Theoretical foundations of ITRF determination. The algebraic and the kinematic approach

Zuheir Altamimi[1], Athanasios Dermanis[2]

[1] *Institut Géographique National, ENSG/LAREG, Champs-sur-Marne, France*
[2] *Department of Geodesy and Surveying, Aristotle University of Thessaloniki*

Abstract: The minimal constraints required for the definition of the reference system in the ITRF formulation are derived for both the one-step approach (simultaneous stacking of coordinate time series for all available space techniques) as well as for the currently implemented two-step approach of stacking for each technique separately and combination of the derived station initial coordinates and velocities. Two types of minimal constraints are studied, the usual algebraic ones related to the variation of the unknown parameters under a change of the reference system known as inner constraints and kinematic ones based on the choice of a reference system with respect to which the apparent variation of station coordinates is minimized. It is shown that under mild conditions on the approximate parameter values used in the linearization of the observation equations the two approaches, algebraic and kinematic one, become identical.

1. Introduction

The implementation of an International Terrestrial Reference System (ITRS) by means of an International Terrestrial Reference Frame (ITRF) involves three elements:

(a) The adoption of a model of the general form $\mathbf{x}_i(t) = \mathbf{F}(\mathbf{a}_i, t)$ where the coordinates $\mathbf{x}_i$ of any ITRF station i are expressed as functions of time t through a set of parameters $\mathbf{a}_i$ particular to each station.

(b) Sets of available data one from each particular technique (T) in the form of coordinates $\mathbf{x}_{T,i}(t_k)$ of stations i at various "observation" epochs t_k, referring to a reference system S_T particular to each technique T different from the reference system S_{ITRF} of the final ITRF coordinates.

(b) The adoption of a particular criterion for the choice of the "optimal" ITRF reference system among all equivalent ones giving rise to the same time sequence of shapes for the ITRF station network.

The objective is to develop a strategy for the determination of the ITRF parameters $\mathbf{a}_i$ from the available data which has a sound theoretical basis and it is also numerically feasible in view of the enormous amount of data involved.

The basic characteristic of all the above elements of ITRF implementation is the use of coordinates while the analyzed geodetic data are not capable of determining point coordinates but only the shape of a network of points at each observation epoch. And this is so because in this case coordinates have no physical substance whatsoever; they are only a mathematical device for determining shape via positions. Coordinates are entering the game only after we introduce an arbitrary reference system, a mathematical apparatus also without physical substance. For the shake of convenience the term “shape” is used here in a loose sense which becomes more concrete when applied to different data sets. If the scale in a technique (defined by means of a particular set of clocks) is entering in the ITRF formulation, then “shape” means shape and size. In the SLR data where the system of reference to be introduced lacks only orientation since the origin is taken to be the geocenter a point “provided” by nature (although not without involving a mathematical definition!). In this case “shape” is the extended shape (or shape and size accordingly) of an extended set of points, the network points and the geocenter. As a consequence of the use of coordinates with coordinate-free (coordinate invariant) observations the formulated least squares (or best estimation) problem does not have a unique solution. All sets of coordinates giving rise to the same network shape are equally valid solutions. From the algebraic point of view the design matrix transforming unknown parameters into observables has a column rank deficiency d, equal to the number of parameters in the coordinate transformations leaving the observables invariant. The same rank deficiency appears in the coefficient matrix of the normal equations formed under the least-squares principle and a unique solution requires the introduction of an arbitrary reference system, a choice which is algebraically expressed by means of a set of minimal (d in number and independent) additional constraints. Any other set of non-minimal constraints may resolve the uniqueness problem but it distorts the network “shape” from its optimal form implied by the least squares principle. The estimates of the ITRF parameters obtained in this way are (or should be) accompanied by covariance matrices which demonstrate the same rank deficiency.

In formulating a strategy for the ITRF parameter estimation we may depart from theoretical optimality for the sake of numerical feasibility. Note that this has already done before the very formulation of the ITRF implementation problem: the rigorously optimal solution calls for the common analysis of all available original primary data, an approach which is beyond reach, not only because of the immense amount of the data but also because of the peculiarities in the models, parameters and processing strategies of each technique.

We now return to the particular characteristics of each of the three elements of the ITRF implementation problem (Altamimi et al., 2002, 2007, 2008, 2011, Sillard, 1999, Sillard & Boucher, 2001). Currently the model has the form

$$\mathbf{x}_i(t) = \mathbf{F}(\mathbf{a}_i, t) = \mathbf{x}_{0i} + (t - t_0)\mathbf{v}_i, \tag{1}$$

where the station parameters $\mathbf{a}_i$ consist of their coordinates $\mathbf{x}_{0i}$ at the ITRF reference epoch t_0 and their constant velocities $\mathbf{v}_i$. The next model will probably be of the form

$$\mathbf{F}(\mathbf{a}_i, t) = \mathbf{x}_{0i} + (t - t_0)\mathbf{v}_i + \begin{bmatrix} b_{i1}\cos(2\pi t / T_0 + \varphi_{i1}) \\ b_{i2}\cos(2\pi t / T_0 + \varphi_{i2}) \\ b_{i3}\cos(2\pi t / T_0 + \varphi_{i3}) \end{bmatrix} \tag{2}$$

involving an additional annual periodic term (T_0 = 1 year) with the amplitudes b_{ik} and phases φ_{ik}, $k = 1,2,3$, as additional elements of each station parameters $\mathbf{a}_i$. Note however that such a model cannot efficiently fit the observed "annual" variations because b_{ik} and φ_{ik} vary from year to year due to the varying weather conditions.

The main role of the coordinate model is to secure that the final time sequence of coordinates is a smooth one, reflecting the general fact that the earth surface and hence the ITRF network changes shape with time in a smooth way, apart from isolated discontinuities due to earthquakes, landslides, etc. Thus any non-linear variations present in the available data are attributed to observational errors and are removed. Note that such an analytic deterministic model is just one way to impose smoothness; alternatively ITRF coordinates can be modelled as stochastic processes: the high correlation of nearby values imposes smoothness. A combination of a deterministic parameterized analytical trend and a residual zero mean stochastic process is also possible. The model can also rise from geophysical reasoning: the current linear in time coordinate model reflects the motion of tectonic plates which rotate around a pole with constant angular velocity. Thus stations on a plate move along circular arcs which in the relatively small period of validity of an ITRF solution appear as straight lines. Note however that this justifies the model only with respect to the horizontal coordinate components; the vertical ones remain unaffected by plate rotation.

Although all data $\mathbf{x}_{T,i}(t_k)$ from each particular technique T refer nominally to the same reference system S_T, their actual variation in time cannot be attributed to the effect of observational errors only. A significant part is due to instability in the reference frame definition and can be removed accordingly by assuming that for each epoch t_k the data refer to a separate reference system $S_T(t_k)$. Since these reference systems are different from the desired ITRF system S_{ITRF}, the data $\mathbf{x}_{T,i}(t_k)$ can be deprived from their dependence on $S_T(t_k)$ by introducing the parameters $\mathbf{q}_{T,k} = \mathbf{q}_T(t_k)$ of the transformation connecting the different reference systems $T(\mathbf{q}_{T,k}) : S_{\text{ITRF}} \to S_T(t_k)$, which are additional nuisance parameters in the problem. Combining the transformation

$$\mathbf{x}_{T,i}(t_k) = \Phi\left(\mathbf{q}_{T,k}, \mathbf{x}_i(t_k)\right) = (1 + s_{T,k})\mathbf{R}(\boldsymbol{\theta}_{T,k})\mathbf{x}_i(t_k) + \mathbf{d}_{T,k} \ , \tag{3}$$

$$\mathbf{q}_{T,k} = [s_{T,k} \ \ \theta_{1,T,k} \ \ \theta_{2,T,k} \ \ \theta_{3,T,k} \ \ d_{1,T,k} \ \ d_{2,T,k} \ \ d_{3,T,k}]^T \tag{4}$$

with the ITRF coordinate model $\mathbf{x}_i(t_k) = \mathbf{F}(\mathbf{a}_i, t_k)$ we arrive at the observation equations model of the form

$$\mathbf{x}_{T,i}(t_k) = \Phi\left(\mathbf{q}_{T,k}, \mathbf{F}(\mathbf{a}_i, t_k)\right) = \mathbf{f}(\mathbf{q}_{T,k}, \mathbf{a}_i, t_k) \tag{5}$$

to be solved on the basis of the observational outcomes $\mathbf{x}'_{T,i}(t_k) = \mathbf{x}_{T,i}(t_k) + \mathbf{e}_{T,i}(t_k)$ affected by random zero-mean errors $\mathbf{e}_{T,i}(t_k)$, by applying the least squares principle $\mathbf{e}^T\mathbf{P}\mathbf{e} = \min$ to all data from all techniques T at all the corresponding epochs t_k and all stations i.

Once the data have been deprived from their own reference through the transformation parameters $\mathbf{p}_{T,k}$, we can reformulate the ITRF problem in an (almost) coordinate free way: Given the time discrete sequence of "shapes" of the sub-networks from each techniques at particular epochs t_k, connect them through their common tie points in order to construct a time continuous sequence of shapes of the total network, which when expressed in a particular reference system S_{ITRF} as coordinate functions $\mathbf{x}_i(t)$ should comply with the ITRF model $\mathbf{x}_i(t) = \mathbf{F}(\mathbf{a}_i, t)$. We cannot have a genuine coordinate-free version of the problem because the smoothness of the final ITRF time-continuous sequence of shapes is defined via a particular reference system. For example for the linear in time coordinate model $\mathbf{x}_i(t) = \mathbf{x}_{0i} + (t - t_0)\mathbf{v}_i$ it is not possible to define a linear in time sequence in shapes without resorting to coordinates. There is no such thing as a linear in time deformation for a network. This assertion reveals the somewhat arbitrary character of any coordinate based model, which is a matter of algebraic convenience rather than a demonstration of a physical fact. Linearity in time is not preserved under a change of coordinate system which is an absolutely permissible action from the physical point of view:

$$\begin{aligned}\tilde{\mathbf{x}}_i(t) &= \Phi\left(\mathbf{q}(t), \mathbf{x}_i(t)\right) = \left(1 + s(t)\right)\mathbf{R}\left(\boldsymbol{\theta}(t)\right)\mathbf{x}_i(t) + \mathbf{d}(t) = \\ &= \left(1 + s(t)\right)\mathbf{R}\left(\boldsymbol{\theta}(t)\right)[\mathbf{x}_{i0} + (t - t_0)\mathbf{v}_i] + \mathbf{d}(t)\end{aligned} \tag{6}$$

is a non linear function of time and does not have the form $\tilde{\mathbf{x}}_i(t) = \tilde{\mathbf{x}}_{i0} + (t - t_0)\tilde{\mathbf{v}}_i$. Despite this fact the adopted coordinate model allows the determination of a unique "optimal" time-continuous sequence of shapes from the available data through the least-squares principle. On the contrary, the model parameters $\mathbf{a}_i$ (e.g. $\mathbf{x}_{0i}$, $\mathbf{v}_i$) cannot be uniquely determined since they are not estimable quantities for the data in hand.

We may select one of the reference systems compatible with the model by ignoring at first the requirement of an optimal system and choose instead a more-or-less arbitrary one, through the use of appropriate additional minimal constraints, which resolve the reference system choice problem without affecting the optimal network shape at any epoch. In this case the a-posteriori application of the optimality criterion with the use of the general transformation law (6) will lead to optimal coordinate functions which do not comply with the selected model $\tilde{\mathbf{x}}_i(t) = \tilde{\mathbf{x}}_{i0} + (t - t_0)\tilde{\mathbf{v}}_i$.

To overcome this serious problem a compromise must be made. Noting that all reference systems involved (per technique and epoch or for the ITRF) are closed to each other we may restrict our choice to coordinate transformations "close to the identity" and apply the ITRF reference system optimality criterion not to all possible transformations, not even to all possible close to the identity ones, but rather to the restricted class of transformations $\mathbf{x}_i(t) \leftrightarrow \tilde{\mathbf{x}}_i(t)$ which approximately preserve the model form, i.e. those for which if $\mathbf{x}_i(t) = \mathbf{x}_{i0} + (t - t_0)\mathbf{v}_i$ then $\tilde{\mathbf{x}}_i(t) \approx \tilde{\mathbf{x}}_{i0} + (t - t_0)\tilde{\mathbf{v}}_i$.

With respect to the optimality criterion for the choice of the ITRF reference system, there are two possible alternatives. The first, which we will call "algebraic", is based on the analytical form of the problem and seeks a unique solution by minimizing the sum of squares of the deviations of all or some selected subset of the unknowns from prescribed approximate values. This approach has been introduced in geodesy by Meissl (1965, 1969) with his famous inner constraints (see also, Blaha, 1971, 1982, Grafarend and Schaffrin, 1973, 1976, Dermanis, 2003). The inner constraints are additional constraints which help realize a unique solution namely the one minimizing the above sum of squares. The second choice, which we will call "kinematic", is seeking directly conditions (constraints) which will lead to a reference system such that the apparent motion of the network points with respect to it will be minimized in some prescribed sense.

