"Mathematische Methoden der Geodäsie"
Mathematisches Forschunginstitut Oberwolfach
01.10 - 07.10.1995

The Nonlinear and Space-time Geodetic Datum Problem

Athanasios Dermanis

Department of Geodesy and Surveying
University of Thessaloniki

CONTENTS

A. The nonlinear datum problem

1. Introduction

The datum problem or zero order design problem (Grafarend, 1974) arising in the adjustment of observations related to geodetic networks has received considerable attention since the pioneering work of Meissl (1965, 1969) and its popularization by Blaha (1971). The nature of the problem has been clarified in two important papers by Grafarend & Schaffrin (1974, 1976) while the relation of various solutions to Meissl's "inner" solution has been established with the introduction of the famous Baarda *S*-transformation (Baarda, 1973, see also van Mierlo, 1980, Koch, 1982). It might therefore seem that every aspect of this problem has been fully investigated and well understood for quite some time now, though some recent work (Xu, 1995) might point to the contrary. However, with the exception of some simple cases, such as that of leveling in a small area, the mathematical treatment of the problem is confined to its linearized version, although it is well understood as a nonlinear problem.
Nonlinear adjustment has been treated in its deterministic aspect by Teunissen (1985, 1990) while Dermanis and Sanso (1995) have looked into the stochastic aspects. In both treatments the datum problem is solved in an implicit way. In the nonlinear least squares adjustment solved by an iteration scheme, the choice of datum depends on the choice of the initial parameter values used for starting the iteration process, as well as, on the principle used for "improving" these parameters. In the nonlinear estimation case, the need to adopt a Bayesian point of view in order to obtain a meaningful estimation independent of the "true" parameter values, also solves the datum problem in an implicit way where the datum choice is hidden in the choice of the prior probability distribution of the parameters.
The solution to the nonlinear datum problem presented here is based in the concept of the *S*-transformation and it has the form of such a similarity (or rigid) transformation with parameters which come from the solution of a system of nonlinear equations.
Another important aspect of the geodetic datum problem in its linear form is its relation to generalized inverses of matrices (linear operators) which led Bjerhammar (1951) to an independent introduction of the Moore-Penrose generalized inverse, later that Moore but before Penrose. The question arises if the nonlinear version of the geodetic datum problem bears a similar relation to some type of generalized inverses of nonlinear operators. An attempt will be made to look mainly into the geometric aspects of such nonlinear generalized inverses, although the building of a concrete mathematical theory requires a more rigorous treatment which is beyond our present scope. We shall base our investigation on the representation theory of various types of linear generalized inverses which has been introduced by Takos (1976) from an algebraic point of view and especially by Teunisssen (1985) from a geometric point of view.
The datum problem is always a part of the problem of the adjustment of redundant observations which are related to a set of parameters (coordinates in the geodetic case) which in fact cannot be determined from the available observations. The reason is that the information contained in the angle and distance (angle) observations relates only to the shape and size (shape) of the network, while coordinates relate in addition to its position (position and size) with respect to a certain reference frame. The problem of placing the network in relation to a given reference frame can be also seen from the inverse point of view of placing a reference frame in relation to a given (i.e. physically existing) network. This choice of reference frame (or datum in geodetic terminology) poses a problem, the "datum problem", which must be solved in an arbitrary but consistent way based in the introduction of additional information not contained in the observations.
The usual approach to the description of the adjustment and choice of datum problem is to consider the n observed parameters ($n>r$) as the coordinates of a linear space Y called the ***observation*** space, in which the r-dimensional manifold M modeling the physical system, called the ***model manifold***, is lying. (We confine ourselves to the case of discrete and finite observations.) The model manifold is described as the range of a nonlinear operator f from an m-dimensional ***parameter space*** X ($m>r$) into the observation space Y. The mapping f is established by the mathematical equations which relate all observables to the unknown parameters. In a "normal" situation, which is almost never the case in geodesy, the number of parameters and the dimension of the model manifold are equal ($m=r$), in which case the restriction $f_{|\mathcal{M}}$ of f to M has an inverse which may serve as a coordinate mapping from M to $X=R^m$. In other words the chosen parameters can serve as a particular system of coordinates on the model manifold and the only problem to be solved is the adjustment problem. In the linear case the rank of f is $r=m$ and the corresponding model is called a ***full-rank model***.
The unavoidable observational errors are added to the true values of the observables (which correspond to a point on the model manifold) and give observations corresponding to a known point outside the manifold. The ***adjustment problem*** can be simply defined as the problem of finding an optimal way to

"return to the model manifold". In the case where $m>r$ r (***model without full rank*** in the linear case) the determination of parameter values is not trivial anymore, because there is an infinite set of parameter values which *f* maps on the same manifold point corresponding to the adjusted observations. The ***datum problem*** is exactly the problem of choosing one out of all possible parameter sets.
We must point out that the above point of view is not coordinate-free, since it depends on the choice of a specific set of m unknown parameters. This is perhaps of little concern to geodesy where the choice of coordinates as parameters is imposing itself as a matter of convenience. It is possible however to establish a theory where both the adjustment and the datum problem are treated in a coordinate-free way in the spirit of modern differential geometry.
A more in depth introduction to the (nonlinear) datum problem, especially in relation to the modeling problem can be found in Dermanis (1991).

2. Generalized inverses of nonlinear operators

Let *X*, *Y* be finite dimensional spaces and $s:X\to Y$ a (nonlinear) mapping from *X* to *Y*. The question arises whether it is possible to define, under some appropriate conditions, a class of mappings $g:Y\to X$ which can serve as generalized inverses of the mapping *f*. The similar theory developed for linear mappings may serve as a guide to a certain extend, although the non-linearity of *f* poses problems which do not allow for a direct extension of the existing theory.

Definition: A partition of a set *A* into disjoint sets A_t, $A=\bigcup_t A_t$ is called a ***fibering*** of *A* and its elements A_t are called ***fibers***. Such a partition can arise from an equivalence relation on *A*, each fiber consisting of equivalent elements of *A*. (Loomis & Sternberg, 1980.)

A mapping *f* defined on *X* gives rise to a fibering $\mathscr{F}$ of *X* through the equivalence relation $x_1\sim x_2$ if $f(x_1)=f(x_2)$. To every element $y\in R(f)$ in the range $R(f)$ of *f* corresponds a fiber $F_y\in\mathscr{F}$ defined by

$$F_y=\{x\in X\,|\,f(x)=y\}\,. \tag{1}$$

The function *f* gives rise to a bijection $\bar{f}:R(f)\to\mathscr{F}:y\to F_y$. Obviously for every $x\in X$, $x\in F_{f(x)}$.

Definition: The mapping $\pi:X\to\mathscr{F}:x\to F_{f(x)}$ is called the ***projection mapping*** from *X* to the fibering $\mathscr{F}$. It holds that $\pi=\bar{f}\circ f$.

Definition: A ***section*** of a fibering $\mathscr{F}$ of *X* is the range $S=R(s)$ of a mapping $s:\mathscr{F}\to X$ such that $\pi\circ s=id_{\mathscr{F}}$, where $id_{\mathscr{F}}$ is the identity mapping in $\mathscr{F}$.

