Gravimetry, Relativity, and the Global Navigation Satellite Systems*

Albert Tarantola†
Ludek Klimeš‡
José Maria Pozo & Bartolomé Coll§

May 27, 2009

Abstract

Relativity is an integral part of positioning systems, and this is taken into account in today's practice by applying many 'relativistic corrections' to computations performed using concepts borrowed from Galilean physics. A different, fully relativistic paradigm can be developed for operating a positioning system. This implies some fundamental changes. For instance, the basic coordinates are four times (with a symmetric meaning, not three space coordinate and one time coordinate) and the satellites *must* have cross-link capabilities. Gravitation must, of course, be taken into account, but not using the Newtonian theory: the gravitation field is, and only is, the space-time metric. This implies that the positioning problem and the gravimetry problem can not be separated. An optimization theory can be developed that, because it is fully relativistic, does not contain any 'relativistic correction'. We suggest that all positioning satellite systems should be operated in this way. The first benefit of doing so would be a clarification and a simplification of the theory. We also expect, at the end, that positioning systems will provide increased positioning accuracy.

Contents

*Lesson delivered at the School *Relativistic Coordinates, Reference and Positioning Systems* (Salamanca, January 2005).

†Institut de Physique du Globe de Paris, albert.tarantola@ipgp.org, `http://www.ipgp.jussieu.fr/~tarantola/`.
‡Charles University, Prague, `http://sw3d.cz/staff/klimes.htm`.
§Observatoire de Paris, `http://syrte.obspm.fr/~coll/`.

1 Introduction

Many relativistic corrections are applied to the Global Navigation Satellite Systems (GNSS). Neil Ashby presents in Physics Today (May 2002) a good account of how these relativistic corrections are applied, why, and which are their orders of magnitude. Unfortunately, it is generally proposed that relativity is only a correction to be applied to Newtonian physics. We rather believe that there is a fully relativistic way to understand a GNSS system, this leading to a new way of operating it.

As gravitation has to be taken into account, it is inside the framework of general relativity that the theory must be developed. The shift from a Newtonian viewpoint (relativistic corrections included or not) into a relativistic framework requires some fundamental conceptual changes. Perhaps the most important concerns the operational definition of a system of four space-time coordinates. We reach the conclusion that there is an (essentially unique) coordinate system that, while being consistent with a relativistic formulation, allows an *immediate* positioning of observers (the traditional Minkowski coordinates $\{t, x, y, z\}$ of flat space-time do not allow such an immediate positioning).

These coordinates are defined as follows[1]. If four clocks —having an arbitrary space-time trajectory— broadcast their proper time —using electromagnetic signals,— then, any observer receives, at any point along his personal space-time trajectory, four times, corresponding to the four signals arriving at that space-time point. These four times, say $\{\tau^1, \tau^2, \tau^3, \tau^4\}$, *are*, by definition, the coordinates of the space-time point. One doesn't has one time coordinate and three space coordinates, as usual, but a 'symmetric' coordinate system with four time coordinates.

The space-time having been endowed with those coordinates, any observer with a receiver may obtain (in real time) his personal trajectory. This is true, in particular, for the four clocks themselves: each clock constantly receives three of the coordinates and it defines the fourth. Therefore, *each clock knows its own trajectory* in this self-consistent coordinate system. Note that even if the clocks are satellites around the Earth, the coordinates and the orbits are defined without any reference to an Earth based coordinate system: this allows to achieve maximum precision for this *primary* reference system. Of course, for applications on the Earth's surface, the primary coordinates must be attached

[1] Coll (1985, 2000, 2001a, 2001b, 2002), Rovelli (2001), Coll and Morales (1992), Blagojević et al. (2001), Coll and Tarantola (2003).

to some terrestrial coordinate system, but this is just an attachment problem that should not interfere with the problem of defining the primary system itself.

In general relativity, the gravity field is the space-time metric. Should this metric be exactly known (in any coordinate system), the system just described would constitute an ideal positioning system (and the components of the metric could be expressed in these coordinates). In practice the space-time metric (i.e., the gravity field) is not exactly known, and the satellite system itself has to be used to infer it. This article is about the problem of using a satellite system for both, positioning, and measuring the space-time metric.

Information on the space-time metric may come from different sources. First, any satellite system has more than four clocks. While four of the clocks define the coordinates, the redundant clocks can be used to monitor the space-time metric. The considered satellites may have more that a clock: they may have an accelerometer (this providing information on the space-time connection), a gradiometer (this providing information on the Riemann), etc. Our theory will provide seamless integration of positioning systems with systems designed for gravimetry.

In the "post Newtonian" paradigm used today for operating positioning and gravimetry systems, the ever increasing accuracy of clocks makes that more and more "relativistic effects" have to be taken into account. On the contrary, the fully relativistic theory here developed will remain valid as long as relativity itself remains valid.

It is our feeling that when GNSS and gravimetry systems will be operated using the principles here exposed, new experimental possibilities will appear. One must realize that with the optical clocks being developed may one day have a relative accuracy of 10^{18}. The possibility that some day we may approach this accuracy for positioning immediately suggests extraordinarily interesting applications.

These applications would simply be impossible if sticking to the present-day paradigm. To realize how deeply nonrelativistic this paradigm is, consider that GPS clocks are kept synchronized. In this year 2005, when we celebrate the centenary of relativity, this sounds strange: is there anything less relativistic than the obstination to keep synchronized a system of clocks in relative movement?

There is one implication of the theory here developed for the Galileo positioning system now being developed by the European Union. Our theory requires, as a fundamental fact, that the GNSS satellites exchange signals. The most recent GPS satellites (from the USA) do have this "cross-link" capability. One could, in principle, use the cross-link data (or an ameliorated version of it) to operate the system in the way here proposed. Unfortunately, the Galileo satellites will not have, to our knowledge, this cross-link capability. This is a serious limitation that will complicate the evolution of the system towards a more precise one.

Finally, we need to write a disclaimer here. None of the algorithms proposed below are intended to be practical. They are the simplest algorithms that would be fully consistent with relativity theory. Passing from these to actually implementable algorithms will require some developments in numerical analysis. Finally, some of the simplifying hypotheses made below are not necessary and are only intended to start with a theory that is as simple as possible.

2 Setting of the Problem

2.1 Model Parameters and Observable Parameters

Imagine that four clocks (called below the *basic clocks*) broadcast their proper time using light signals. Any observer in space-time may receive, at any point along its space-time trajectory, four times $\{\tau^\alpha\} = \{\tau^1, \tau^2, \tau^3, \tau^4\}$: these are, by definition, the (space-time) coordinates of the space-time point where the observer is. If the observer had his own clock, with proper time denoted τ, then he would know his trajectory

$$\tau^\alpha = \tau^\alpha(\tau) \quad , \tag{1}$$

i.e., he would know four functions (of his proper time τ). He could, in addition, evaluate his four-velocity, and his four-acceleration (see below).

In an actual experiment, the clocks are never ideal, and the reception of times implies a measurement that has always attached uncertaities. We must, therefore, carefully distinguish between model parameters and observable parameters. This places the present discussion into the usual conceptual frame of inverse problem theory[2].

