}$$

The acceleration vector is the time derivative of (3.3)

$$\ddot{\vec{x}} = \dot{v}\,\vec{t} + v\ \dot{\vec{t}} \tag{3.4}$$

and since $\vec{t} = v\ \frac{d\vec{t}}{dS}$, equation (3.4) with the help of (3.1), (3.2) is written

$$\ddot{\vec{x}} = \dot{v}\,\vec{t} + v^2\ \kappa\ \vec{n} \tag{3.5}$$

where $\dot{v}$ is the tangential acceleration and $v^2\kappa$ the normal, or centripetal, acceleration. Differentiating (3.5), along the orbit S, we obtain, with the help of (3.1), (3.2),

$$\frac{d\ddot{\vec{x}}}{dS} = \frac{d\dot{v}}{dS}\vec{t} + (\ 3\ \dot{v}\,\kappa + v^2\ \frac{d\kappa}{dS}\ )\ \vec{n} - v^2\ \kappa\,\tau\ \vec{b} \tag{3.6}$$

which in matrix form is written

$$\frac{d\ddot{\vec{x}}}{dS} = \frac{d\ddot{\mathbf{x}}_F^T}{dS}\,\mathbf{e}_F \tag{3.7}$$

where

$$\frac{d\ddot{\mathbf{x}}_F}{dS} = \begin{bmatrix} 0 \\ 3\dot{v}\kappa \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & v^2 & 0 \\ 0 & 0 & -v^2\kappa \end{bmatrix} \begin{bmatrix} \frac{d\dot{v}}{dS} \\ \frac{d\kappa}{dS} \\ \tau \end{bmatrix} \tag{3.8}$$

Combining with (1.2) and (2.6) we obtain

$$\frac{d\ddot{\mathbf{x}}}{dS} = \mathbf{R}_3(-\Omega)\,\mathbf{R}_1(-i)\,\mathbf{R}_3(-q)\,\mathbf{R}_1(-\beta)\ \frac{d\ddot{\mathbf{x}}_F}{dS} \tag{3.9}$$

from which, with the approximations i=90°, β=0°, ζ=90°, it is

$$\begin{bmatrix} \frac{d\dot{x}}{dS} \\ \frac{d\dot{y}}{dS} \\ \frac{d\dot{z}}{dS} \end{bmatrix} = 3\,\dot{v}\,\kappa \begin{bmatrix} \cos\Omega\cos(\omega+f) \\ \sin\Omega\cos(\omega+f) \\ \sin(\omega+f) \end{bmatrix} + \tag{3.10}$$

$$+ \begin{bmatrix} \cos\Omega\sin(\omega+f) & v^2\cos\Omega\cos(\omega+f) & -v^2\kappa\sin(\omega+f) \\ \sin\Omega\sin(\omega+f) & v^2\sin\Omega\cos(\omega+f) & v^2\kappa\cos(\omega+f) \\ -\cos(\omega+f) & v^2\sin(\omega+f) & 0 \end{bmatrix} \begin{bmatrix} \frac{d\dot{v}}{dS} \\ \frac{d\kappa}{dS} \\ \tau \end{bmatrix}$$

4. INTRINSIC PROPERTIES OF THE ORBIT

Differentiating (2.6) and due to (2.1), recalling the relevant transformations, we obtain

$$\begin{aligned} \mathbf{F} = {} & \mathbf{P}_1 \frac{d\beta}{dS} \\ & + \mathbf{R}_1(\beta)\mathbf{P}_3\mathbf{R}_1(-\beta) \quad \frac{dq}{dS} \\ & + \mathbf{R}_1(\beta)\mathbf{R}_3(q)\mathbf{P}_1\mathbf{R}_3(-q)\mathbf{R}_1(-\beta) \quad \frac{di}{dS} \\ & + \mathbf{R}_1(\beta)\mathbf{R}_3(q)\mathbf{R}_1(i)\mathbf{P}_3\mathbf{R}_1(-i)\mathbf{R}_3(-q)\mathbf{R}_1(-\beta) \quad \frac{d\Omega}{dS} \end{aligned} \tag{4.1}$$

where

$$\mathbf{P}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix};\ \mathbf{P}_3 = \begin{bmatrix} 0 & 1,0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & \end{bmatrix}. \tag{4.2}$$

