J. Optimization Problems in Geodetic Networks with Signals
A. Dermanis

1. Introduction

The most important task of applied science is to obtain information about the real world from the analysis of observational data. In the case of the physical sciences, geodesy being one of them, the relative simplicity of the considered part of the physical world allows the construction of models, i.e., creations of the human mind which with the help of the language of mathematics are more or less valid images of the relevant phenomena in the real world.

Observation, modelling and data analysis is what applied science must do and when one tries to do his best, he is naturally facing problems of optimization. Optimization, in general, is a systematic effort to answer questions about what to observe, how to model and how to analyze data in order to obtain information about some part of the physical world, depending on the specific scientific field, in the best possible way, within the framework of a priori defined possibilities for action.

Geodetic networks are a means for obtaining information about the geometry of a certain area on the earth surface and the associated optimization problems are:
- what to observe (first order design, the configuration problem)
- how to observe, i.e., with what accuracy to observe (second order design, the weight problem)
- how to best analyze data in their adjustment stage (zero order design, the datum problem)
- how to improve existing information with the help of additional data (third order design, the improvement problem).

It must be noted that the problem of optimum modelling has not attracted the attention it deserves, at least within the framework of optimization theory.

All the design-optimization problems are related to the adjustment of the relevant observations and are in fact formulated with the help of the same terms which appear in the adjustment (design matrix, weight matrix, covariance matrix of coordinates), see e.g. Grafarend (1974) and Schmitt (1982).

In the classical network concept, the network is just a set of points or, more precisely, the relative positions of a finite number of points. A finite set of observations is related to a finite set of unknown parameters, almost exclusively the coordinates of the network points, with the help of a simple Euclidean model expressed in terms of simple mathematical relations from analytic geometry. Any recognizable inconsistencies between the available observations and this simplified model which Grafarend (1982) calls kernel model are removed
by means of reductions which follow from the comparison of the kernel model with an original more realistic periphery model. All remaining inconsistencies are attributed to random errors and the adjustment aims at minimizing their effect with the help of a stochastic model of their behaviour.

The use of such a simplified notion of a network is strongly related to the traditional separation of obviously three-dimensional networks into their horizontal and vertical parts. The reasons for this separation are two. First, for small surveying networks, horizontal angle measurements refer to the same horizontal datum for all points, since the direction of the vertical is practically parallel within a small area. The second reason is that different types of observations are sensitive to different geometric characteristics of the network. Angular and distance observations on the earth surface mainly provide information about the two-dimensional geometry of the network, while levelling provides information about the remaining third dimension. This led to a separate adjustment of the triangulation from that of levelling and when the need for more extended networks arose, the horizon was replaced by a more suitable reference surface, the sphere and at a second stage the ellipsoid.

Theodolite and level observations always refer to the local vertical direction, the direction of the force of gravity. Their reduction with respect to a reference surface made necessary the knowledge of the gravity field. Deflections of the vertical are needed for the reduction of angular measurements and geoid heights are needed for the reduction of observed heights above the geoid to heights above the ellipsoid. As a consequence of the involvement of the gravity field, the subject matter of geodesy is no more the original purely geometric one, but the study of both the shape and the gravity field of the earth.

The pioneering ideas of Antonio Marussi, as well as, the advent of artificial satellites which provided geodesy with completely different types of observational data, led to the popularization of the concept of three-dimensional (3D) geodetic networks, although the original concept dates back to Bruns (1878) and his famous polyhedron. In terrestrial 3D networks the gravitational field of the earth is either ignored by treating the direction of the vertical independently from its relation to gravity, or it is assumed to be known from the independent analysis of other types of observations, mainly gravimetric ones.

The success of the implementation of the technique of collocation in gravimetric work led Eeg and Krarup (1975) to the suggestion of treating the gravity potential as an unknown in the analysis of 3D networks within the framework of the so called integrated geodesy. The operational geodesy has also been introduced by Grafarend (1978) and Moritz (1978a), thus emphasizing the significance of this new approach for practical geodetic work too.

The geodetic network has now lost its finite-parameter character, since in the analysis of the relevant observations the gravity potential function appears as an unknown and a function is an infinite-parameter mathematical entity. The improvement in the accuracy of geodetic observations and relevant ideas from geodynamics led to the concept of a deformable geodetic network where the point coordinates become also functions of time.

In a paper (Dermanis, 1979) presented at the previous Erice School, it
has been emphasized that there is no need to bring up the unknown function itself in the adjustment procedure. All that is needed is to consider only a finite set of parameters depending on the unknown function, the so-called *signals*.

The presence of signals in the adjustment of geodetic observations, in connection with the problem of how to treat them in the adjustment, does not allow the direct formulation of the standard optimization problems which have been developed for geodetic networks without signals.

In this work an attempt will be made to look into the problems of network optimization arising from the presence of signals.

2. DATA ANALYSIS AND SIGNALS

As we have already mentioned, network optimization problems are formulated in strict reference to the adjustment performed in the analysis of the relevant geodetic data. It is therefore necessary to look into the adjustment problems arising when treating geodetic observations depending on signals. This can be better done in contrast to the classical type of adjustment without signals.

The adjustment of the observations in a classical geodetic network falls within the following more general context. The specialization to the case of geodetic networks is given in parentheses.

A part of the physical world (N points) is artificially isolated from the rest of it. We call this conceptually isolated part a physical system. A model (Euclidean geometry) is constructed about the interrelations among the various elements of the physical system. A finite set of parameters (coordinates of points) is selected for the description of the physical system in such a way that every other element of the system is uniquely defined from the values of these parameters.

A set of observations is performed in the real world which assign values to various observable elements of the system (directions, angles, distances). With the help of appropriate mathematical tools (analytic geometry) mathematical equations are derived interrelating observables with parameters.

Since observations, i.e., the outcomes of the observational procedure, are part of the real world, while observables are elements of the man-made model, there are always inconsistencies between observations and observables. Such inconsistencies are unavoidable and reflect the shortcomings of a more or less simple model in describing a complicated real world situation. Any recognizable part of these inconsistencies is calculated when this is possible with the help of additional information and it is removed from the observations which become reduced observations (e.g., reductions for atmospheric refraction effects). The total effect of the remaining inconsistencies is characterized as observational error.

The determination of a finite number of parameters requires the knowledge of the values of at least as many other elements of the physical system. These can be the values of observables, but available observations differ from observables by unknown error values. Each observation brings in an equation which the unknown parameters must satisfy, but it also brings in a new unknown the observational error. Thus for n unknown parameters and m observations (with m observational errors) a system of m equations is formulated with n + m unknowns. Such a
system has an infinity of solutions.

In making a choice among infinite possible solutions additional information about the observational errors must be used. Since measuring instruments are constructed with care so that they give values of observations as close as possible to the corresponding observables and reductions are assumed to remove the difference between model observables and real world observables, errors must have small absolute values. One looks therefore for solutions with small error values.

The deterministic principle of least squares provides a way for obtaining a unique solution with small error values but it is not the only one. Minimization of the sum of the absolute values and minimization of the maximum absolute error (minimax) are two other less usual in practice alternatives. Weighted least squares where the weighted sum of the squares of the error estimates is minimized is a way to obtain a solution where the errors, beyond being all small, they are comparatively smaller for the more accurate observations.

Looking for the best choice of weights one must define a measure for the accuracy of the observations. At this point statistics and probability theory enter into the picture. The errors are assumed to be random outcomes out of a collection with zero mean and normal distribution. The dispersion or variance of this distribution becomes an inverse measure of accuracy.

The effort to obtain parameter estimates which are as accurate as possible leads to (linear) minimum variance estimation within the well known linear Gauss-Markov model. It is with respect to this linear Gauss-Markov model with a finite number of unknown parameters and a finite number of observations corrupted by zero-mean errors, that the various optimization problems are formulated.

The simplicity of the Euclidean model for network geometry and its convenient description by means of analytic geometry make it so easy to work with, that one would hardly replace it by a more complicated model. This even leads to procrustean solutions where reality is adapted to a simple model by means of not always permissible reductions. These reductions require additional information, such as the index of refraction or the deflections of the vertical at the network area. Insufficient information leads to erroneous reductions, but this is of no practical significance as far as the errors in the reductions remain well bellow the magnitude level of the observational errors.

The need for more accurate parameter estimates leads to the introduction of more precise instrumentation and eventually the errors in the above reductions become significant in comparison to the declining observational errors. These significant model errors can not be ignored and the original simplified model must be replaced by a more appropriate one. A more efficient model requires a larger number of parameters even an infinite one.

The idea of integrated or operational geodesy where the gravity potential function itself becomes one of the unknowns of the problem, set up the need for adjustment models with infinite number of parameters and finite number of observations. In such a case the least squares principle which involves only the observational errors is not sufficient for obtaining a unique solution. The remaining unknown parameters are too many (even infinite) and a modified least squares principle is needed where the parameters or a part of them participate together with the observational errors.
A more careful examination of such problems with a finite number of observations and an infinite number of unknown parameters (in fact an unknown function) shows that there is actually no need to include the unknown function itself in the problem. A finite number of parameters which depend on the underlying unknown function is sufficient for the formulation of the observational model. For example, horizontal and vertical angle observations in 3D networks do not depend directly on the gravity potential function, but only on two parameters for each station (longitude and latitude) which define the local direction of the vertical and depend in turn on the gravity potential function.

Such intermediate function-dependent parameters are called signals. Their introduction allows the replacement of the original infinite-to-finite adjustment problem (infinite parameters and finite observations) by two separate but related problems: a finite-to-finite problem (finite number of signals and other parameters to finite number of observations) and an infinite-to-finite problem (infinite function to finite signals). The second problem need not necessarily be solved.

The problem of how to treat the signals in the adjustment will be examined in connection with the most usual examples of geodetic networks where signals arise.

3. GEODETIC NETWORKS WITH SIGNALS

Signals in geodetic networks are first of all characterized by the particular functions they depend on.

Horizontal and vertical angle observations depend on the direction of the vertical and therefore on the gravity potential function. Thus three-dimensional networks are the first type of geodetic networks with signals. Assuming a rigid earth model, the network coordinates are unknown parameters and the gravity potential is a function of position only.

When a classical type of network is set up for the study of crustal deformation with the traditional separation into a horizontal and a vertical part, a deformable earth model is assumed and the plane coordinates or heights become functions of time. Their values at the observation epochs are the signals in the relevant geodetic network. If the coordinates at some reference epoch are used as unknown parameters their variation with time gives rise to displacements which are functions of time. The components of the displacement vector at the observation epochs can be alternatively used as signals.

When a deformable earth model is used in a 3D network with potential dependent observations, the underlying unknown functions are on one hand the coordinate or displacement functions of time and on the other hand a gravity potential function which is a function not only of space but also of time. One group of signals are the coordinate or displacement values at the observation epochs and the other one depends on a varying with time gravity potential. The signals in the last group are similar to those arising in a rigid 3D network, the only difference being that they must now also refer to the respective observational epochs.