2.1.1 Complete Model

For us, a *complete model* consists in:

- A **space-time metric field**, denoted $\{ g_{\alpha\beta}(\tau^1, \tau^2, \tau^3, \tau^4) \}$, and some (at least four) **trajectories**, denoted $\{ \tau_1^\alpha(\lambda_1) , \tau_2^\alpha(\lambda_2) , \dots \}$, parameterized by some parameters $\lambda_1 , \lambda_2 , \dots$ Because the metric is given, these parameters can be converted into proper time (by integration of the element $\sqrt{g_{\alpha\beta}\, d\tau^\alpha\, d\tau^\beta}$), so the four trajectories can always be considered to be given as a function of proper time, $\{ \tau_1^\alpha(\tau_1) , \tau_2^\alpha(\tau_2) , \dots \}$. These trajectories shall be the model of the trajectories of "satellites" consisting on physical clocks, and, perhaps, accelerometers, gradiometers, and other measuring instruments. Four of the trajectories are arbitrarily selected as 'basic trajectories, and the **working coordinates** $\{\tau^1, \tau^2, \tau^3, \tau^4\}$ are assumed to be linked to the metric and the four basic trajectories as follows: the four coordinates $\{\tau^1, \tau^2, \tau^3, \tau^4\}$ of a space-time point $\mathcal{P}$ are, by definition, the four emission times (one in each of the four trajectories) of the four light "cones" passing by the point $\mathcal{P}$. This is an idealization of the heuristic protocol suggested in the introduction.
- Associated to each trajectory, a (typically smooth) **clock drift function** $\{f_1(\tau_1) , f_2(\tau_2) , \dots \}$, describing the drift of the physical clock with respect to proper time (if z is clock time, and τ is proper time, then $z(\tau) = \tau + f(\tau)$).

To these fundamental parameters, we need to add another set, that are also necessary for the prediction of observations:

- A **set of instants** $\{\tau_a , \tau_b , \dots \}$ along each trajectory, that represent the nominal instants when a clock is observed, a light signal is emitted (that will be received by some other satellite), or the instant when a measurement (acceleration, gradiometry, etc.) is made.

To simplify the theory, we shall assume that the space-time trajectory of all the satellites has crossed at a given space-time point, and that all the clocks have been synchronized (to zero time) at this point. From then on, all clocks will follow their proper time, without any further synchronization (the drift function just mentioned takes into account that physical clocks never exactly follow proper time).

We choose in this first version of the theory, not to introduce the fact that light does not propagate in absolute vacuum. One very simple model for the upper layers of the atmosphere would be as follows. One could assume that light propagates in a gas that, at rest, is homogeneous and isotropic, with an index of refraction n that, in general, depends on the carrier frequency of the signal. If the

[2]For an introductory text, see Tarantola (2005).

four-velocity of the gas is U^α , then Maxwell equations in the gas take the same form as in vacuum, excepted that one must replace the actual (Riemannian) metric by the (Finslerian) metric (see Gordon (1923) or Pham Mau Quan (1957) for details)

$$g_{\alpha\beta} + \left(1 - \frac{1}{n^2}\right) U_\alpha \, U_\beta \quad . \tag{2}$$

When we will introduce this aspect in the theory, a complete model will also include the field n (index of refraction at different frequencies), and the field U^α (four-velocity of the gas). The a priori information that one has on these quantities is precise enough as for hoping that satellite data will easily be able to refine the model (tomography of the ionosphere using GPS data is already a well developed topic of research).

In a physical implementation of the proposed system, the space-time trajectory of any satellite can be approximately known by just recording the time signals received from the four basic satelites. But the arrival time of the signals must be *measured* and any measurement is subject to experimental uncertainties. It is only through the methodology to be proposed below that an optimal 'model trajectory' is produced.

2.1.2 Observable Parameters

Given a complete model, as just defined, any observation can be predicted, as, for instance:

- The reading of the time of a physical clock (on board of a satellite) can be predicted as $z_i = \int_0^{\tau_i} d\tau \, \sqrt{g_{\alpha\beta} \, (d\tau^\alpha / d\tau) \, (d\tau^\beta / d\tau)} + f(\tau_i)$ along the trajectory;
- The signals sent by the satellites may be observed by other satellites (this measurement being subject to experimental uncertainties). The time of arrival of the signals can be predicted by tracing the zero length geodesic going from the emission point to the reception trajectory (see the methods proposed in appendix 6.2).
- The satellites may have accelerometers, gradiometers, gyroscopes, etc. The observations can also be computed using the given metric and the given trajectories.

The methodology here proposed in based in the assumption that *any observation made by a satellite (the time of the physical clock, the time of arrival of a received signal, the satellite acceleration, etc.) is broadcasted.* The goal of the paper is to propose a methodology that can allow any observer to use all the broadcasted observations to build a complete model that is as good as possible. The model should predict values of the observations that are close to the actual observations (within experimental uncertainties) and it should also have some simple properties (for instance, the metric and the trajectories should have some degree of smoothness). In principle, any observer could run his (inverse) modelization in real (proper) time. As the accessible information will differ for the different observers, the models will also differ. It is only for the part of the space-time that belongs to the common past of two observers that the models may be arbitrarily similar (although not necessarily identical).

Note that the assumed smoothness of the metric and of the trajectories will allow not only to obtain a model of space-time in the past of an observer, but also in his future, although the accuracy of his prediction will rapidly degrade with increasing prediction time (the methodology proposed spontaneously evaluates uncertainties).

In some applications, the observer may not need to make a personal modelization, but just use a simple extrapolation of the information about space-time broadcasted by a central observer in charge of all the computations[3].

2.2 First Constraint on the Metric (Zero Diagonal)

It can be shown (see the lecture by J.M. Pozo in this school) that in the 'light-coordinates' $\{\tau^\alpha\}$ being used, the contravariant components of the metric must are have zeros on the diagonal,

$$\{g^{\alpha\beta}\} \;=\; \begin{pmatrix} 0 & g^{12} & g^{13} & g^{14} \\ g^{12} & 0 & g^{23} & g^{24} \\ g^{13} & g^{23} & 0 & g^{34} \\ g^{14} & g^{24} & g^{34} & 0 \end{pmatrix} \quad , \tag{3}$$

so the basic unknowns of the problem are the six quantities $\{g^{12}, g^{13}, g^{14}, g^{23}, g^{24}, g^{34}\}$. This constraint is imposed exactly, by just expressing all the relations of the theory in terms of these six quantities[4]. The covariant components $g_{\alpha\beta}$ are defined, as usual, by the condition $g_{\alpha\gamma}\, g^{\gamma\beta} = \delta_\alpha^\beta$. The diagonal components of $g_{\alpha\beta}$ are *not* zero.

2.3 Second Constraint on the Metric (Smoothness)

Let $\mathbf{g}_{\text{prior}}$ be some simple initial estimation of the space-time metric field. For instance, we could take for $\mathbf{g}_{\text{prior}}$ the metric of a flat space-time, the Schwarzschild metric of a point mass with the Earth's mass, or a realistic estimation of the actual space-time metric around the Earth.

We wish that our final estimation of the metric, $\mathbf{g}$, is close to the initial estimation. More precisely, letting $\mathbf{C_g}$ be a suitably chosen covariance operator, we are going to impose that the least-squares norm[5]

$$\| \mathbf{g} - \mathbf{g}_{\text{prior}} \|^2_{\mathbf{C_g}} \;\equiv\; \langle\, \mathbf{C}_\mathbf{g}^{-1} \,(\, \mathbf{g} - \mathbf{g}_{\text{prior}} \,)\, ,\, (\, \mathbf{g} - \mathbf{g}_{\text{prior}} \,)\, \rangle \tag{4}$$

is small.