Equation (4.1) gives the curvature and torsion of the orbit, in terms of differential change of parameters ω, i, Ω, f, ζ, β along the orbit,

$$\begin{bmatrix} \kappa \\ \tau - \frac{d\beta}{dS} \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\beta\sin q & \cos\beta\cos i - \sin\beta\cos q\sin i \\ 0 & \cos q & \sin q\sin i \\ \sin\beta & -\cos\beta\sin q & \sin\beta\cos i + \cos\beta\cos q\sin i \end{bmatrix} \begin{bmatrix} \frac{dq}{dS} \\ \frac{di}{dS} \\ \frac{d\Omega}{dS} \end{bmatrix} \tag{4.3}$$

and the inverse relations, for $i \neq 0°$, are

$$\begin{bmatrix} \frac{dq}{dS} \\ \frac{di}{dS} \\ \frac{d\Omega}{dS} \end{bmatrix} = \begin{bmatrix} \cos\beta + \sin\beta\cos q\cot i & -\sin q\cot i & \sin\beta - \cos\beta\cos q\cot i \\ \sin\beta\sin q & \cos q & -\cos\beta\sin q \\ -\sin\beta\frac{\cos q}{\sin i} & \frac{\sin q}{\sin i} & \cos\beta\frac{\cos q}{\sin i} \end{bmatrix} \begin{bmatrix} \kappa \\ \tau - \frac{d\beta}{dS} \\ 0 \end{bmatrix} \tag{4.4}$$

5. INTRINSIC PROPERTIES OF THE GRAVITY FIELD

The differential change of $\mathbf{e}_N$ along the orbit-trace S* on the equipotential surface, is given by

$$\frac{d\mathbf{e}_N}{dS^*} = \mathbf{K}^* \, \mathbf{e}_N \tag{5.1}$$

with

$$\mathbf{K}^* = \begin{bmatrix} 0 & \kappa_g & \kappa_n \\ -\kappa_g & 0 & \tau_g \\ -\kappa_n & -\tau_g & 0 \end{bmatrix} \tag{5.2}$$

where κ_n, κ_g the normal and the geodetic curvatures respectively and τ_g the geodetic torsion of the equipotential surface, namely the second derivatives of the geopotential W,

$$\begin{aligned} \kappa_n &= \frac{1}{g}\frac{\partial^2 W}{\partial \tau^{*2}} = \frac{1}{g} W_{t^*t^*} \\ \kappa_g &= \frac{1}{g}\frac{\partial^2 W}{\partial T^2} = \frac{1}{g} W_{TT} \\ \tau_g &= \frac{1}{g}\frac{\partial^2 W}{\partial \tau^* \partial T} = \frac{1}{g} W_{t^*T} \end{aligned} \tag{5.3}$$

The curvatures and torsion of the equipotential surface, in (5.3), the components of the curvature of the line of force χ, tangent to the vertical direction $\vec{N}$, given by

$$\begin{aligned} \chi_{t^*} &= \frac{1}{g}\frac{\partial^2 W}{\partial \tau^* \partial T} = \frac{1}{g}\frac{\partial g}{\partial \tau^*} = \frac{1}{g} W_{t^*N}, \\ \chi_T &= \frac{1}{g}\frac{\partial^2 W}{\partial T \partial N} = \frac{1}{g}\frac{\partial g}{\partial T} = \frac{1}{g} W_{TN} \end{aligned} \tag{5.4}$$

and the gradient along the vertical

$$\frac{\partial g}{\partial N} = \frac{\partial^2 W}{\partial N^2} = W_{NN} \ . \tag{5.5}$$

form the gravity gradient tensor $\mathbf{W}_N$ with respect to the to the $\mathbf{e}_N$ triad,

$$\mathbf{W}_N = g \begin{bmatrix} \kappa_n & \tau_g & \chi_{t*} \\ \tau_g & \kappa_g & \chi_T \\ \chi_{t*} & \chi_T & \frac{1}{g}\frac{\partial g}{\partial N} \end{bmatrix} = \begin{bmatrix} W_{t*t*} & W_{t*T} & W_{t*N} \\ W_{t*T} & W_{TT} & W_{TN} \\ W_{t*N} & W_{TN} & W_{NN} \end{bmatrix} \tag{5.6}$$