The above three types of geodetic networks with signals will be used as examples for the investigation of some data analysis alternatives in treating this type of problem. However they are not the only ones.
When extraterrestrial observations are analyzed in geodetic networks, signals arise which depend on such functions as the orbit of a satellite, the time varying angles defining earth rotation or moon librations, the lunar, solar or planetary orbits, the atmospheric density function for satellite drag, etc.

Even for networks with classical terrestrial observations, where the study, e.g., of crustal deformation requires very accurate data, it is necessary to correct zenith angles and distances measured with EDM instruments for the effect of refraction. These corrections, or in fact the differences between the corrections applied and those that should be applied, depend on the distribution of the index of refraction along the optical paths at the epoch of the observation. They are therefore signals, the underlying function being the variation of the index of refraction or of the other meteorological parameters which define refraction, with respect to space and time.

4. DIFFERENT APPROACHES FOR THE ADJUSTMENT OF OBSERVATIONS DEPENDING ON SIGNALS

Strictly speaking, signals are additional unknown parameters which together with the original unknown parameters and the unknown errors must be estimated from the observational data by means of an adjustment procedure.

Expressing observables \( y_i^a \) in terms of unknown parameters \( x_i^a \) and signals \( s_i^a \)

\[
y_i^a = f_i(x_1^a, x_2^a, \ldots, x_n^a; s_1^a, s_2^a, \ldots, s_q^a) \quad i=1,2,\ldots,m
\]

or in matrix form

\[
y^a = f(x^a, s^a)
\]

the observation equations are

\[
y^b = y^a + v = f(x^a, s^a) + v
\]

where \( y_i^b \) are the observed values and \( v_i \) the corresponding errors. Linearization with the help of approximate parameter values \( x_0^a \) and approximate signal values \( s_0^a \), gives the linearized observation equations

\[
y^b - f(x_0^a, s_0^a) = \frac{\partial f}{\partial x} \bigg|_0 (x^a - x_0^a) + \frac{\partial f}{\partial s} \bigg|_0 (s^a - s_0^a) + v
\]

or

\[
1 = A x + G s + v
\]

with obvious notations.

Some minor complications arise from the fact that usually signals depend not only on the unknown underlying function but also on the unknown point coordinates which are included in the unknown parameters \( x^a \) (see Dermanis, 1979, for details). In any event, the final linearized observation equations are of the form of equation (5), although a more generalized approach is possible where a matrix of coefficients appears in front of the errors \( v \).
The approximate signals $s^0$ are obtained by using in the place of the unknown function a known approximate function, such as the normal gravity potential, and replacing the unknown coordinates when necessary by their approximate values.

It must be noted that the use of a set of signals is not uniquely determined. One must choose among possible alternatives, although some of them are naturally suggesting themselves. The same is also true for the unknown parameters.

The results from the least squares adjustment without signals can be used for the treatment of the adjustment with signals, if

a) we recognize that in the general adjustment model without signals

$$1 = A \mathbf{x} + B \mathbf{v}, \quad \mathbf{v}^T \mathbf{P} \mathbf{v} = \min$$

there are two types of unknowns. The unknowns $\mathbf{x}$ which do not participate in the minimization principle and the unknown errors $\mathbf{v}$ which are minimized.

b) we decide whether the signals should participate or not in the minimization principle which is necessary in order to arrive at a unique solution.

There are three different fundamental approaches to the treatment of signals, although hybrid approaches resulting from their combination are also possible. We give some "code names" to these approaches which are simply a matter of convenience and should not be taken literally. They have first been introduced in Dermanis and Grafarend (1981).

4.1 The Deterministic Approach

One way of looking upon signals is to consider them as additional unknown parameters without taking into account their common dependence on an underlying function. In this case the signals $s$ are treated in the same way as the original parameters $x$ and they do not participate in the least squares principle. The adjustment problem is

$$1 = [A \ 0] \begin{bmatrix} \mathbf{x} \\ \mathbf{s} \end{bmatrix} + \mathbf{v}, \quad \mathbf{v}^T \mathbf{P} \mathbf{v} = \min$$

with solution provided from the solution of the normal equations

$$\begin{bmatrix} A^T \mathbf{P} & A^T \mathbf{G} \\ G^T \mathbf{P} & G^T \mathbf{G} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} A^T \mathbf{P} \mathbf{1} \\ G^T \mathbf{P} \mathbf{1} \end{bmatrix}$$

and

$$\hat{\mathbf{v}} = 1 - A \hat{\mathbf{x}} - G \hat{\mathbf{s}}.$$  \hspace{1cm} (9)

This approach has the advantage that it does not require any additional assumptions about the signals. However, it is not always applicable in any type of situation. The observations must contain enough information for the estimation of a set of parameters and signals which are not necessarily the ones appearing in equation (7). In fact it is sufficient that

$$m > r = \text{rank}([A \ 0])$$  \hspace{1cm} (10)
The use in equation (7) of parameters and signals which are more than those actually needed for a proper description of the physical system under consideration, does not allow the determination of unique estimates for them. The situation is analogous to the familiar practice of using a redundant number of parameters in the adjustment of geodetic networks without signals, e.g., use of 2N coordinates for a horizontal triangulation-trilateration network instead of the 2N-3 parameters actually needed. The consequence of such an overparametrization, dictated by the convenience in the formulation of the observation equations, is that the normal equations have an infinite number of solutions since the rank of their coefficient matrix is

\[ r < n + q \]  \hspace{1cm} (11)

Additional criteria must be introduced in order to obtain a unique solution which is "best" in a certain sense. We shall look into this matter in the discussion of the Zero Order Design or Datum Problem.

4.2 The Model Function Approach

In this approach the dependence of signals \( s^q \) on an underlying function \( \phi(t) \) is taken into account. \( t \) is a domain index, such as time, point on plane or ellipsoid, point in 3D space, etc. The function \( \phi \) is modeled in a deterministic way so that it is completely specified by a set of real parameters \( a_i \)

\[ \phi = \phi(a, t) \]  \hspace{1cm} (12)

The simplest case is the use of a linear combination

\[ \phi(a, t) = \sum_{i=1}^{P} a_i \psi_i(t) \]  \hspace{1cm} (13)

where \( \psi_i(t) \) are known functions the so called base functions. In a linearized form the observation equations become

\[ 1 = A x + G F a + v \]  \hspace{1cm} (14)

where \( F \) is a matrix with elements depending on the base functions \( \psi_i \) and the way that signals depend on the function \( \phi \).

The adjustment problem becomes

\[ 1 = A x + H a + v \]  \hspace{1cm} (15)

\[ v^T P v = \text{min} \]

where

\[ H = G F \]

and the normal equations are

\[ \begin{bmatrix} A^T P A & A^T P H \\ H^T P A & H^T P H \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} A^T P 1 \\ H^T P 1 \end{bmatrix} \]  \hspace{1cm} (17)

We also have

\[ \hat{v} = 1 - A \hat{x} - H \hat{a} \]

Completely analogous is the case where the signals depend on more than one underlying functions or a vector-valued function.
The choice of the model function (12) plays the most important role in this approach. Such a function must be capable of describing reality with sufficient accuracy in comparison to the order of magnitude of the observational errors. On the other hand the number of the parameters \( \mathbf{a} \) must be small enough, so that the observations have sufficient informational content for their estimation in addition to the estimation of the original parameters \( \mathbf{x} \).

Instead of the parametrization of the function \( \varphi \) itself, it is possible to parametrize only its difference \( \delta \varphi \) from a corresponding approximate function \( \varphi_0 \), in which case

\[
\varphi(t) = \varphi_0(t) + \delta \varphi(a, t) \quad .
\]  

A familiar example of the model function approach is the well known short arc method of satellite geodesy. The signals are the satellite coordinates at the epochs of observation. The underlying function is the satellite orbit which is parametrized with the help of polynomials or other appropriate base functions. In contrast, the deterministic approach is followed in the geometric mode of satellite geodesy where simultaneous observations are performed and satellite coordinates are treated as additional unknowns ignoring the fact that they refer to points on a particular orbit. An example of the deterministic approach similar to the geometric mode of satellite geodesy but referring to the analysis of VLBI observations can be found in Dermawan and Graafarend (1981). The model function approach is also followed in the dynamic mode of satellite geodesy, where the underlying function is not the satellite orbit but the gravitational potential of the earth which is parametrized in terms of the coefficients of a truncated expansion in spherical harmonics.

4.3 The Stochastic Approach

In this approach the underlying function \( \varphi \) is supposed to be a sample function of a corresponding stochastic process with known mean function

\[
\mathbb{E} [\varphi(t)] = \varphi_0(t) \quad .
\]  

and known covariance function

\[
\mathbb{E} [ (\varphi(t) - \varphi_0(t)) (\varphi(t') - \varphi_0(t')) ] = c(t, t') \quad .
\]  

The signals \( \mathbf{s}^0 \) are known random variables with known means and the residual signals \( \mathbf{s} \) appearing in the linearized observation equations are random variables with zero means

\[
\mathbb{E} [\mathbf{s}] = \mathbf{0} \quad .
\]  

and known covariance matrix

\[
\mathbb{E} [\mathbf{s} \mathbf{s}^\top] = \mathbf{C}_{\mathbf{s}s} = \mathbf{K} \quad .
\]  

The adjustment problem becomes

\[
1 = \mathbf{A} \mathbf{x} + [G \mathbf{I}] [\begin{array} {c} \mathbf{s} \\ \mathbf{v} \end{array} ] , \quad \mathbf{s}^\top \mathbf{K}^{-1} \mathbf{s} + \mathbf{v}^\top \mathbf{P} \mathbf{v} = \min
\]  

where \( \mathbf{P} = \Sigma^{-1} \) is the error weight matrix and \( \Sigma \) the error covariance matrix. The corresponding normal equations are
\[(A^TM^{-1}A)^\hat{X} = A^TM^{-1}I\]  \hspace{1cm} (25)

where
\[M = GKG^T + \Sigma\]  \hspace{1cm} (26)

From the solution of the normal equations estimates of the signals and the errors are computed from
\[\hat{\Sigma} = KG^T M^{-1} (I - A\hat{X})\]  \hspace{1cm} (27)
\[\hat{\nu} = \Sigma M^{-1} (I - A\hat{X})\]  \hspace{1cm} (28)

An advantage of the stochastic approach is the possibility of predicting values of signals other than those present in the observations by means of the least squares prediction
\[\hat{s}' = C_s's \Sigma^{-1}s\]  \hspace{1cm} (29)

The same possibility exists in the model function approach where the new signals \(s'\) are directly computed from the estimates of the parameters \(a\) by means of a linearized relation of the form
\[\hat{s}' = F'\hat{a}\]  \hspace{1cm} (30)

where \(F'\) is a matrix with elements depending on the way that the new signals are related to the parametrized unknown function.

In the deterministic approach new signals can be "predicted" only after a reconstruction of the unknown function \(\Phi\) has been obtained by solving an interpolation problem. Equation (29) can also be used in the deterministic approach by using an arbitrary model covariance function \(c(t,t')\). In this case equation (29) must not be interpreted in a stochastic way, but it should be simply seen as a particular way of solving the above mentioned interpolation problem.