The covariance operator, to be discussed later, shall be a 'smoothing operator' this implying, from one side, that at every point of space-time the final metric is close to the initial metric, and, from another side, that the difference of the two metrics is smooth. As the initial metric shall be smooth, this imposes that the final metric is also smooth. In particular, the final metric will be defined 'continously', in spite of the fact that we only 'sample' it along the space-time trajectories of the satellites and of the light signals.

This kind of smoothing, could perhaps be replaced by a criterion imposing that the Riemann tensor should be as 'small' as possible. The two possibilities must be explored.

---

[3] We call this central observer "Houston".

[4] In the same lecture by J.M. Pozo, it is demonstrated that the metric above is Lorentzian if and only if the following condition is satisfied: defining $A = \sqrt{g^{12}\, g^{34}}$, $B = \sqrt{g^{13}\, g^{24}}$, and $C = \sqrt{g^{14}\, g^{23}}$, one must have $A + B > C$, $B + C > A$, and $C + A > B$. This constraint is not yet introduced, as it may be strongly simplified when using the logarithmic metric.

[5] The criterion in equation 4, that is based on a difference of (contravariant) metrics, is only provisional. In a more advanced state of the theory, we should introduce the logarithm of the metric, and base the minimization criterion on the difference of logarithmic metrics.

2.4 Einstein Equation (Stress-Energy Data)

The notations use in this text for the connection $\Gamma^\alpha{}_{\beta\gamma}$, the Riemann $R^\alpha{}_{\beta\gamma\delta}$, and the Ricci $R_{\alpha\beta}$ associated to a metric $g_{\alpha\beta}$, are as follows:

$$\begin{aligned} \Gamma^\alpha{}_{\beta\gamma} &= \tfrac{1}{2}\, g^{\alpha\sigma} \left(\partial_\beta\, g_{\gamma\sigma} + \partial_\gamma\, g_{\beta\sigma} - \partial_\sigma\, g_{\beta\gamma} \right) \\ R^\alpha{}_{\beta\gamma\delta} &= \partial_\gamma\, \Gamma^\alpha{}_{\delta\beta} - \partial_\delta\, \Gamma^\alpha{}_{\gamma\beta} + \Gamma^\alpha{}_{\mu\gamma}\, \Gamma^\mu{}_{\delta\beta} - \Gamma^\alpha{}_{\mu\delta}\, \Gamma^\mu{}_{\gamma\beta} \\ R_{\alpha\beta} &= R^\gamma{}_{\alpha\gamma\beta} \quad . \end{aligned} \tag{5}$$

The Einstein tensor is then

$$E_{\alpha\beta} \;=\; R_{\alpha\beta} - \tfrac{1}{2}\, g_{\alpha\beta}\, R \quad , \tag{6}$$

where $R = g^{\alpha\beta}\, R_{\alpha\beta}$.

The Einstein equation states that, at every point of the space-time, the Einstein tensor $E_{\alpha\beta}$ (associated to the metric) is proportional to the stress-energy tensor $t_{\alpha\beta}$ describing the matter at this space-time point:

$$E_{\alpha\beta} \;=\; \chi\, t_{\alpha\beta} \quad , \tag{7}$$

where the proportionality constant is $\chi = 8\pi G / c^4$. For instance, in vacuum, $t_{\alpha\beta} = 0$, and, therefore, $E_{\alpha\beta} = 0$. When solving the Einstein equation for $t_{\alpha\beta}$,

$$t_{\alpha\beta} \;=\; \frac{1}{\chi}\, E_{\alpha\beta} \tag{8}$$

we obtain (when replacing $E_{\alpha\beta}$ by the expressions 6–5) the application

$$\mathbf{g} \quad \mapsto \quad \mathbf{t}_{\text{computed}} \;=\; \mathbf{t}(\mathbf{g}) \quad , \tag{9}$$

associating to any metric field $\mathbf{g}$ the corresponding stress-energy field $\mathbf{t}$.

Let $\mathbf{t}_{\text{obs}}$ be our estimation of the stress-energy of the space-time. It could, for instance, be zero, if we take for the space-time the model of vacuum. More realistically, we may take a simple model of the rarefied gas that constitutes the high atmosphere. We wish that the space-time metric $\mathbf{g}$ is such that the associated stress-energy $\mathbf{t}(\mathbf{g})$ is close to $\mathbf{t}_{\text{obs}}$.

More precisely, we are going to impose that the $\mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}}$ is small in the sense of a least-squares norm

$$\| \, \mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}} \, \|^2_{\mathbf{C_t}} \;\equiv\; \langle \, \mathbf{C_t}^{-1} \left(\mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}} \right) , \, \left(\mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}} \right) \rangle \quad , \tag{10}$$

where $\mathbf{C_t}$ is a covariance operator to be discussed later. The notation $\langle \, \cdot \, , \, \cdot \, \rangle$ stands for a duality product.

We shall later need the tangent linear application, $\mathbf{T}$, to the application $\mathbf{t}(\mathbf{g})$. By definition,

$$\mathbf{t}(\, \mathbf{g} + \delta\mathbf{g} \,) \;=\; \mathbf{t}(\mathbf{g}) + \mathbf{T}\, \delta\mathbf{g} + \dots \tag{11}$$

As demonstrated in appendix 6.1 (see equations 90–91), this linear tangent application is the (linear) application that to every $\delta g_{\alpha\beta}$ associates the $\delta t_{\alpha\beta}$ given by

$$\delta t_{\alpha\beta} \;=\; \frac{1}{\chi} \left(\, A_{\alpha\beta}{}^{\mu\nu,\rho\sigma}\, \nabla_{(\rho} \nabla_{\sigma)}\, \delta g_{\mu\nu} + B_{\alpha\beta}{}^{\mu\nu}\, \delta g_{\mu\nu} \, \right) \quad , \tag{12}$$

where

$$\begin{aligned} A_{\alpha\beta}{}^{\mu\nu,\rho\sigma} &= 2\, g^{(\mu|(\sigma}\, \delta^{\rho)}_{(\alpha}, \delta^{|\nu)}_{\beta)} - \tfrac{1}{2}\, g^{\mu\nu}\, \delta^\rho_{(\alpha}\, \delta^\sigma_{\beta)} - \tfrac{1}{2}\, g^{\rho\sigma}\, \delta^\mu_{(\alpha}\, \delta^\nu_{\beta)} + \tfrac{1}{2}\, g^{\mu\nu}\, g^{\rho\sigma}\, g_{\alpha\beta} - \tfrac{1}{2}\, g^{\mu(\rho}\, g^{\sigma)\nu}\, g_{\alpha\beta} \\ B_{\alpha\beta}{}^{\mu\nu} &= \tfrac{1}{2} \left(R^\mu{}_{(\alpha\beta)}{}^\nu + R^{(\mu}{}_{(\alpha}\, \delta^{\nu)}_{\beta)} + R^{\mu\nu}\, g_{\alpha\beta} - R\, \delta^\mu_{(\alpha}\, \delta^\nu_{\beta)} \right) \quad . \end{aligned} \tag{13}$$

2.5 Proper Time Data

Assume that there is a physical clock (i.e., a perhaps accurate, but certainly imperfect clock) on board of some of the satellites. At some instant along the trajectory of such a satelite, a reading of the physical clock is made (i.e., the proper time is measured), this giving some value $z_{\rm obs}$ with some associated uncertainty.