The solution departs from the observation equations (6) having the form

$$\mathbf{x}_{T,i}(t_k) = (1 + s_{T,k})\mathbf{R}(\boldsymbol{\theta}_{T,k})[\mathbf{x}_{i0} + (t_k - t_0)\mathbf{v}_i] + \mathbf{d}_{T,k} + \mathbf{e}_{T,k},$$

$$T = \text{VLBI,SLR,GPS,DORIS}, \qquad i = 1,2,\ldots,N, \quad k = 1,2,\ldots,n_T, \tag{7}$$

where different transformation parameters $\mathbf{q}_{T,k}$ according to (4) are considered for each technique and each observation epoch within any particular technique, while $\mathbf{e}_{T,k}$ are the observational errors. Following the standard procedure in ITRF formulation the problem is separated into two sequential steps for computational convenience. In the first step, which is called "stacking", the observation equations (7) are analyzed separately for each technique in order to obtain the optimal parameters $\mathbf{x}_{T,i0}$, $\mathbf{v}_{T,i}$ for the relevant stations. In the second step, which is called "combination", the ITRF optimal values $\mathbf{x}_{i0}$, $\mathbf{v}_i$ are obtained from the previous estimates $\mathbf{x}_{T,i0}$, $\mathbf{v}_{T,i}$, with the help of an "observation equations" model of the form

$$\begin{bmatrix} \mathbf{x}_{T,0i} \\ \mathbf{v}_{T,i} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_{\mathbf{x}}(\mathbf{q}_T, \mathbf{x}_{0i}, \mathbf{v}_i) \\ \mathbf{g}_{\mathbf{v}}(\mathbf{q}_T, \mathbf{x}_{0i}, \mathbf{v}_i) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{\mathbf{x}_{T,0i}} \\ \mathbf{e}_{\mathbf{v}_{T,i}} \end{bmatrix}, \tag{8}$$

which involves the parameters $\mathbf{q}_T$ transforming coordinates and velocities from the reference system adopted in each technique to the final ITRF system. In both steps we have two common characteristics (a) the presence of "nuisance" transformation parameters $\mathbf{q}_{T,k}$ or $\mathbf{q}_T$ and (b) the need to select among all permissible reference systems the "optimal" one. Note that although the transformation parameters are of no direct interest for the ITRF implementation problem they are needed for the a posteriori analysis of Earth Orientation Parameter (EOP) series which will relate the ITRF reference system with the corresponding system of the ICRF (International Celestial Reference Frame) involving the direction parameters of extragalactic radio sources participating in VLBI observations. In fact EOP data can also be treated simultaneously with the coordinate time series data, since the two correlated sets are connected by their cross-covariance matrix. We neglect here the simultaneous treatment of EOP series, since it has no effect on the derived partial inner constraints with respect to coordinates-velocities and/or to nuisance transformation parameters.

The presence of nuisance parameters and the reference system choice problem does not allow the easy separation of the original "simultaneous stacking" problem of equations (7) for all techniques, into a rigorous separation in two stages utilizing e.g. the addition theorem of the normal equations formed separately from the data of each technique. The determination of such a rigorous two step numerically feasible approach remains an open problem.

Another problem that takes us even further from the traditional data adjustment problem with no rank deficiency, is the fact that the covariance matrix of the data is (or ought to be) singular and cannot be inverted to obtain the weight matrix entering in the least square principle. Indeed coordinates $\mathbf{x}_{T,i}(t_k)$ provided from each technique are derived from singular normal equations with the use of some additional minimal constraints, i.e. constrains which lead to the choice of a reference system without affecting the optimal shape of the network. In such a case the covariance matrix of the coordinates is a generalized inverse of the normal equations coefficient matrix and has the same rank deficiency d. We shall assume for the time being that the proper weight matrices for the coordinate observations are available, although this is a problem that requires further theoretical investigation. A convenient choice of coordinate weight matrices are the normal equations coefficient matrix in the relevant data adjustment. The more general class of weight matrices leaving estimable parameters invariant and equal to their optimal values is given by Rao's unified theory (Rao, 1971, 1973).

2. The classical theory of inner and partial inner constraints

We shall give a short presentation of the classical theory of Meissl in a slightly generalized way that it can be applied to our problem (Dermanis, 1994a, 1994b). Consider a set of observables $\mathbf{y}^a$, which are invariant under certain coordinate transformations and are modelled as functions $\mathbf{y}^a = \mathbf{f}(\mathbf{x}^a)$ of parameters $\mathbf{x}^a$ which depend on the use of a particular coordinate system. In such a case it is not possible to recover the unknowns $\mathbf{x}^a$ by only applying the least squares principle $\mathbf{e}^T\mathbf{P}\mathbf{e} = \min$ to the observations $\mathbf{y}^b = \mathbf{y}^a + \mathbf{e} = \mathbf{f}(\mathbf{x}^a) + \mathbf{e}$ corrupted by additive noise $\mathbf{e}$. Upon linearization the problem transforms to $\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where, for some selected approximate values $\mathbf{x}^0$, $\mathbf{A} = \left(\partial\mathbf{f}/\partial\mathbf{x}^a\right)\Big|_{\mathbf{x}^0}$ is the design matrix, $\mathbf{b} = \mathbf{y}^b - \mathbf{y}^0 = \mathbf{y}^b - \mathbf{f}(\mathbf{x}^0)$ are the reduced observations and $\mathbf{x} = \mathbf{x}^a - \mathbf{x}^0$ are the corrections to the approximate values serving as the new unknowns. Let $\tilde{\mathbf{x}}^a = \mathbf{G}(\mathbf{p}, \mathbf{x}^a)$ be the transformation of the unknowns caused by a change of reference system involving transformation parameters $\mathbf{p} = [\lambda \;\; \psi_1 \;\; \psi_2 \;\; \psi_3 \;\; g_1 \;\; g_2 \;\; g_3]^T$. Note that we have changed the notation to $\mathbf{p}$, λ, $\boldsymbol{\psi}$, $\mathbf{g}$ in order to distinguish between the arbitrary change of the reference system to which the unknowns refer from the previous $\mathbf{q}$, s, $\boldsymbol{\theta}$, $\mathbf{d}$, which refer to the nuisance unknown transformation parameters of the various adjustment models, which connect the reference system of the unknowns to that of the observations. Then for any value of $\mathbf{p}$, if $\mathbf{x}^a$ is a solution to the least squares problem then $\tilde{\mathbf{x}}^a$ is also a solution, since observables are invariant under coordinate transformations and $\tilde{\mathbf{y}}^a = \mathbf{f}(\tilde{\mathbf{x}}^a) = \mathbf{f}(\mathbf{x}^a) = \mathbf{y}^a$. Taking into account that for $\mathbf{p} = \mathbf{0}$, $\mathbf{G}(\mathbf{p}, \mathbf{x}^a)\Big|_{\mathbf{p}=\mathbf{0}} = \mathbf{G}(\mathbf{0}, \mathbf{x}^a) = \mathbf{x}^a$ we have in the linearized case

$$\tilde{\mathbf{x}}^a = \mathbf{G}(\mathbf{p}, \mathbf{x}^a) \approx \mathbf{G}(\mathbf{0}, \mathbf{x}^a) + \left.\frac{\partial\mathbf{G}}{\partial\mathbf{p}}\right|_{\mathbf{p}=0,\mathbf{x}^a} \mathbf{p} \approx \mathbf{x}^a + \left(\frac{\partial\mathbf{G}}{\partial\mathbf{p}}\right)_{\mathbf{p}=0,\mathbf{x}^0} \mathbf{p} \equiv \mathbf{x}^a + \mathbf{E}\mathbf{p}\,. \tag{9}$$

Upon subtraction of $\mathbf{x}^0$ from both sides, the transformation becomes $\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{E}\mathbf{p}$ where $\mathbf{E} = \left(\partial\mathbf{G}/\partial\mathbf{p}\right)\Big|_{\mathbf{p}=\mathbf{0},\mathbf{x}^0}$. Since $\tilde{\mathbf{y}}^a = \mathbf{y}^a$ it also holds that

$\tilde{\mathbf{y}} \equiv \tilde{\mathbf{y}}^a - \mathbf{y}^0 = \mathbf{y}^a - \mathbf{y}^0 \equiv \mathbf{y}$ and $\tilde{\mathbf{y}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{E}\mathbf{p} = \mathbf{y} = \mathbf{A}\mathbf{x}$ for every $\mathbf{p}$ so that

$$\mathbf{A}\mathbf{E} = \mathbf{0}\,. \tag{10}$$

The above relation states that every column $\mathbf{e}_k$, $k = 1,...,d$, of $\mathbf{E} = \left[\mathbf{e}_1 \cdots \mathbf{e}_d\right]$ belongs to the null space of $\mathbf{A}$, i.e. to the set $N(\mathbf{A}) = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$ of all $\mathbf{x}$ which $\mathbf{A}$ maps to zero. Furthermore since the dimension of $N(\mathbf{A})$ is equal to the rank deficiency d, the d columns of $\mathbf{E}$ form a basis for $N(\mathbf{A})$. Hence every $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$ can be uniquely expressed as a linear combination $\mathbf{x} = \sum_{k=1}^{d} p_k \mathbf{e}_k = \mathbf{E}\mathbf{p}$. Any solution of the normal equations $\mathbf{N}\hat{\mathbf{x}} = \mathbf{u}$ ($\mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}$, $\mathbf{u} = \mathbf{A}^T\mathbf{P}\mathbf{b}$) derived from the application of the least squares principle gives rise to the same value

$\hat{\mathbf{y}} = \mathbf{A}\hat{\mathbf{x}}$ of the "adjusted" observables $\hat{\mathbf{y}}$. Let us assume now that we have at hand one solution $\hat{\mathbf{x}}$ of the normal equations and we seek to determine the unique one $\hat{\mathbf{x}}_+$, which satisfies the "minimum norm" principle $\|\hat{\mathbf{x}}_+\|^2 = \hat{\mathbf{x}}_+^T \hat{\mathbf{x}}_+ = \min$ among all solutions of the normal equations. In such a case $\mathbf{A}\hat{\mathbf{x}}_+ = \mathbf{A}\hat{\mathbf{x}}$, $\mathbf{A}(\hat{\mathbf{x}}_+ - \hat{\mathbf{x}}) = \mathbf{0}$ and $\hat{\mathbf{x}}_+ - \hat{\mathbf{x}} = \mathbf{E}\mathbf{p}_+ \in N(\mathbf{A})$ for an appropriate set of transformation parameters $\mathbf{p}_+$. Therefore the solution $\mathbf{p}_+$ satisfying $\varphi(\mathbf{p}_+) = \hat{\mathbf{x}}_+^T \hat{\mathbf{x}}_+ = (\hat{\mathbf{x}} + \mathbf{E}\mathbf{p}_+)^T (\hat{\mathbf{x}} + \mathbf{E}\mathbf{p}_+) = \min$ is derived from $\partial\varphi / \partial\mathbf{p}_+ = 2(\hat{\mathbf{x}} + \mathbf{E}\mathbf{p}_+)^T \mathbf{E} = \mathbf{0}$, yielding $\mathbf{p}_+ = -(\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}_+ = \hat{\mathbf{x}} + \mathbf{E}\mathbf{p}_+ = \hat{\mathbf{x}} - \mathbf{E}(\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\hat{\mathbf{x}}$. This minimum norm solution $\hat{\mathbf{x}}_+$ obviously satisfies $\mathbf{E}^T\hat{\mathbf{x}}_+ = \mathbf{E}^T\hat{\mathbf{x}} - \mathbf{E}^T\mathbf{E}(\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\hat{\mathbf{x}} = \mathbf{0}$.

Thus an alternative approach to directly obtain the minimum norm solution $\hat{\mathbf{x}}_+$ of the least squares problem, without the need for any prior solution $\hat{\mathbf{x}}$ of the normal equations, is to apply the least squares principle under the limitations posed by the additional "inner constraints" $\mathbf{E}^T\mathbf{x} = \mathbf{0}$. The matrix $\mathbf{E}$ can be determined from the procedure implicit in equation (9), i.e. by $\mathbf{E} = \left(\partial\mathbf{G} / \partial\mathbf{p}\right)\big|_{\mathbf{p}=\mathbf{0},\mathbf{x}^0}$, where $\tilde{\mathbf{x}}^a = \mathbf{G}(\mathbf{p}, \mathbf{x}^a)$ is the transformation of the unknown parameters under an arbitrary change of the reference system with transformation parameters $\mathbf{p}$.

In the geodetic case where the unknowns are coordinate corrections, the inner constraints lead to a solution depending on the approximate coordinates, which define an approximate shape with a reference system attached to it. The inner constraints take the optimal network shape uniquely defined by the least squares criterion and position it in a way that it best fits the approximate network. Then the coordinate system of the approximate coordinates passes on to that of the optimal shape and thus "optimal" coordinates are obtained. The name "inner constraints" refers to the difference with the previous external constraints of fixing the necessary number of point coordinates, e.g. two points in a triangulation planar network or one point and one azimuth in a planar network with also distance observations.

In classical triangulation networks coordinates are not the only parameters. Instead of isolated angles the directions were observed from each station to all others that were visible, as readings on the horizontal theodolite circle. Thus the azimuth of the zero reading appeared as an additional unknown per station. These nuisance unknowns did not participate in the inner constraints and in fact were eliminated from the very beginning by replacing directions with their differences. It is also possible to apply the best fitting of the optimal network shape not to the whole approximate network but only to a chosen approximate sub-network. For example when a new network overlapping with a higher order one is adjusted, a best fit with respect to the common points can be implemented. In such cases the (total) inner constraints $\mathbf{E}^T\mathbf{x} = \begin{bmatrix} \mathbf{E}_1^T & \mathbf{E}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{0}$ are replaced (rearranging the order of the unknowns as necessary) by the "partial inner constraints" (Fritz & Schaffrin, 1981)

$\begin{bmatrix} \mathbf{E}_1^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{E}_1^T \mathbf{x}_1 = \mathbf{0}$. Accordingly the minimization principle $\mathbf{x}^T \mathbf{x} = \mathbf{x}_1^T \mathbf{x}_1 + \mathbf{x}_2^T \mathbf{x}_2 = \min$ is replaced by $\mathbf{x}_1^T \mathbf{x}_1 = \min$.