with the condition

$$\kappa_n + \kappa_g = \frac{1}{g}(2\omega^2 - \frac{\partial g}{\partial N}) \tag{5.7}$$

where ω the Earth rotation. Equation (5.1) is written as

$$\frac{d\mathbf{e}_N}{dS}\frac{dS}{dS^*} = \mathbf{K}^* \, \mathbf{e}_N \tag{5.8}$$

which, with the help of (2.8) and (2.11), it is

$$\frac{d\mathbf{e}_N}{dS} = \sin Z \, \mathbf{K}^* \, \mathbf{e}_N = \mathbf{K} \, \mathbf{e}_N \tag{5.9}$$

where obviously

$$\mathbf{K} = \sin Z \, \mathbf{K}^* \, . \tag{5.10}$$

6. ORBIT AND GRAVITY FIELD CURVATURES AND TORSIONS

Differentiating (2.6) and due to (3.1), (5.1), with the relevant transformations, we obtain the relation between **F** and **K** matrices of orbit and equipotential surface curvatures and torsions.

$$\begin{aligned} \mathbf{F} = {} & \mathbf{R}_1(\beta)\mathbf{R}_3(-\varepsilon)\mathbf{R}_1(-\theta)\mathbf{K}\mathbf{R}_1(\theta)\mathbf{R}_3(\varepsilon)\mathbf{R}_1(-\beta) \\ & - \mathbf{R}_1(\beta)\mathbf{R}_3(-\varepsilon)\mathbf{P}_1\mathbf{R}_3(\varepsilon)\mathbf{R}_1(-\beta) \quad \frac{d\theta}{dS} \\ & - \mathbf{R}_1(\beta)\mathbf{P}_3\mathbf{R}_1(-\beta) \quad \frac{d\varepsilon}{dS} \\ & + \mathbf{P}_1 \quad \frac{d\beta}{dS} \end{aligned} \tag{6.1}$$

from which we obtain

$$\begin{bmatrix} \kappa \\ \tau - \frac{d\beta}{dS} \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\beta\sin Z & \sin\beta\sin^2 Z & -\sin\beta\cos Z \\ 0 & \sin Z\cos Z & \sin Z \\ \sin\beta\sin Z & -\cos\beta\sin^2 Z & \cos\beta\cos Z \end{bmatrix} \begin{bmatrix} \kappa_g\cos\theta - \kappa_n\sin\theta \\ \kappa_g\sin\theta + \kappa_n\cos\theta \\ \tau_g\sin Z - \frac{d\theta}{dS} \end{bmatrix} + \begin{bmatrix} \cos\beta \\ 0 \\ \sin\beta \end{bmatrix} \frac{dZ}{dS} \tag{6.2}$$

and due to (4.3), for $i \neq 0°$, we obtain

$$\begin{bmatrix} \frac{dq}{dS} \\ \frac{di}{dS} \\ \frac{d\Omega}{dS} \end{bmatrix} = \begin{bmatrix} \sin Z & \sin(Z-q)\cot i \sin Z & -\cos(Z-q)\cot i \\ 0 & \cos(Z-q)\sin Z & \sin(Z-q) \\ 0 & \frac{-\sin(Z-q)}{\sin i}\sin Z & \frac{\cos(Z-q)}{\sin i} \end{bmatrix} \begin{bmatrix} \kappa_g \cos\theta - \kappa_n \sin\theta \\ \kappa_g \sin\theta + \kappa_n \cos\theta \\ \tau_g \sin Z - \frac{d\theta}{dS} \end{bmatrix} + \begin{bmatrix} \frac{dZ}{dS} \\ 0 \\ 0 \end{bmatrix} \qquad (6.3)$$

References

Grafarend, E. (1974 -): *All Erik Grafarend's work on differential geodesy.*

Hotine, M. (1969): *Mathematical geodesy.* ESSA Monograph 2, US Dept.of Commerce. Washington DC.

Marussi, A. (1985): *Intrinsic geodesy.* Springer-Verlag.

Marussi, A. (1962): Intrinsic geodesy and earth satellites. II Symp. on *Geodesy in three dimensions.* Cortina d' Ampezzo, May 1962.

Rummel, R. (1986): Satellite gradiometry. In: *Mathematical and numerical methods in geodesy,* H. Sünkel (ed.), Springer-Verlag.