The theoretical prediction of the observation is made via the integration of the space-time length element along trajectory. So, given a model of the metric $\mathbf{g} = \{g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\}$, a model of the trajectory, $\boldsymbol{\lambda} = \{\tau^\alpha(\lambda)\}$, and a model of the clock drift, $\mathbf{f} = \{f(\tau)\}$, the prediction of the clock reading is

$$z(\tau) \;=\; \underbrace{\int_0^\tau d\lambda}_{\lambda} \; \sqrt{g_{\alpha\beta}\, u^\alpha\, u^\beta} + f(\tau) \quad , \tag{14}$$

where the integral is performed along the trajectory where the vector u^α is defined (along the trajectory) as

$$u^\alpha \;=\; \frac{d\tau^\alpha}{d\lambda} \quad , \tag{15}$$

and where λ is a parameter along the trajectory.

Expression 14 is written for the evaluation of one single time, while we shall typically many times evaluated along the trajectory, $\mathbf{z} = \{z_1, z_2, \dots\}$. The application defined by expression 14, but considered for all times, shall be written as

$$\{\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}\} \;\mapsto\; \mathbf{z}_{\rm computed} \;=\; \mathbf{z}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}) \quad . \tag{16}$$

Let $\mathbf{z}_{\rm obs}$ the set of observed values, with experimental uncertainties represented by a covariance matrix $\mathbf{C_z}$. We wish that our final model $\{\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}\}$ is such that the (least-squares) norm

$$\| \,\mathbf{z}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}) - \mathbf{z}_{\rm obs}\, \|^2_{\mathbf{C_z}} \;\equiv\; \langle\, \mathbf{C_z}^{-1}\,(\, \mathbf{z}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}) - \mathbf{z}_{\rm obs}\,)\,,\, (\, \mathbf{z}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}) - \mathbf{z}_{\rm obs}\,)\,\rangle \tag{17}$$

is small.

For later use, we shall need the three (partial) linear tangent applications to the application so defined. They are defined through the series development

$$\mathbf{z}(\mathbf{g}+\delta\mathbf{g}, \boldsymbol{\lambda}+\delta\boldsymbol{\lambda}, \mathbf{f}+\delta\mathbf{f}) \;=\; \mathbf{z}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f}) + \mathbf{Z_g}\,\delta\mathbf{g} + \mathbf{Z}_\lambda\,\delta\boldsymbol{\lambda} + \mathbf{Z_f}\,\delta\mathbf{f} + \dots \quad , \tag{18}$$

where, for short, we write $\mathbf{Z_g}$, $\mathbf{Z}_\lambda$, and $\mathbf{Z_f}$, instead of $\mathbf{Z_g}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f})$, $\mathbf{Z}_\lambda(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f})$, and $\mathbf{Z_f}(\mathbf{g}, \boldsymbol{\lambda}, \mathbf{f})$. One asily sees that $\mathbf{Z_g}$ is the (linear) operator that to any metric perturbation $g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \mapsto g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) + \delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)$, associates, at each measure point, the time perturbation

$$\delta z \;=\; \frac{1}{2} \int d\lambda\; \frac{\delta g_{\alpha\beta}\, u^\alpha\, u^\beta}{\sqrt{g_{\mu\nu}\, u^\mu\, u^\nu}} \quad . \tag{19}$$

$\mathbf{Z}_\lambda$ is the (linear) operator that to any trajectory perturbation $\tau^\alpha(\lambda) \mapsto \tau^\alpha(\lambda) + \delta\tau^\alpha(\lambda)$, associates, at each measure point, the time perturbation

$$\delta z \;=\; \int d\lambda\; \frac{g_{\alpha\beta}\, u^\alpha\, \delta u^\beta}{\sqrt{g_{\mu\nu}\, u^\mu\, u^\nu}} \quad , \tag{20}$$

where

$$\delta u^\alpha \;=\; \frac{d\,\delta\tau^\alpha}{d\lambda} \quad . \tag{21}$$

Finally, $\mathbf{Z_f}$ is the (linear) operator that to any perturbation $\delta f(\tau)$ of the clock drift function associates

$$\delta z \;=\; \delta f \quad . \tag{22}$$

2.6 Arrival Time Data

At some instant τ_e along its trajectory, a satellite emits a time signal, that is received by another satellite at (proper) time σ .

Given a model of the metric $\mathbf{g} = \{g_{\alpha\beta}(\tau^1, \tau^2, \tau^3, \tau^4)\}$, a model of the trajectory that emits the signal, $\boldsymbol{\tau}_e = \{\tau_e^\alpha(\tau)\}$, a model of the emission time τ_e along this trajectory, a model of the trajectory that receives the signal, $\boldsymbol{\tau}_r = \{\tau_r^\alpha(\tau)\}$, and a model of the clock drift of the receiving satellite, $\mathbf{f}_r = \{f_r(\tau)\}$, we can predict the reception time by tracing the zero-length geodesic that connects the emission point to the receiving trajectory. We shall write this theoretical prediction as

$$\sigma_{\text{computed}} \; = \; \sigma(\mathbf{g}, \boldsymbol{\tau}_e, \tau_e, \boldsymbol{\tau}_r, \mathbf{f}_r) \quad , \tag{23}$$

In the real situation, the metric is only known approximately, and the computed value of the arrival time will not be identical to the time actually observed time, say σ_{obs} .

Roughly speaking, our goal is going to be to determine the space-time metric that minimizes the differences between calculated and observed arrival times.

Our data, therefore, consists on a set of values

$$\{ \, \sigma^i_{\text{obs}} \, \} \quad , \tag{24}$$

assumed to be subjected to some observational uncertainties. Letting $\mathbf{C}_\sigma$ denote the covariance operator describing experimental uncertainties, we wish the (least-squares) norm

$$\| \, \boldsymbol{\sigma}(\mathbf{g}, \boldsymbol{\tau}_e, \tau_e, \boldsymbol{\tau}_r, \mathbf{f}_r) - \boldsymbol{\sigma}_{\text{obs}} \, \|^2_{\mathbf{C}_\sigma} \; \equiv \; \langle \; \mathbf{C}_\sigma^{-1} \, (\, \boldsymbol{\sigma}(\mathbf{g}, \boldsymbol{\tau}_e, \tau_e, \boldsymbol{\tau}_r, \mathbf{f}_r) - \boldsymbol{\sigma}_{\text{obs}} \,) \; , \; (\, \boldsymbol{\sigma}(\mathbf{g}, \boldsymbol{\tau}_e, \tau_e, \boldsymbol{\tau}_r, \mathbf{f}_r) - \boldsymbol{\sigma}_{\text{obs}} \,) \; \rangle \tag{25}$$

to be small.

Below, we shall need the (partial) tangent linear operators to the operator $\boldsymbol{\sigma}$, defined as follows,

$$\begin{aligned} &\boldsymbol{\sigma}(\, \mathbf{g} + \delta\mathbf{g} \, , \, \boldsymbol{\tau}_e + \delta\boldsymbol{\tau}_e \, , \, \tau_e + \delta\tau_e \, , \, \boldsymbol{\tau}_r + \delta\boldsymbol{\tau}_r \, , \, \mathbf{f}_r + \delta\mathbf{f}_r \,) \; = \\ &\boldsymbol{\sigma}(\mathbf{g}, \boldsymbol{\tau}_e, \tau_e, \boldsymbol{\tau}_r, \mathbf{f}_r) + \boldsymbol{\Sigma}_{\mathbf{g}} \, \delta\mathbf{g} + \boldsymbol{\Sigma}_{\boldsymbol{\tau}_e} \, \delta\boldsymbol{\tau}_e + \boldsymbol{\Sigma}_{\tau_e} \, \delta\tau_e + \boldsymbol{\Sigma}_{\boldsymbol{\tau}_r} \, \delta\boldsymbol{\tau}_r + \boldsymbol{\Sigma}_{\mathbf{f}_r} \, \delta\mathbf{f}_r + \dots \quad . \end{aligned} \tag{26}$$

Let us evaluate them.