We assume for the moment that *f* is defined on the whole of *X* it is continuous and has a differential mapping df_x at each *x* which is a linear mapping from the tangent vector space T_x at *x* to the tangent vector space $T_{f(x)}$ at $f(x)$. The mapping df_x is defined by

$$df_x:T_x\to T_{f(x)}:v\to u\,,\qquad u(f\circ\phi)=v(\phi)\quad\text{for every}\quad\phi:X\to R \tag{2}$$

(where the tangent vectors *v* and *u* are visualized as "directional derivatives" acting on functions).
We further assume that df_x has constant rank over *X*, $\operatorname{rank}(df_x)=r\le\min(n,m)$, where $m=\dim X$ and $n=\dim Y$. As a consequence $R(f)$ is an *r*-dimensional submanifold of *Y*.
We now come to the possibility of defining a generalized inverse *g* of the mapping *f* by a similar way as in the linear case.

Definition: A mapping $g:Y\to X$ is called a ***generalized inverse*** of a given mapping $f:X\to Y$ when

(G1) $f \circ g \circ f = f$. (3)

This means that for every $x \in X$, $(f \circ g \circ f)(x) = f(x)$, so that for every $y = f(x) \in R(f)$ it holds that $(f \circ g)(y) = f\big(g(y)\big) = y$. As a consequence $g(y) \in F_y$, i.e. g must map every element of $R(f)$ into an element of its corresponding fiber. In other words g maps $R(f)$ onto a section C of the fibering $\mathscr{F}$ where $C \subset R(g)$. The restriction of g on $R(f)$ is a bijection between $R(f)$ and C which are both r-dimensional manifolds.
The relation (G1) satisfied by the generalized inverse g implies a corresponding relation between the differentials df_x and $dg_{f(x)}$ of f and g respectively:

$$df_x \circ dg_{f(x)} \circ df_x = df_x \,. \tag{4}$$

Lemma: $R(f \circ g) = R(f)$.

Proof: $y \in R(f) \;\Rightarrow\; \exists x \in X_{\cdot} :\; y = f(x) = (f \circ g \circ f)(x) = (f \circ g)\big(f(x)\big) = (f \circ g)(y) \in R(f \circ g)$
$\Rightarrow \quad R(f) \subset R(f \circ g)$
which combined with the obvious relation $R(f \circ g) \subset R(f)$ implies that in fact $R(f \circ g) = R(f)$. □

Assuming that g has a constant rank $r_g = \mathrm{rank}(dg_y)$ for every y, application of the property of the rank of a linear map composition $r(A \circ B) \le \min\big(r(A), r(B)\big)$ to (3), gives $r \le \min(r, r_g) \le r_g$ so that

$$r \le r_g \le \min(n, m) \tag{5}$$

Once C is given g is defined on $R(f) \subset Y$. The question is how this particular g is extended outside $R(f)$. A step in this direction is to note that as a consequence of (G1) both $q \equiv g \circ f$ and $p \equiv f \circ g$ are idempotent mappings ($q^2 = q$, $p^2 = p$)

$$f \circ g \circ f = f \quad \Rightarrow \quad (g \circ f) \circ (g \circ f) = g \circ f \quad \text{and} \quad (f \circ g) \circ (f \circ g) = f \circ g \tag{6}$$

which can be considered as "nonlinear projections"

$$q = g \circ f : X \to R(g \circ f) \subset R(g) \subset X \,, \qquad p = f \circ g : Y \to R(f \circ g) = R(g) \subset Y \,. \tag{7}$$

We have set $R(g \circ f) \subset R(g)$ which is obvious, and $R(f \circ g) = R(f)$ according to the previous lemma.
Elements belonging to the range of idempotent mappings are invariant under the mapping. Idempotent mappings induce also a fibering of the space they act on. We denote with $\mathscr{P}$, $\mathscr{Q}$ the fiberings induced by p and q, respectively, with corresponding elements $P_y \in R(p) = R(f)$ and $Q_x \in R(q) \subset R(g)$.
The mapping g gives rise to a fibering $\mathscr{G}$ of Y with elements $G_x = \{y \in Y \,|\, g(y) = x\}$. Obviously for every element y of Y it holds that $y \in G_{g(y)}$. If $y \in R(f)$ then $y \in G_x$, where $\{x\} = F_y \cap C$.

Definition: Let $\mathscr{F}$ and $\mathscr{F}'$ be two fiberings of the same set A. $\mathscr{F}'$ is called a ***refinement*** of $\mathscr{F}$ if every fiber F'_t of $\mathscr{F}'$ is a subset of some fiber F_s of $\mathscr{F}$: $\forall F'_t \in \mathscr{F}' :\; \exists F_s \in \mathscr{F}$ such that $F'_t \subset F_s$.

Lemma: For the fiberings of the space Y it holds that the fibering $\mathscr{G}$ induced by g is a refinement of the fibering $\mathscr{P}$ induced by $p = f \circ g$.

Proof: For a fixed $z \in Y$, $G_{g(z)} \in G$ and $P_{(f \circ g)(z)} \in P$. For every $y \in G_{g(z)}$ it holds that $g(y) = g(z)$ and $f\big(g(y)\big) = f\big(g(z)\big)$, or $(f \circ g)(y) = (f \circ g)(z)$ and $y \in P_{(f \circ g)(z)}$. Consequently $G_{g(z)} \subset P_{(f \circ g)(z)}$ and $\mathscr{G}$ is a refinement of $\mathscr{P}$. □

Proposition: A generalized inverse *g* of *f* is uniquely defined if the following are specified:

(1) The section $C=g(R(f))\subset R(g)$ of the fibering $\mathscr{F}$ of *X* induced by *f*.
(2) The fibering $\mathscr{P}$ to be induced by $p=f\circ g$, i.e. all the fibers P_y corresponding to every $y\in R(f)$.
(3) The refinement of the fibering $\mathscr{P}$ by the fibering $\mathscr{G}$, i.e. for every fiber $P_y\in\mathscr{P}$ its own fibering with members G_x corresponding to all $x\in F_y$.

A special case of interest is the case of a minimal rank generalized inverse *g* where $r_g=r$. In this case $r_g=\dim(R(g))=r=\dim(R(f))$ which combined with $C\subset R(g)$ implies that $C=R(g)$.

Lemma: When $r_g=r$ the fibering $\mathscr{G}$ induced by *g* coincides with the fibering $\mathscr{P}$ induced by $p=f\circ g$.

Proof: Since $C=R(g)$ the restriction $g_{|R(f)}$ of *g* to the range of *f*, is in this case a bijection between $R(f)$ and $R(g)$. For every $z\notin R(f)$ it holds that $g(z)\in R(g)=C$ and since *C* is a section of $\mathscr{F}$, $\{g(z)\}=C\cap F_y$ for a unique fiber F_y corresponding to a unique $y\in R(f)$.
Therefore $y=f(g(z))=(f\circ g)(z)=p(z)$ and $z\in P_y=P_{(f\circ g)(z)}$. Consequently the fibering $\mathscr{G}$ induced by *g* coincides with the fibering $\mathscr{P}$ induced by $p=f\circ g$. □

For the particular choice $r_g=r$, the requirement (3) in the above proposition can be relaxed. The same simplified situation can arise in a different way by requiring that in addition to *g* being a generalized inverse of *f*, at the same time *f* is a generalized inverse of g. Repeating property (G1) by interchanging the roles of *f* and *g* leads to the following definition.