3. Approximate coordinate and velocity transformations and model preserving transformations

The general similarity transformation $\tilde{\mathbf{x}}(t) = (1+\lambda(t))\mathbf{R}(\boldsymbol{\psi}(t))\mathbf{x}_i(t) + \mathbf{g}(t)$ from the coordinates of a reference system to those of another with scale factor $1+\lambda$, small angles ψ_i and small displacements g_i, can be simplified using the sufficient approximation by $\mathbf{R}(\boldsymbol{\psi}(t)) = \mathbf{I} - [\boldsymbol{\psi}(t)\times]$ so that

$$\tilde{\mathbf{x}}(t) = \mathbf{x}(t) + \lambda(t)\mathbf{x}(t) + [\mathbf{x}(t)\times]\boldsymbol{\psi}(t) + \mathbf{g}(t) . \tag{11}$$

Here $\mathbf{A} = [\mathbf{a}\times]$ denotes the antisymmetric matrix having $\mathbf{a}$ as its axial vector, completely defined by the relations $A_{32} = a_1$, $A_{13} = a_2$, $A_{21} = a_3$. Second order terms in the small quantities, λ, ψ_1, ψ_2, and ψ_3 have been ignored. The corresponding transformation for velocities follows by differentiation

$$\dot{\tilde{\mathbf{x}}}(t) = \dot{\lambda}(t)\mathbf{x}(t) + (1+\lambda(t))\dot{\mathbf{x}}(t) + [\dot{\mathbf{x}}(t)\times]\boldsymbol{\psi}(t) + [\mathbf{x}(t)\times]\dot{\boldsymbol{\psi}}(t) + \dot{\mathbf{g}}(t) , \tag{12}$$

or after setting $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$, $\tilde{\mathbf{v}}(t) = \dot{\tilde{\mathbf{x}}}(t)$ and taking into account that the components of $\mathbf{v}(t)$ are also small quantities the corresponding transformation for velocities becomes

$$\tilde{\mathbf{v}}(t) = \mathbf{v}(t) + \dot{\lambda}(t)\mathbf{x}(t) + [\mathbf{x}(t)\times]\dot{\boldsymbol{\psi}}(t) + \dot{\mathbf{g}}(t) . \tag{13}$$

From the linearization of $\mathbf{x} = (1+\lambda)^{-1}\mathbf{R}^T(\tilde{\mathbf{x}} - \mathbf{g})$, the inverses of the transformations (13) become

$$\mathbf{x}(t) \approx \tilde{\mathbf{x}}(t) - [\tilde{\mathbf{x}}(t)\times]\boldsymbol{\psi}(t) - \lambda(t)\tilde{\mathbf{x}}(t) - \mathbf{g}(t) , \tag{14}$$

$$\mathbf{v}(t) \approx \tilde{\mathbf{v}}(t) - [\tilde{\mathbf{x}}(t)\times]\dot{\boldsymbol{\psi}}(t) - \dot{\lambda}(t)\tilde{\mathbf{x}}(t) - \dot{\mathbf{g}}(t) . \tag{15}$$

In the ITRF case the coordinate functions must conform with the model $\mathbf{x}(t) = \mathbf{x}_0 + (t-t_0)\mathbf{v}$ where $\mathbf{x}_0 = \mathbf{x}(t_0)$ and $\mathbf{v}$ are constants, so that after removing second order terms in λ, $\dot{\lambda}$, $\boldsymbol{\psi}$, $\dot{\boldsymbol{\psi}}$ and $\mathbf{v}$ the transformation into an other coordinate system becomes

$$\tilde{\mathbf{x}}(t) = \mathbf{x}_0 + (t-t_0)\mathbf{v} + \lambda(t)\mathbf{x}_0 + [\mathbf{x}_0\times]\boldsymbol{\psi}(t) + \mathbf{g}(t) \tag{16}$$

$$\tilde{\mathbf{v}}(t) = \mathbf{v} + \dot{\lambda}(t)\mathbf{x}_0 + [\mathbf{x}_0\times]\dot{\boldsymbol{\psi}}(t) + \dot{\mathbf{g}}(t) . \tag{17}$$

In order for the model to be preserved we must have $\tilde{\mathbf{x}}(t) = \tilde{\mathbf{x}}_0 + (t - t_0)\tilde{\mathbf{v}}$ and $\tilde{\mathbf{v}}$ constant. But $\tilde{\mathbf{v}}(t)$ is constant if $\dot{\lambda}(t)$, $\dot{\boldsymbol{\psi}}(t)$ and $\dot{\mathbf{g}}(t)$ are constant, i.e. when the transformation parameter functions have the form

$$\boldsymbol{\psi}(t) = \boldsymbol{\psi}_0 + (t - t_0)\dot{\boldsymbol{\psi}}\,, \quad \lambda(t) = \lambda_0 + (t - t_0)\dot{\lambda}\,, \quad \mathbf{g}(t) = \mathbf{g}_0 + (t - t_0)\dot{\mathbf{g}} \tag{18}$$

which are also sufficient conditions, with constant $\boldsymbol{\psi}_0 = \boldsymbol{\psi}(t_0)$, $\lambda_0 = \lambda(t_0)$, $\mathbf{g}_0 = \mathbf{g}(t_0)$, $\dot{\boldsymbol{\psi}}$, $\dot{\lambda}$ and $\dot{\mathbf{g}}$. Under these restrictions we arrive at the smaller class of model preserving transformations

$$\tilde{\mathbf{x}}(t) \approx \left(\mathbf{x}_0 + \lambda_0 \mathbf{x}_0 + [\mathbf{x}_0 \times]\boldsymbol{\psi}_0 + \mathbf{g}_0\right) + (t - t_0)\left(\mathbf{v} + \dot{\lambda}\mathbf{x}_0 + [\mathbf{x}_0 \times]\dot{\boldsymbol{\psi}} + \dot{\mathbf{g}}\right) \equiv \tilde{\mathbf{x}}_0 + (t - t_0)\tilde{\mathbf{v}} \tag{19}$$

$$\tilde{\mathbf{v}}(t) \approx \mathbf{v} + \dot{\lambda}\mathbf{x}_0 + [\mathbf{x}_0 \times]\dot{\boldsymbol{\psi}} + \dot{\mathbf{g}} \equiv \tilde{\mathbf{v}}\,. \tag{20}$$

Therefore the transformation laws for the ITRF model parameters under a change of coordinate system are

$$\tilde{\mathbf{x}}_0 \approx \mathbf{x}_0 + \lambda_0 \mathbf{x}_0 + [\mathbf{x}_0 \times]\boldsymbol{\psi}_0 + \mathbf{g}_0 \tag{21}$$

$$\tilde{\mathbf{v}} \approx \mathbf{v} + \dot{\lambda}\mathbf{x}_0 + [\mathbf{x}_0 \times]\dot{\boldsymbol{\psi}} + \dot{\mathbf{g}} \tag{22}$$

The corresponding inverse transformations are

$$\mathbf{x}_0 \approx \tilde{\mathbf{x}}_0 - [\tilde{\mathbf{x}}_0 \times]\boldsymbol{\psi}_0 - \lambda_0 \tilde{\mathbf{x}}_0 - \mathbf{g}_0 \tag{23}$$

$$\mathbf{v} \approx \tilde{\mathbf{v}} - [\tilde{\mathbf{x}}_0 \times]\dot{\boldsymbol{\psi}} - \dot{\lambda}\tilde{\mathbf{x}}_0 - \dot{\mathbf{g}} \tag{24}$$

In data analysis we deal not directly with $\mathbf{x}_0$ and $\mathbf{v}$, but rather with corrections $\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^{\mathrm{ap}}$ and $\delta\mathbf{v} = \mathbf{v} - \mathbf{v}^{\mathrm{ap}}$ to their approximate values $\mathbf{x}_0^{\mathrm{ap}}$ and $\mathbf{v}^{\mathrm{ap}}$. Assuming common approximate values (e.g. from a previous version of the ITRF) for both coordinate systems we obtain the transformation laws for corrections

$$\delta\tilde{\mathbf{x}}_0 = \delta\mathbf{x}_0 + \lambda_0 \mathbf{x}_0^{\mathrm{ap}} + [\mathbf{x}_0^{\mathrm{ap}} \times]\boldsymbol{\psi}_0 + \mathbf{g}_0\,, \tag{25}$$

$$\delta\tilde{\mathbf{v}} \approx \delta\mathbf{v} + \dot{\lambda}\mathbf{x}_0^{\mathrm{ap}} + [\mathbf{x}_0^{\mathrm{ap}} \times]\dot{\boldsymbol{\psi}} + \dot{\mathbf{g}}\,, \tag{26}$$

and their respective inverses

$$\delta\mathbf{x}_0 \approx \delta\tilde{\mathbf{x}}_0 - [\mathbf{x}_0^{\mathrm{ap}} \times]\boldsymbol{\psi}_0 - \lambda_0 \mathbf{x}_0^{\mathrm{ap}} - \mathbf{g}_0\,, \tag{27}$$

$$\delta\mathbf{v} \approx \delta\tilde{\mathbf{v}} - [\mathbf{x}_0^{\mathrm{ap}} \times]\dot{\boldsymbol{\psi}} - \dot{\lambda}\mathbf{x}_0^{\mathrm{ap}} - \dot{\mathbf{g}}\,. \tag{28}$$

4. The stacking problem and the corresponding inner constraints

As already mentioned the stacking problem deals with the estimation of initial coordinates and velocities of the ITRF model from "observed" coordinates at different epochs, assuming that they refer to a different coordinate system for each epoch. Stacking is performed for each technique separately but it has the same structure as the simultaneous stacking for data from all techniques, which is an alternative to the two-step approach of "stackings per technique" and "combination". For this reason we shall drop in our rotation the dependence to any particular technique and write the observation equations (7) in the simpler form

$$\mathbf{x}_i^k = \mathbf{x}_i(t_k) = (1+s_k)\mathbf{R}(\boldsymbol{\theta}_k)[\mathbf{x}_{i0} + (t_k - t_0)\mathbf{v}_i] + \mathbf{d}_k + \mathbf{e}_i^k \,, \tag{29}$$

where $\boldsymbol{\theta}_k$, $\mathbf{d}_k$, s_k are the transformation parameters from the ITRF reference system S_{ITRF} to the particular technique epoch system $S_T(t_k)$. In the linear approximation of the form of equation (16) the observation equations become

$$\mathbf{x}_i^k = \mathbf{x}_{i0} + (t_k - t_0)\mathbf{v}_i + s_k \mathbf{x}_{i0} + [\mathbf{x}_{i0}\times]\boldsymbol{\theta}_k + \mathbf{d}_k + \mathbf{e}_i^k \tag{30}$$

With the use of approximate values $\mathbf{x}_{0i}^{\text{ap}}$, $\mathbf{v}_i^{\text{ap}}$ ($s_k^{\text{ap}} = 0$, $\boldsymbol{\theta}_k^{\text{ap}} = \mathbf{0}$, $\mathbf{d}_k^{\text{ap}} = \mathbf{0}$), we may switch to the corrections $\delta\mathbf{x}_{i0}$, $\delta\mathbf{v}_i$ and the reduced observations $\delta\mathbf{x}_i^k = \mathbf{x}_i^k - (\mathbf{x}_i^k)^{\text{ap}} = \mathbf{x}_i^k - [\mathbf{x}_{i0}^{\text{ap}} + (t_k - t_0)\mathbf{v}_i^{\text{ap}}]$, so that the observation equations become

$$\delta\mathbf{x}_i^k = \delta\mathbf{x}_{i0} + (t_k - t_0)\delta\mathbf{v}_i + s_k \mathbf{x}_{i0}^{\text{ap}} + [\mathbf{x}_{i0}^{\text{ap}}\times]\boldsymbol{\theta}_k + \mathbf{d}_k + \mathbf{e}_i^k \tag{31}$$

We may rewrite the above observation equations in the more compact form

$$\delta\mathbf{x}_i^k = \begin{bmatrix} \mathbf{I} & (t_k - t_0)\mathbf{I} \end{bmatrix} \begin{bmatrix} \delta\mathbf{x}_{i0} \\ \delta\mathbf{v}_i \end{bmatrix} + \begin{bmatrix} [\mathbf{x}_{i0}^{ap}\times] & \mathbf{I} & \mathbf{x}_{i0}^{ap} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_k \\ \mathbf{d}_k \\ s_k \end{bmatrix} + \mathbf{e}_i^k = \mathbf{A}_{\mathbf{a}_i}\mathbf{a}_i + \mathbf{A}_{\mathbf{z}i}\mathbf{z}_k + \mathbf{e}_i^k \tag{32}$$

where $\mathbf{a}_i = \begin{bmatrix} \delta\mathbf{x}_{i0} \\ \delta\mathbf{v}_i \end{bmatrix}$ stands now for the station i ITRF parameters $\delta\mathbf{x}_{i0}$ and $\delta\mathbf{v}_i$, while $\mathbf{z}_k$ are the transformation parameters from the ITRF reference system to the one of each particular epoch t_k (and technique). Note that some stations may be missing for particular epochs.

For the determination of the inner constraints matrix we shall look at the transformation $\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{E}\mathbf{p}$ of the ITRF parameters under a model-preserving change of coordinates with rotation angles $\boldsymbol{\psi}(t) = \boldsymbol{\psi}_0 + (t - t_0)\dot{\boldsymbol{\psi}}$, scale parameter $\lambda(t) = \lambda_0 + (t - t_0)\dot{\lambda}$ and displacements $\mathbf{g}(t) = \mathbf{g}_0 + (t - t_0)\dot{\mathbf{g}}$. The relevant transformation equations for each station are given from equations (25) and (26) which can be combined into

$$\tilde{\mathbf{a}}_i = \begin{bmatrix} \delta\tilde{\mathbf{x}}_{0i} \\ \delta\tilde{\mathbf{v}}_i \end{bmatrix} \approx \begin{bmatrix} \delta\mathbf{x}_{0i} + [\mathbf{x}_{0i}^{\mathrm{ap}}\times]\boldsymbol{\psi}_0 + \lambda_0\mathbf{x}_{0i}^{\mathrm{ap}} + \mathbf{g}_0 \\ \delta\mathbf{v}_i + [\mathbf{x}_{0i}^{\mathrm{ap}}\times]\dot{\boldsymbol{\psi}} + \dot{\lambda}\mathbf{x}_{0i}^{\mathrm{ap}} + \dot{\mathbf{g}} \end{bmatrix} =$$

$$= \begin{bmatrix} \delta\mathbf{x}_{0i} \\ \delta\mathbf{v}_i \end{bmatrix} + \begin{bmatrix} [\mathbf{x}_{0i}^{\mathrm{ap}}\times] & \mathbf{I} & \mathbf{x}_{0i}^{\mathrm{ap}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{x}_{0i}^{\mathrm{ap}}\times] & \mathbf{I} & \mathbf{x}_{0i}^{\mathrm{ap}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi}_0 \\ \mathbf{g}_0 \\ \lambda_0 \\ \dot{\boldsymbol{\psi}} \\ \dot{\mathbf{g}} \\ \dot{\lambda} \end{bmatrix} \equiv \mathbf{a}_i + \mathbf{E}_{\mathbf{a}_i}\mathbf{p} \tag{33}$$