When the metric is perturbed from $g_{\alpha\beta}$ to $g_{\alpha\beta} + \delta g_{\alpha\beta}$, the computed arrival times are perturbed from σ to $\sigma + \delta\sigma$, where (see equation 120 in appendix 6.2)

$$\delta\sigma \; = \; -\frac{1/2}{g_{\mu\nu} \, u^\mu \, \ell^\nu} \int_{\lambda(\mathbf{g})} d\lambda \; \ell^\alpha \, \ell^\beta \, \delta g_{\alpha\beta} \quad , \tag{27}$$

where u^α is the tangent vector to the trajectory of the receiver, $u^\alpha = d\tau^\alpha / d\tau$, and ℓ^α is the tangent vector to the trajectory of the light ray, $\ell^\alpha = d\tau^\alpha / d\lambda$ (where λ is an affine parameter along the ray). Therefore, $\boldsymbol{\Sigma}_{\mathbf{g}}$ is the linear operator that to any metric perturbation $\delta\mathbf{g}$ associates the $\delta\sigma$ in equation 27.

To evaluate the operator $\boldsymbol{\Sigma}_{\boldsymbol{\tau}_e}$ we have to solve the following problem: which is the first order perturbation $\delta\sigma$ to the arrival time when the trajectory of the emitter is perturbed from $\tau_e^\alpha(\tau)$ to $\tau_e^\alpha(\tau) + \delta\tau_e^\alpha(\tau)$? ($\rightarrow$ computation being performed).

To evaluate the operator $\boldsymbol{\Sigma}_{\tau_e}$ we have to solve the following problem: which is the first order perturbation $\delta\sigma$ to the arrival time when the (proper) time the emitter is perturbed from τ_e to $\tau_e + \delta\tau_e$? ($\rightarrow$ computation being performed).

To evaluate the operator $\mathbf{\Sigma}_{\tau_r}$ we have to solve the following problem: which is the first order perturbation $\delta\sigma$ to the arrival time when the trajectory of the receiver is perturbed from $\tau_r^\alpha(\tau)$ to $\tau_r^\alpha(\tau) + \delta\tau_r^\alpha(\tau)$? (→ computation being performed).

The operator $\mathbf{\Sigma_f}$ is the linear operator that to any perturbation $\delta f_r(\tau)$ of the receiver clock drift functions associates (check the notations)

$$\delta\sigma \; = \; \delta f_r(\sigma) \quad . \tag{28}$$

2.7 Accelerometer Data

We have to explore here the case where each 'satellite' has an accelerometer. The acceleration along a trajectory is

$$a^\alpha \; = \; u^\beta \, \frac{\partial u^\alpha}{\partial x^\beta} + \Gamma^\alpha{}_{\beta\gamma} \, u^\beta \, u^\gamma \; = \; \frac{du^\alpha}{d\tau} + \Gamma^\alpha{}_{\beta\gamma} \, u^\beta \, u^\gamma \quad , \tag{29}$$

where τ is the proper time along the trajectory.

The easiest way to 'measure' the acceleration on-board would be, of course, to force the satellite (or its clock) to be in free-fall (i.e., to follow a geodesic of the spacetime metric). Then, one would have $a^\alpha = 0$. Let us keep considering here the general case where the acceleration may be nonzero (because, for instance, by residual drag by the high atmosphere), but it is measured.

The measure of the acceleration provides information on the connection, i.e., in fact, on the gradients of the metric.

Given a model $\mathbf{g}$ of the metric field and a model $\boldsymbol{\lambda}$ of the trajectory, equation 29 allows to compute the acceleration at all the space-time points when it is measured. We write

$$\{\mathbf{g}, \boldsymbol{\lambda}\} \quad \mapsto \quad \mathbf{a}_{\text{computed}} \; = \; \mathbf{a}(\mathbf{g}, \boldsymbol{\lambda}) \tag{30}$$

the application so defined. We wish that the computed accelerations, $\mathbf{a}(\mathbf{g})$, are close to the observed ones, say $\mathbf{a}_{\text{obs}}$. More precisely, we wish the (least-squares) norm

$$\| \, \mathbf{a}(\mathbf{g}, \boldsymbol{\lambda}) - \mathbf{a}_{\text{obs}} \, \|^2_{\mathbf{C_a}} \; \equiv \; \langle \; \mathbf{C_a}^{-1} \, (\, \mathbf{a}(\mathbf{g}, \boldsymbol{\lambda}) - \mathbf{a}_{\text{obs}} \,) \; , \; (\, \mathbf{a}(\mathbf{g}, \boldsymbol{\lambda}) - \mathbf{a}_{\text{obs}} \,) \; \rangle \tag{31}$$

to be small, where $\mathbf{C_a}$ is a covariance operator describing the experimental uncertainties in the measured acceleration values.

We introduce the tangent linear operators

$$\mathbf{a}(\, \mathbf{g} + \delta\mathbf{g} \, , \, \boldsymbol{\lambda} + \delta\boldsymbol{\lambda} \,) \; = \; \mathbf{a}(\mathbf{g}, \boldsymbol{\lambda}) + \mathbf{A_g} \, \delta\mathbf{g} + \mathbf{A}_\lambda \, \delta\boldsymbol{\lambda} + \ldots \tag{32}$$

It follows from equation 29 that a perturbation of the metric $g_{\alpha\beta} \mapsto g_{\alpha\beta} + \delta g_{\alpha\beta}$, produces a perturbation of the computed acceleration given by $\delta a^\alpha = \delta\Gamma^\alpha{}_{\beta\gamma} \, u^\beta \, u^\gamma$. The expression for $\delta\Gamma^\alpha{}_{\beta\gamma}$ is in appendix 6.1 (see equation 79, page 25), $\delta\Gamma^\alpha{}_{\beta\gamma} \; = \; g^{\alpha\sigma} \, \delta\Gamma_{\sigma\beta\gamma}$, with $\delta\Gamma_{\alpha\beta\gamma} \; = \; \frac{1}{2} \, (\nabla_\gamma \delta g_{\alpha\beta} + \nabla_\beta \delta g_{\alpha\gamma} - \nabla_\alpha \delta g_{\beta\gamma})$. This gives

$$\delta a^\alpha \; = \; \tfrac{1}{2} \, g^{\alpha\sigma} \left(\nabla_\gamma \, \delta g_{\sigma\beta} + \nabla_\beta \, \delta g_{\sigma\gamma} - \nabla_\sigma \, \delta g_{\beta\gamma} \right) u^\beta \, u^\gamma \quad . \tag{33}$$

The linear operator so defined was denoted $\mathbf{A_g}$ in equation 32. To any metric field perturbation $\delta g_{\alpha\beta}$ this operator associates, at every point of a space-time trajectory where the acceleration was measured, the values δa^α just written.

We leave to the reader the characterization of the operator $\mathbf{A}_\lambda$.

2.8 Gyroscope Data

A gyroscope is described by its *spin vector* (or angular momentum vector) s^α , a four-vector that is orthogonal to the four-velocity u^α of the rotating particle: $g_{\alpha\beta}\, u^\alpha\, s^\beta$.