Definition: A generalized inverse *g* of a given mapping *f* is called a ***reflexive*** generalized inverse, if in addition to (G1) satisfies:

(G2) $\qquad g\circ f\circ g=g$. (8)

Lemma: a generalized inverse *g* of *f* is a reflexive generalized inverse of *f* if and only if $r_g=r$.

Proof: *Necessity*: An immediate consequence of *g* being a generalized inverse of *f* is that $R(g\circ f)=R(g)$ and $r_g\le\min(r,r_g)$ so that $r_g\le r\le\min(n,m)$ which combined with $r\le r_g$ from (5) implies that $r_g=r$.
Sufficiency: If $r=r_g$ then $\dim(R(g))=\dim(R(f))=\dim(C)$ and $C=R(g)$. If $x\in C$ then $\{x\}=C\cap F_{f(x)}$ and $g(f(x))=x$. Let *y* be an arbitrary element of *Y*; then $g(y)\in R(g)=C$ and setting $g(y)$ in the place of the previous *x* we obtain $g(f(g(y)))=g(y)$ or $(g\circ f\circ g)(y)=g(y)$ and since *y* is arbitrary $g\circ f\circ g=g$. □

Proposition: A reflexive generalized inverse *g* of *f* is uniquely defined if the following are specified:

(1) The section $C=g(R(f))=R(g)$ of the fibering $\mathscr{F}$ of *X* induced by *f*.
(2) The fibering $\mathscr{P}$ to be induced by $p=f\circ g$, i.e. all the fibers P_y corresponding to every $y\in R(f)$.

The question what are the fibers G_x of a reflexive generalized inverse is answered by the following lemma.

Lemma: For a reflexive generalized inverse *g*, the $(f\circ g)$-induced fibers are identical to the *g*-induced fibers, i.e. $\mathscr{P}=\mathscr{G}$.

Proof: Let $z\in P_{(f\circ g)(z)}=P_y$ where $y=(f\circ g)(z)\in R(f)$.

Then $x=g(z)=(g\circ f\circ g)(z)=g\big((f\circ g)(z)\big)=g(y) \;\Rightarrow\; y\in G_x=G_{g(z)}$. Of course $z\in G_{g(z)}$ which means that every $z\in P_y$ belongs to the same fiber G_x as y, so that $P_y=G_x$, i.e., $P_{(f\circ g)(z)}=G_{g(z)}$. Therefore the $(f\circ g)$-induced fibers are identical to the g-induced fibers, i.e. $\mathscr{P}=\mathscr{G}$. □

For reflexive generalized inverses we have a possibility of choices corresponding to the choices of the section C of $\mathscr{F}$ and of the fibering $\mathscr{P}$ of Y over the elements of $R(f)$. In analogy to the uniquely defined pseudoinverse of a linear operator we seek two conditions, additional to (G1) and (G2), such that (G3) specifies the fibering $\mathscr{P}$ and (G4) specifies the section C of $\mathscr{F}$.

(G3) - ***Minimum distance generalized inverse***.
Let ρ_Y be the distance function of the metric space Y. For any fixed element $\hat{y}\in R(f)$ define

$$R_{\hat{y}}=\{z\in Y\,|\,\rho_Y(z,\hat{y})=\min_{y\in R(f)}\rho_Y(z,y)\,. \tag{9}$$

If $R_{\hat{y}}$ are fibers of Y, as $\hat{y}$ varies over $R(f)$, then g is a ***minimum distance*** generalized inverse of f, if $\mathscr{P}=\mathscr{R}$, i.e. when the fibering $\mathscr{P}$ induced by $p=f\circ g$ is identical with the fibering $\mathscr{R}$ with fibers $R_{\hat{y}}$. If Y is an inner product vector space with metric induced by a quadratic norm, we may use the term "***least squares***" in place of "minimum distance".

(G4) - x_0 ***-nearest generalized inverse***. ######
Let ρ_X be the distance function of X, $x_0\in X$ be fixed and let

$$x_{F_y}=\min_{z\in F_y}\rho_X(x_0,z)\,. \tag{10}$$

If the set of all x_{F_y} constitutes a section C of $\mathscr{F}$, then g is a x_0 ***-nearest*** generalized inverse of f if C is taken as the image of $R(f)$ under g, i.e. $C=g\big(R(f)\big)$.

Remark: If X is an inner product vector space with metric induced by the norm, while $\mathbf{x}_0=\mathbf{0}$, we may use the term "minimum norm" in place of "minimum distance". However, in the geodetic case $\mathbf{0}$ is a "prohibited" point, since it corresponds to a network with all its points coinciding!

Definition: The unique reflexive generalized inverse g of a given mapping f which is also a minimum distance and x_0-nearest generalized inverse is called the ***pseudoinverse*** of f.

Various other non-unique generalized inverses can be defined by combination of (G1) with the some of the others properties (G2), (G3) and (G4). The inverses within the same class, i.e. the ones satisfying the same set of properties, may differ in the following aspects:

In the section C when (G4) is not satisfied.
In the fibering $\mathscr{P}$ induced by $p=f\circ g$ when (G3) is not satisfied.
In the refinement $\mathscr{G}$ of $\mathscr{P}$ when (G2) is not satisfied.

In the geodetic case property (G3) is a "must", since it solves the adjustment problem. Reflexivity (G2) is not necessary but convenient. The section C is specified either directly by a set of minimal constraints or (G4) is introduced. (G4) is introduced either directly (pseudoinverse solution) or indirectly by a set of inner constraints which describe the particular section C corresponding to (G4).

Definition: A set of (nonlinear) equations $h(x)=d$ is called a set of ***minimal constraints*** (with respect to f) if the corresponding mapping $h:X\to R^{m-r}$ gives rise to a fibering $\mathscr{H}$ of X such that its member H_d corresponding to d (i.e such that $h(x)=d$ for every $x\in H_d$) is a section of the fibering $\mathscr{F}$ induced by f.

Definition: A set of minimal constraints $h(x)=d$ are called ***inner constraints*** (with respect to a given point x_0 of X) when the fiber H_d coincides with the section of $\mathscr{F}$ specified by the property (G4), i.e.,

when for every $F_y \in \mathscr{F}$ the unique element x such that $F_y \cap H_d = \{x\}$ is the one closest to x_0 among all elements of F_y.

Remark: In the special case that f and its generalize inverse are linear or affine mappings, $R(f)$, the fibers F_y, P_y, G_x, Q_x and the section C are all affine subspaces of the corresponding spaces Y and X. The fibers $\mathscr{F}$, $\mathscr{P}$, $\mathscr{G}$, $\mathscr{Q}$ consist of parallel affine subspaces, i.e. they are equivalent to quotient spaces of Y and X. Each quotient space is uniquely determined by its member passing through zero, i.e. by their "modulo" linear subspace in X or Y. The fiber $\mathscr{F}$, e.g., is equivalent to the quotient space of affine subspaces parallel to the null space of f. Thus the linear of affine generalized inverses are determined by a set of linear subspaces, instead of the fiberings (see Teunissen, 1985). By the way, there is no reason to restrict the class of generalized inverses of a linear operator to be itself linear, as is usually done within the linear theory. A linear operator may well have a non-linear generalized inverse. A geodetic example is the affine generalized inverse obtain implicitly by imposing a set of inhomogeneous linear constrains $\mathbf{Hx}=\mathbf{d}$, $\mathbf{d}\neq\mathbf{0}$, on the solution of the least squares adjustment problem.