In order to see how the remaining "nuisance" parameters $\mathbf{z}_k$ transform under a change of coordinate system we will apply the inverse transformations (27), (28) to each station to obtain

$$\delta\mathbf{x}_{0i} \approx \delta\tilde{\mathbf{x}}_{0i} - [\mathbf{x}_{0i}^{ap}\times]\boldsymbol{\psi}_0 - \lambda_0\mathbf{x}_{0i}^{ap} - \mathbf{g}_0\,, \tag{34}$$

$$\delta\mathbf{v}_i \approx \delta\tilde{\mathbf{v}}_i - [\mathbf{x}_{0i}^{ap}\times]\dot{\boldsymbol{\psi}} - \dot{\lambda}\mathbf{x}_{0i}^{ap} - \dot{\mathbf{g}}\,, \tag{35}$$

and replace these in the observation equations (31) to obtain

$$\begin{aligned} \delta\mathbf{x}_i^k &= \delta\mathbf{x}_{i0} + (t_k - t_0)\delta\mathbf{v}_i + s_k\mathbf{x}_{i0}^{\mathrm{ap}} + [\mathbf{x}_{i0}^{\mathrm{ap}}\times]\boldsymbol{\theta}_k + \mathbf{d}_k + \mathbf{e}_i^k = \\ &= \delta\tilde{\mathbf{x}}_{0i} - [\mathbf{x}_{0i}^{ap}\times]\boldsymbol{\psi}_0 - \lambda_0\mathbf{x}_{0i}^{ap} - \mathbf{g}_0 + (t_k - t_0)\{\delta\tilde{\mathbf{v}}_i - [\mathbf{x}_{0i}^{ap}\times]\dot{\boldsymbol{\psi}} - \dot{\lambda}\mathbf{x}_{0i}^{ap} - \dot{\mathbf{g}}\} + \\ &\qquad + s_k\mathbf{x}_{i0}^{\mathrm{ap}} + [\mathbf{x}_{i0}^{\mathrm{ap}}\times]\boldsymbol{\theta}_k + \mathbf{d}_k + \mathbf{e}_i^k = \\ &= \delta\tilde{\mathbf{x}}_{0i} + (t_k - t_0)\delta\tilde{\mathbf{v}}_i + [\mathbf{x}_{i0}^{\mathrm{ap}}\times][\boldsymbol{\theta}_k - \boldsymbol{\psi}_0 - (t_k - t_0)\dot{\boldsymbol{\psi}}] + [s_k - \lambda_0 - (t_k - t_0)\dot{\lambda}]\mathbf{x}_{0i}^{ap} + \\ &\qquad + [\mathbf{d}_k - \mathbf{g}_0 - (t_k - t_0)\dot{\mathbf{g}}] + \mathbf{e}_i^k \end{aligned} \tag{36}$$

If we on the other hand we write directly the observation equations (31) in the new frame we get

$$\delta\mathbf{x}_i^k = \delta\tilde{\mathbf{x}}_{i0} + (t_k - t_0)\delta\tilde{\mathbf{v}}_i + \tilde{s}_k\mathbf{x}_{i0}^{\mathrm{ap}} + [\mathbf{x}_{i0}^{\mathrm{ap}}\times]\tilde{\boldsymbol{\theta}}_k + \tilde{\mathbf{d}}_k + \mathbf{e}_i^k \tag{37}$$

Comparison of the two above forms reveals the transformation rule for the remaining parameters

$$\tilde{\boldsymbol{\theta}}_k = \boldsymbol{\theta}_k - \boldsymbol{\psi}_0 - (t_k - t_0)\dot{\boldsymbol{\psi}}\,, \quad \tilde{\mathbf{d}}_k = \mathbf{d}_k - \mathbf{g}_0 - (t_k - t_0)\dot{\mathbf{g}}\,, \quad \tilde{s}_k = s_k - \lambda_0 - \dot{\lambda}(t_k - t_0)\,, \tag{38}$$

which combined give

$$\tilde{\mathbf{z}}_k = \begin{bmatrix} \tilde{\boldsymbol{\theta}}_k \\ \tilde{\mathbf{d}}_k \\ \tilde{s}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}_k - \boldsymbol{\psi}_0 - (t_k - t_0)\dot{\boldsymbol{\psi}} \\ \mathbf{d}_k - \mathbf{g}_0 - (t_k - t_0)\dot{\mathbf{g}} \\ s_k - \lambda_0 - \dot{\lambda}(t_k - t_0) \end{bmatrix} =$$

$$
=\begin{bmatrix}\boldsymbol{\theta}_k\\ \mathbf{d}_k\\ s_k\end{bmatrix}+\begin{bmatrix}-\mathbf{I} & \mathbf{0} & \mathbf{0} & -(t_k-t_0)\mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} & -(t_k-t_0)\mathbf{I} & \mathbf{0}\\ 0 & 0 & -1 & 0 & 0 & -(t_k-t_0)\mathbf{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{\psi}_0\\ \mathbf{g}_0\\ \lambda_0\\ \dot{\boldsymbol{\psi}}\\ \dot{\mathbf{g}}\\ \dot{\lambda}\end{bmatrix}\equiv \mathbf{z}_k+\mathbf{E}_{\mathbf{z}_k}\mathbf{p} \quad (39)
$$

Combining (33) for each station and (39) for each observation epoch t_k we arrive at the inner constraints matrix $\mathbf{E}$ as follows

$$
\begin{bmatrix}\tilde{\mathbf{a}}\\ \tilde{\mathbf{z}}\end{bmatrix}\equiv\begin{bmatrix}\tilde{\mathbf{a}}_1\\ \vdots\\ \tilde{\mathbf{a}}_N\\ -\\ \mathbf{z}_1\\ \vdots\\ \mathbf{z}_M\end{bmatrix}=\begin{bmatrix}\mathbf{a}_1+\mathbf{E}_{\mathbf{a}_1}\mathbf{p}\\ \vdots\\ \mathbf{a}_N+\mathbf{E}_{\mathbf{a}_N}\mathbf{p}\\ -\\ \mathbf{z}_1+\mathbf{E}_{\mathbf{z}_1}\mathbf{p}\\ \vdots\\ \mathbf{z}_M+\mathbf{E}_{\mathbf{z}_M}\mathbf{p}\end{bmatrix}=\begin{bmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_N\\ -\\ \mathbf{z}_1\\ \vdots\\ \mathbf{z}_M\mathbf{p}\end{bmatrix}+\begin{bmatrix}\mathbf{E}_{\mathbf{a}_1}\\ \vdots\\ \mathbf{E}_{\mathbf{a}_N}\\ -\\ \mathbf{E}_{\mathbf{z}_1}\\ \vdots\\ \mathbf{E}_{\mathbf{z}_M}\end{bmatrix}\mathbf{p}\equiv\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}+\begin{bmatrix}\mathbf{E}_{\mathbf{a}}\\ \mathbf{E}_{\mathbf{z}}\end{bmatrix}\mathbf{p}\equiv\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}+\mathbf{E}\mathbf{p}\,. \qquad (40)
$$

The inner constraints themselves become

$$
\mathbf{0}=\mathbf{E}^T\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}=\begin{bmatrix}\mathbf{E}_{\mathbf{a}}^T & \mathbf{E}_{\mathbf{z}}^T\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}=\mathbf{E}_{\mathbf{x}}^T\mathbf{x}+\mathbf{E}_{\mathbf{a}}^T\mathbf{a}=
$$

$$
=\begin{bmatrix}\mathbf{E}_{\mathbf{a}_1}^T & \cdots & \mathbf{E}_{\mathbf{a}_N}^T\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_N\end{bmatrix}+\begin{bmatrix}\mathbf{E}_{\mathbf{z}_1}^T & \cdots & \mathbf{E}_{\mathbf{z}_M}^T\end{bmatrix}\begin{bmatrix}\mathbf{z}_1\\ \vdots\\ \mathbf{z}_M\end{bmatrix}=
$$

$$
=\sum_{i=1}^{N}\mathbf{E}_{\mathbf{a}_i}^T\mathbf{a}_i+\sum_{k=1}^{M}\mathbf{E}_{\mathbf{z}_k}^T\mathbf{z}_k=\begin{bmatrix}-\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{x}_{0i}-\sum_{k=1}^{M}\boldsymbol{\theta}_k\\ \sum_{i=1}^{N}\delta\mathbf{x}_{0i}-\sum_{k=1}^{M}\mathbf{d}_k\\ \sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T\delta\mathbf{x}_{0i}-\sum_{k=1}^{M}s_k\\ -\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{v}_i-\sum_{k=1}^{M}(t_k-t_0)\boldsymbol{\theta}_k\\ \sum_{i=1}^{N}\delta\mathbf{v}_i-\sum_{k=1}^{M}(t_k-t_0)\mathbf{d}_k\\ \sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T\delta\mathbf{v}_i-\sum_{k=1}^{M}(t_k-t_0)s_k\end{bmatrix}=\mathbf{0}\,. \qquad (41)
$$

The above inner constraints can be separated into 6 groups:

The 3 *orientation* inner constraints

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{x}_{0i}+\sum_{k=1}^{M}\boldsymbol{\theta}_k=\mathbf{0}\,, \tag{42a}$$

the 3 *translation* inner constraints

$$\sum_{i=1}^{N}\delta\mathbf{x}_{0i}-\sum_{k=1}^{M}\mathbf{d}_k=\mathbf{0}\,, \tag{42b}$$

the *scale* inner constraint

$$\sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T\delta\mathbf{x}_{0i}-\sum_{k=1}^{M}s_k=0\,, \tag{42c}$$

the 3 *orientation rate* inner constraints

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{v}_i+\sum_{k=1}^{M}(t_k-t_0)\boldsymbol{\theta}_k=\mathbf{0}\,, \tag{42d}$$

the 3 *translation rate* inner constraints

$$\sum_{i=1}^{N}\delta\mathbf{v}_i-\sum_{k=1}^{M}(t_k-t_0)\mathbf{d}_k=\mathbf{0}\,, \tag{42e}$$

and the *scale rate* inner constraint

$$\sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T\delta\mathbf{v}_i-\sum_{k=1}^{M}(t_k-t_0)s_k=0\,. \tag{42f}$$

These are the "total" inner constraints leading to a minimum norm solution satisfying

$$\begin{aligned}\left\|\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}\right\|^2&=\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}^T\begin{bmatrix}\mathbf{a}\\ \mathbf{z}\end{bmatrix}=\mathbf{a}^T\mathbf{a}+\mathbf{z}^T\mathbf{z}=.\\ &=\sum_i(\delta\mathbf{x}_{i0}^T\delta\mathbf{x}_{i0}+\delta\mathbf{v}_i^T\delta\mathbf{v}_i)+\sum_k(\boldsymbol{\theta}_k^T\boldsymbol{\theta}_k+\mathbf{d}_k^T\mathbf{d}_k+s_k^2)=\min\end{aligned} \tag{43}$$

Two sets of partial inner constraints derive from the above total ones:

Considering only coordinates and velocities the partial inner constraints $\mathbf{E}_{\mathbf{a}}^T\mathbf{a}=\Sigma_{i=1}^{N}\mathbf{E}_{\mathbf{a}_i}^T\mathbf{a}_i=\mathbf{0}$ satisfying $\sum_i(\delta\mathbf{x}_{i0}^T\delta\mathbf{x}_{i0}+\delta\mathbf{v}_i^T\delta\mathbf{v}_i)=\min$, take the form

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{x}_{0i}=\mathbf{0}, \qquad \sum_{i=1}^{N}\delta\mathbf{x}_{0i}=\mathbf{0}, \qquad \sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^{T}\delta\mathbf{x}_{0i}=0, \tag{44a}$$

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{v}_{i}=\mathbf{0}, \qquad \sum_{i=1}^{N}\delta\mathbf{v}_{i}=\mathbf{0}, \qquad \sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^{T}\delta\mathbf{v}_{i}=0. \tag{44b}$$

Note that the three first constraints (44a) are the familiar ones for non-deforming networks.

Considering only transformation parameters the partial inner constraints

$\mathbf{E}_{\mathbf{z}}^{T}\mathbf{z}=\sum_{k=1}^{M}\mathbf{E}_{\mathbf{z}_k}^{T}\mathbf{z}_k=\mathbf{0}$ satisfying $\sum_k(\boldsymbol{\theta}_k^T\boldsymbol{\theta}_k+\mathbf{d}_k^T\mathbf{d}_k+s_k^2)=\min$, become

$$\sum_{k=1}^{M}\boldsymbol{\theta}_k=\mathbf{0}, \qquad \sum_{k=1}^{M}\mathbf{d}_k=\mathbf{0}, \qquad \sum_{k=1}^{M}s_k=0, \tag{45a}$$

$$\sum_{k=1}^{M}(t_k-t_0)\boldsymbol{\theta}_k=\mathbf{0}, \qquad \sum_{k=1}^{M}(t_k-t_0)\mathbf{d}_k=\mathbf{0}, \qquad \sum_{k=1}^{M}(t_k-t_0)s_k=0. \tag{45b}$$

We have derived here the most complete set of constraints, referring to the rank deficiency with respect to scale, position and orientation. In the particular case that scale and/or position (e.g., geocentric reference system from SLR) can be determined from the available data the corresponding constraints must be removed. The above constraints (42), (44), (45) apply to the stacking for any particular space technique T, where the index T has been removed for simplicity from $\delta\mathbf{x}_{T,0i}$, $\delta\mathbf{v}_{T,i}$, $\boldsymbol{\theta}_{T,k}$, $\mathbf{d}_{T,k}$, $s_{T,k}$. The same holds true for the simultaneous stacking of all space techniques in the one-step approach to the ITRF formulation, with parameters $\delta\mathbf{x}_{0i}$, $\delta\mathbf{v}_i$, $\boldsymbol{\theta}_{T,k}$, $\mathbf{d}_{T,k}$, $s_{T,k}$. In this case however, an additional summation $\sum_{T=1}^{K}$ over all the K space techniques must be applied to the constraints (42) and (45) and their corresponding optimality minimization criteria, replacing the summation $\sum_{k=1}^{M}$ with the double summation $\sum_{T=1}^{K}\sum_{k=1}^{M}$.