Assume that the gyroscope follows a trajectory $x^\alpha = x^\alpha(\tau)$, whose velocity is $u^\alpha = dx^\alpha/d\tau$ and whose acceleration a^α is that expressed in equation 29. Then, the evolution of the spin vector along the trajectory is described[6] by the so-called Fermi-Walker transport:

$$\frac{Ds^\alpha}{d\tau} \;\equiv\; \frac{ds^\alpha}{d\tau} + \Gamma^\alpha{}_{\beta\gamma}\, u^\beta\, s^\gamma \;=\; s_\beta\, (a^\beta\, u^\alpha - a^\alpha\, u^\beta) \quad . \tag{34}$$

Should the gyroscope be in free fall, $a^\alpha = 0$, and $ds^\alpha/d\tau + \Gamma^\alpha{}_{\beta\gamma}\, u^\beta\, s^\gamma = 0$, this meaning that the spin vector would be transported by parallelism.

In our case, the monitoring of the spin vector $s^\alpha(\tau)$ (besides the monitoring of the acceleration a^α) would provide the values $\Gamma^\alpha{}_{\beta\gamma}\, u^\beta\, s^\gamma$, an information complementary to that provided by the monitoring of the acceleration (that provides the values $\Gamma^\alpha{}_{\beta\gamma}\, u^\beta\, u^\gamma$).

Consider that our data is

$$\pi^\alpha \;=\; \frac{ds^\alpha}{d\tau} \quad . \tag{35}$$

Then we have

$$\pi^\alpha \;=\; s_\beta\, (a^\beta\, u^\alpha - a^\alpha\, u^\beta) - \Gamma^\alpha{}_{\beta\gamma}\, u^\beta\, s^\gamma \quad . \tag{36}$$

Given the metric field model $\mathbf{g}$ and the trajectory model $\boldsymbol{\lambda}$, this equation allows to compute the vector π^α at all the space-time points when it is measured. We write

$$\{\mathbf{g}, \boldsymbol{\lambda}\} \quad\mapsto\quad \boldsymbol{\pi}_{\text{computed}} \;=\; \boldsymbol{\pi}(\mathbf{g}, \boldsymbol{\lambda}) \tag{37}$$

the application so defined. We wish that the computed values, $\boldsymbol{\pi}(\mathbf{g})$, are close to the observed ones, say $\boldsymbol{\pi}_{\text{obs}}$. More precisely, we wish the (least-squares) norm

$$\|\, \boldsymbol{\pi}(\mathbf{g}, \boldsymbol{\lambda}) - \boldsymbol{\pi}_{\text{obs}} \,\|^2_{\mathbf{C}_\pi} \;\equiv\; \langle\, \mathbf{C}_\pi^{-1}\, (\, \boldsymbol{\pi}(\mathbf{g}, \boldsymbol{\lambda}) - \boldsymbol{\pi}_{\text{obs}} \,)\, ,\, (\, \boldsymbol{\pi}(\mathbf{g}, \boldsymbol{\lambda}) - \boldsymbol{\pi}_{\text{obs}} \,)\, \rangle \tag{38}$$

to be small, where $\mathbf{C}_\pi$ is a covariance operator describing the experimental uncertainties in the measured values.

Of course, one may not wish to measure the evolution of the spin vector to provide information on the connection, but to 'test' general relativity, as in the Gravity Probe B experiment. From the viewpoint of the present work, the detection of any inconsistency in the data would put relativity theory in jeopardy.

Let us introduce the linear tangent operators

$$\boldsymbol{\pi}(\, \mathbf{g} + \delta\mathbf{g}\, ,\, \boldsymbol{\lambda} + \delta\mathbf{g}\,) \;=\; \boldsymbol{\pi}(\mathbf{g}, \boldsymbol{\lambda}) + \boldsymbol{\Pi}_{\mathbf{g}}\, \delta\mathbf{g} + \boldsymbol{\Pi}_{\boldsymbol{\lambda}}\, \delta\boldsymbol{\lambda} + \dots \tag{39}$$

The application $\mathbf{g} \mapsto \boldsymbol{\pi}(\mathbf{g})$ is given in equation 36. To compute the first order perturbation $\boldsymbol{\pi} \mapsto \boldsymbol{\pi} + \delta\boldsymbol{\pi}$ produced by a perturbation $\mathbf{g} \mapsto \mathbf{g} + \delta\mathbf{g}$, we must, in this equation, make the replacements $\Gamma^\alpha{}_{\beta\gamma} \mapsto \Gamma^\alpha{}_{\beta\gamma} + \delta\Gamma^\alpha{}_{\beta\gamma}$ and $s^\alpha \mapsto s^\alpha + \delta s^\alpha$, with subsequent expression of $\delta\Gamma^\alpha{}_{\beta\gamma}$ and δs^α in terms of δg_α . We obtain

$$\delta\pi^\alpha \;=\; -\delta\Gamma^\alpha{}_{\beta\gamma}\, u^\beta\, s^\gamma \quad . \tag{40}$$

[6]For details on the relativistic treatment of a spinning test particle, see Papatetrou (1951), Weinberg (1972), or Hernández-Pastora et al. (2001).

Using the expression for $\delta\Gamma^{\alpha}{}_{\beta\gamma}$ in appendix 6.1, we are immediately left to an expression similar to 33:

$$\delta\pi^{\alpha} \;=\; -\tfrac{1}{2}\, g^{\alpha\sigma}\left(\nabla_{\gamma}\,\delta g_{\sigma\beta} + \nabla_{\beta}\,\delta g_{\sigma\gamma} - \nabla_{\sigma}\,\delta g_{\beta\gamma}\right) u^{\beta}\,\sigma^{\gamma} \quad . \tag{41}$$

The operator that to any $\delta g_{\alpha\beta}$ associates the $\delta\pi^{\alpha}$ given by this equation is the operator $\mathbf{\Pi_g}$, we were searching for.

We leave to the reader the evaluation of the operator $\mathbf{\Pi}_{\lambda}$.

2.9 Gradiometer Data

To study the gravity field around the Earth, different satellite missions are on course or planned[7]. Of particular importance are the *gradiometers* with which modern gravimetric satellites are equipped. In the GOCE satellite, there are three perpendicular "gradiometer arms", each arm consisting in two masses (50 cm apart) that are submitted to electrostatic forces to keep each of them at the center of a cage. These forces are monitored, thus providing the accelerations. The basic data are the half-sum and the difference of these accelerations (for each of the three gradiometer arms).

The half-sum of the accelerations gives what a simple accelerometer would give. The difference corresponds to the "tidal forces" in the region where the satellite operates.

A simple model for the gradiometry data is as follows. A mass follows some space-time line that, to simplify the discussion, is assumed to be a geodesic (i.e., the mass is assumed to be in free-fall, but taking into account its possible acceleration would be simple). (We leave to the reader the writing of the general formulas for the case when the initial trajectory is not a geodesic.) This geodesic is represented at the left in figure 1. Let u^{α} be the unit vector tangent to this geodesic trajectory. Consider, at some initial point along the geodesic, a "small" space-time vector δv^{α} that, to fix ideas, may be assumed to be a space-like vector. By parallel transport of δv^{α} along the geodesic one defines a second trajectory, that is not necessarily a geodesic (the line at the right in figure 1). Let us denote w^{α} the tangent vector to this trajectory, and δa^{α} the acceleration along it. Note that, as the trajectory is close to being geodesic, the acceleration δa^{α} is small (and would vanish if $\delta v^{\alpha} = 0$).