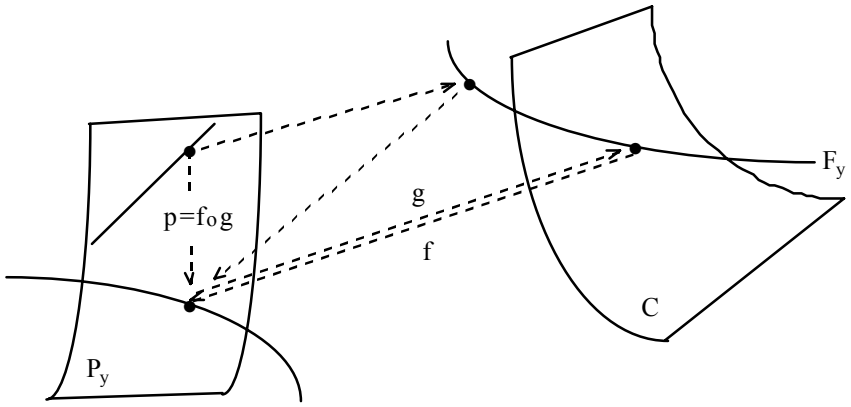

Fig. 1: The geometry of the generalized inverse

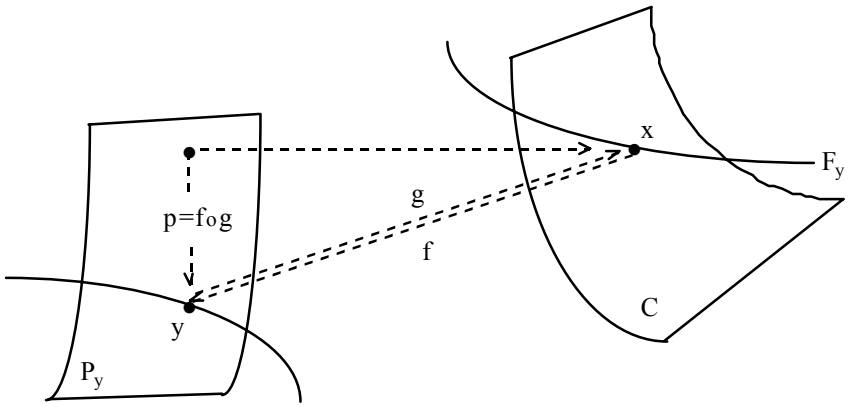

Fig. 2: The geometry of the reflexive generalized inverse

The approach followed above can be characterized as naive and falls sort of a proper mathematical theory of generalized inverses of nonlinear mappings, even in the relatively simple case of finite-dimensional

spaces. The reason is that the assumptions made are too strong to build a general theory on, a fact which will become obvious even in the simple geodetic case. Some of the assumptions that do not generally hold are the following:

- f may not be defined at all points of the space X but only on an open subset $U \subset X$.
- g may not be possible to define at all points of the space Y but only on an open subset $V \subset Y$.
- The minimum distance property (G3) may not suffice for the determination of a fibering of Y or even V. This has to do with the differential geometric properties of $R(f)$ as a submanifold of Y and in particular with its curvature tensor. There might be points in V such that their minimum distance from $R(f)$ does not correspond to a unique point of $R(f)$ but either on a set of discrete points or even a non-discrete subset of $R(f)$. (Think e.g. of the case where $R(f)$ is a ball in Y and consider its center which belongs to any fiber induced by the minimum distance principle.)
- The x_0 -nearest property (G4) may not suffice for the determination of a fibering of X or even U. This has to do with the differential geometric properties of the fibers F_y as a submanifolds of X. There might be one or more fibers F_y on which there are more than one points attaining the minimum distance from x_0 .
- The rank of f (= rank of df_x) may not be the same at all points x of U. The consequence is that $R(f)$ is no more a smooth submanifold of Y and its behavior at singular points (points with $\text{rank}(df_x) < r$) may pose additional problems.

We shall keep these problems in mind when studying the particular case of the mappings f arising in geodetic problems.

3. The geodetic non-linear mapping

The non-linear mapping f in the geodetic case arises from an individual mappings f_k , $k = 1, ..., n$, which map coordinates of points into various types of observables. We consider a geodetic network of N points with cartesian coordinates $\mathbf{x}_i$, $i = 1, ..., N$, where $\mathbf{x}_i$ is a vector of dimension either 2 (plane networks) or 3 (spatial networks). These vectors constitute a coordinate vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} \tag{11}$$

of dimension $2N$ or $3N$, which is considered as an element of the m-dimensional Euclidean space ($m = 2N$ or $m = 3N$), $X = E^{2N} = (\times E^2)^N$ or $X = E^{3N} = (\times E^3)^N$, equipped with the simple inner product $(\mathbf{x}_\alpha, \mathbf{x}_\beta) = \mathbf{x}_\alpha^T \mathbf{x}_\beta$.

The geodetic observables we will examine here can be distinguished into "angular" observables ψ_{ijk} and "distance" observables d_{ik} , which differ with respect to their invariance characteristics under coordinate transformations. Angular observables are invariant under similarity coordinates transformations $\mathbf{x}' = S(\mathbf{x})$ while distance observables are invariant under rigid coordinate transformations $\mathbf{x}' = R(\mathbf{x})$, which are defined point-wise by

$$\mathbf{x}_i' = S(\mathbf{x}_i) = \lambda \mathbf{R}\mathbf{x}_i + \mathbf{t} , \qquad \mathbf{x}_i' = R(\mathbf{x}_i) = \mathbf{R}\mathbf{x}_i + \mathbf{t} . \tag{12}$$

$\mathbf{t}$ is a displacement vector, $\lambda > 0$ is a scale parameter and $\mathbf{R}$ is a proper orthogonal matrix ($\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$ and $|\mathbf{R}| = +1$) depending on a single rotational parameter θ in the 2-dimensional case or on 3 rotational parameters θ_1 , θ_2 , θ_3 in the 3-dimensional case. The observables are functions of the coordinates

$$\psi_{ijk} = \psi(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) = \psi_{ijk}(\mathbf{x}) , \qquad d_{ik} = d(\mathbf{x}_i, \mathbf{x}_k) = d_{ik}(\mathbf{x}) , \tag{13}$$

which have the invariance properties

$$\psi(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)=\psi\big(S(\mathbf{x}_i),S(\mathbf{x}_j),S(\mathbf{x}_k)\big),\qquad d(\mathbf{x}_i,\mathbf{x}_k)=d\big(R(\mathbf{x}_i),R(\mathbf{x}_k)\big), \tag{14}$$

or

$$\psi_{ijk}(\mathbf{x})=\psi_{ijk}\big(S(\mathbf{x})\big),\qquad d_{ik}(\mathbf{x})=d_{ik}\big(R(\mathbf{x})\big). \tag{15}$$