5. The combination problem and the corresponding inner constraints

In the combination problem the estimates of initial coordinates and velocities $\delta\mathbf{x}_{T0i}$, $\delta\mathbf{v}_{Ti}$, $i=1,\ldots,n_T$, obtained in the previous stacking step for each technique T separately, are used as "observations" in order to obtain common estimates $\delta\mathbf{x}_{0i}$, $\delta\mathbf{v}_i$ of the corresponding ITRF model parameters. As already shown, in order to preserve the linear-in-time coordinate model the transformation parameters from the ITRF reference system S_{ITRF} to the system S_T of each technique must have the form

$$\boldsymbol{\theta}_T(t) = \boldsymbol{\theta}_{T0} + (t - t_0)\dot{\boldsymbol{\theta}}_T, \qquad \mathbf{d}_T(t) = \mathbf{d}_{T0} + (t - t_0)\dot{\mathbf{d}}_T, \quad s_T(t) = s_{T0} + (t - t_0)\dot{s}_T. \tag{46}$$

Applying equations (25) and (26) to each station i with $\delta\tilde{\mathbf{x}}_0 \to \delta\mathbf{x}_{T0i}$, $\delta\tilde{\mathbf{v}} \to \delta\mathbf{v}_{Ti}$, $\delta\mathbf{x}_0 \to \delta\mathbf{x}_{0i}$, $\delta\mathbf{v} \to \delta\mathbf{v}_i$, $\mathbf{x}_0^{\text{ap}} \to \mathbf{x}_{0i}^{\text{ap}}$ and $\boldsymbol{\psi}_0 \to \boldsymbol{\theta}_{T0}$, $\mathbf{g}_0 \to \mathbf{d}_{T0}$, $\lambda_0 \to s_{T0}$, $\dot{\boldsymbol{\psi}} \to \dot{\boldsymbol{\theta}}_T$, $\dot{\mathbf{g}} \to \dot{\mathbf{d}}_T$, $\dot{\lambda} \to \dot{s}_T$, and taking into account the presence of observational noise we obtain the observation equations

$$\delta\mathbf{x}_{T0i} = \delta\mathbf{x}_{0i} + [\mathbf{x}_{0i}^{\text{ap}}\times]\boldsymbol{\theta}_{T0} + s_{T0}\mathbf{x}_{0i}^{\text{ap}} + \mathbf{d}_{T0} + \mathbf{e}_{\mathbf{x}_{T0i}}, \tag{47}$$

$$\delta\mathbf{v}_{Ti} = \delta\mathbf{v}_i + [\mathbf{x}_{0i}^{\text{ap}}\times]\dot{\boldsymbol{\theta}}_T + \dot{s}_T\mathbf{x}_{0i}^{\text{ap}} + \dot{\mathbf{d}}_T + \mathbf{e}_{\mathbf{v}_{Ti}}, \tag{48}$$

On the basis of the above observation equation for all techniques and all stations in each technique estimates will be obtained for the unknown parameters $\delta\mathbf{x}_{0i}$, $\delta\mathbf{v}_i$, as well as of the "nuisance" parameters $\boldsymbol{\theta}_{T0}$, $\dot{\boldsymbol{\theta}}_T$, $\mathbf{d}_{T0}$, $\dot{\mathbf{d}}_T$, s_{T0}, $\dot{s}_T$, for all techniques.

In order to obtain the inner constraints for the removal of the rank deficiency, we introduce an arbitrary model-preserving coordinate transformation with parameters $\boldsymbol{\psi}_0$, $\mathbf{g}_0$, λ_0, $\dot{\boldsymbol{\psi}}$, $\dot{\mathbf{g}}$, $\dot{\lambda}$. Applying the inverse transformations (27) and (28) for each station, namely

$$\delta\mathbf{x}_{0i} \approx \delta\tilde{\mathbf{x}}_{0i} - [\mathbf{x}_{0i}^{\text{ap}}\times]\boldsymbol{\psi}_0 - \lambda_0\mathbf{x}_{0i}^{\text{ap}} - \mathbf{g}_0, \tag{49}$$

$$\delta\mathbf{v}_i \approx \delta\tilde{\mathbf{v}}_i - [\mathbf{x}_{0i}^{\text{ap}}\times]\dot{\boldsymbol{\psi}} - \dot{\lambda}\mathbf{x}_{0i}^{\text{ap}} - \dot{\mathbf{g}}. \tag{50}$$

to the observation equations (47), (48) we obtain

$$\begin{aligned}\delta\mathbf{x}_{T0i} &= \delta\mathbf{x}_{0i} + [\mathbf{x}_{0i}^{\text{ap}}\times]\boldsymbol{\theta}_{T0} + s_{T0}\mathbf{x}_{0i}^{\text{ap}} + \mathbf{d}_{T0} + \mathbf{e}_{\mathbf{x}_{T0i}} = \\ &= \delta\tilde{\mathbf{x}}_{0i} - [\mathbf{x}_{0i}^{\text{ap}}\times]\boldsymbol{\psi}_0 - \lambda_0\mathbf{x}_{0i}^{\text{ap}} - \mathbf{g}_0 + [\mathbf{x}_{0i}^{\text{ap}}\times]\boldsymbol{\theta}_{T0} + s_{T0}\mathbf{x}_{0i}^{\text{ap}} + \mathbf{d}_{T0} + \mathbf{e}_{\mathbf{x}_{T0i}} = \\ &= \delta\tilde{\mathbf{x}}_{0i} + [\mathbf{x}_{0i}^{\text{ap}}\times](\boldsymbol{\theta}_{T0} - \boldsymbol{\psi}_0) + (s_{T0} - \lambda_0)\mathbf{x}_{0i}^{\text{ap}} + (\mathbf{d}_{T0} - \mathbf{g}_0) + \mathbf{e}_{\mathbf{x}_{T0i}},\end{aligned} \tag{51}$$

$$\begin{aligned}\delta\mathbf{v}_{Ti} &= \delta\mathbf{v}_i + [\mathbf{x}_{0i}^{\text{ap}}\times]\dot{\boldsymbol{\theta}}_T + \dot{s}_T\mathbf{x}_{0i}^{\text{ap}} + \dot{\mathbf{d}}_T + \mathbf{e}_{\mathbf{v}_{Ti}} = \\ &= \delta\tilde{\mathbf{v}}_i - [\mathbf{x}_{0i}^{\text{ap}}\times]\dot{\boldsymbol{\psi}} - \dot{\lambda}\mathbf{x}_{0i}^{\text{ap}} - \dot{\mathbf{g}} + [\mathbf{x}_{0i}^{\text{ap}}\times]\dot{\boldsymbol{\theta}}_T + \dot{s}_T\mathbf{x}_{0i}^{ap} + \dot{\mathbf{d}}_T + \mathbf{e}_{\mathbf{v}_{Ti}} = \\ &= \delta\tilde{\mathbf{v}}_i + [\mathbf{x}_{0i}^{\text{ap}}\times](\dot{\boldsymbol{\theta}}_T - \dot{\boldsymbol{\psi}}) + (\dot{s}_T - \dot{\lambda})\mathbf{x}_{0i}^{\text{ap}} + (\dot{\mathbf{d}}_T - \dot{\mathbf{g}}) + \mathbf{e}_{\mathbf{v}_{Ti}},\end{aligned} \tag{52}$$

If on the other hand we write the observation equations directly into the new system as

$$\delta\mathbf{x}_{T0i} = \delta\tilde{\mathbf{x}}_{0i} + [\mathbf{x}_{0i}^{\text{ap}}\times]\tilde{\boldsymbol{\theta}}_{T0} + \tilde{s}_{T0}\mathbf{x}_{0i}^{\text{ap}} + \tilde{\mathbf{d}}_{T0} + \mathbf{e}_{\mathbf{x}_{T0i}}, \tag{53}$$

$$\delta\mathbf{v}_{Ti} = \delta\tilde{\mathbf{v}}_i + [\mathbf{x}_{0i}^{\text{ap}}\times]\dot{\tilde{\boldsymbol{\theta}}}_T + \dot{\tilde{s}}_T\mathbf{x}_{0i}^{\text{ap}} + \dot{\tilde{\mathbf{d}}}_T + \mathbf{e}_{\mathbf{v}_{Ti}}, \tag{54}$$

comparison of (51) with (53) and of (52) with (54) reveals the transformation laws

$$\tilde{\boldsymbol{\theta}}_{T0} = \boldsymbol{\theta}_{T0} - \boldsymbol{\psi}_0 , \qquad \tilde{\mathbf{d}}_{T0} = \mathbf{d}_{T0} - \mathbf{g}_0 , \qquad \tilde{s}_{T0} = s_{T0} - \lambda_0 , \tag{55a}$$

$$\dot{\tilde{\boldsymbol{\theta}}}_{T} = \dot{\boldsymbol{\theta}}_{T} - \dot{\boldsymbol{\psi}} , \qquad \dot{\tilde{\mathbf{d}}}_{T} = \dot{\mathbf{d}}_{T} - \dot{\mathbf{g}} , \qquad \dot{\tilde{s}}_{T} = \dot{s}_{T} - \dot{\lambda} . \tag{55b}$$

These together with the direct transformations (25) and (26) applied to each station, namely

$$\delta\tilde{\mathbf{x}}_{0i} = \delta\mathbf{x}_{0i} + [\mathbf{x}_{0i}^{\mathrm{ap}}\times]\boldsymbol{\psi}_0 + \lambda_0\mathbf{x}_{0i}^{\mathrm{ap}} + \mathbf{g}_0 , \tag{56}$$

$$\delta\tilde{\mathbf{v}}_{i} \approx \delta\mathbf{v}_{i} + [\mathbf{x}_{0i}^{\mathrm{ap}}\times]\dot{\boldsymbol{\psi}} + \dot{\lambda}\mathbf{x}_{0i}^{\mathrm{ap}} + \dot{\mathbf{g}} , \tag{57}$$

form the basis for the determination of the inner constraint matrix. Combining (56) and (57) we have

$$\tilde{\mathbf{a}}_i \equiv \begin{bmatrix} \delta\tilde{\mathbf{x}}_{0i} \\ \delta\tilde{\mathbf{v}}_i \end{bmatrix} \approx \begin{bmatrix} \delta\mathbf{x}_{0i} + [\mathbf{x}_{0i}^{\mathrm{ap}}\times]\boldsymbol{\psi}_0 + \lambda_0\mathbf{x}_{0i}^{\mathrm{ap}} + \mathbf{g}_0 \\ \delta\mathbf{v}_i + [\mathbf{x}_{0i}^{\mathrm{ap}}\times]\dot{\boldsymbol{\psi}} + \dot{\lambda}\mathbf{x}_{0i}^{\mathrm{ap}} + \dot{\mathbf{g}} \end{bmatrix} =$$

$$= \begin{bmatrix} \delta\mathbf{x}_{0i} \\ \delta\mathbf{v}_i \end{bmatrix} + \begin{bmatrix} [\mathbf{x}_{0i}^{\mathrm{ap}}\times] & \mathbf{I} & \mathbf{x}_{0i}^{\mathrm{ap}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{x}_{0i}^{\mathrm{ap}}\times] & \mathbf{I} & \mathbf{x}_{0i}^{\mathrm{ap}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi}_0 \\ \mathbf{g}_0 \\ \lambda_0 \\ \dot{\boldsymbol{\psi}} \\ \dot{\mathbf{g}} \\ \dot{\lambda} \end{bmatrix} \equiv \mathbf{a}_i + \mathbf{E}_{\mathbf{a}_i}\mathbf{p} . \tag{58}$$

and jointly for all stations

$$\tilde{\mathbf{a}} \equiv \begin{bmatrix} \tilde{\mathbf{a}}_1 \\ \vdots \\ \tilde{\mathbf{a}}_N \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_N \end{bmatrix} + \begin{bmatrix} \mathbf{E}_{\mathbf{a}_1} \\ \vdots \\ \mathbf{E}_{\mathbf{a}_N} \end{bmatrix} \mathbf{p} \equiv \mathbf{E}_{\mathbf{a}}\mathbf{p} \tag{59}$$

Combining (55a) and (55b) we have

$$\tilde{\mathbf{z}}_T \equiv \begin{bmatrix} \tilde{\boldsymbol{\theta}}_{T0} \\ \tilde{\mathbf{d}}_{T0} \\ \tilde{s}_{T0} \\ \dot{\tilde{\boldsymbol{\theta}}}_{T} \\ \dot{\tilde{\mathbf{d}}}_{T} \\ \dot{\tilde{s}}_{T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}_{T0} \\ \mathbf{d}_{T0} \\ s_{T0} \\ \dot{\boldsymbol{\theta}}_{T} \\ \dot{\mathbf{d}}_{T} \\ \dot{s}_{T} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\psi}_0 \\ \mathbf{g}_0 \\ \lambda_0 \\ \dot{\boldsymbol{\psi}} \\ \dot{\mathbf{g}} \\ \dot{\lambda} \end{bmatrix} = \mathbf{z}_T - \mathbf{p} \equiv \mathbf{z}_T + \mathbf{E}_{\mathbf{z}_T}\mathbf{p} , \qquad (\mathbf{E}_{\mathbf{z}_T} = -\mathbf{I}). \tag{60}$$

or jointly for all techniques

$$\tilde{\mathbf{z}} \equiv \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_K \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_K \end{bmatrix} + \begin{bmatrix} \mathbf{E}_{\mathbf{z}_1} \\ \vdots \\ \mathbf{E}_{\mathbf{z}_K} \end{bmatrix} \mathbf{p} \equiv \mathbf{E}_{\mathbf{z}} \mathbf{p} \tag{61}$$