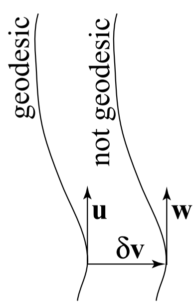

Figure 1: For the incorporation of gradiometry data, we consider a geodesic space-time trajectory, and the trajectory defined by transporting a small vector along the geodesic (see text for details).

[7]The LAGEOS (LAser GEOdynamics Satellites) are passive spherical bodies covered with retroreflectors. Note that Ciufolini and Pavlis (2004) have recently been able to confirm the Lense-Thirring effect using LAGEOS data. The CHAMP (CHAllenging Minisatellite Payload) satellite is equipped with a precise orbit determination and an accelerometer. The GRACE (GRAvity recovery and Climate Experiment) consists in two satellites with precise orbit determination, accelerometers and measure of their mutual distance with an accuracy of a few microns. The GOCE (Gravity Field and Steady-State Ocean Circulation Explorer) satellite has recently been launched. It consists in a three axis *gradiometer*: six accelerometers in a so-called diamond configuration. The observables are the differences of the accelerations.

Notice that one has

$$g_{\mu\nu}\, u^\mu\, u^\nu \;=\; 1 \qquad ; \qquad g_{\mu\nu}\, w^\mu\, \delta a^\nu \;=\; 0 \quad , \tag{42}$$

and that (i) the tangent vector w^α is obtained, all along the trajectory, by parallel transport of u^α along δv^α, and (ii) at this level of approximation, the proper time along the second trajectory is identical to the proper time along the first trajectory.

A mass can be forced to follow this line, and the forces required to do this can be monitored, this giving a measurement of the acceleration δa^α of the mass.

We do not need to exactly evaluate the theoretical relation expressing δa^α, the approximation that is first order in δv^α will be sufficient (because δv^α is small). As demonstrated by Pozo et al. (2005), one has $\delta a^\alpha \;=\; R^\alpha{}_{\mu\nu\rho}\, u^\mu\, u^\rho\, \delta v^\nu + \dots$, where the remaining terms are at least second order in δv^α. Then, with a sufficient approximation, we use below the expression

$$\delta a^\alpha \;=\; R^\alpha{}_{\mu\nu\rho}\, u^\mu\, u^\rho\, \delta v^\nu \quad . \tag{43}$$

As the three vectors a^α, u^α, and δv^α are known, we have a direct information on the components of the Riemann tensor.

A typical gradiometer contains three arms (in three perpendicular directions in space). This means that we have three different vectors δv^α with which to apply equation 43. The vector u^α is unique (fixed by the trajectory of the satellite). Should one have different satellites at approximately the same space-time point, with significantly different trajectories, one would have extra constraints on the Riemann tensor (at the given space-time point).

In order to simplify the notations in later sections of the paper, we drop the δ for the vector δv^α, and we write ω^α instead of δa^α. Then, equation 43 becomes

$$\omega^\alpha \;=\; R^\alpha{}_{\mu\nu\rho}\, u^\mu\, u^\rho\, v^\nu \quad . \tag{44}$$

Given a model metric field $\mathbf{g}$ and a model trajectory $\boldsymbol{\lambda}$, the theoretical values of the tidal acceleration are detoted $\boldsymbol{\omega}_{\text{computed}}$, and we write

$$\{\mathbf{g}, \boldsymbol{\lambda}\} \quad \mapsto \quad \boldsymbol{\omega}_{\text{computed}} \;=\; \boldsymbol{\omega}(\mathbf{g}, \boldsymbol{\lambda}) \quad , \tag{45}$$

where $\omega^\alpha_{\text{computed}} \;=\; R^\alpha{}_{\mu\nu\rho}(\mathbf{g})\, u^\mu\, u^\rho\, v^\nu$. The gradiometer provides the 'observed acceleration' $\boldsymbol{\omega}_{\text{obs}}$, with observational uncertainties represented by a covariance operator $\mathbf{C}_\omega$. We wish that the tidal accelerations, $\boldsymbol{\omega}(\mathbf{g}, \boldsymbol{\lambda})$, are close to the observed ones, $\boldsymbol{\omega}_{\text{obs}}$. More precisely, we wish the (least-squares) norm

$$\|\, \boldsymbol{\omega}(\mathbf{g}, \boldsymbol{\lambda}) - \boldsymbol{\omega}_{\text{obs}} \,\|^2_{\mathbf{C}_\omega} \;\equiv\; \langle\, \mathbf{C}_\omega^{-1}\, (\, \boldsymbol{\omega}(\mathbf{g}, \boldsymbol{\lambda}) - \boldsymbol{\omega}_{\text{obs}}\,)\, ,\, (\, \boldsymbol{\omega}(\mathbf{g}, \boldsymbol{\lambda}) - \boldsymbol{\omega}_{\text{obs}}\,)\, \rangle \tag{46}$$

to be small.

We introduce the linear tangent operators

$$\boldsymbol{\omega}(\, \mathbf{g} + \delta\mathbf{g}\, ,\, \boldsymbol{\lambda} + \delta\boldsymbol{\lambda}\,) \;=\; \boldsymbol{\omega}(\mathbf{g}, \boldsymbol{\lambda}) + \boldsymbol{\Omega}_{\mathbf{g}}\, \delta\mathbf{g} + \boldsymbol{\Omega}_{\boldsymbol{\lambda}}\, \delta\boldsymbol{\lambda} + \dots \tag{47}$$

In view of equation 44, a perturbation of the metric field will produce the perturbation

$$\delta\omega^\alpha \;=\; \delta R^\alpha{}_{\beta\gamma\delta}\, u^\beta\, u^\delta\, v^\gamma \tag{48}$$

of the tidal acceleration, where $\delta R^\alpha{}_{\mu\nu\rho}$ is the first order perturbation to the Riemann tensor. This perturbation is obtained as a by product in our computation of the perturbation of the Einstein tensor in appendix 6.1 (see equation 81, page 25). The result is

$$\delta R^\alpha{}_{\beta\gamma\delta} \;=\; 2\nabla_{[\gamma}\Omega^\alpha{}_{\delta]\beta} \quad , \tag{49}$$

where

$$\Omega^{\alpha}{}_{\beta\gamma} \; = \; g^{\alpha\sigma}\Omega_{\sigma\beta\gamma} \qquad \text{with} \qquad \Omega_{\alpha\beta\gamma} \; = \; \tfrac{1}{2}\left(\nabla_\gamma \delta g_{\alpha\beta} + \nabla_\beta \delta g_{\alpha\gamma} - \nabla_\alpha \delta g_{\beta\gamma}\right) \quad . \tag{50}$$

This characterizes the operator $\mathbf{\Omega}_{\mathbf{g}}$.

We leave to the reader the characterization of the operator $\mathbf{\Omega}_{\lambda}$.

2.10 Total Misfit

It is clear that we need to optimize for all the components of a complete model, and these include, in particular, the space-time metric and the space-time trajectories. In the context of this school, we choose to present a simplified version of the theory, were only the space-time metric is optimized. The students should be able to complete the expression of the total misfit, as an exercise.