The mapping f which maps the network coordinates $\mathbf{x}$ to the observables $\mathbf{y}=f(\mathbf{x})\in R(f)\subset Y$ consists in general of n_1 angular mappings and n_2 distance mappings ($n_1+n_2=n$) and we can distinguish two cases with respect to its invariance properties. When $n_2=0$ (only angular observations) then

$$f(\mathbf{x})=f\big(S(\mathbf{x})\big), \tag{16}$$

while when $n_2\neq 0$ (observations of distances or both distances and angles) then

$$f(\mathbf{x})=f\big(R(\mathbf{x})\big). \tag{17}$$

The presence of angular observations has as a consequence that f is not defined at every point of X. For an angular observable $\psi_{ijk}=\psi(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)$ to be defined, it is necessary that the relevant network points are distinct, i.e. we must exclude from X points $\mathbf{x}$ with $\mathbf{x}_i=\mathbf{x}_j$, or $\mathbf{x}_j=\mathbf{x}_k$, or $\mathbf{x}_k=\mathbf{x}_i$. Although the same problem does not appear in the case of distance observations, we will stick to the restriction that those network points which are joined by observations must be distinct. Let J be the index set of those pairs (i,k) of point indices corresponding to points joined by observations and let D_{ik} be the diagonal subspace of X defined by

$$D_{ik}=\{\mathbf{x}\in X\,|\,\mathbf{x}_i=\mathbf{x}_k\}. \tag{18}$$

We must exclude all inappropriate diagonal subspaces which leaves as the domain of definition of f the open subset $V\subset X$ defined by its complement V^c with respect to X

$$V^c=\bigcup_{(i,k)\in J}D_{ik}=D. \tag{19}$$

Since all D_{ik} are closed the same holds for $D=V^c$ and V is an open subset of X.

Turning to the range space Y we note that distances are mappings from X to R^+, since they obtain only positive values, while angles are mappings from X to the unit circle S. Thus, strictly speaking, f is a mapping

$$f:V\to(\times S)^{n_1}\times(\times R^+)^{n_2}. \tag{20}$$

However it is standard practice to replace S with an interval of R, our choice being $I=(0,2\pi]$, in which case f becomes a mapping

$$f:V\to(\times I)^{n_1}\times(\times R^+)^{n_2}\subset Y=E^{n_1+n_2}=E^n. \tag{21}$$

Since I does not contain its limit point 0 and the same holds true for R^+, U is an open subset of Y. We have assumed that $Y=E^n$ is R^n equipped with the euclidean inner product, though a more general inner product $(\mathbf{y}_\alpha,\mathbf{y}_\beta)=\mathbf{y}_\alpha^T\mathbf{P}\mathbf{y}_\beta$ may be used, in which case $Y=E^n_{\mathbf{P}}$.

The above restriction of f from a mapping $f:X\to Y$ to a mapping $f:V\to U$, poses some serious problems as far as the properties (G3) and (G4) of generalized inverses is concerned. Now the range $R(f)$ is

restricted from $f(X)$ to $U \cap f(X)$ and further to $U \cap f(V)$. The adjustment problem may have a solution $\hat{\mathbf{y}} \in f(X)$ closest to a given $\mathbf{y} \in U$, such that $\hat{\mathbf{y}} \notin U \cap f(V)$. This means that adjusted observations $\hat{\mathbf{y}}$ may include negative or zero distances, or angles outside the prescribed interval.
Turning to the domain V we must distinguish between the similarity transformation and the rigid transformation case. When distances are observed and $\hat{\mathbf{y}} \in f(X)$, the fibers $F_{\hat{y}}$ may be initially defined with the help of any element $\mathbf{x}' \in X$, such that $\hat{\mathbf{y}} = f(\mathbf{x}')$ by means of

$$F_{\hat{y}} = \{\mathbf{x} \in X | \exists R: \mathbf{x} = R(\mathbf{x}')\}. \tag{22}$$

This means that if $\mathbf{x}' \in V$ the same holds for any $\mathbf{x} = R(\mathbf{x}')$ and therefore $F_{\hat{y}} \subset V$. Thus $F_{\hat{y}}$ is either included in V or lies completely outside V. This means that the problem of finding an element $\hat{\mathbf{x}}$ of a fiber closest to a given element $\mathbf{x}_0 \in V$ has a solution in V provided that $F_{\hat{y}} \subset V$, i.e., that $\hat{\mathbf{y}} \in f(V)$. When distance measurements are not included the similarity transformation S gives rise to fibers

$$F_{\hat{y}} = \{\mathbf{x} \in X | \exists S: \mathbf{x} = S(\mathbf{x}')\}, \tag{23}$$

which are not anymore closed subsets of X, because they "converge" at the zero element $\mathbf{0} \in X$. Although the scale is restricted to positive values $\lambda > 0$, it is possible to consider a sequence $\lambda_i \to 0$, giving rise to a sequence of similarity transformations S_i such that $\mathbf{x}_i = S_i(\mathbf{x}') \in F_{\hat{y}} \to \mathbf{0} \notin F_{\hat{y}}$ (convergence in norm). As a consequence the problem of finding an element $\hat{\mathbf{x}}$ of a fiber closest to a given element $\mathbf{x}_0 \in V$ may have no solution (consider the case where $\mathbf{0}$ is the closest to $\mathbf{x}_0$ element from the closure $\overline{F}_{\hat{y}}$ of the fiber $F_{\hat{y}}$).
These are problems of great concern to the mathematician, but of little or no concern at all to the geodesist. The reason is that in geodetic applications we are confined to a small neighborhood $N_{y^*} \subset U \cap f(V)$ of the true observables $\mathbf{y}^* \in U \cap f(V)$, such that the observations $\mathbf{y}$ and the adjusted observations $\hat{\mathbf{y}}$ also belong to N_{y^*}. In the same way it is possible to choose an element $\mathbf{x}_0 \in V$ such that $F_{\hat{y}}$ crosses a small neighborhood $N_{\mathbf{x}_0} \subset V$ of $\mathbf{x}_0$ and further more the element $\hat{\mathbf{x}}$ of $F_{\hat{y}}$ closest to $\mathbf{x}_0$ also lies in the same neighborhood $N_{\mathbf{x}_0}$.
With these considerations in mind we turn to the nonlinear datum problem. We assume that the adjustment problem has already been solved and its solution $\hat{\mathbf{y}}$ defines a specific fiber $F_{\hat{y}} \subset V \subset X$, described by either (22) or (23).

4. Solution to the nonlinear geodetic datum problem

The datum problem can be viewed as the problem of completing the projection mapping $p: Y \to R(f)$ which solved the adjustment problem into a (minimum distance) generalized inverse $g = c \circ p$ of f, with the choice of a complementary mapping $c: R(f) \to C$, where C is a section of the fibering $\mathcal{F}$ induced by f in V. In fact it is sufficient to determine only C, in which case c (and thus g) is automatically defined as the mapping of any $\mathbf{y} \in R(f)$ into $c(\mathbf{y}) = \mathbf{x}$, where $\mathbf{x}$ is the unique element of $C \cap F_y$. One way of describing C is by means of a set of minimal constraints $\mathbf{h}(\mathbf{x}) = \mathbf{d}$. A particular choice for C is as the set of points $\mathbf{x}$ such that if $C \cap F_y = \{\mathbf{x}\}$ then $\mathbf{x}$ is the closest point to a given $\mathbf{x}_0 \in V$ among all elements of F_y.
The description of this particular section by the corresponding set of minimal constraints (***inner constraints***) is not as easy in the linear case. Instead we shall follow an approach similar to that of Baarda's S-transformation in the linear case, where a known reference "minimal" solution is transformed into the desired "inner" solution.
In order to proceed with the solution we need an analytical description of the solution fibers as submanifolds of X, e.g. by appropriate curvilinear coordinates.