For all the unknown parameters for the stations $i = 1, 2, ..., N$ and techniques $T = 1, 2, ..., K$, the transformation law is

$$\begin{bmatrix} \tilde{\mathbf{a}} \\ \tilde{\mathbf{z}} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{z} \end{bmatrix} + \begin{bmatrix} \mathbf{E}_{\mathbf{a}} \\ \mathbf{E}_{\mathbf{z}} \end{bmatrix} \mathbf{p} \equiv \begin{bmatrix} \mathbf{a} \\ \mathbf{z} \end{bmatrix} + \mathbf{E}\mathbf{p}\,. \tag{62}$$

and the inner constraints become

$$\mathbf{0} = \mathbf{E}^T \begin{bmatrix} \mathbf{a} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{E}_{\mathbf{a}}^T & \mathbf{E}_{\mathbf{z}}^T \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{z} \end{bmatrix} = \mathbf{E}_{\mathbf{a}}^T \mathbf{a} + \mathbf{E}_{\mathbf{z}}^T \mathbf{z} = \sum_{i=1}^{N} \mathbf{E}_{\mathbf{a}_i}^T \mathbf{a}_i + \sum_{T=1}^{K} \mathbf{E}_{\mathbf{z}_T}^T \mathbf{z}_T =$$

$$= \begin{bmatrix} -\sum_{i=1}^{N} [\mathbf{x}_{0i}^{\mathrm{ap}} \times] \delta \mathbf{x}_{0i} - \sum_{T=1}^{K} \boldsymbol{\theta}_{T0} \\ \sum_{i=1}^{N} \delta \mathbf{x}_{0i} - \sum_{T=1}^{K} \mathbf{d}_{T0} \\ \sum_{i=1}^{N} (\mathbf{x}_{0i}^{\mathrm{ap}})^T \delta \mathbf{x}_{0i} - \sum_{T=1}^{K} s_{T0} \\ -\sum_{i=1}^{N} [\mathbf{x}_{0i}^{\mathrm{ap}} \times] \delta \mathbf{v}_i - \sum_{T=1}^{K} \dot{\boldsymbol{\theta}}_T \\ \sum_{i=1}^{N} \delta \mathbf{v}_i - \sum_{T=1}^{K} \dot{\mathbf{d}}_T \\ \sum_{i=1}^{N} (\mathbf{x}_{0i}^{\mathrm{ap}})^T \delta \mathbf{v}_i - \sum_{T=1}^{K} \dot{s}_T \end{bmatrix} = \mathbf{0}\,. \tag{63}$$

The above inner constraints can be separated into 6 groups:

The 3 *orientation* inner constraints

$$\sum_{i=1}^{N} [\mathbf{x}_{0i}^{\mathrm{ap}} \times] \delta \mathbf{x}_{0i} + \sum_{T=1}^{K} \boldsymbol{\theta}_{T0} = \mathbf{0}\,, \tag{64a}$$

the 3 *translation* inner constraints

$$\sum_{i=1}^{N} \delta \mathbf{x}_{0i} - \sum_{T=1}^{K} \mathbf{d}_{T0} = \mathbf{0}\,, \tag{64b}$$

the *scale* inner constraint

$$\sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T \delta\mathbf{x}_{0i} - \sum_{T=1}^{K} s_{T0} = 0\,, \tag{64c}$$

the 3 *orientation rate* inner constraints

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{v}_i + \sum_{T=1}^{K}\dot{\boldsymbol{\theta}}_T = \mathbf{0}\,, \tag{64d}$$

the 3 *translation rate* inner constraints

$$\sum_{i=1}^{N}\delta\mathbf{v}_i - \sum_{T=1}^{K}\dot{\mathbf{d}}_T = \mathbf{0}\,, \tag{64e}$$

and the *scale rate* inner constraint

$$\sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T \delta\mathbf{v}_i - \sum_{T=1}^{K}\dot{s}_T = 0\,. \tag{64f}$$

These are the "total" inner constraints leading to a minimum norm solution satisfying

$$\left\|\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix}\right\|^2 = \begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix}^T\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \mathbf{x}^T\mathbf{x} + \mathbf{z}^T\mathbf{z} =$$
$$= \sum_i(\delta\mathbf{x}_{i0}^T\delta\mathbf{x}_{i0} + \delta\mathbf{v}_i^T\delta\mathbf{v}_i) + \sum_T(\boldsymbol{\theta}_{T0}^T\boldsymbol{\theta}_{T0} + \mathbf{d}_{T0}^T\mathbf{d}_{T0} + s_{T0}^2 + \dot{\boldsymbol{\theta}}_T^T\dot{\boldsymbol{\theta}}_T + \dot{\mathbf{d}}_T^T\dot{\mathbf{d}}_T + \dot{s}_T^2) = \min \tag{65}$$

Considering only coordinates and velocities, the partial inner constraints $\mathbf{E}_{\mathbf{a}}^T\mathbf{a} = \Sigma_{i=1}^N \mathbf{E}_{\mathbf{a}_i}^T\mathbf{a}_i = \mathbf{0}$ satisfying $\Sigma_i(\delta\mathbf{x}_{i0}^T\delta\mathbf{x}_{i0} + \delta\mathbf{v}_i^T\delta\mathbf{v}_i) = \min$ take the familiar form

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{x}_{0i} = \mathbf{0}\,, \qquad \sum_{i=1}^{N}\delta\mathbf{x}_{0i} = \mathbf{0}\,, \qquad \sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T\delta\mathbf{x}_{0i} = 0 \tag{66a}$$

$$\sum_{i=1}^{N}[\mathbf{x}_{0i}^{\mathrm{ap}}\times]\delta\mathbf{v}_i = \mathbf{0}\,, \qquad \sum_{i=1}^{N}\delta\mathbf{v}_i = \mathbf{0}\,, \qquad \sum_{i=1}^{N}(\mathbf{x}_{0i}^{\mathrm{ap}})^T\delta\mathbf{v}_i = 0\,. \tag{66b}$$

Considering only transformation parameters, the partial inner constrains $\mathbf{E}_{\mathbf{z}}^T\mathbf{z} = \Sigma_{T=1}^K \mathbf{E}_{\mathbf{z}_T}^T\mathbf{z}_T = \mathbf{0}$ satisfying $\Sigma_T(\boldsymbol{\theta}_{T0}^T\boldsymbol{\theta}_{T0} + \mathbf{d}_{T0}^T\mathbf{d}_{T0} + s_{T0}^2 + \dot{\boldsymbol{\theta}}_T^T\dot{\boldsymbol{\theta}}_T + \dot{\mathbf{d}}_T^T\dot{\mathbf{d}}_T + \dot{s}_T^2) = \min$, become

$$\sum_{T=1}^{K}\boldsymbol{\theta}_{T0} = \mathbf{0}\,, \qquad \sum_{T=1}^{K}\mathbf{d}_{T0} = \mathbf{0}\,, \qquad \sum_{T=1}^{K}s_{T0} = 0\,, \tag{67a}$$

$$\sum_{T=1}^{K} \dot{\boldsymbol{\theta}}_T = \mathbf{0}\,, \qquad \sum_{T=1}^{K} \dot{\mathbf{d}}_T = \mathbf{0}\,, \qquad \sum_{T=1}^{K} \dot{s}_T = 0\,. \tag{67b}$$

Again when scale and/or position are provided by the available data the corresponding constraints must be removed.

6. Optimal minimal constraints based on kinematic principles

The above inner and partial constraints based on algebraic considerations are not the only possible ones. In fact any set of d constraints, equal in number to the rank deficiency, which provide a unique solution can serve a set of "minimal constraints" (Dermanis, 1995, 2000, 2008). The solution $\hat{\mathbf{x}}_C$ from any set of minimal constraints $\mathbf{C}^T\mathbf{x} = \mathbf{0}$ and its covariance matrix can be easily transformed to the solution of the (partial) inner constraints $\hat{\mathbf{x}}_+$ and vice-versa. Note that any solution $\hat{\mathbf{x}}$ of the normal equations can be transformed into the solution $\hat{\mathbf{x}}_C = \hat{\mathbf{x}} + \mathbf{E}\mathbf{p}$ satisfying the minimal constraints by adding an appropriate term $\mathbf{E}\mathbf{p} \in \mathbf{R}(\mathbf{E}) = N(\mathbf{A})$. Then $\mathbf{C}^T\hat{\mathbf{x}}_C = \mathbf{C}^T\hat{\mathbf{x}} + \mathbf{C}^T\mathbf{E}\mathbf{p} = \mathbf{0}$ yielding $\mathbf{p} = -(\mathbf{C}^T\mathbf{E})^{-1}\mathbf{C}^T\hat{\mathbf{x}}$ and

$$\hat{\mathbf{x}}_C = \hat{\mathbf{x}} - \mathbf{E}(\mathbf{C}^T\mathbf{E})^{-1}\mathbf{C}^T\hat{\mathbf{x}}\,.$$

The transformation of any solution $\hat{\mathbf{x}}$ to the inner solution $\hat{\mathbf{x}}_+$ is just the special case $\hat{\mathbf{x}}_+ = \hat{\mathbf{x}} - \mathbf{E}(\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\hat{\mathbf{x}}$, with $\mathbf{C} = \mathbf{E}$. Therefore the required transformations are

$$\hat{\mathbf{x}}_+ = \hat{\mathbf{x}}_C - \mathbf{E}(\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\hat{\mathbf{x}}_C \tag{68}$$

$$\hat{\mathbf{x}}_C = \hat{\mathbf{x}}_+ - \mathbf{E}(\mathbf{C}^T\mathbf{E})^{-1}\mathbf{C}^T\hat{\mathbf{x}}_+\,. \tag{69}$$

A particular choice, which may resolve a lot of inherent difficulties in ITRF implementation, are the "trivial" minimal constraints where d coordinates distributed over at least 3 points are fixed to their approximate values (i.e. their unknown coordinate corrections are set to zero). By rearranging the order of the unknowns $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$, the trivial constraints have the form $\mathbf{x}_2 = \mathbf{0}$. Their main advantage is that they can be applied a priori at the level of the observation equations and thus lead to uniquely solvable normal equations. Their disadvantage is that numerical instability may result if the d coordinates are not properly selected.

A different approach in seeking minimal constraints for use in ITRF implementation is the use of kinematic considerations by asking the question what is a best choice for attaching a reference system (i.e. a smooth with respect to time set of instantaneous reference systems, one for every epoch t) to a deformable network of points. The answer is easy in loose terms: we want a reference system such that the apparent motion of the network stations, expressed through their coordinate and apparent velocity functions, $\mathbf{x}_i(t)$, $(d\mathbf{x}_i/dt)(t)$ are as small as possible. The question thus reduces in choosing criteria quantifying the expression "as small as possible". Fortunately this problem has been addressed a long time ago by Tisserand in his famous "Mécanique Céleste" (Tisserand, 1889, see also Munk & MacDonald, 1960), though not for a discrete set of points but rather for the continuum of the material points of the earth. He proposed a geocentric reference system with directions of the axes such that the relative angular momentum $\mathbf{h}_R$ vanishes, a condition which satisfies the minimization of the relative kinetic energy T_R:

$$\mathbf{x}_C \equiv \frac{1}{M}\int_E \mathbf{x}\,dm = \mathbf{0}, \quad \mathbf{h}_R = \int_E [\mathbf{x}\times]\frac{d\mathbf{x}}{dt} = \mathbf{0} \quad \Leftrightarrow \quad T_R = \int_E \left(\frac{d\mathbf{x}}{dt}\right)^T \frac{d\mathbf{x}}{dt}dm = \min . \qquad (70)$$

Above dm is the mass element of the earth, M its total mass and $\mathbf{x}_C$ the coordinates of the center of mass of the earth. Since $\mathbf{x}_C(t)$, $\mathbf{h}_R(t)$ and $T_R(t)$ are functions of time, the above relations must hold for every epoch t. The Tisserand condition of vanishing angular momentum does not lead to a unique reference system with respect to its orientation, but rather to a family of Tisserant axes. All the members of this family vary in orientation by an orthogonal rotation matrix which is time independent. Any member of the family is thus uniquely defined by its orientation at an initial epoch t_0 which must be independently chosen.

We may imitate these conditions by considering our network as a set of mass points with equal masses $m_i = m$, which we may take to be equal to one $m_i = 1$, without loss of generality. In this case the "discrete Tisserant conditions" become

$$\mathbf{x}_B(t) \equiv \frac{1}{N}\sum_i \mathbf{x}_i(t) = \mathbf{c}_T = \text{const.}, \qquad \mathbf{h}_R(t) = \sum_i [\mathbf{x}_i(t)\times]\frac{d\mathbf{x}_i}{dt}(t) = \mathbf{0}, \qquad (71)$$

where the second one minimizes $T_R(t) = \sum_i \left(\frac{d\mathbf{x}_i}{dt}(t)\right)^T \frac{d\mathbf{x}_i}{dt}(t) = \min$.

The first condition is putting the origin at the network barycenter $\mathbf{x}_B(t)$ and the second controls the variation of the direction of the axes with time. For the ITRF model where $\mathbf{x}_i(t) = \mathbf{x}_{0i} + (t-t_0)\mathbf{v}_i$ and $d\mathbf{x}_i/dt = \mathbf{v}_i$, the above conditions take the corresponding forms

$$\mathbf{x}_B(t) \equiv \frac{1}{N}\sum_i \mathbf{x}_{0i} + (t-t_0)\frac{1}{N}\sum_i \mathbf{v}_i = \mathbf{c}_T \qquad (72)$$

$$\mathbf{h}_R = \sum_i [\mathbf{x}_{0i}\times]\mathbf{v}_i + (t-t_0)\sum_i [\mathbf{v}_i\times]\mathbf{v}_i = \mathbf{0} \qquad (73)$$

which must be satisfied for any epoch t. Since $[\mathbf{v}_i\times]\mathbf{v}_i=\mathbf{0}$ this leads to only three constraints

$$\frac{1}{N}\sum_i \mathbf{x}_{0i}=\mathbf{c}_T\,,\qquad \sum_i \mathbf{v}_i=\mathbf{0}\,,\qquad \sum_i [\mathbf{x}_{0i}\times]\mathbf{v}_i=\mathbf{0}\,. \tag{74}$$

Expressed in terms of the corrections $\delta\mathbf{x}_{0i}=\mathbf{x}_{0i}-\mathbf{x}_{0i}^{\text{ap}}$, $\delta\mathbf{v}_i=\mathbf{v}_i-\mathbf{v}_i^{\text{ap}}$ to approximate values $\mathbf{x}_{0i}^{\text{ap}}$, $\mathbf{v}_i^{\text{ap}}$ and neglecting the second order terms $\Sigma_i[\delta\mathbf{x}_i\times]\delta\mathbf{v}_i$ and $\Sigma_i[\mathbf{v}_i^{\text{ap}}\times]\delta\mathbf{x}_{0i}$ the kinematically derived constraints become

$$\frac{1}{N}\sum_i \delta\mathbf{x}_{0i}=\mathbf{c}_T-\overline{\mathbf{x}}_0^{\text{ap}}\,,\qquad \overline{\mathbf{x}}_0^{\text{ap}}=\frac{1}{N}\sum_i \mathbf{x}_{0i}^{\text{ap}} \tag{75a}$$

$$\frac{1}{N}\sum_i \delta\mathbf{v}_i=-\overline{\mathbf{v}}^{\text{ap}}\qquad \overline{\mathbf{v}}^{\text{ap}}=\frac{1}{N}\sum_i \mathbf{v}_i^{\text{ap}} \tag{75b}$$

$$-\sum_i [\mathbf{x}_{0i}^{\text{ap}}\times]\delta\mathbf{v}_i=\sum_i [\mathbf{x}_{0i}^{\text{ap}}\times]\mathbf{v}_i^{\text{ap}}\equiv\mathbf{h}_R^{\text{ap}} \tag{75c}$$

where we have used the initial epoch barycenter $\overline{\mathbf{x}}_0^{\text{ap}}=\frac{1}{N}\Sigma_i\mathbf{x}_{0i}^{\text{ap}}$ and the approximate relative angular momentum $\mathbf{h}_R^{\text{ap}}=\Sigma_i[\mathbf{x}_{0i}^{\text{ap}}\times]\mathbf{v}_i^{\text{ap}}$ of the "approximate" network. The first constraint defines the origin of the system at t_0 by setting the barycenter to that of the approximate network ($\mathbf{x}_B(t_0)=\mathbf{x}_B^{\text{ap}}$) and is thus a "translation" constraint . The second constraints the network barycenter at any epoch to remain in its initial position ($\mathbf{x}_B(t)=\mathbf{x}_B(t_0)$) and is thus a "translation rate" constraint . The third constraint deriving from $\mathbf{h}_R(t)=\mathbf{0}$ controls the time evolution of the orientation of the system and is thus an "orientation rate" constraint.