Schriftenreihe Institut für Photogrammetrie der Universität Stuttgart, Stuttgart University, Geschwister-Scholl-Str. 24/D, D-70174 Stuttgart
Fax: ++49 711 121 3297; Internet: http://www.ifp.uni-stuttgart.de

---

Nr.	1	(1976)	Vorträge des Lehrgangs Numerische Photogrammetrie (III), Esslingen 1975 - vergriffen
Nr.	2	(1976)	Vorträge der 35. Photogrammetrischen Woche Stuttgart 1975
Nr.	3	(1976)	Contributions of the XIIIth ISP-Congress of the Photogrammetric Institute, Helsinki 1976 - vergriffen
Nr.	4	(1977)	Vorträge der 36. Photogrammetrischen Woche Stuttgart 1977
Nr.	5	(1979)	E. Seeger: Das Orthophotoverfahren in der Architekturphotogrammetrie, Dissertation
Nr.	6	(1980)	Vorträge der 37. Photogrammetrischen Woche Stuttgart 1979
Nr.	7	(1981)	Vorträge des Lehrgangs Numerische Photogrammetrie (IV): Grobe Datenfehler und die Zuverlässigkeit der photogrammetrischen Punktbestimmung, Stuttgart 1980 - vergriffen
Nr.	8	(1982)	Vorträge der 38. Photogrammetrischen Woche Stuttgart 1981
Nr.	9	(1984)	Vorträge der 39. Photogrammetrischen Woche Stuttgart 1983
Nr.	10	(1984)	Contributions to the XVth ISPRS-Congress of the Photogrammetric Institute, Rio de Janeiro 1984
Nr.	11	(1986)	Vorträge der 40. Photogrammetrischen Woche Stuttgart 1985
Nr.	12	(1987)	Vorträge der 41. Photogrammetrischen Woche Stuttgart 1987
Nr.	13	(1989)	Vorträge der 42. Photogrammetrischen Woche Stuttgart 1989
Nr.	14	(1989)	Festschrift - Friedrich Ackermann zum 60. Geburtstag, Stuttgart 1989
Nr.	15	(1991)	Vorträge der 43. Photogrammetrischen Woche Stuttgart 1991
Nr.	16	(1992)	Vorträge zum Workshop "Geoinformationssysteme in der Ausbildung", Stuttgart 1992

Technical Reports Department of Geodetic Science, Stuttgart University, Geschwister-Scholl-Str. 24/D, D-70174 Stuttgart
Fax: ++49 711 121 3285; Internet: http://www.uni-stuttgart.de/gi/research

---

Nr.	1	(1987)	K. Eren: Geodetic Network Adjustment Using GPS Triple Difference Observations and a Priori Stochastic Information
Nr.	2	(1987)	F.W.O. Aduol: Detection of Outliers in Geodetic Networks Using Principal Component Analysis and Bias Parameter Estimation
Nr.	3	(1987)	M. Lindlohr: SIMALS; SIMulation, Analysis and Synthesis of General Vector Fields
Nr.	4	(1988)	W. Pachelski, D. Lapucha, K. Budde: GPS-Network Analysis: The Influence of Stochastic Prior Information of Orbital Elements on Ground Station Position Measures
Nr.	5	(1988)	W. Lindlohr: PUMA; Processing of Undifferenced GPS Carrier Beat Phase Measurements and Adjustment Computations
Nr.	6	(1988)	R.A. Snay, A.R. Drew: Supplementing Geodetic Data with Prior Information for Crustal Deformation in the Imperial Valley, California 1988
Nr.	7	(1989)	H.-W. Mikolaiski, P. Braun: Dokumentation der Programme zur Behandlung beliebig langer ganzer Zahlen und Brüche
Nr.	8	(1989)	H.-W. Mikolaiski: Wigner 3j Symbole, berechnet mittels Ganzzahlarithmetik
Nr.	9	(1989)	H.-W. Mikolaiski: Dokumentation der Programme zur Multikplikation nach Kugelfunktionen entwickelter Felder
Nr.	10	(1989)	H.-W. Mikolaiski, P. Braun: Dokumentation der Programme zur Differentiation und zur Lösung des Dirichlet-Problems nach Kugelfunktionen entwickelter Felder
Nr.	11	(1990)	L. Kubácková, L. Kubácek: Elimination Transformation of an Observation Vector preserving Information on the First and Second Order Parameters
Nr.	12	(1990)	L. Kubácková: Locally best Estimators of the Second Order Parameters in Fundamental Replicated Structures with Nuisance Parameters
Nr.	13	(1991)	G. Joos, K. Jörg: Inversion of Two Bivariate Power Series Using Symbolic Formula Manipulation
Nr.	14	(1991)	B. Heck, K. Seitz: Nonlinear Effects in the Scalar Free Geodetic Boundary Value Problem
Nr.	15	(1991)	B. Schaffrin: Generating Robustified Kalman Filters for the Integration of GPS and INS
Nr.	16	(1992)	Z. Martinec: The Role of the Irregularities of the Earth's Topography on the Tidally Induced Elastic Stress Distribution within the Earth
Nr.	17	(1992)	B. Middel: Computation of the Gravitational Potential of Topographic-Isostatic Masses
Nr.	18	(1993)	M.I. Yurkina, M.D. Bondarewa: Einige Probleme der Erdrotationsermittlung
Nr.	19	(1993)	L. Kubácková: Multiepoch Linear Regression Models
Nr.	20	(1993)	O.S. Salychev: Wave and Scalar Estimation Approaches for GPS/INS Integration