Let G denote the group of applicable transformations (i.e. either similarity or rigid) with elements $g: X \to X$. If $\mathbf{z}$ belongs to a specific fiber $\mathbf{z} \in F_y$ then it holds true that $\mathbf{x} = g(\mathbf{z}) \in F_y$ for every $g \in G$.
This establishes for every fiber F_y of X a natural correspondence $h: F_y \times F_y \to G$:

$$g = h(\mathbf{z}, \mathbf{x}) \qquad \Leftrightarrow \qquad \mathbf{x} = g(\mathbf{z}) \,. \tag{24}$$

This means that by fixing one of the arguments of h we can obtain a mapping $h_{\mathbf{z}}: \mathbf{x} \to h_{\mathbf{z}}(\mathbf{x}) \equiv h(\mathbf{z}, \mathbf{x}): F_y \to G$, which is invertible. If further $\phi_G: G \to R^d$ is a coordinate system on G (i.e. a parameterization of G) the composite mapping

$$\phi \equiv \phi_G \circ h_{\mathbf{z}}: \mathbf{x} \to \mathbf{p}(\mathbf{x}): F_y \to R^d \tag{25}$$

establishes a coordinate system on F_y. We have already introduced the coordinate systems ϕ and the corresponding coordinates $\mathbf{p}$ in the descriptions (12) of the similarity and rigid transformations. The dimension d of F_y depends on the applicable transformation group and the dimension of the network. For 2-dimensional networks $d = 4$ for the similarity case (coordinates θ, t_1, t_2, λ) and $d = 3$ for the rigid case (coordinates θ, t_1, t_2). For 3-dimensional networks $d = 7$ for the similarity case (coordinates θ_1, θ_2, θ_3, t_1, t_2, t_3, λ) and $d = 6$ for the rigid case (coordinates θ_1, θ_2, θ_3, t_1, t_2, t_3). In fact we have introduced not one but a family of coordinate systems depending on the "reference point" $\mathbf{z}$ chosen for any fiber F_y. This choice can be made by introducing a section Z of the fibering $\mathcal{F}$ of X, e.g. by the use of easy to handle computationally minimal constraints, e.g of the form $x_k^i = 0$ (***trivial constraints***), where k is the network point index and i the coordinate index (i = 1, 2 for plane and i = 1, 2, 3 for 3D networks).
Once a reference solution $\mathbf{z}$ is established all other solutions in the same fiber $\mathbf{x} = \mathbf{x}(\mathbf{z}, \mathbf{p})$ depend on the coordinates $\mathbf{p}$ which are identical to the parameters of the transformation $g(\mathbf{p}): \mathbf{z} \to \mathbf{x}$. To obtain the solution specified by a given set of d minimal constraints $\mathbf{h}(\mathbf{x}) = \mathbf{d}$ we have to solve a system of d nonlinear equations with d unknowns

$$\mathbf{h}(\mathbf{p}) = \mathbf{h}\big(\mathbf{x}(\mathbf{z}, \mathbf{p})\big) = \mathbf{d} \,. \tag{26}$$

The solution $\mathbf{p}$ specifies the transformation $g_{\mathbf{p}} = g(\mathbf{p})$ which maps the reference solution $\mathbf{z}$ into the minimal constraint solution $\mathbf{x} = g_{\mathbf{p}}(\mathbf{z})$.
The solution $\mathbf{x}$ closest to a given element $\mathbf{x}_0 \in X$ is obtained either by minimizing

$$\Phi(\mathbf{p}) = \big[\mathbf{x}(\mathbf{z}, \mathbf{p}) - \mathbf{x}_0\big]^T \big[\mathbf{x}(\mathbf{z}, \mathbf{p}) - \mathbf{x}_0\big] \tag{27}$$

with respect to $\mathbf{p}$, or by imposing the condition that the vector $\mathbf{x}_0 - \mathbf{x}$ is orthogonal to the tangent space to the fiber $F_{\hat{y}}$ at the point $\mathbf{x}$. Following the first approach we are led to a system of d nonlinear equations with d unknowns

$$\frac{\partial \Phi}{\partial \mathbf{p}}(\mathbf{p}) = \mathbf{0} \qquad \Rightarrow \qquad \left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}(\mathbf{p})\right)^T \big[\mathbf{x}(\mathbf{p}) - \mathbf{x}_0\big] = \mathbf{0} \,. \tag{28}$$

The solution $\hat{\mathbf{p}}$ of the above system specifies the transformation $g_{\hat{\mathbf{p}}} = g(\hat{\mathbf{p}})$ which maps the reference solution $\mathbf{z}$ into the solution $\mathbf{x} = g_{\hat{\mathbf{p}}}(\mathbf{z})$ closest to $\mathbf{x}_0$. We may call $g_{\hat{\mathbf{p}}}$ the ***nonlinear Baarda S-transformation*** in the similarity, and the *nonlinear Baarda R-transformation* in the rigid case. The detailed derivation of the specific form of the above nonlinear system is given in Appendix A for the 4 special cases corresponding to 2- or 3-dimensional networks and similarity or rigid transformations. The results are:

3-dimensional networks - similarity transformation:

$$\mathbf{x}_i = \bar{\mathbf{x}}_0 + \frac{\sum_i (\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)^T \mathbf{R}(\mathbf{z}_i - \bar{\mathbf{z}})}{\sum_i (\mathbf{z}_i - \bar{\mathbf{z}})^T \mathbf{R}(\mathbf{z}_i - \bar{\mathbf{z}})} \mathbf{R}(\boldsymbol{\theta})(\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{29}$$

where

$$\bar{\mathbf{z}} \equiv \frac{1}{N} \sum_i \mathbf{z}_i , \qquad \bar{\mathbf{x}}_0 \equiv \frac{1}{N} \sum_i \mathbf{x}_{0i} \tag{30}$$

and $\boldsymbol{\theta}$ is the solution of the 3 nonlinear equations

$$\frac{1}{N} \sum_i [(\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)\times] \mathbf{R}(\boldsymbol{\theta})(\mathbf{z}_i - \bar{\mathbf{z}}) = \mathbf{0} . \tag{31}$$

3-dimensional networks - rigid transformation:

$$\mathbf{x}_i = \bar{\mathbf{x}}_0 + \mathbf{R}(\boldsymbol{\theta})(\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{32}$$

where $\boldsymbol{\theta}$ is the solution of the 3 nonlinear equations

$$\frac{1}{N} \sum_i [(\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)\times] \mathbf{R}(\boldsymbol{\theta})(\mathbf{z}_i - \bar{\mathbf{z}}) = \mathbf{0} . \tag{33}$$

2-dimensional networks - similarity transformation:

$$\mathbf{x}_i = \bar{\mathbf{x}}_0 + \frac{\sum_i (\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)^T \mathbf{R}(\mathbf{z}_i - \bar{\mathbf{z}})}{\sum_i (\mathbf{z}_i - \bar{\mathbf{z}})^T \mathbf{R}(\mathbf{z}_i - \bar{\mathbf{z}})} \mathbf{R}(\theta)(\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{34}$$

where θ is the solution of the single nonlinear equation

$$\frac{1}{N} \sum_i (\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)^T \mathbf{R}(\theta)(\mathbf{z}_i - \bar{\mathbf{z}}) = 0 . \tag{35}$$