The stochastic approach is the one most widely used in practice. It is also the one whose theoretical foundation has been strongly questioned especially when it is used for phenomena not having a clearly stochastic character, such as the gravity field of the earth (Moritz and Sansó, 1978), Moritz (1978b).

4.4 Hybrid Approaches

In addition to the above three fundamental approaches to the treatment of signals, hybrid approaches which lie midway among them are also possible.

Since, at least according to the stochastic interpretation, the signals \(s\) must have zero means, it is possible to try to detect the unknown generally non-zero mean function of \(\Phi\) by means of a technique known as trend determination or detrending, a term borrowed from time series analysis. The function \(\Phi\) is written
\[\Phi(t) = \Phi_0(t) + \sum_{i=1}^{P} a_i \psi_i(t) + \delta\Phi(t)\]  \hspace{1cm} (31)

where \(a_i\) are real non-stochastic parameters which are added to the original set of unknowns \(x\) and \(\delta\Phi\) is a zero mean stochastic process with known covariance function.
Setting \( \psi^0 \) for the part of the signals obtained from the non-stochastic part

\[
\psi(t) + \sum_{i=1}^{p} a_i \psi_i(t)
\]

of \( \psi \), the linearized observation equations take the form

\[
\begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + \begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + \begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix} + \begin{bmatrix} \mathbf{v} \\ \mathbf{v} \end{bmatrix}.
\]  

(32)

The adjustment problem is

\[
\begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + \begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix} \quad \text{with normal equations}
\]

\[
\begin{bmatrix} \begin{bmatrix} \mathbf{A}^T \mathbf{M}^{-1} \mathbf{A} \\ \mathbf{H}^T \mathbf{M}^{-1} \mathbf{A} \\ \mathbf{A}^T \mathbf{M}^{-1} \mathbf{H} \\ \mathbf{H}^T \mathbf{M}^{-1} \mathbf{H} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T \mathbf{M}^{-1} \mathbf{I} \\ \mathbf{H}^T \mathbf{M}^{-1} \mathbf{I} \end{bmatrix}
\]  

(34)

with normal equations

\[
\begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix} = \mathbf{G} \mathbf{K} \mathbf{G}^T + \Sigma
\]  

(35)

and

\[
\mathbf{G} \mathbf{K} \mathbf{G}^T \mathbf{M}^{-1} (I - \mathbf{A} \mathbf{\hat{x}} - \mathbf{H} \mathbf{\hat{a}})
\]  

\[
\mathbf{v} = \Sigma \mathbf{M}^{-1} (I - \mathbf{A} \mathbf{\hat{x}} - \mathbf{H} \mathbf{\hat{a}})
\]  

(36)

It is also possible to treat some of the parameters \( a_i \) as non-stochastic parameters with no weights and some of them as stochastic ones with zero mean and known covariance matrix \( \mathbf{R}^{-1} \) (weight matrix \( \mathbf{R} \)). When all \( a_i \) are treated as stochastic parameters the adjustment problem becomes

\[
\begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} = \mathbf{A} \mathbf{x} + \mathbf{H} \mathbf{a} + \mathbf{G} \mathbf{s} + \mathbf{v}, \quad \mathbf{a}^T \mathbf{R} \mathbf{a} = \mathbf{s}^T \mathbf{K}^{-1} \mathbf{s} + \mathbf{v}^T \mathbf{P} \mathbf{v} = \min
\]  

(38)

A stochastic version of the model function approach is to set

\[
\psi(t) = \psi(t) + \sum_{i=1}^{p} a_i \psi_i(t)
\]  

(39)

with no restriction on the magnitude of \( p \), in which case the adjustment problem becomes

\[
\begin{bmatrix} \mathbf{a} \end{bmatrix} = \mathbf{[G F I]} \begin{bmatrix} \mathbf{a} \end{bmatrix}, \quad \mathbf{a}^T \mathbf{R} \mathbf{a} + \mathbf{v}^T \mathbf{P} \mathbf{v} = \min
\]  

(40)

This case is equivalent to the stochastic approach of equation (24) with

\[
\mathbf{s} = \mathbf{F} \mathbf{a}
\]  

(41)

and

\[
\mathbf{K} = \mathbf{G} \mathbf{R}^{-1} \mathbf{G}^T
\]  

(42)
A final case is when estimates $x'$ of $x^a$ are a priori available with covariance matrix $\Sigma_x$. Using these estimates as approximate values in the linearization $x$, the following adjustment problem is formulated

$$1 = A x + G s + v, \quad x^T \Sigma_x^{-1} x + s^T k^{-1} s + v^T \Sigma^{-1} v = \min$$

(43)

provided that $\det(\Sigma_x) \neq 0$.

A hybrid approach of special interest is when the parameters $a_i$ are related to the datum definition. An example can be found in Grønland (1978) where corrections are introduced for the parameters defining the shape, size, position and orientation of the reference ellipsoid associated with the Somigliana-Pizzetti normal gravity field. An other example is the inclusion of datum shift parameters in problems of prediction of gravity related quantities, in order to remove systematic trends (see, e.g., Tscherning, 1978).

5. ZERO ORDER DESIGN WITH SIGNALS

5.1 General Remarks

The zero order optimal design is not really an optimization problem in the sense of optimizing the network itself; it can be rather characterized as the search for an optimal coordinate reference frame or datum.

In classical geodetic networks without signals, geodetic observations of distances and angles are sufficient for the determination of the shape and size of the network. A proper parametrization would be to choose a set of independent quantities whose number is the minimum required for the determination of shape and size. This minimum number of parameters, which we call the rank of the network, is $2N - 3$ for two-dimensional $3N - 6$ for three-dimensional networks with $N$ stations. When no distances are observed only the shape of the network can be determined and the network rank is $2N - 4$ and $3N - 7$ for 2D and 3D networks respectively.

The most convenient choice of parameters is to use the cartesian coordinates of the network points in view of the simplicity in constructing the observation equations which relate observed quantities to the coordinates. Coordinates, although they are independent quantities, they are more ($2N$ for 2D and $3N$ for 3D networks) than the minimum number required for a proper parametrization. This redundancy is due to the fact that coordinates contain more information than the shape and size of the network. They also define the position of the network with respect to a reference frame or, to be more precise, the position of a reference frame with respect to the network. The number $d$ of extra parameters (number of coordinates minus network rank) is equal to the number of parameters required for the placement of the network with respect to the reference frame. These are 3 parameters in the two-dimensional case (2 for translation and 1 for rotation) and 6 parameters in the three-dimensional case (3 for translation and 3 for rotation). When no distances are observed the above numbers become 4 and 7 respectively by the inclusion of one more parameter for scale. Equivalently, $d$ is the number of the degrees of freedom in coordinate transformations which leave the observations invariant.

The standard least squares problem

$$1 = A x + v, \quad v^T P v = \min$$

(44)
has an infinity of solutions and this fact is reflected in the rank deficiency of the \( m \times n \) design matrix \( A \) whose rank cannot exceed the rank of the network. In fact

\[
\text{rank}(A) = r = n - d < m \quad .
\]  

(45)

The normal equations have a parameter coefficient matrix with the same rank deficiency \( d \) and they accept an infinite number of solutions. Each solution corresponds to a choice of reference frame.

There are two ways to overcome this datum defect problem. One way is to include additional observations which relate the network to an existing datum. Such observations usually remove some of the \( d \) degrees of freedom but not all of them. For example, azimuth observations in 2D networks remove the rotation defect but the translation defect remains. In 3D networks, astronomic longitude, latitude and azimuth observations at one point at least define the orientation of the network with respect to a geocentric frame, but the translation defect remains.

The other way to overcome the datum defect problem is to arbitrarily define a reference frame having no physical significance, by means of a number of constraints on the parameters (coordinate corrections). Such constraints which define the reference frame without distorting the shape and size of the network are called minimal constraints. They are mutually independent and equal in number to the datum defect \( d \). In linear form they are written

\[
C \mathbf{x} = \mathbf{b} \quad .
\]  

(46)

When added to the normal equations they reduce the number of independent parameters from \( n \) to \( n - d = r \) and allow only one unique solution.

In such a solution, the position of the reference frame depends on the approximate coordinate values used for the linearization and on the particular minimal constraints.

Among the various possible minimal constraints, the so-called inner constraints (Blaha, 1971) lead to an optimal solution \( \mathbf{x} \), in the sense that

\[
\mathbf{x}^\top \mathbf{x} = \min
\]  

(47)

and

\[
\text{trace}(\Sigma_\mathbf{x}) = \min
\]  

(48)

where \( \Sigma_\mathbf{x} \) is the covariance matrix of \( \mathbf{x} \). The inner constraints have the form

\[
E \mathbf{x} = 0
\]  

(49)

where

\[
A E = 0 \implies N E = A^\top P A E = 0 \quad .
\]  

(50)

In this case

\[
\Sigma_\mathbf{x} = \sigma^2 N^\top
\]  

(51)

where \( \sigma^2 \) is the a posteriori variance of unit weight and \( N^\top \) the pseudoinverse of \( N \).
When parameters other than coordinate corrections are included in the parameter vector \( \mathbf{x} \), such as orientation unknowns in direction observations, they must be first eliminated before applying the pseudoinverse solution. In this way they do not participate in the optimality criteria (47) and (48). For a more detailed discussion on this problem see Frisch and Schaffrin (1981).

It is not possible to extend the above results to the case of geodetic networks with signals in a generalized way. Each type of network must be examined separately, depending on the type of included signals. In addition different formulations of the datum problem correspond to different approaches to the treatment of signals. We shall only look into the deterministic and the stochastic approach. The model function approach and related hybrid approaches depend strongly on the particular parametrization chosen for the underlying function and thus no general results can be obtained.

5.2 Three-dimensional Networks

We first consider 3D networks where the observables are horizontal and vertical angles measured with theodolites and distances. The case where only distances are observed (spatial trilateration) falls within the classical case of geodetic networks without signals.

The signals are two parameters defining the direction of the vertical at each station and they depend on the gravity potential \( W \) of the earth. The original signals are the astronomic longitude \( \lambda \) and latitude \( \phi \), while the signals appearing in the linearized observation equations are either the corresponding disturbances \( \Delta \lambda \), \( \Delta \phi \) or the anomalies \( \lambda_0 \), \( \phi_0 \). However \( \lambda \) and \( \phi \) need not refer to the geocentric conventional frame defined by the CIO and the Greenwich meridian. They may refer to any cartesian frame of reference fixed with respect to the rigid earth, provided the normal gravity potential \( U \) is modified accordingly when necessary.

Some problems result when a frame is used with its third axis being close to the mean vertical direction at the network area. This is analogous to the problem of weak definition of \( \lambda \) in the vicinity of the poles. Use of a different type of angles instead of \( \lambda \), \( \phi \) takes care of this problem.

The observation equations for three-dimensional geodetic networks and their linearization are given in Appendix A.

We shall determine the ranks of the matrices \( \mathbf{A} \) and \( \mathbf{G} \) appearing in the linearized observation equations in a heuristic way.