In order to compare these results with the partial inner constraints for coordinates and velocities obtained in the (simultaneous) stacking problem we shall examine the special case where $\mathbf{v}_i^{\text{ap}}=\mathbf{0}$, a permissible choice since $\mathbf{v}_i$ are already small quantities, and we also choose $\mathbf{x}_{0i}^{\text{ap}}$ in such a way that $\Sigma_i\mathbf{x}_{0i}^{\text{ap}}=\mathbf{0}$ and set $\mathbf{c}_T=\mathbf{0}$. In such a case the constraints simplify to

$$\sum_i \delta\mathbf{x}_{0i}=\mathbf{0}\,,\qquad \sum_i \delta\mathbf{v}_i=\mathbf{0}\,,\qquad \sum_i [\mathbf{x}_{0i}^{\text{ap}}\times]\delta\mathbf{v}_i=\mathbf{0} \tag{76a}$$

which are indeed three of the partial inner constraints. As expected from the very nature of the Tisserant condition the constraint which selects the orientation of the reference system is missing and must be independently introduced. We may borrow the "orientation" constraint from the relevant partial inner constraints

$$\sum_{i=1}^{N} [\mathbf{x}_{0i}^{\text{ap}}\times]\delta\mathbf{x}_{0i}=\mathbf{0} \tag{76b}$$

and thus have a complete set except from the scale problem.

To obtain scale and scale rate constraints in a kinematic way we impose the condition that the overall scale of the network does not vary. The issue is again the choice of a quantity quantifying the average scale of the network at every epoch. We propose to use the mean quadratic length $S(t)$ of the network, defined by

$$Q(t) = S^2(t) = \frac{1}{N}\sum_i S_{Bi}^2(t) = \frac{1}{N}\sum_i \left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right]^T \left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right] \tag{77}$$

where $S_{Bi}(t) = \sqrt{\left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right]^T \left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right]}$ is the distance of each network point $\mathbf{x}_i(t)$ from the network barycenter $\mathbf{x}_B(t) = \frac{1}{N}\sum_i \mathbf{x}_i(t)$ which is uniquely defined for each epoch. Setting $Q(t) = S^2(t)$ we may fix the scale and the scale rate by setting $S(t_0) = \text{constant}$ and $S(t) = S(t_0)$ for every t. This is equivalent to setting $Q(t_0) \equiv Q_0 = \text{constant}$ and $\frac{d}{dt}Q(t) = 0$. The scale condition is therefore

$$Q(t_0) \equiv S^2(t_0) = \frac{1}{N}\sum_i \left[\mathbf{x}_i(t_0) - \mathbf{x}_B(t_0)\right]^T \left[\mathbf{x}_i(t_0) - \mathbf{x}_B(t_0)\right] = Q_0 \tag{78}$$

Using the mean value $\overline{\mathbf{x}}_0 = \frac{1}{N}\sum_i \mathbf{x}_{0i}$ and $\mathbf{x}_i(t_0) = \mathbf{x}_{0i}$ it follows that $\mathbf{x}_B(t_0) = \overline{\mathbf{x}}_0$ and (78) becomes

$$Q(t_0) = \frac{1}{N}\sum_i (\mathbf{x}_{0i} - \overline{\mathbf{x}}_0)^T (\mathbf{x}_{0i} - \overline{\mathbf{x}}_0) = \frac{1}{N}\sum_i \mathbf{x}_{0i}^T \mathbf{x}_{0i} - \overline{\mathbf{x}}_0^T \overline{\mathbf{x}}_0 = Q_0 . \tag{79}$$

Replacing $\mathbf{x}_{0i} = \mathbf{x}_{0i}^{\text{ap}} + \delta\mathbf{x}_i$, using $\overline{\mathbf{x}}_0^{\text{ap}} = \frac{1}{N}\sum_i \mathbf{x}_{0i}^{\text{ap}}$ and noting that $\overline{\mathbf{x}}_0 = \overline{\mathbf{x}}_0^{\text{ap}} + \frac{1}{N}\sum_i \delta\mathbf{x}_{0i}$ we arrive after some algebraic manipulation at

$$\frac{2}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T \delta\mathbf{x}_{0i} = Q_0 - \frac{1}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}})^T \mathbf{x}_{0i}^{\text{ap}} + (\overline{\mathbf{x}}_0^{\text{ap}})^T \overline{\mathbf{x}}_0^{\text{ap}} . \tag{80}$$

The approximate value of $Q(t_0)$ becomes

$$Q_0^{\text{ap}} = \frac{1}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}}) = \frac{1}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}})^T \mathbf{x}_{0i}^{\text{ap}} - (\overline{\mathbf{x}}_0^{\text{ap}})^T \overline{\mathbf{x}}_0^{\text{ap}} . \tag{81}$$

and (80) simplifies to

$$\frac{2}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T \delta\mathbf{x}_{0i} = Q_0 - Q_0^{\text{ap}} . \tag{82}$$

Comparing the previous two equations we conclude that an advantageous choice for the so far unspecified constant Q_0 is to set $Q_0 = Q^{\text{ap}}$. This condition is equivalent to setting the mean quadratic scale S at the initial epoch t_0, equal to the quadratic scale S^{ap} of the approximate network

$$S(t_0) = S^{\text{ap}} = \sqrt{\frac{1}{N}\sum_i (\mathbf{x}_i^{ap} - \mathbf{x}_B^{ap})^T (\mathbf{x}_i^{ap} - \mathbf{x}_B^{ap})}\,, \qquad \mathbf{x}_B^{ap} = \frac{1}{N}\sum_i \mathbf{x}_i^{ap}\,. \tag{83}$$

For this particular choice (82) simplifies to the scale constraint

$$\sum_i (\mathbf{x}_{0i}^{ap} - \overline{\mathbf{x}}_0^{ap})^T \delta\mathbf{x}_{0i} = 0\,. \tag{84}$$

For the scale rate we have the condition

$$\begin{aligned}\frac{d}{dt}Q(t) &= \frac{1}{N}\frac{d}{dt}\sum_i \left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right]^T \left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right] = \\ &= \frac{2}{N}\sum_i \left[\mathbf{x}_i(t) - \mathbf{x}_B(t)\right]^T \left[\dot{\mathbf{x}}_i(t) - \dot{\mathbf{x}}_B(t)\right] = 0\,.\end{aligned} \tag{85}$$

The model $\mathbf{x}_i(t) = \mathbf{x}_{0i} + (t - t_0)\mathbf{v}_i$ gives $\dot{\mathbf{x}}_i(t) = \mathbf{v}_i$, $\mathbf{x}_B(t) = \overline{\mathbf{x}}_0 + (t - t_0)\overline{\mathbf{v}}$, $\dot{\mathbf{x}}_B(t) = \overline{\mathbf{v}} = \frac{1}{N}\sum_i \mathbf{v}_i$ and the above condition becomes

$$\begin{aligned}0 &= \sum_i [\mathbf{x}_{0i} + (t - t_0)\mathbf{v}_i - \overline{\mathbf{x}}_0 - (t - t_0)\overline{\mathbf{v}}]^T (\mathbf{v}_i - \overline{\mathbf{v}}) = \\ &= \sum_i (\mathbf{x}_{0i} - \overline{\mathbf{x}}_0)^T (\mathbf{v}_i - \overline{\mathbf{v}}) + (t - t_0)\sum_i (\mathbf{v}_i - \overline{\mathbf{v}})^T (\mathbf{v}_i - \overline{\mathbf{v}}) = 0\end{aligned} \tag{86}$$

which must hold for every t. Thus two conditions must be satisfied

$$\sum_i (\mathbf{x}_{0i} - \overline{\mathbf{x}}_0)^T (\mathbf{v}_i - \overline{\mathbf{v}}) = 0\,, \qquad \sum_i (\mathbf{v}_i - \overline{\mathbf{v}})^T (\mathbf{v}_i - \overline{\mathbf{v}}) = 0\,. \tag{87}$$

These conditions after some algebraic manipulation become

$$\sum_i \mathbf{x}_{0i}^T \mathbf{v}_i = N\overline{\mathbf{x}}_0^T \overline{\mathbf{v}}\,, \qquad \sum_i \mathbf{v}_i^T \mathbf{v}_i = N\,\overline{\mathbf{v}}^T \overline{\mathbf{v}} \tag{88}$$

In terms of approximate values and correction they take the form

$$\sum_i (\mathbf{v}_i^{\text{ap}} - \overline{\mathbf{v}}^{\text{ap}})^T \delta\mathbf{x}_{0i} + \sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T \delta\mathbf{v}_i = N(\overline{\mathbf{x}}_0^{ap})^T \overline{\mathbf{v}}^{ap} - \sum_i (\mathbf{x}_{0i}^{ap})^T \mathbf{v}_i^{ap} \tag{89}$$

$$2\sum_i (\mathbf{v}_i^{\text{ap}} - \overline{\mathbf{v}}^{\text{ap}})^T \delta\mathbf{v}_i = N(\overline{\mathbf{v}}^{\text{ap}})^T \overline{\mathbf{v}}^{\text{ap}} - \sum_i (\mathbf{v}_i^{\text{ap}})^T \mathbf{v}_i^{\text{ap}} \tag{90}$$

Taking into account that velocities are small quantities, second order quantities in velocities and/or corrections can be ignored, in which case the (90) degenerates to $0 = 0$, while (89) gives the required “scale rate constraint”

$$\frac{1}{N}\sum_i (\mathbf{x}_{0i}^{ap} - \overline{\mathbf{x}}_0^{ap})^T \delta\mathbf{v}_i = (\overline{\mathbf{x}}_0^{ap})^T \overline{\mathbf{v}}^{ap} - \frac{1}{N}\sum_i (\mathbf{x}_{0i}^{ap})^T \mathbf{v}_i^{ap} \tag{91}$$

to be applied together with the already derived in (89) scale constraint $\sum_i (\mathbf{x}_{0i}^{ap} - \overline{\mathbf{x}}_0^{ap})^T \delta\mathbf{x}_{0i} = 0$. In the special case that zero approximate values $\mathbf{v}_i^{ap} = \mathbf{0}$ are used for the small velocities (89) remains unchanged while (91) simplifies to $\sum_i (\mathbf{x}_{0i}^{ap} - \overline{\mathbf{x}}_0^{ap})^T \delta\mathbf{v}_i = \mathbf{0}$. They both differ from corresponding partial inner constraints $\sum_{i=1}^N (\mathbf{x}_{0i}^{\text{ap}})^T \delta\mathbf{x}_{0i} = 0$ (scale) and $\sum_{i=1}^N (\mathbf{x}_{0i}^{\text{ap}})^T \delta\mathbf{v}_i = 0$ (scale rate) obtained in the (simultaneous) stacking problem. They may coincide under the additional condition that $\overline{\mathbf{x}}_0^{ap} = \mathbf{0}$, i.e. when care has been taken to have the origin of approximate network at its barycenter. This is easily done by subtracting from any set of approximate coordinates their corresponding mean values.