When $\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

$$\mathbf{x}_i = \bar{\mathbf{x}}_0 + \frac{1}{\sigma_z^2} \begin{bmatrix} b & a \\ -a & b \end{bmatrix} (\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{36}$$

where

$$a \equiv \frac{1}{N} \sum_i (\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)^T \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} (\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{37}$$

$$b \equiv \frac{1}{N} \sum_i (\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)^T (\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{38}$$

$$\sigma_z^2 \equiv \frac{1}{N}\sum_i (\mathbf{z}_i - \bar{\mathbf{z}})^T (\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{39}$$

2-dimensional networks - rigid transformation:

$$\mathbf{x}_i = \bar{\mathbf{x}}_0 + \mathbf{R}(\boldsymbol{\theta})(\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{40}$$

where θ is the solution of the single nonlinear equation

$$\frac{1}{N}\sum_i (\mathbf{x}_{0i} - \bar{\mathbf{x}}_0)^T \, \mathbf{R}(\theta)(\mathbf{z}_i - \bar{\mathbf{z}}) = 0 . \tag{41}$$

When $\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

$$\mathbf{x}_i = \bar{\mathbf{x}}_0 \pm \frac{1}{\sqrt{a^2 + b^2}} \begin{bmatrix} b & a \\ -a & b \end{bmatrix} (\mathbf{z}_i - \bar{\mathbf{z}}) , \tag{42}$$

where *a* and *b* are defined as in the previous case.

5. Differential Geometry of Solution Fibers

We shall investigate some of the differential geometric characteristics of the solution fibers viewed as (curved) submanifolds of dimension $d = m - r$ of the $3N$ euclidean space of the coordinates of the N points of a three-dimensional geodetic network, i.e. $X = E^{3N} = (\times E^3)^N$ equipped with the simple inner product $(\mathbf{x}_\alpha, \mathbf{x}_\beta) = \mathbf{x}_\alpha^T \mathbf{x}_\beta$.

The points $\mathbf{x}$ of the fiber are identified with the help of a fixed element $\mathbf{z}$ belonging to the fiber through the similarity transformation

$$\mathbf{x}_i = \lambda \, \mathbf{R}(\boldsymbol{\theta})\mathbf{z}_i + \mathbf{t} = \mathbf{x}_i(\mathbf{z}_i, \mathbf{p}) , \qquad \mathbf{p} = \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{t} \\ \lambda \end{bmatrix} \tag{43}$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} , \qquad \mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_N \end{bmatrix} . \tag{44}$$

Since $\mathbf{x} = \mathbf{x}(\mathbf{p}, \mathbf{z}) = \mathbf{x}(\mathbf{p})$, $\mathbf{z}$ being fixed, the transformation parameters $\mathbf{p}$ serve as a set of curvilinear coordinates on the fiber. They consist of 3 rotational parameters $\boldsymbol{\theta} = [\theta_1 \; \theta_2 \; \theta_3]^T$ defining the orthogonal matrix $\mathbf{R} = \mathbf{R}(\boldsymbol{\theta})$, 3 parallel displacement parameters $\mathbf{t} = [t_1 \; t_2 \; t_3]^T$ and a scale parameter λ . The point $\mathbf{z}$ in particular has coordinates $\boldsymbol{\theta}_\mathbf{z} = \mathbf{0}$, where $\mathbf{R}(\mathbf{0}) = \mathbf{I}$, $\mathbf{t}_\mathbf{z} = \mathbf{0}$ and $\lambda_\mathbf{z} = 1$. Since

$$ds^2 = d\mathbf{x}^T d\mathbf{x} = d\mathbf{p}^T \left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}} \right)^T \frac{\partial \mathbf{x}}{\partial \mathbf{p}} d\mathbf{p} = d\mathbf{p}^T \mathbf{G} \, d\mathbf{p} \tag{45}$$

the metric coefficients on the surface are the elements of the matrix

$$\mathbf{G}=\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T \frac{\partial \mathbf{x}}{\partial \mathbf{p}}=\begin{bmatrix} \lambda^2 \mathbf{\Omega}^T \mathbf{R} \mathbf{C}_z \mathbf{R}^T \mathbf{\Omega} & N\lambda \mathbf{\Omega}^T \mathbf{R}[\bar{\mathbf{z}}\times]\mathbf{R}^T & \mathbf{0} \\ & N\mathbf{I} & N\mathbf{R}\bar{\mathbf{z}} \\ \textit{symm.} & & \mathbf{z}^T\mathbf{z} \end{bmatrix} \tag{46}$$

where

$$\frac{\partial}{\partial \theta_k}\mathbf{R}=[\boldsymbol{\omega}_k\times]\mathbf{R}\,, \qquad \mathbf{\Omega}=[\boldsymbol{\omega}_1\ \boldsymbol{\omega}_2\ \boldsymbol{\omega}_3] \tag{47}$$

$$\mathbf{C}_z=-\sum_i [\mathbf{z}_i\times][\mathbf{z}_i\times]=\left(\sum_i \mathbf{z}_i^T\mathbf{z}_i\right)\mathbf{I}-\sum_i \mathbf{z}_i\mathbf{z}_i^T=(\mathbf{z}^T\mathbf{z})\,\mathbf{I}-\sum_i \mathbf{z}_i\mathbf{z}_i^T \tag{48}$$

is the matrix of inertia of the network (Goldstein, 1950, ch. 5.1) and

$$\bar{\mathbf{z}}\equiv\frac{1}{N}\sum_i \mathbf{z}_i\ . \tag{49}$$

The detailed derivations of **G** is given in Appendix B.
The $d=7$ rows of the matrix

$$\frac{\partial \mathbf{x}}{\partial \mathbf{p}}=\left[\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \quad \frac{\partial \mathbf{x}}{\partial \mathbf{t}} \quad \frac{\partial \mathbf{x}}{\partial \lambda}\right] \tag{50}$$

constitute the local coordinate basis for the tangent space at any point **x** of the fiber.
The connection coefficients Γ^k_{mi} are defined by

$$\frac{D}{dp_m}\left(\frac{\partial \mathbf{x}}{\partial p_i}\right)=\Gamma^k_{mi}\frac{\partial \mathbf{x}}{\partial p_k}\,, \tag{51}$$

where the covariant (intrinsic) derivative $\frac{D\mathbf{v}}{dp_m}$ of any tangent vector **v** equals the projection of the ordinary (extrinsic) derivative $\frac{\partial \mathbf{v}}{\partial p_m}$ on the local tangent space (= span of the columns of $\frac{\partial \mathbf{x}}{\partial \mathbf{p}}$), see e.g. Hicks (1965), ch. 6.4, do Carmo (1992), ch. 6.2. Using the (well known from least squares adjustment) projection operator we obtain

$$\frac{\partial \mathbf{x}}{\partial p_k}\Gamma^k_{mi}=\left[\frac{\partial}{\partial p_m}\left(\frac{\partial \mathbf{x}}{\partial p_i}\right)\right]_T=\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\left[\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right]^{-1}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T\frac{\partial}{\partial p_m}\left(\frac{\partial \mathbf{x}}{\partial p_i}\right). \tag{52}$$