Let us assume for a moment that the signals \( \mathbf{s} \) are known. This means that the orientation of the network is known and writing

\[
\mathbf{l} - \mathbf{G} \mathbf{s} = \mathbf{A} \mathbf{x} + \mathbf{v}
\]

with known left side, \( \mathbf{A} \) must have a rank deficiency of 3 corresponding to the remaining parameters of network translation

\[
\text{rank}(\mathbf{A}) = r = n - 3 < m
\]

If no distances are observed the rank deficiency is 4 because scale is also not defined.

In the particular case where the directions of the vertical at all
stations are mutually parallel, a rotation of the network around this common vertical direction is possible with no effect on the observations. In this case where one out of the three orientation parameters is not defined, the rank deficiency of \( A \) is 4 (5 if no distances are observed). From now on we refer to networks where one distance at least has been observed.

In practice, the vertical directions are not exactly the same but they are practically the same for small local networks. Then, the theoretical rank deficiency is 3 but in practice, as far as numerical operations are concerned, the rank deficiency is 4. Failure to recognise this simple fact will result in computational problems in the adjustment.

Let us assume now that all station coordinates are known. Observations of vertical angles from each station to at least two others are sufficient for the determination of the direction of the vertical at network points. Writing

\[
1 - A x = G s + v
\]  
(54)

with known left side, \( G \) must be of full column rank

\[
\text{rank}(G) = q < m
\]  
(55)

If \( x \) contains parameters other than coordinate corrections, such as orientation unknowns for horizontal direction observations, we assume they are also known in the argument above.

A singular case appears when the vertical angles observed at some stations are in the same azimuth or in two azimuths differing by 200°.

**Deterministic approach:** Since orientation and position of the network is unknown the combined design matrix in the observation equations (7) must have a rank deficiency of 6

\[
\text{rank}(\begin{bmatrix} A & G \end{bmatrix}) = r = n + q - 6
\]  
(56)

The simplest way of defining the reference frame is by means of minimal constraints. One way is to fix 6 coordinates divided over at least 3 points not on the same line, to their approximate values. Another way is to fix an azimuth from one station together with two vertical angles observed from the same station at their approximate values computed from the approximate coordinates and fix the coordinates of the same point to its approximate ones.

For the optimal datum definition the situation differs from the classical case because of the presence of signals. For example a pseudo-inverse solution of the normal equations

\[
\begin{bmatrix}
  x \\
  s
\end{bmatrix} = \begin{bmatrix}
  A^T A & A^T G \\
  G^T A & G^T G
\end{bmatrix} \begin{bmatrix}
  A^T P \\
  G^T P
\end{bmatrix}
\]  
(57)

satisfies

\[
x^T x + s^T s = \min
\]  
(58)

which means that the frame definition depends not only on the approximate values of station coordinates (\( x^T x \) part) but also on the normal gravity potential (\( s^T s \) part) used in the linearization. A similar
problem involving station coordinates and orientation unknowns has been treated by Fritsch and Schaffrin (1981).

In order to have an optimal datum definition in the classical sense
\[
x^T x = \min, \quad \text{trace}(\mathbf{\Sigma}_x) = \min \tag{59}
\]
assuming that \( \hat{x} \) contains only coordinate corrections, the signals \( \hat{s} \) must first be eliminated from the normal equations. Since \( \mathbf{G} \) is full column rank \( \mathbf{G}^T \mathbf{P} \mathbf{G} \) is invertible and the reduced normal equations are
\[
[\mathbf{A}^T \mathbf{P} \mathbf{A} - \mathbf{A}^T \mathbf{P} \mathbf{G} (\mathbf{G}^T \mathbf{P} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{P} \mathbf{A}] \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{P} \mathbf{1} - \mathbf{A}^T \mathbf{P} \mathbf{G} (\mathbf{G}^T \mathbf{P} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{P} \mathbf{1} \quad \tag{60}
\]
After the normal equations are solved the signals can be computed by
\[
\hat{s} = (\mathbf{G}^T \mathbf{P} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{P} (\mathbf{1} - \mathbf{A} \hat{\mathbf{x}}) \quad \tag{61}
\]
In practice the reduced normal equations (60) are solved with the help of the standard inner constraints for 3D networks

\[
\mathbf{E} \mathbf{x} = \sum_{i=1}^{N} \mathbf{E}_i \mathbf{r}_i = 0 \tag{62}
\]
where \( \mathbf{r}_i \) is the vector of coordinate corrections for the \( i \)th station,

\[
\mathbf{E}_i = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & -z_i^0 & y_i^0 \\
z_i^0 & 0 & -x_i^0 \\
y_i^0 & x_i^0 & 0
\end{bmatrix} \quad \tag{63}
\]
and \( x_i^0, y_i^0, z_i^0 \) are the corresponding approximate coordinates.

The estimates \( \hat{x} \) and their covariance matrix \( \mathbf{\Sigma}_x \) can be obtained with the help of \( \mathbf{E} \) either by augmentation
\[
\begin{bmatrix}
\hat{x} \\
\hat{k}
\end{bmatrix} = \begin{bmatrix}
\mathbf{N}_x & \mathbf{E}^T \\
\mathbf{E} & \mathbf{0}
\end{bmatrix}^{-1} \begin{bmatrix}
\mathbf{u}_x \\
\mathbf{0}
\end{bmatrix} \tag{64}
\]
where \( \mathbf{N}_x \) and \( \mathbf{u}_x \) are the coefficient and the constant term matrices of the reduced normal equations (60) respectively, or directly by
\[
\hat{x} = (\mathbf{N}_x + \mathbf{E}^T \mathbf{E})^{-1} \mathbf{u}_x \quad \tag{65}
\]
see, e.g., Koch (1980). The covariance matrix \( \mathbf{\Sigma}_x \), apart from the a posteriori variance of unit weight scale factor, is found in the augmentation approach in the left upper part of the inverse of the augmented matrix in equation (64), or directly by
\[
\mathbf{\Sigma}_x = (\mathbf{N}_x + \mathbf{E}^T \mathbf{E})^{-1} - \mathbf{E}^T (\mathbf{E} \mathbf{E}^T \mathbf{E})^{-1} \mathbf{E} \quad \tag{66}
\]
The situation is different when the datum orientation is defined with the help of additional observations related to the signals, such as observations of astronomic longitude, latitude and azimuth. Three such observations at one station fully determine the orientation of the
network with respect to the standard geocentric equatorial frame. A
translation defect of 3 remains and must be removed for the definition
of a reference frame parallel to the geocentric one but with arbitrary
origin.

When additional observations related to the signals \( s \) are available,
which are more than the minimum required number, they do not only de-
fine the network orientation but they also contribute to its shape and
size. In this case the linearized observation equations have the form

\[
\begin{align*}
L_1 &= A x + G s + v_1 \\
L_2 &= B s + v_2
\end{align*}
\]  

(67) (68)

where azimuth observations are contained in \( L_1 \), longitude and latitude
ones in \( L_2 \). The corresponding adjustment problem is

\[
\begin{align*}
L &= \begin{bmatrix} A & G \\ 0 & B \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} + v, \\
\nu^T v &= \nu_1^T \nu_1 + \nu_2^T \nu_2 = \min
\end{align*}
\]  

(69)

with normal equations

\[
\begin{align*}
\begin{bmatrix} A^T P_1 A & A^T P_1 G \\ G^T P_1 A & G^T P_1 G + B^T P_2 B \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{s} \end{bmatrix} &= \begin{bmatrix} A^T P_1 L_1 \\ G^T P_1 L_1 + B^T P_2 L_2 \end{bmatrix}
\end{align*}
\]  

(70)

Elimination of the signals leads to the reduced normal equations

\[
\begin{align*}
\begin{bmatrix} A^T P_1 A - A^T P_1 G (G^T P_1 G + B^T P_2 B)^{-1} G^T P_1 A \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{s} \end{bmatrix} &= \\
&= A^T P_1 L_1 - A^T P_1 G (G^T P_1 G + B^T P_2 B)^{-1} (G^T P_1 L_1 + B^T P_2 L_2)
\end{align*}
\]  

(71)

The optimal solution \( \hat{x} \) can be obtained with the help of inner con-
straints referring only to the datum origin

\[
E x = \sum_{i=1}^{N} r_i = 0
\]  

(72)
i.e., with

\[
E = \begin{bmatrix} I & I & \ldots & I \end{bmatrix}
\]  

(73)
The signal estimates are computed from

\[
\hat{s} = (G^T P_1 G + B^T P_2 B)^{-1} \begin{bmatrix} G^T P_1 (I_1 - A) \hat{x} + B^T P_2 L_2 \end{bmatrix}
\]  

(74)

Astronomic observations are directly related to the stars and relation
to the earth fixed frame is achieved by means of reductions where
estimates of precession-nutation and polar motion parameters are used.
When these estimates do not conform with their nominal statistics
(variances and covariances, means) the inclusion of astronomic obser-
vations may cause a decline in network quality by introducing deforma-
tions which are not compatible with the accuracy of the geometric ob-
servations. Such an effect can be detected by means of statistical
tests on the results of solutions with and without the astronomic ob-
servations.
When the accuracy of the astrometric observations is not satisfactory, their use may give a bad picture in terms of coordinate variances or error ellipsoids as a consequence of a weak definition of network orientation. To overcome this problem we suggest the following procedure: Refer signals (corrections for $\lambda$ and $\omega$) not to the geocentric but to an arbitrary frame. Relate astrometric observations to their counterparts with respect to an arbitrary frame including 3 additional parameters as additional unknowns. Define both orientation and position of the network with 6 inner constraints to obtain also estimates for the 3 rotational parameters. Such an approach allows a complete, i.e., with respect to both translation and rotation, optimum datum definition, with parallel estimation of the rotational relation between the geocentric frame present in the astrometric observations and the optimal frame used in the adjustment. Low accuracy astrometric observations will simply give low accuracies for the 3 rotational parameters between the two frames.

Examples of the use of the deterministic approach in 3D networks can be found, e.g., in Engler et al. (1981) and Hradil (1984).

**Stochastic approach:** In the stochastic approach the adjustment problem is

$$1 = A x + [G I] \begin{bmatrix} s \\ v \end{bmatrix}, \quad s^T K^{-1} s + v^T \Sigma^{-1} v = \min \tag{75}$$

with

$$\text{rank}(A) = r = n - 3 < m \tag{76}$$

$$\text{rank}(G) = m < q + m \tag{77}$$

The participation of $s$ in the minimization condition defines the network orientation. The signals are deviations of vertical directions from their normal counterparts and the term $s^T K^{-1} s$ gives an orientation to the network close to the orientation of the normal vertical directions, which depend on the normal gravity potential used. A normal potential $U$, such as the Somigliana-Pizzetti normal field, implies the intrinsic definition of a reference frame which must be the same one present in the approximate station coordinates.

It is assumed here that at least one astrometric azimuth has been observed. Otherwise attention must be paid to the close to singularity situation discussed in the previous section.

In the normal equations given by equations (25) and (26) a rank deficiency of 3 remains. An optimal datum definition with respect to translation can be obtained with the help of the inner constraints of equation (72) and signal estimates can be obtained by means of equation (27).