Schriftenreihe der Institute des Studiengangs Geodäsie und Geoinformatik (ehemals Fachbereich Vermessungswesen); Technical Reports Department of Geodesy and GeoInformatics, Stuttgart University, Geschwister-Scholl-Str. 24/D, D-70174 Stuttgart
Fax: ++49 711 121 3285; Internet: http://www.uni-stuttgart.de/gi/research

---

Nr. 1994.1	(1994)	H.-J. Euler: Generation of Suitable Coordinate Updates for an Inertial Navigation System
Nr. 1994.2	(1994)	W. Pachelski: Possible Uses of Natural (Barycentric) Coordinates for Positioning
Nr. 1995.1	(1995)	J. Engels, E.W. Grafarend, P. Sorcik: The Gravitational Field of Topographic-Isostatic Masses and the Hypothesis of Mass Condensation - Part I & II
Nr. 1995.2	(1995)	Minutes of the ISPRS Joint Workshop on Integrated Acquisition and Interpretation of Photogrammetric Data
Nr. 1996.1	(1996)	Festschrift für Klaus Linkwitz anläßlich der Abschiedsvorlesung im Wintersemester 1995/96; herausgegeben von Eberhard Baumann, Ulrich Hangleiter und Wolfgang Möhlenbrink
Nr. 1996.2	(1996)	J. Shan: Edge Detection Algorithms in Photogrammetry and Computer Vision
Nr. 1997.1	(1997)	Erste Geodätische Woche Stuttgart, 7.-12. Oktober 1996; herausgegeben von A. Gilbert und E.W. Grafarend
Nr. 1997.2	(1997)	U. Kälberer: Untersuchungen zur flugzeuggetragenen Radaraltimetrie
Nr. 1998.1	(1998)	L. Kubácek, L. Kubácková: Regression Models with a weak Nonlinearity
Nr. 1999.1	(1999)	GIS-Forschung im Studiengang Geodäsie und Geoinformatik der Universität Stuttgart; herausgegeben von M. Sester und F. Krumm
Nr. 1999.2	(1999)	Z. Martinec: Continuum Mechanics for Geophysicists and Geodesists. Part I: Basic Theory
Nr. 1999.3	(1999)	J. H. Dambeck: Diagnose und Therapie geodätischer Trägheitsnavigationssysteme. Modellierung – Systemtheorie – Simulation – Realdatenverarbeitung
Nr. 1999.4	(1999)	G. Fotopoulos, C. Kotsakis, M. G. Sideris: Evaluation of Geoid Models and Their Use in Combined GPS/Levelling/Geoid Height Network Adjustment
Nr. 1999.5	(1999)	Ch. Kotsakis, M. G. Sideris: The Long Road from Deterministic Collocation to Multiresolution Approximation
Nr. 1999.6	(1999)	Quo vadis geodesia...? Festschrift for Erik W. Grafarend on the occasion of his 60th birthday; herausgegeben von F. Krumm und V.S. Schwarze - vergriffen, out of stock -