This relation can be written in matrix notation as

$$\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\mathbf{\Gamma}_m=\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\left[\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right]^{-1}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T\frac{\partial}{\partial p_m}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)=\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\mathbf{G}^{-1}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T\frac{\partial}{\partial p_m}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right). \tag{53}$$

where the matrix $\mathbf{\Gamma}_m$ has elements $(\mathbf{\Gamma}_m)_{ki}=\Gamma^k_{mi}$, in its k row and i column. Multiplying from the left with $\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T$ we get

$$\mathbf{K}_m = \mathbf{G}\boldsymbol{\Gamma}_m = \left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right)^T \frac{\partial}{\partial p_m}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{p}}\right), \qquad \boldsymbol{\Gamma}_m = \mathbf{G}^{-1}\mathbf{K}_m \,. \tag{54}$$

Since $\mathbf{K}_m = \mathbf{G}\boldsymbol{\Gamma}_m$ the elements $(\mathbf{K}_m)_{ik} = K_{mi,k}$ of the $\mathbf{K}_m$ matrices are related to the connection coefficients Γ^k_{mi} by

$$K_{mi,k} = \sum_n G_{kn}\Gamma^n_{mi} \,, \tag{55}$$

and they may be called "connection coefficients of the first kind". The matrices $\mathbf{K}_m$ are derived in Appendix B and have the form

$$\mathbf{K}_{\theta_m} = \begin{bmatrix} [\mathbf{K}_{\theta_m}]_{\theta\theta} & \mathbf{0} & [\mathbf{K}_{\theta_m}]_{\theta\lambda} \\ [\mathbf{K}_{\theta_m}]_{t\theta} & \mathbf{0} & [\mathbf{K}_{\theta_m}]_{t\lambda} \\ [\mathbf{K}_{\theta_m}]_{\lambda\theta} & \mathbf{0} & [\mathbf{K}_{\theta_m}]_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} [\mathbf{K}_{\theta_m}]_{\theta\theta} & \mathbf{0} & \lambda\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\omega}_m \\ [\mathbf{K}_{\theta_m}]_{t\theta} & \mathbf{0} & -N\mathbf{R}[\bar{\mathbf{z}}\times]\mathbf{R}^T\boldsymbol{\omega}_m \\ -\lambda\boldsymbol{\omega}_m^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} & \mathbf{0} & 0 \end{bmatrix} \tag{56}$$

$$\mathbf{K}_{t_m} = \mathbf{0} \tag{57}$$

$$\mathbf{K}_\lambda = \begin{bmatrix} [\mathbf{K}_\lambda]_{\theta\theta} & \mathbf{0} & \mathbf{0} \\ [\mathbf{K}_\lambda]_{t\theta} & \mathbf{0} & \mathbf{0} \\ [\mathbf{K}_\lambda]_{\lambda\theta} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \lambda\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} & \mathbf{0} & \mathbf{0} \\ -N\mathbf{R}[\bar{\mathbf{z}}\times]\mathbf{R}^T\boldsymbol{\Omega} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{58}$$

where the non-zero submatrices are

$$[\mathbf{K}_{\theta_m}]_{\theta\theta} = -\lambda^2\boldsymbol{\Omega}^T[\boldsymbol{\omega}_m\times]\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} + \lambda^2(\mathbf{z}^T\mathbf{z})\boldsymbol{\Omega}^T[\boldsymbol{\omega}_m\times]\boldsymbol{\Omega} + \lambda^2\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\frac{\partial\boldsymbol{\Omega}}{\partial\theta_m} \tag{59}$$

$$[\mathbf{K}_{\theta_m}]_{t\theta} = N\lambda\boldsymbol{\omega}_m\bar{\mathbf{z}}^T\mathbf{R}^T\boldsymbol{\Omega} - N\lambda\mathbf{R}\bar{\mathbf{z}}\boldsymbol{\omega}_m^T\boldsymbol{\Omega} - N\lambda\mathbf{R}[\bar{\mathbf{z}}\times]\mathbf{R}^T\frac{\partial\boldsymbol{\Omega}}{\partial\theta_m} \,. \tag{60}$$

When the reference solution $\mathbf{z}$ is centered, i.e., $\bar{\mathbf{z}} = \mathbf{0}$, then the above matrices reduce to

$$\mathbf{G} = \begin{bmatrix} \lambda^2\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & N\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{z}^T\mathbf{z} \end{bmatrix} \tag{61}$$

$$[\mathbf{K}_{\theta_m}]_{\theta\theta} = -\lambda^2\boldsymbol{\Omega}^T[\boldsymbol{\omega}_m\times]\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} + \lambda^2(\mathbf{z}^T\mathbf{z})\boldsymbol{\Omega}^T[\boldsymbol{\omega}_m\times]\boldsymbol{\Omega} + \lambda^2\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\frac{\partial\boldsymbol{\Omega}}{\partial\theta_m} \tag{62}$$

$$\mathbf{K}_{\theta_m} = \begin{bmatrix} [\mathbf{K}_{\theta_m}]_{\theta\theta} & \mathbf{0} & \lambda\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\omega}_m \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\lambda\boldsymbol{\omega}_m^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} & \mathbf{0} & 0 \end{bmatrix} \tag{63}$$

$$\mathbf{K}_{t_m} = \mathbf{0} \tag{64}$$

$$\mathbf{K}_\lambda = \begin{bmatrix} \lambda\boldsymbol{\Omega}^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\boldsymbol{\Omega} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} . \tag{65}$$

In this case we can use the relation $\mathbf{\Gamma}_m = \mathbf{G}^{-1}\mathbf{K}_m$ with

$$\mathbf{G}^{-1} = \begin{bmatrix} \frac{1}{\lambda^2}\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{C}_z^{-1}\mathbf{R}^T\mathbf{\Omega}^{-T} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{N}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\mathbf{z}^T\mathbf{z}} \end{bmatrix} \tag{66}$$

in order to compute the matrices of the connection coefficients

$$\mathbf{\Gamma}_{\theta_m} = \begin{bmatrix} [\mathbf{\Gamma}_{\theta_m}]_{\theta\theta} & \mathbf{0} & \frac{1}{\lambda}\mathbf{\Omega}^{-1}\boldsymbol{\omega}_m \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\frac{\lambda}{\mathbf{z}^T\mathbf{z}}\boldsymbol{\omega}_m^T\mathbf{R}\mathbf{C}_z\mathbf{R}^T\mathbf{\Omega} & \mathbf{0} & 0 \end{bmatrix} \tag{67}$$

$$[\mathbf{\Gamma}_{\theta_m}]_{\theta\theta} = \frac{1}{\lambda^2}\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{C}_z^{-1}\mathbf{R}^T\mathbf{\Omega}^{-T}[\mathbf{K}_{\theta_m}]_{\theta\theta} = \mathbf{\Omega}^{-1}\mathbf{R}\mathbf{C}_z^{-1}\mathbf{R}^T[\boldsymbol{\omega}_m\times]\left((\mathbf{z}^T\mathbf{z})\mathbf{I} - \mathbf{R}\mathbf{C}_z^{-1}\mathbf{R}^T\right)\mathbf{\Omega} + \mathbf{\Omega}^{-1}\frac{\partial\mathbf{\Omega}}{\partial\theta_m} \tag{68}$$

$$\mathbf{\Gamma}_{t_m} = \mathbf{0} \tag{69}$$

$$\mathbf{\Gamma}_{\lambda} = \begin{bmatrix} \frac{1}{\lambda}\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{70}$$