When additional observations related to the signals are available the observation equations are those of equations (67) and (68). However the adjustment problem now is

$$1 = \begin{bmatrix} A \\ 0 \end{bmatrix} x + \begin{bmatrix} G & I & 0 \\ 0 & B & I \end{bmatrix} \begin{bmatrix} s \\ v_1 \\ v_2 \end{bmatrix}, \quad s^T K^{-1} s + v_1^T \Sigma^{-1} v_1 + v_2^T \Sigma^{-1} v_2 = \min \tag{78}$$

Again
\[
\text{rank}\left(\begin{bmatrix}
A \\
0
\end{bmatrix}\right) = r = n - 3 < m = m_1 + m_2
\] (79)

where \(m_1, m_2\) are the number of observations in \(l_1\) and \(l_2\) respectively, while

\[
\text{rank}\left(\begin{bmatrix}
G & I & 0 \\
B & 0 & I
\end{bmatrix}\right) = m < q + m.
\] (80)

The reduced normal equations are

\[
(A^T Q A) \hat{x} = A^T Q \left[ l_1 - \frac{G}{K} B^T (B K B^T + \Sigma_2)^{-1} l_2 \right]
\] (81)

where

\[
Q = \left[ \begin{bmatrix} G & K G^T + \Sigma_1 - G & K B^T (B K B^T + \Sigma_2)^{-1} B & K G^T \end{bmatrix} \right]^{-1}
\] (82)

and have a translation rank defect of 3. Use of the inner constraints of equation (72) defines the optimal datum position. Signal estimates are given by

\[
\hat{s} = K [G B^T] \left[ \begin{bmatrix} G & K G^T + \Sigma_1 & G & K B^T \\
B & K G^T & B & K B^T + \Sigma_2 \end{bmatrix}^{-1} \begin{bmatrix} l_1 - A \hat{x} \\
l_2 \end{bmatrix} \right].
\] (83)

In the stochastic approach it is also possible to include observations related to signals which only partly coincide, or are even completely different from the signals appearing in the original observation equations. The observation equations in this case have the form

\[
l_1 = A x + G s_1 + \nu_1
\] (84)

\[
l_2 = B s_2 + \nu_2
\] (85)

and the adjustment problem is

\[
L = \begin{bmatrix} A \\
0 \end{bmatrix} x + \begin{bmatrix} G & 0 & I & 0 \\
0 & B & 0 & I \end{bmatrix} \begin{bmatrix} s_1 \\
0 \\
v_1 \\
v_2 \end{bmatrix}
\]

\[
\begin{bmatrix} s_1^T \\
s_2^T \end{bmatrix} \begin{bmatrix} K_1 & K_{12} \\
K_{12} & K_2 \end{bmatrix} \begin{bmatrix} s_1 \\
s_2 \end{bmatrix} + v_1 \Sigma_{1}^{-1} v_1 + v_2 \Sigma_{2}^{-1} v_2 = \text{min}
\] (86)

The normal equations are

\[
(A^T Q A) \hat{x} = A^T Q \left[ l_1 - \frac{G K_{12}}{K_2} B^T (B K_2 B^T + \Sigma_2)^{-1} l_2 \right]
\] (87)

where now

\[
Q = \left[ \begin{bmatrix} G K_1 G^T + \Sigma_1 - G K_{12} B^T (B K_2 B^T + \Sigma_2)^{-1} B & K_{12} G^T \end{bmatrix} \right]^{-1}
\] (88)
Again, application of the inner constraints (72) defines the optimum reference frame position and leads to unique estimates for x and s.

Additional observations related to the gravity field such as observations of longitude, latitude and gravity at points other than the network stations, have in fact a linearized form

$$l_2 = A_2 x_2 + B s_2 + v_2$$  \hspace{1cm} (89)

since they also depend on the unknown coordinates of the points where the observations are performed. For longitude and latitude observations the A_2 terms are negligible (see Appendix A). For gravity observations we have to use x_2=0, since we cannot hope to determine these coordinates from gravity observations alone. Thus equations (89) reduce to equations (85).

Longitude and latitude observations at the network points contain signals included in s_1 and they can be considered within l_1 with corresponding zero rows of A in equation (84).

Gravity observations at network points or potential differences between network points from levelling, give rise to linearized observation equations

$$l_2 = A_2 x + B s_2 + v$$  \hspace{1cm} (90)

with A_2≠0. Looking at the explicit form of these equations (Appendix A) it follows that they are essentially sensitive only to translations along the vertical direction. The same is true for the signals in s_2, which are gravity disturbances 6g and disturbing potential values 6T.

From the computational point of view a defect of 2 for translations within the horizontal plane remains in the total matrix A of the system

$$\begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x + \begin{bmatrix} G \\ 0 \end{bmatrix} s_1 + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$  \hspace{1cm} (91)

in the case of small regional networks. This defect can be removed with the help of two appropriate inner constraints. In addition, inner constraints are needed for the rotation defect when no longitude, latitude or azimuth observations are contained in l_1.


6. DEFORMABLE NETWORKS

In the case of deformable networks the station coordinates appearing in the observation equations are the instantaneous coordinates at the epoch of observation. As such, they are signals depending on the unknown coordinate functions x(t), y(t), z(t), t being the time.

For the definition of the coordinate functions a reference frame must be specified, in principle for every epoch. The network station coordinates x_i(t_0), y_i(t_0), z_i(t_0) at some initial epoch t_0 can be considered as unknown parameters, while the displacements
\[ p_i(t) = \begin{bmatrix} u_i(t) \\ v_i(t) \\ w_i(t) \end{bmatrix} = \begin{bmatrix} x_i(t) - x_i(t_o) \\ y_i(t) - y_i(t_o) \\ z_i(t) - z_i(t_o) \end{bmatrix} = r_i(t) - r_i(t_o) \]  

(92)

are signals depending on the displacement functions

\[ p(r,t) = \begin{bmatrix} u(r,t) \\ v(r,t) \\ w(r,t) \end{bmatrix} \]  

(93)

which are functions of space and time.

The reference frame at the initial epoch \( t_o \) has to be defined, as well as, its subsequent motion.

In practice, observations falling within a small time interval, say a week, are considered to be simultaneous referring to a common mean epoch. This is equivalent to a model function approach where the coordinate functions are modeled as step functions. Other type of parameterizations such as time polynomials are also used, especially for the height component in studies of crustal uplift from repeated levelings, see, e.g., Vanicek et al. (1979).

A usual approach to the study of deformations is the comparison of network geometry at two epochs \( t \) and \( t' \) for the derivation of relevant crustal deformation parameters. The station coordinates \( r_i(t) \) at the first epoch are considered as unknown parameters while the displacements

\[ p_i = r_i(t') - r_i(t) \]  

(94)

are signals depending on the unknown displacement function \( p(r) \) which is now only a function of space, the two epochs being fixed.

The horizontal and the vertical parts of the network are usually analyzed separately. For the sake of simplicity we shall consider only horizontal networks for the case of two-epoch comparisons.

In two epoch comparisons two reference frames must be defined, one for each epoch. The two frames can be identified only when two at least network points are assumed to be motionless, thus providing a connection between the two network configurations. For the most general case of unconnected configurations the observation equations have the form

\[ l_1 = A_1 x + v_1 \quad v_1^T p_1 v_1 = \text{min} \]  

\[ l_2 = A_2 x' + v_2 \quad v_2^T p_2 v_2 = \text{min} \]  

(95) \hspace{1cm} (96)

where \( l_1, l_2 \) are the observations and \( x, x' \) the vectors of network coordinates at epochs \( t \) and \( t' \) respectively. The observations are adjusted separately at each epoch and the coordinate estimates \( x, x' \) are compared a posteriori.

If we write

\[ x' = x + s \]  

(97)

where the signal vector \( s \) contains the displacements at network points the observation equations become
\[ l_1 = A_1 x + v_1 \]  \hspace{1cm} \text{(98)}
\[ l_2 = A_2 x + A_2 s + v_2 \]  \hspace{1cm} \text{(99)}

In the deterministic approach for equations (98) and (99) the adjustment results are exactly the same as those obtained from the separate adjustments for (95) and (96). The adjustment problem is

\[ \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ A_2 & A_2 \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad v_1^T P_1 v_1 + v_2^T P_2 v_2 = \min \]  \hspace{1cm} \text{(100)}

and the design matrix above has a rank deficiency corresponding to the definition of two datums, one for each epoch.

In practice, the adjustment for the first epoch is carried out first and estimates \( \hat{x} \) are obtained after the arbitrary definition of a datum. Using the adjusted coordinates \( \hat{x}^* \) of the first epoch as approximate values for the coordinates of the second epoch we have

\[ l_2 - A_2 \hat{x} = A_2 s + v_2, \quad v_2^T P_2 v_2 = \min \]  \hspace{1cm} \text{(101)}

It must be clarified that the covariance matrix of the estimates \( \hat{x}^* \) or \( \hat{x} \) is ignored. The values of these estimates are only used as approximate coordinates for linearization purposes in a completely deterministic way, the same as any other approximate values.

The datum problem at the second epoch is solved by requiring that

\[ s^T s = \sum_{i=1}^n s_i^2 = \min \]  \hspace{1cm} \text{(102)}

i.e., by means of inner constraints (Brunner et al., 1981). This approach minimizes the motion between the two datums in a least squares sense, filtering out any undetectable by the available observations common translation and rotation of all network points.

In the model function approach the displacements \( s \) are expressed in terms of fewer parameters in accordance with simple deformation models. An example of this approach can be found in Chrzanoski et al. (1983). In such a case the choice of a particular parametrization usually takes care of the choice of datum for the second epoch with respect to the first. The datum defect for the first epoch still remains and must be removed with the usual means.

In the stochastic approach the displacements are assumed to be spacewise correlated with known covariance functions. This is a reasonable assumption since one naturally expects similar displacements for neighbour points. The adjustment problem becomes

\[ \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} + \begin{bmatrix} 0 & I & 0 \\ A_2 & 0 & I \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad s^T K^{-1} s + v_1^T P_1 v_1 + v_2^T P_2 v_2 = \min \]  \hspace{1cm} \text{(103)}

The participation of the signals in the minimization condition takes automatically care of the datum definition problem at the second epoch.
with respect to the first. The two datums are brought in relative
optimal fit and any change from this optimal situation increases the
term $s_1 K s$, while leaving the two other terms $v_1 p_1 v_1$, $v_2 p_2 v_2$ invari-
ant. The normal equations

$$
\left[ A_1^T P_1 A_1 + A_2^T (A_2 K A_2^T + P_2^{-1})^{-1} A_2 \right] \tilde{x} = 0
$$

$$
= A_1^T P_1 1_1 + A_2^T (A_2 K A_2^T + P_2^{-1})^{-1} 1_2
$$

(104)

still have a rank defect corresponding to the lack of datum definition
at the first epoch. Removing this defect in the usual way, the signals
$s$ are computed from the estimates $\tilde{x}$ by means of

$$
s = K A_2^T (A_2 K A_2^T + P_2^{-1})^{-1} (1_2 - A_2 \tilde{x})
$$

(105)

The stochastic approach finds application also when the comparison is
not between two epochs, but observations performed at various epochs
are to be analyzed simultaneously. Selecting a reference epoch $t_0$,
which is not necessarily one of the observation epochs, we let $x$ be
the vector of corrections to station coordinates at $t_0$ and $s$ the vec-
tor of displacements from $t_0$ to the particular epoch of each observa-
tion. In this case correlations must be assumed for the displacements
not only at different positions but also at different epochs. The ma-
trix $K$ is computed from a displacement covariance function which in-
volves time in addition to position. An example of such an approach
can be found in Kaniugleser (1983).