Summarizing our kinematically derived constraints (in comparison with the algebraically derived partial inner constraints) are:

Translation:

$$\frac{1}{N}\sum_i \delta\mathbf{x}_{0i} = \mathbf{c}_T - \overline{\mathbf{x}}_0^{\text{ap}}, \qquad \overline{\mathbf{x}}_0^{\text{ap}} = \frac{1}{N}\sum_i \mathbf{x}_{0i}^{\text{ap}} \qquad (\text{versus } \sum_{i=1}^N \delta\mathbf{x}_{0i} = \mathbf{0}). \tag{92}$$

or when the arbitrary constant is $\mathbf{c}_T$ set to $\mathbf{c}_T = \overline{\mathbf{x}}_0^{\text{ap}}$

$$\frac{1}{N}\sum_i \delta\mathbf{x}_{0i} = -\overline{\mathbf{x}}_0^{\text{ap}}, \qquad (\text{versus } \sum_{i=1}^N \delta\mathbf{x}_{0i} = \mathbf{0}). \tag{92′}$$

Translation rate:

$$\frac{1}{N}\sum_i \delta\mathbf{v}_i = -\overline{\mathbf{v}}^{\text{ap}}, \qquad \overline{\mathbf{v}}^{\text{ap}} = \frac{1}{N}\sum_i \mathbf{v}_i^{\text{ap}} \qquad (\text{versus } \sum_{i=1}^N \delta\mathbf{v}_i = \mathbf{0}). \tag{93}$$

Orientation (borrowed from algebraic):

$$\sum_{i=1}^N [\mathbf{x}_{0i}^{\text{ap}}\times]\delta\mathbf{x}_{0i} = \mathbf{0}. \tag{94}$$

Orientation rate:

$$-\sum_i [\mathbf{x}_{0i}^{\text{ap}}\times]\delta\mathbf{v}_i = \mathbf{h}_R^{\text{ap}} \equiv \sum_i [\mathbf{x}_{0i}^{\text{ap}}\times]\mathbf{v}_i^{\text{ap}} \qquad (\text{versus } \sum_{i=1}^N [\mathbf{x}_{0i}^{\text{ap}}\times]\delta\mathbf{v}_i = \mathbf{0}). \tag{95}$$

Scale:

$$\frac{2}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T \delta\mathbf{x}_{0i} = Q_0 - Q_0^{\text{ap}}. \tag{96}$$

or when the arbitrary constant Q_0 is set to $Q_0 = Q_0^{\text{ap}} \equiv \frac{1}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})$

$$\sum_i (\mathbf{x}_{0i}^{ap} - \overline{\mathbf{x}}_0^{ap})^T \delta\mathbf{x}_{0i} = 0\,, \qquad \text{(versus } \sum_{i=1}^{N} (\mathbf{x}_{0i}^{\text{ap}})^T \delta\mathbf{x}_{0i} = 0\,\text{)}. \tag{96′}$$

Scale rate:

$$\frac{1}{N}\sum_i (\mathbf{x}_{0i}^{\text{ap}} - \overline{\mathbf{x}}_0^{\text{ap}})^T \delta\mathbf{v}_i = (\overline{\mathbf{x}}_0^{ap})^T \overline{\mathbf{v}}^{ap} - \frac{1}{N}\sum_i (\mathbf{x}_{0i}^{ap})^T \mathbf{v}_i^{ap}\,, \text{ (versus } \sum_{i=1}^{N} (\mathbf{x}_{0i}^{\text{ap}})^T \delta\mathbf{v}_i = 0\,\text{)}. \tag{97}$$

Unlike the corresponding algebraic partial inner constraints the kinematic constraints do not depend on the approximate values of the coordinates and the velocities. Note however that in replacing (80) with (82), (92) with (92′) and (96) with (96′) we have replaced the arbitrary constants Q_0 and $\mathbf{c}_T$ with values depending on approximate values so that the simplified constraints (92′) and (96′) are no more independent of the choice of these values. The two sets of the simplified (by the special choices of Q_0 and $\mathbf{c}_T$) kinematic minimal constraints and the corresponding algebraic partial inner constraints become identical when the approximate values satisfy following conditions:

$$\overline{\mathbf{x}}_0^{\text{ap}} = \mathbf{0}\,, \qquad \mathbf{v}_i^{\text{ap}} = \mathbf{0}\,. \tag{98}$$

If the translation and translation rate constraints are to be incorporated together with the scale and scale rate ones (i.e. for techniques other than the “geocentric” SLR where translation constraints do not apply) we may take advantage of (92) and (93) to bring the scale related constraints (96) and (97) to the simpler form

$$\frac{1}{N}\sum_i (\mathbf{x}_{0i}^{ap})^T \delta\mathbf{x}_{0i} = -(\overline{\mathbf{x}}_0^{ap})^T \overline{\mathbf{x}}_0^{\text{ap}}\,, \tag{99}$$

$$\sum_i (\mathbf{x}_{0i}^{ap})^T \delta\mathbf{v}_i = -\sum_i (\mathbf{x}_{0i}^{ap})^T \mathbf{v}_i^{ap}\,. \tag{100}$$

Unlike the (total) inner or the partial inner constraints, which are homogeneous, the kinematic constraints are inhomogeneous with non-zero right-side terms of the general form $\mathbf{C}^T\mathbf{x} = \mathbf{d}$. The solution $\hat{\mathbf{x}}_C$ of the least squares problem with normal equations $\mathbf{N}\mathbf{x} = \mathbf{u}$ under additional inhomogeneous minimal constraints $\mathbf{C}^T\mathbf{x} = \mathbf{d}$ and its corresponding covariance cofactor matrix $\mathbf{Q}_{\hat{\mathbf{x}}_C}$ are given by (Dermanis, 1987)

$$\hat{\mathbf{x}} = (\mathbf{N} + \mathbf{C}\mathbf{C}^T)^{-1}(\mathbf{u} + \mathbf{C}\mathbf{d})\,, \qquad \mathbf{Q}_{\hat{\mathbf{x}}_C} = (\mathbf{N} + \mathbf{C}\mathbf{C}^T)^{-1}\mathbf{N}(\mathbf{N} + \mathbf{C}\mathbf{C}^T)^{-1}. \tag{101}$$

Note however that if $\mathbf{R}$ is a $d \times d$ regular matrix then $\mathbf{R}^T\mathbf{C}^T\mathbf{x} = \mathbf{R}^T\mathbf{d}$ or $\tilde{\mathbf{C}}^T\mathbf{x} = \tilde{\mathbf{d}}$ with $\tilde{\mathbf{C}} = \mathbf{C}\mathbf{R}$ and $\tilde{\mathbf{d}} = \mathbf{R}^T\mathbf{d}$, is an equivalent set of minimal constraints. Replacing $\mathbf{C}$ with $\tilde{\mathbf{C}}$ and $\mathbf{d}$ with $\tilde{\mathbf{d}}$ in the above formulas and setting $\mathbf{S} = \mathbf{R}\mathbf{R}^T$ we arrive at

$$\hat{\mathbf{x}} = (\mathbf{N} + \mathbf{CSC}^T)^{-1}(\mathbf{u} + \mathbf{CSd}), \qquad \mathbf{Q}_{\hat{\mathbf{x}}_C} = (\mathbf{N} + \mathbf{CSC}^T)^{-1}\mathbf{N}(\mathbf{N} + \mathbf{CSC}^T)^{-1}, \tag{102}$$

where $\mathbf{S}$ is an arbitrary regular symmetric matrix, to be chosen in a way that it improves computational efficiency. A set of alternative formulas utilizing the matrix $\mathbf{E}$ of the total inner constraints $\mathbf{E}^T\mathbf{x} = \mathbf{0}$ (Koch, 1999) are

$$\hat{\mathbf{x}} = (\mathbf{N} + \mathbf{CC}^T)^{-1}\mathbf{u} + \mathbf{E}(\mathbf{C}^T\mathbf{E})^{-1}\mathbf{d}, \quad \mathbf{Q}_{\hat{\mathbf{x}}_C} = (\mathbf{N} + \mathbf{CC}^T)^{-1} - \mathbf{E}(\mathbf{E}^T\mathbf{CC}^T\mathbf{E})^{-1}\mathbf{E}^T. \tag{103}$$

Following the previous argument these can also be generalized to

$$\hat{\mathbf{x}} = (\mathbf{N} + \mathbf{CSC}^T)^{-1}\mathbf{u} + \mathbf{E}(\mathbf{C}^T\mathbf{E})^{-1}\mathbf{d}, \quad \mathbf{Q}_{\hat{\mathbf{x}}_C} = (\mathbf{N} + \mathbf{CSC}^T)^{-1} - \mathbf{E}(\mathbf{E}^T\mathbf{CSC}^T\mathbf{E})^{-1}\mathbf{E}^T. \tag{104}$$

Conclusions

The problem of choosing an optimal reference system for the International Terrestrial Reference Frame (ITRF) has been studied for two alternative approaches: (a) the one-step approach of simultaneous stacking (removal of the reference system at each data epoch and implementation of a linear in time coordinate model) for all techniques and (b) the two-step approach where stacking is first applied to each technique separately followed by a combination step where the obtained estimates of station initial coordinates and velocities per technique are combined to obtain their final ITRF estimates. Two different in principle approaches have been followed for the definition of the optimal ITRF reference system, the classical algebraic one of inner constraints and a new approach for the derivation of optimal minimal constraints based on kinematic principles where the apparent variation of coordinates with time is minimized. To achieve these goals to compromises had to be made. The permissible reference system changes leading to the optimal one have been restricted (a) in transformations close to the identity (very small changes of the reference system, where second order terms in the transformation parameters can be neglected) and (b) in transformations which preserve the linear-in-time form of the ITRF model. The second restriction leads to transformation parameters which are linear functions of time and lead to a sub-optimal solution to the choice of reference system problem. Absolute optimality requires transformation parameters which are arbitrary functions of time which nevertheless lead to ITRF models which are no more linear functions of time as imposed a priori.

The main difference between algebraic and kinematic constraints is that they are based on different optimality principles. The algebraic ones minimize the sum of squares of the unknowns, which are corrections to the approximate values of all or a selected subset of the original unknowns (station initial coordinates and velocities and nuisance transformation parameters from the ITRF reference system to that of the available observations). The kinematic constraints preserve the barycenter of the network, its mean quadratic scale and minimize its relative (apparent) kinetic energy by causing its relative (apparent) angular momentum to vanish (discrete Tisserand condition). Relative kinetic energy and relative angular momentum refer to the ITRF network viewed as a set of material points with equal mass.

As a consequence of their very definition the algebraic constraints depend on the more-or-less arbitrary approximate values of the unknown parameters. On the contrary the kinematic constraints are independent of the choice of approximate values and are thus uniquely defined.

Despite the above differences the kinematic minimal constraints coincide with a particular set of the algebraic partial inner constraints, the one referring to only the initial values and velocities of the ITRF stations, under mild conditions on the choice of approximate values. Indeed such a coincidence results by choosing approximate values of initial coordinates which have zero mean and zero approximate values for velocities.

References

Altamimi Z., P. Sillard, C. Boucher (2002). ITRF2000: A New Release of the International Terrestrial Reference Frame for Earth Science Applications, Journal of Geophysical Research, 107 (B10), p. 2214.

Altamimi, Z., X. Collilieux, C. Boucher (2008). Accuracy Assessment of the ITRF Datum Definition. In: Peiliang Xu, Jingnan Liu and Athanasios Dermanis (eds), *VI Hotine-Marussi Symposium on Theoretical and Computational Geodesy*, International Association of Geodesy Symposia, Volume 132, pp. 101-110, Springer Berlin, 2008.

Altamimi, Z., X. Collilieux, J. Legrand, B. Garayt, C. Boucher (2007). ITRF2005: A new release of the International Terrestrial Reference Frame based on time series of station positions and Earth Orientation Parameters, *Journal of Geophysical Research*, vol. 112, B09401.

Altamimi Z., X. Collilieux, L. Métivier (2011). ITRF2008: an improved solution of the international terrestrial reference frame. *Journal of Geodesy*, Vol. 85, 475-485.

Blaha, G. (1971), *Inner adjustment constraints with emphasis on range observations*. Department of Geodetic Science, Report 148, The Ohio State University, Columbus.

Blaha, G. (1982). Free networks: Minimum norm solution as obtained by the inner adjustment constraint method. *Bulletin Géodésique*, 56, 209-219.

Dermanis, A. (1987). *Adjustment of Observations and Estimation Theory* (in Greek). Editions Ziti, Thessalomiki.

Dermanis, A. (1994a). Inner constraints in photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 49 (1), 25-39.

Dermanis, A. (1994b). Free network solutions with the Direct Linear Transformation method. *ISPRS Journal of Photogrammetry and Remote Sensing*, 49 (2), 2-12.

Dermanis, A. (1995). The Non-Linear and the Space-Time Datum problem. Presented at the Meeting "Mathematische Methoden der Geodäsie", Mathematisches Forschungsinstitut Oberwolfach, October 1-7, 1995.

Dermanis, A. (2000). Establishing Global Reference Frames. Nonlinear, Temporal, Geophysical and Stochastic Aspects. IAG Inter. Symp. Banff, Alberta, Canada, July 31-Aug. 4, 2000. In: M.G. Sideris, ed. "*Gravity, Geoid and Geodynamics 2000*", IAG Symposia vol. 123, p. 35-42, Springer, Berlin 2002.

Dermanis, A. (2003). The rank deficiency in estimation theory and the definition of reference systems. In: Sansò, F. (ed.), 2003: *V Hotine-Marussi Symposium on Mathematical Geodesy*, Matera, Italy June 17-21, 2003. International Association of Geodesy Symposia, Vol. 127, pp. 145-156. Springer Verlag, Heidelberg.

Dermanis, A. (2008). The ITRF Beyond the "Linear" Model. Choices and Challenges. In: Peiliang Xu, Jingnan Liu and Athanasios Dermanis (eds), *VI Hotine-Marussi Symposium on Theoretical and Computational Geodesy*, International Association of Geodesy Symposia, Volume 132, pp. 111-118, Springer Berlin, 2008.

Fritz, D. and B. Schaffrin (1981). The "choice of norm" problem for the free net adjustment with orientation parameters. *Proceedings 8th Symposium on Mathematical Geodesy*, Como, September7-9, 1981, Istituto Geografico Militare, Firenze, 1983. Also in: *Bollettino di Geodesia e Sienze Affini*, Anno XLI (1982), N. 3, 259-282.

Grafarend, E., and Schaffrin, B. (1974). Unbiased free net adjustment. *Survey Review*, 22 (171): 200-218.

Grafarend, E. and B. Schaffrin (1976). Equivalence of estimable quantities and invariants in geodetic networks. *Zeitschrift für Vemessungswesen*, 101, 11, 485-491.

Gérard Petit and Brian Luzum, eds. (2010). IERS Conventions. IERS Technical Note No. 36. Available at:
http://www.iers.org/nn_11216/IERS/EN/Publications/TechnicalNotes/tn36.html

Koch, K.-R. (1999). Parameter *Estimation and Hypothesis Testing in Linear Models*, 2nd Ed., Springer.

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

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

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

Rao, C.R. (1971).Unified Theory of Linear Estimation. *Sankhyā, The Indian Journal of Statistics*, Series A, 33, 371-394.

Rao, C.R. (1973). Unified theory of least squares. *Communications in Statistics - Theory and Methods*, 1 (1), 1-8.

Sillard, P. (1999). *Modélisation des systèmes de référence terrestres. Contribution théorique et méthodologique*. Doctoral Thesis, Observatoire de Paris.

Sillard, P. and C. Boucher (2001). Review of algebraic constraints in terrestrial reference frame datum definition. *Journal of Geodesy*, 75, 63–73.

Tisserand, F. (1889). *Traité de mécanique céleste*. Gauthier-Villars et fils, Paris.