The only datum defect in this case is the one referring to the datum
definition at the reference epoch $t_0$. The further motion of the refer-
ence frame, i.e., its position at any other epoch $t$, is taken care by
the stochastic assumption that the signals have zero means.

The case of three-dimensional deformable networks is the most compli-
cated one. Two types of signals are present. The displacement signals
depend on the displacement function of space and time and refer to the
geometric deformation of the crust. Furthermore, the same signals as
in the case of rigid 30 networks are present, the only difference
being that they now depend on a gravity potential function which is a
function of both space and time. Disturbances in longitude, latitude,
gravity, etc., are now different not only at different points but also
for different epochs at the same point. When a stochastic approach is
followed the relevant covariance function of the gravity potential
must involve time in addition to position.

When comparisons of 30 networks at two epochs are made the observation
equations have the form

$$
l_1 = A_1 x + G_1 s_1 + v_1
$$

(106)

$$
l_2 = A_2 x + A_2 s + G_2 s_2 + v_2
$$

(107)

considering only geometric observations for simplicity. $s$ are the dis-
placement signals while $s_1$, $s_2$ are the gravity dependent signals which
are different for the two different epochs respectively.

In addition to a completely deterministic approach and a completely
stochastic approach, intermediate approaches are also possible. For
example, one can treat $s$ in a stochastic and $s_1$, $s_2$ in a deterministic way, or treat $s$ as stochastic and $s_1$, $s_2$ stochastic with respect to position and non-stochastic with respect to time, and so on.

The systematic study of deformable three-dimensional networks is the subject of four-dimensional integrated geodesy.

7. ESTIMABILITY PROBLEMS

It is well known that the problem of datum definition is related to the problem of estimability of network related quantities (Grafarend and Schaffrin, 1973, 1976, Grafarend and Richter, 1978, Grafarend et al., 1982). Estimable quantities in classical geodetic networks without signals are equivalent to quantities invariant under coordinate transformations that also leave the relevant available geodetic observations invariant.

Coordinates are typical non-estimable quantities since they refer to an arbitrarily chosen frame. A quantity that can be expressed as a function of the coordinates is estimable if its estimated value computed from the coordinate estimates is independent of the particular choice of coordinate system.

Within the stochastic Gauss-Markov model, estimability is defined as the property of a quantity to have an unbiased estimate. However estimability is not a stochastic property. It depends only on how the observables and the quantity in question are related to the coordinates.

If we assume that observations without errors were available, the estimate of an estimable quantity computed on the basis of errorless observations would coincide with its true value. This associates estimability with determinability, i.e., possibility to determine true values from true values of the observed quantities. Of course this is impossible in reality because of the unavoidable addition of random errors. The term identifiability is sometimes used with the context very similar to the one of determinability used here.

A quantity is estimable on the basis of observations with zero mean errors, if it is determinable on the basis of errorless observations.

For a scalar quantity $q$ which is a linear function of coordinates

$$ q = b^T x $$

(108)

to be estimable, the estimate

$$ \hat{q} = b^T \hat{x} $$

(109)

must be unbiased for any solution $\hat{x}$ of the consistent normal equations

$$ (A^T P A) \hat{x} = A^T P 1 $$

(110)

corresponding to the usual least squares adjustment without signals of equation (44). Any such solution can be represented by

$$ x = (A^T P A)^{-} A^T P 1 $$

(111)

where $R^{-}$ denotes the simple generalized inverse of any matrix $R$, i.e., a matrix satisfying $R R^{-} R = R$. Replacing from (111) into (109) we obtain
\[ \hat{q} = b^T (A^T P A)^{-1} A^T P 1 \]  \hspace{1cm} (112)

Recalling that
\[ E(1) = A x \]  \hspace{1cm} (113)

and taking into account equation (108) the estimability condition
\[ E(q) = q \]  \hspace{1cm} (114)

becomes
\[ b^T (A^T P A)^{-1} A^T P A x = b^T x \]  \hspace{1cm} (115)

Since the particular values in \( x \) play no role, provided some singular cases are excluded, a necessary and sufficient condition for estimability is
\[ b^T (A^T P A)^{-1} A^T P A = b^T \]  \hspace{1cm} (116)

for any generalized inverse. In particular the obvious estimability of the observed quantities themselves gives
\[ A (A^T P A)^{-1} A^T P A = A \]  \hspace{1cm} (117)

With the help of equation (117) which can also be rigorously proved (Rao, 1973, p. 26) it is easy to show that
\[ b^T (A^T P A)^{-1} A^T P A = b^T \iff b^T A^T A = b^T \]  \hspace{1cm} (118)

Indeed from (117)
\[ b^T A^T A = b^T \implies b^T A^T A (A^T P A)^{-1} A^T P A = b^T \implies b^T (A^T P A)^{-1} A^T P A = b^T \]  \hspace{1cm} (119)

and from \( A A^T A = A \)
\[ b^T (A^T P A)^{-1} A^T P A = b^T \implies b^T (A^T P A)^{-1} A^T P A A = b^T A^T A \implies b^T A^T A = b^T. \]  \hspace{1cm} (120)

With the help of (118) the estimability condition (116) can be replaced by the equivalent simpler one
\[ b^T A^T A = b^T \]  \hspace{1cm} (121)

The estimability condition (121) is independent of the stochastic part of the model. It depends only on the matrix \( A \) which relates the observables with the unknown parameters.

These results can be directly applied to the case of geodetic networks with signals when the deterministic approach is followed. Estimability is now related with quantities which are functions of coordinates, of signals, or of both coordinates and signals. For example, consider the case of a 3D network where only distances, horizontal angles and vertical angles have been observed. Distances between any two network points or angles formulated by any three points are examples of estimable quantities which depend only on coordinates. The angle between the directions of the vertical at any two points is an estimable quantity depending only on the signals. The angle between the direction of the vertical at one point and the line joining any two other points is an estimable quantity depending on both coordinates and signals.
If \( q \) is any scalar quantity which is a linear function of both coordinates and signals

\[
q = b^\top x + c^\top s
\]  

(122)

the estimability condition for \( q \) analogous to (121) can directly be written by replacing the matrix \( A \) with the new design matrix \([A \ G]\)

appearing in equation (7)

\[
[b^\top c^\top] \ [A \ G]^{-1} [A \ G] = [b^\top c^\top]
\]  

(123)

The extension of the estimability results to the case where the stochastic approach is followed, is not so straightforward. As a consequence of the stochastic character of the signals \( s \) the previous results hold now only for the estimability of quantities which are functions of \( x \) only. In order to obtain estimability conditions for functions of the signals too, we introduce the following device.

Consider the fictitious case where the original signals \( s^a \) are non-stochastic unknown parameters while estimates \( s' \) of \( s^a \) are available as the result of "observations"

\[
s' = s^a + v'
\]  

(124)

where \( v' \) are zero mean errors with covariance matrix \( K \). Since \( s^a = s^o + s \)

we obtain the observation equations

\[
l_s = s' - s^o = s + v'
\]  

(125)

with \( E(l_s) = s \) as a consequence of \( E(v') = 0 \). Combining equation (125) with equation (5) and choosing \( s^o = s' \) so that \( l_s = 0 \) we obtain the least squares adjustment problem

\[
\begin{bmatrix}
1 \\
0
\end{bmatrix} = \begin{bmatrix} A & G \end{bmatrix} \begin{bmatrix} x \\
0 \\
s \\
v'
\end{bmatrix} + \begin{bmatrix} v \\
v'
\end{bmatrix}^\top \begin{bmatrix} P & 0 \\
0 & K^{-1}
\end{bmatrix} \begin{bmatrix} v \\
v'
\end{bmatrix}
\]

\[
= \min
\]

(126)

where \( s \) is not stochastic but a vector of unknown parameters.

It can be shown that the results of the adjustment problem of equation (126) are identical with the results of the stochastic approach as formulated in equation (24). This establishes a connection with the ideas of adjustment in the presence of prior information in Schaffrin (1983) where a more generalized in many respects point of view is taken.

The above equivalence allows us to obtain estimability conditions by replacing the matrix \( A \) in equation (121) with matrix

\[
\begin{bmatrix}
A & G \\
0 & I
\end{bmatrix}
\]

which is the design matrix in equation (126). The estimability condition for a scalar quantity \( q \) as given by equation (122) becomes

\[
[b^\top c^\top] \ [A \ G]^{-1} [A \ G] = [b^\top c^\top]
\]  

(127)

Looking at equations (125) and (126) it is realized that the estimable
quantities are those which are invariant under coordinate transformations which leave invariant both the original observations \( s^0 \) as well as the "observations" \( s' = s^0 \). This means that the choice of approximate values for the signals in the stochastic approach, which impose restrictions on the permissible coordinate transformations, enlarges the class of estimable quantities. For example, in a 3D network the original estimable quantities in the deterministic approach are those invariant under translations and rotations of the reference frame. In the stochastic approach the approximate values for longitude and latitude introduce frame orientation and make estimable quantities also those which are invariant under translations only and not under rotations. However this gain in estimable quantities is only formal because it depends on the arbitrary approximate values \( s^0 \) resulting from an underlying approximate model function \( \psi_0 \). Only when \( \psi_0 \) has the stochastic meaning of being the known mean function of the corresponding stochastic process \( \psi \) has this gain in estimability a physical meaning.

The above results can also be extended to the case where additional observations are available which depend either on the same signals (deterministic or stochastic approach) or on other signals (stochastic approach only). In the stochastic approach it is also possible to raise questions about the estimability of other signals not present in the observations but predicted by means of equation (29).

The same can be done in the deterministic approach for any new signals \( s' \) with estimates \( \hat{s}' \) computed on the basis of a function \( \psi \) which is obtained from the estimates \( \hat{s} \) of the signals \( s \) present in the observations by means of an arbitrary interpolation technique. However the estimability of \( s' \) has not the meaning of comparison with their true values, i.e., \( E(s') \neq s' \), but simply the meaning of invariance of the used interpolation technique with respect to the permissible transformations of the reference frame. For an example, where the predicted signals are strain parameters from the comparison of network configurations at two epochs see Dermanis (1984).

8. OTHER OPTIMIZATION PROBLEMS

8.1 General Remarks

The remaining optimization problems are formulated in terms of the elements of the network adjustment problem by interchange between given elements and elements to be found. In the classical adjustment without signals the given elements are the network configuration specified by the design matrix \( A \) and the observational accuracies specified by the error covariance matrix \( \Sigma \) or equivalently by the error weight matrix \( P = \Sigma^{-1} \). The network configuration concept includes not only the relative point positions (configuration in the narrow sense) but also the kind of observations performed (configuration in the wide sense). The element to be found is the accuracy of the coordinate estimates \( \hat{x} \) specified by their covariance matrix \( \Sigma_{\hat{x}} \) or the corresponding weight matrix \( P_{\hat{x}} = \Sigma_{\hat{x}}^{-1} \).

In the First Order Design (configuration problem) the given element is the accuracy of the observations \( P_{\hat{x}} \) and an optimal configuration \( A \) is sought so that the resulting coordinate accuracy is maximized. The optimality condition is the minimization of an appropriate risk function which in addition to the coordinate accuracy \( P_{\hat{x}} \) may involve the cost of the observations which depends on the design matrix \( A \).
In the Second Order Design (weight problem) the given elements are the configuration A and an idealized coordinate accuracy \( P_x \) or \( \Sigma_x \) (criterion matrix). The optimal observation accuracy \( P \) or \( \Sigma \) is to be found in a way that the resulting coordinate accuracy \( P_x \) or \( \Sigma_x \) is as close as possible to its idealized counterpart \( P_x \) or \( \Sigma_x \).

Finally in the Third Order Design (improvement problem) the configuration A and the observational accuracy \( P \) are partly known and their modification is sought by additional network points and/or observations so that the resulting coordinate accuracy \( P_x \) becomes satisfactory.

Turning to the adjustment of geodetic networks with signals, the corresponding elements for the formulation of the optimization problems now are the configuration expressed by the matrices A and S plus the matrix B when additional signal depending observations are available, the accuracy of the observations expressed by \( P \) or \( \Sigma \) plus the signal weight matrix \( K^{-1} \) when the stochastic approach is followed and the resulting accuracy of both coordinates and signals expressed by the covariance matrices \( \Sigma_x \), \( \Sigma_s \) and the cross-covariance matrix \( \Sigma_{xs} \).

For the extension of the formulations of the various optimization problems to the case of geodetic networks with signals, some questions must be first answered concerning the new elements \( \Sigma_s \) and \( K \) which do not appear in the classical case.

The first question relates to the formulation of the risk function. Should the risk function involve only coordinate accuracies \( \Sigma_x \), only signal accuracies \( \Sigma_s \), or both?

The second question relates to the a priori signal covariance matrix \( K \) when the stochastic approach is followed. Should \( K \) be classified together with the error covariance matrix \( \Sigma \), or not?

The answers to the above questions and the formulation of the optimization problems varies according to the adjustment approach followed (deterministic or stochastic) and the type of the network with signals considered (rigid three-dimensional or deformable).

Within the framework of these lecture notes it is only possible to give an outline of the various optimization problems for geodetic networks with signals. The solution of these problems follows in general the same patterns as the solution of the classical optimization problems. In particular, any results obtained in optimization without signals where discrimination is made between coordinates \( x_1 \) and orientation unknowns \( x_2 \) contained in \( x \) apply directly to the optimization problems with signals when the deterministic approach is followed. Simply \( x_1 \), \( x_2 \) should be replaced by \( x \) and \( s \) respectively.

For classical optimization problems we refer to Grafarend (1974), Grafarend et al. (1979) and Schmitt (1982) where further references to the literature can be found. An extension to three-dimensional networks can be found in Grafarend and Krumm (1983).

8.2 First Order Design

In the First Order Design problem, when the deterministic approach is followed, the formulation is directly analogous to the case without signals. The only problem is the proper formulation of the risk function.

In three-dimensional rigid networks the signals (vertical directions)
play a secondary role. The main objective is the determination of station coordinates $x$, while the signals $s$ arise only because observations are used which depend on them. It is therefore reasonable to ignore the accuracy of signal estimates and use a risk function which depends only on the accuracy of coordinate estimates $\Sigma_x$.

Refering back to equations (60) and (66) the formulation of the First Order Design problem could be, for example, the following: Find $A$ and $G$ for given $P$ so that

$$\text{trace}(\Sigma_x) = \min$$

(128)

where

$$\Sigma_x = [A^T P A - A^T P G (G^T P G)^{-1} G^T P A + E^T E]^{-1} - E^T (E E^T E^T)^{-1} E$$

(129)

$E$ being the coefficient matrix of the inner constraints given by equation (63). Of course other risk functions than $\text{trace}(\Sigma_x)$ could be used. As a matter of fact, for three dimensional networks only the matrix $A$ needs to be found since the matrix $G$ is uniquely determined from $A$, a consequence of the presence of signals in observations which correspond to the rows of $A$.

A particular problem is the improvement of coordinate accuracy with the inclusion of additional observations (longitude, latitude, azimuth) depending on the signals. Looking back to equations (67) and (68) azimuths contribute to the matrices $A$ and $G$, longitude and latitude to the new matrix $B$ which is now added to the unknowns of the optimization problem. This problem can be also characterized as a Third Order Design problem when it concerns additional observations in a network already observed. One should be critical of nominal improvement of accuracies in this case because of the possibility of systematic effects from the part of the reductions in the astronomic observations, as we have already discussed.

In the case of deformable networks the parameters of interest are not the coordinates themselves but rather their variations with time as expressed by the displacements which are now classified as signals. Indeed, a geophysicist would be happy to obtain from the geodesist an estimate of a displacement at a certain point of, say, a few centimeters with a millimeter accuracy, without caring too much whether there is an uncertainty of a hundred meters in the position of the particular point. This is why isolated observations such as tiltmeter observations and repeated accurate measurements of baselines and angles not belonging to a geodetic network, give valuable results. The network can be seen in this case not as a means of obtaining the absolute positions of the displaced points, but rather as a means of improving the overall geometric accuracy and controlling systematic effects with the help of the network adjustment. In fact, it is not even the accuracy of the displacements $s$ which is of primary interest here, but rather the accuracy of geodynamically meaningful frame-invariant strain parameters which are computed from the discrete displacements by prediction or interpolation techniques (Dermanis, 1984).

It is therefore reasonable to use a risk function which depends only on the accuracy of the signals $\Sigma_x$. For three-dimensional deformable networks where the signals are not only the displacements but also the ones related to the gravity field, the risk function should depend only on the accuracy of a part of the signals, the part containing the displacements.
When the stochastic approach is followed for the First Order Design problem, the a priori signal covariance matrix $K$ is given in addition to the a priori error covariance matrix $\Sigma$. The matrix $\Sigma$ describes the given accuracy of specific observations carried out with specific instrumentation and is therefore uniquely determined. On the contrary, the matrix $K$ is not uniquely determined since it depends on the chosen covariance function $c(t,t')$ for the underlying unknown function or, with a deterministic interpretation, on the choice of norm for the underlying function.

The a posteriori covariance matrices $\Sigma_X, \Sigma_S$ which enter in the risk function according to the type of network considered, depend on the matrix $K$, i.e., on the choice of covariance function $c(t,t')$. It must be therefore emphasized that the solution to the First Order Design problem depends on the choice of covariance function and it is not unique. A different covariance function leads to a different optimal configuration.

However, in a well designed network, changes in the covariance function within reasonable limits must have small influence on the adjustment results. The main source of information must be the observations and not the stochastic model of the underlying function. In such a case, variations in the covariance function cause small variations in the geometric form of the optimal configuration. Such small variations are of no practical importance anyway. In the configuration problem one really looks on whether some stations will be included or not and not on the specific positions of the network stations. These are rather suggested by the landform in the network area in connection with mutual observability conditions.

Since the stochastic approach allows the prediction of signals other than those present in the observation equations, it is possible to use risk functions depending on the accuracy of predicted signals, provided that these accuracies depend mainly on the accuracy of the observations and not on the chosen covariance function. For example, in deformable networks the risk function may depend on the accuracy of predicted frame-invariant strain parameters and not in the accuracy of the displacements.

8.3 Second Order Design

When the deterministic approach is followed for the treatment of signals, the Second Order Design problem can be formulated in complete analogy with the one in the classical case without signals. Given the configuration of the network find the accuracy of the observations $\Sigma$, or the weight matrix $P$, which leads to accuracies of estimates of all unknown parameters (both coordinates and signs)

$$\begin{bmatrix} \Sigma_X & \Sigma_{XS} \\ \Sigma_{SX} & \Sigma_S \end{bmatrix}$$

which, or the corresponding weight matrix, are as close as possible to some given idealized counterparts, e.g., criterion matrices (see Grafarend and Schaffrin, 1979).

This formulation has the same problems as the choice of a risk function in the First Order Design problem depending on both $\Sigma_X$ and $\Sigma_S$.

In the case of three-dimensional rigid networks, a criterion matrix for the coordinate part only $\Sigma_X$ may be used, letting the signal
accuracies $\Sigma_5$ free to take any possibly large values.

In the case of deformable networks, a criterion matrix can be used for the signal part $\Sigma_5$ only, or for the part of signals referring to displacements for deformable three-dimensional networks.

When the stochastic approach is followed, a formulation of the Second Order Design problem completely analogous to the classical case without signals would require the determination of both $\Sigma$ and $K$. While the matrix $\Sigma$ derived from the solution of the Second Order Design problem may give directions for the choice of instrumentation or observational procedures, a derived matrix $K$ has no similar practical importance. The matrix $K$ must reflect the structure of the unknown function on which the signals depend (especially its spectral-smoothness characteristics) and of course one cannot affect this structure at all.

It is therefore more reasonable to keep the matrix $K$ fixed within the Second Order Design problem and solve only for the matrix $\Sigma$ reflecting the observational accuracies. It must be noted that the results of such an approach depend on the fixed value of $K$, i.e., on the chosen covariance function $c(t,t')$. It is hoped that for a well designed network this dependance will be small and anyway of no practical importance according to the same arguments as those discussed in the First Order Design problem. Indeed one does not change the accuracy of the observations in a continuous sense but he is discretely moving to a different accuracy by choosing, e.g., a higher class instrument.

There is a particular case however, where the derivation of the matrix $K$ within the Second Order Design problem is meaningful. This is the case where in the linearization of the observation equations the approximate signals $s_a$ are not derived from a more or less arbitrary model function $s_0$, but they are estimates obtained from the analysis of independent observations, e.g., deflection of the vertical estimates obtained from the analysis of gravimetric observations.

Formally this case can be treated by including the additional observations in the problem, according to equations (84) and (85). This option however is practically impossible to implement, since a large number of additional observations (large matrix $B$) makes the already difficult to solve Second Order Design problem, even more difficult.

The derived matrix $K$, being in this case the covariance matrix of signal estimates to be determined from observations, gives useful directions on how to perform these observations, e.g., on the accuracy and density of gravity measurements in an area for the prediction of deflections of the vertical at network points.

For the sake of completeness we give one more possible case of a Second Order Design problem. This is to consider the matrix $\Sigma$ fixed and derive only the matrix $K$. For this case to be meaningful, the matrix $K$ must be considered again as the covariance matrix of signal estimates to be obtained by additional observations.

8.4 Third Order Design

The Third Order Design problem of improvement of existing networks appears to be of little importance for geodetic networks with signals by the same arguments that justify its significance for networks without signals (Grafarend, 1974): The majority of classical networks, such as national horizontal and vertical networks of all orders, have already been designed and observed. As far as optimization is
concerned, one can only look for their improvement by additional points and/or observations.

The situation is completely different for geodetic networks with signals, such as three-dimensional networks or networks established for the study of crustal deformations. These are usually small local networks designed for special purposes, where the optimization can naturally precede the establishment of the network and the observations.

As far as existing large networks are concerned, their adjustment in a three-dimensional mode seems to be difficult from the computational point of view in comparison to the classical separation in two parts, a two-dimensional on the ellipsoid and an one-dimensional vertical one. On the other hand, the order of magnitude of the distances in large networks does not allow the introduction of meaningful correlations between the corresponding displacements. Any study of deformation must be therefore carried out in the standard way of comparison after independent adjustments at different epochs.

One particular situation where Third Order Design optimization might be useful is in the case of deformable networks where at some epoch it is realized that displacement accuracies are weak for some part of the network. Since it is impossible to go back in time and reobserve, the only thing to do is to look for the improvement of the accuracies by additional observations or additional points in future epochs.

Appendix: OBSERVATION EQUATIONS OF THREE-DIMENSIONAL NETWORKS

We give the observation equations and their linearization in a somewhat compact form. More detailed expressions can be found in the literature, see, e.g., Grafarend (1978, 1981), Hein (1981a).

The following notation is used:

- \( i \): station point
- \( j \): target point
- \( \mathbf{r}_i \): position vector of point \( i \) in global frame
- \( \mathbf{r}_{ij} \): position vector of \( j \) with respect to local frame at \( i \)
- \( \mathbf{R}(\alpha_i, \beta_i) \): rotation matrix relating local frame at \( i \) with global frame
- \( \alpha_i, \beta_i \): angles determining the direction of the vertical at \( i \) with respect to global frame
- \( s_{ij} \): distance between points \( i \) and \( j \)
- \( \delta_{ij} \): horizontal direction of \( j \) observed from \( i \)
- \( A_{ij} \): azimuth of \( j \) from \( i \)
- \( Z_{ij} \): zenith angle of \( j \) from \( i \)
- \( \theta_i, \lambda, \phi \): orientation unknown at \( i \), astronomic longitude and latitude
- \( g \): gravity
- \( W, U, \lambda, \varphi, \gamma \): a normal counterparts of \( W, \lambda, \phi, g, \theta \) respectively.

The global cartesian frame is any frame common for all network points and not necessarily the conventional geocentric frame. The local cartesian frame has its 3rd axis in the direction of the local vertical and its 2nd axis coplanar with the vertical and the 3rd axis of the...
global frame. The relation between the two frames is
\[ r^*_i = R(\alpha_i, \beta_i) (r_j - r_i) \]  \hspace{1cm} (A1)

**Nonlinear Observation Equations of Geometric Type:**

- **distance** \[ s_{ij} = \sqrt{(x_j^* - x_i)^2 + (y_j^* - y_i)^2 + (z_j^* - z_i)^2} \]  \hspace{1cm} (A2)
- **azimuth** \[ A_{ij} = \arctan(x^*_{ij} / y^*_{ij}) \]  \hspace{1cm} (A3)
- **horizontal direction** \[ \delta_{ij} = \arctan(x^*_{ij} / y^*_{ij}) - \theta_i \]  \hspace{1cm} (A4)
- **zenith angle** \[ Z_{ij} = \arctan(\sqrt{(x^*_{ij})^2 + (y^*_{ij})^2}) - \frac{z^*_{ij}}{s^*_{ij}} \]  \hspace{1cm} (A5)

**Linearization:**
\[ dr^*_i = R (dr_j - dr_i) + R_\alpha (r_j - r_i) \, d\alpha + R_\beta (r_j - r_i) \, d\beta \]  \hspace{1cm} (A6)

\[ R_\alpha = \frac{\partial R}{\partial \alpha}, \quad R_\beta = \frac{\partial R}{\partial \beta} \]  \hspace{1cm} (A7)

\[ \delta r_i = r_i - r_i^*, \quad \delta r_j = r_j - r_j^* \]  \hspace{1cm} (A8)

\[ \delta \alpha = \alpha - \alpha^*, \quad \delta \beta = \beta - \beta^*, \quad \delta \theta = \theta - \theta^* \]  \hspace{1cm} (A9)

\[ s_{ij} - s_{ij}^* = \begin{bmatrix} x^*_{ij} - x_i^* \\ y^*_{ij} - y_i^* \\ z^*_{ij} - z_i^* \end{bmatrix} \begin{bmatrix} s_{ij} \\ s_{ij} \\ s_{ij} \end{bmatrix} \]  \hspace{1cm} (A10)

\[ \delta_{ij} - \delta_{ij}^* = [a^T R]^* (\delta r_j - \delta r_i) - \delta \theta_i + [a^T R_\alpha (r_j - r_i)] a^T R_\beta (r_j - r_i) \]  \hspace{1cm} (A11)

\[ a^T = \begin{bmatrix} \frac{y^*_{ij}}{(p^*_{ij})^2} \\ \frac{x^*_{ij}}{(p^*_{ij})^2} \\ 0 \end{bmatrix} \]  \hspace{1cm} (A12)

\[ A_{ij} - a_{ij} = [a^T R]^* (\delta r_j - \delta r_i) + [a^T R_\alpha (r_j - r_i)] a^T R_\beta (r_j - r_i) \]  \hspace{1cm} (A13)

\[ z_{ij} - z_{ij}^* = [b^T R]^* (\delta r_j - \delta r_i) + [b^T R_\alpha (r_j - r_i)] b^T R_\beta (r_j - r_i) \]  \hspace{1cm} (A14)

\[ b^T = \begin{bmatrix} x^*_{ij} & z^*_{ij} & y^*_{ij} \\ z^*_{ij} & s^*_{ij} & 0 \\ p^*_{ij} & s^*_{ij} & 0 \end{bmatrix} \]  \hspace{1cm} (A15)

In the case where the global frame is the conventional geocentric frame \((\alpha = \lambda, \beta = \phi)\) we have...
\[ R = R(\Lambda, \phi) = R_1(90^\circ - \phi) \quad R_3(90^\circ + \Lambda), \quad R_\Lambda = \frac{2}{\partial \Lambda} \quad R, \quad R_\phi = \frac{2}{\partial \phi} \quad R \]  

(A16)

The signals are
\[ \delta \alpha = \Delta \Lambda = \Lambda - \lambda \quad \delta \beta = \Delta \phi = \phi - \phi \quad \text{(anomalies)} \]  

(A17)

\[ \delta \Lambda_0 = \Lambda_0 - \lambda \quad \delta \phi_0 = \phi_0 - \phi \quad \text{(disturbances)} \]  

(A18)

\[
\begin{bmatrix}
\Delta \Lambda \\
\Delta \phi
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial }{\partial y}
\end{bmatrix}_0 \cdot \begin{bmatrix}
\delta r + \begin{bmatrix}
\delta \Lambda_0 \\
\delta \phi_0
\end{bmatrix}
\end{bmatrix}
\]  

(A19)

where the term \[ \begin{bmatrix}
\begin{array}{c}
\delta \lambda \\
\delta \phi
\end{array}
\end{bmatrix} \] is contained in the Jacobian matrix

\[
\begin{bmatrix}
\frac{\partial }{\partial y}
\end{bmatrix}_0 =
\begin{bmatrix}
-\gamma \cos \phi & 0 & 0 \\
0 & -\gamma & 0 \\
0 & 0 & -1
\end{bmatrix}
\]  

(A20)

Additional Observations (linearized):

astronomic longitude \( \Lambda \), latitude \( \phi \) and gravity \( g \)

\[
\begin{bmatrix}
\Lambda - \lambda_0 \\
\phi - \phi_0 \\
g - \gamma_0
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial }{\partial y}
\end{bmatrix}_0 \cdot \begin{bmatrix}
\delta r + \begin{bmatrix}
\delta \Lambda_0 \\
\delta \phi_0
\end{bmatrix}
\end{bmatrix}
\]  

(A21)

potential differences \( \Delta W_{ij} = W_j - W_i \) (from levelling)

\[
\Delta W_{ij} - \Delta U_{ij} = \begin{bmatrix} \gamma \\ \tau_j \end{bmatrix}_0 \begin{bmatrix} \delta r_j + T_{0j} - T_{0i} \\ \delta \tau_i \\
\end{bmatrix}
\]  

(A22)

astronomic azimuth \( A_{ij} \)

\[
A_{ij} - a_{ij} = \left[ a^T R \right] \cdot \left[ \delta r_j - \delta r_i \right] + \left[ a^T R_\Lambda \left( \mathbf{r}_j - \mathbf{r}_i \right) \right] \cdot a^T R_\phi \left( \mathbf{r}_j - \mathbf{r}_i \right) \times
\]  

\[
\left[ \begin{bmatrix}
\frac{\partial }{\partial y}
\end{bmatrix}_0 \cdot \begin{bmatrix}
\delta r_j + T_{0j} - T_{0i} \\
\delta \tau_i \\
\end{bmatrix}
\end{bmatrix}
\]  

(A23)

The additional types of signals are \( T_0 \) and \( \delta g_0 \).

The terms containing the matrix \( \frac{\partial }{\partial y} \) are very small and can be neglected in the observation equations for longitude, latitude and azimuth. This matrix is a representation of the Bruns transformation (Grafarend, 1980) from geometry to gravity space (see also Moritz, 1980, p. 232, Rummel and Teunissen, 1982 and Engler et al., 1981).

The approximate coordinates used here are supposed to be close to the true ones but otherwise arbitrary. When \( \Lambda, \phi, \psi \) or \( \lambda, \phi, g \) are all observed at network points, one can use as approximate coordinates those obtained through a telluroid mapping. In this case the approximate points are points on the telluroid and there is a consistency with the geodesic boundary value problem. For a discussion see Grafarend (1981).
REFERENCES

Blaha G (1971) Inner adjustment constraints with emphasis on range observations. Department of Geodetic Science Report no. 148, The Ohio State University, Columbus Ohio.


Dermanis A (1984) The role of frame definitions in the geodetic determination of crustal deformation parameters, to be published.


Moritz H (1978a) The operational approach to physical geodesy. Department of Geodetic Science Report no. 277, The Ohio State University, Columbus, Ohio
Moritz H (1978b) Statistical foundations of collocation. Department of Geodetic Science Report no. 272, The Ohio State University, Columbus, Ohio
Schaffrin B (1983) Model choice and adjustment techniques in the presence of prior information. Department of Geodetic Science and Surveying Report no. 351, The Ohio State University, Columbus, Ohio.
Tscherning CC (1978) A users guide to geopotential approximation by stepwise collocation on the RC 4000 - computer. Danish Geodetic Institute, Med. no. 53, Kopenhagen
Optimization and Design of Geodetic Networks

Edited by E. W. Grafarend and F. Sansò

With Contributions by
B. Benciolini  F. Crosilla  P. A. Cross  D. Delikaraoglou
A. Dermanis  D. Fritsch  E. W. Grafarend  K. R. Koch
F. W. Krumm  F. Sansò  B. Schaffrin  G. Schmitt
W. D. Schuh  H. Sünkel  P. J. G. Teunissen

With 139 Figures

Springer-Verlag
Berlin Heidelberg New York Tokyo