Using standard arguments from least-squares theory (see Tarantola [2005]), we shall define here the 'best metric field' as the field $\mathbf{g}$ that minimizes the sum of all the misfit terms introduced above (equations 10, 4, 17, 25, 31, 38, and 46). The total misfit function, that we denote $S(\mathbf{g})$, is, therefore, given by

$$\boxed{\begin{aligned} 2\,S(\mathbf{g}) \; &= \| \, \mathbf{g} - \mathbf{g}_{\text{prior}} \, \|^2_{\mathbf{C}_\mathbf{g}} + \| \, \mathbf{z}(\mathbf{g}) - \mathbf{1} \, \|^2_{\mathbf{C}_\mathbf{z}} + \| \, \mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}} \, \|^2_{\mathbf{C}_\mathbf{t}} + \| \, \boldsymbol{\sigma}(\mathbf{g}) - \boldsymbol{\sigma}_{\text{obs}} \, \|^2_{\mathbf{C}_\sigma} \\ &+ \| \, \mathbf{a}(\mathbf{g}) - \mathbf{a}_{\text{obs}} \, \|^2_{\mathbf{C}_\mathbf{a}} + \| \, \boldsymbol{\pi}(\mathbf{g}) - \boldsymbol{\pi}_{\text{obs}} \, \|^2_{\mathbf{C}_\pi} + \| \, \boldsymbol{\omega}(\mathbf{g}) - \boldsymbol{\omega}_{\text{obs}} \, \|^2_{\mathbf{C}_\omega} \quad , \end{aligned}} \tag{51}$$

i.e.,

$$\begin{aligned} 2\,S(\mathbf{g}) \; = \; & \langle \, \mathbf{C}_\mathbf{g}^{-1} \, (\, \mathbf{g} - \mathbf{g}_{\text{prior}} \,) \; , \; (\, \mathbf{g} - \mathbf{g}_{\text{prior}} \,) \, \rangle \\ & + \langle \, \mathbf{C}_\mathbf{z}^{-1} \, (\, \mathbf{z}(\mathbf{g}) - \mathbf{1} \,) \; , \; (\, \mathbf{z}(\mathbf{g}) - \mathbf{1} \,) \, \rangle \\ & + \langle \, \mathbf{C}_\mathbf{t}^{-1} \, (\, \mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}} \,) \; , \; (\, \mathbf{t}(\mathbf{g}) - \mathbf{t}_{\text{obs}} \,) \, \rangle \\ & + \langle \, \mathbf{C}_\sigma^{-1} \, (\, \boldsymbol{\sigma}(\mathbf{g}) - \boldsymbol{\sigma}_{\text{obs}} \,) \; , \; (\, \boldsymbol{\sigma}(\mathbf{g}) - \boldsymbol{\sigma}_{\text{obs}} \,) \, \rangle \\ & + \langle \, \mathbf{C}_\mathbf{a}^{-1} \, (\, \mathbf{a}(\mathbf{g}) - \mathbf{a}_{\text{obs}} \,) \; , \; (\, \mathbf{a}(\mathbf{g}) - \mathbf{a}_{\text{obs}} \,) \, \rangle \\ & + \langle \, \mathbf{C}_\pi^{-1} \, (\, \boldsymbol{\pi}(\mathbf{g}) - \boldsymbol{\pi}_{\text{obs}} \,) \; , \; (\, \boldsymbol{\pi}(\mathbf{g}) - \boldsymbol{\pi}_{\text{obs}} \,) \, \rangle \\ & + \langle \, \mathbf{C}_\omega^{-1} \, (\, \boldsymbol{\omega}(\mathbf{g}) - \boldsymbol{\omega}_{\text{obs}} \,) \; , \; (\, \boldsymbol{\omega}(\mathbf{g}) - \boldsymbol{\omega}_{\text{obs}} \,) \, \rangle \quad . \end{aligned} \tag{52}$$

Sometimes, in least-squares theory it is allowed for these different terms to have different 'weights', by multiplying them by some ad-hoc numerical factors. This is not necessary if all the covariance operators are chosen properly. In any case, adding some extra numerical factors is a trivial task that we do not contemplate here.

Although in this paper we limit our scope to providing the simplest method that could be used to actually find the metric field $\mathbf{g}$ that minimizes the misfit function, it is interesting to know that the function $S(\mathbf{g})$ carries a more fundamental information. In fact, as shown, for instance, in Tarantola (2005), the expression

$$\varphi(\mathbf{g}) \; = \; k \, \exp(\, - S(\mathbf{g}) \,) \tag{53}$$

defines a probability density (infinite-dimensional) that represents the information we have on the actual metric field, i.e., in fact, the respective 'likelihoods' of all possible metric fields.

3 Optimization

We are going to present here a very plain optimization algorithm, based on the Newton's method. This is not a valid candidate for a practical algorithm. Let us see why.

The Newton algorithm, as proposed here, produces an ab-initio calculation: one takes all the observations, the a priori information on the complete model and produces the a posteriori information on the complete model. All the data are treated together. Of course, any practical algorithm should rather use the basic idea of a *Kalman filter*: data are integrated into the computation as they are available (see appendix 6.5 for a short description of the linear Kalman filter).

In itself, this in an important topic for our future research: Kalman filter computations are made in real time, but what is 'real time' in the context of the relativistic physics used here? It turns out that 'real time' is replaced in this context by 'proper observer time' and that the data integrated by an observer in his Kalman filter are the data arriving to him at this moment, i.e., the data belonging to the 'surface' of his past light cone. We leave this advanced topic out from the teachings of this school.

Therefore, we proceed with the analysis of a simple steepest-descent algorithm.

3.1 Iterative Algorithm

Once the misfit function $S(\mathbf{g})$ as been defined (equation 52), and the associated probability distribution $f(\mathbf{g})$ has been introduced (equation 53), the ideal (although totally impractical) approach for extracting all the information on $\mathbf{g}$ brought by the data of our problem would be to sample the probability distribution[8] $f(\mathbf{g})$. Examples of the sampling of a probability distribution in the context of inverse problems can be found in Tarantola (2005).

In the present problem, where the initial metric shall not be too far from the actual metric, the nonlinearities of the problem are going to be weak. This implies that the probability distribution $f(\mathbf{g})$ is monomodal, i.e., the misfit function $S(\mathbf{g})$ has a unique minimum (in the region of interest of the parameter space). Therefore, the general sampling techniques can here be replaced by the much more efficient optimization techniques. The basic question becomes: *for which metric field $\mathbf{g}$ the misfit function $S(\mathbf{g})$ attains its minimum?*

This problem can be solved using gradient-based techniques. These techniques are quite sophisticated, and require careful adaptation to the problem at hand if they have to work with acceptable efficiency. As we do not wish to develop this topic in this paper, we just choose here to explore the simple steepest descent algorithm, while we explore the more complete Newton algorithm in appendix 6.4.

To run a steepest descent optimization algorithm, there is only one evaluation that must be done extremely accurately, the evaluation of the direction of steepest ascent. For this one may use the

[8] Sampling an infinite-dimensional probability distribution is not possible, but we could define a (dense enough) grid in the space-time where the values of $\mathbf{g}$ are considered, this discretization rendering the probability distribution finite-dimensional.

formulas developed in Tarantola (2005). One obtains the following direction of steepest ascent,

$$\boxed{\begin{aligned} \gamma_k \; = \; (\mathbf{g}_k - \mathbf{g}_{\text{prior}}) &+ (\mathbf{Z}_k\,\mathbf{C_g})^t\,\mathbf{C_z}^{-1}\,(\mathbf{z}(\mathbf{g}_k) - \mathbf{1}) \\ &+ (\mathbf{T}_k\,\mathbf{C_g})^t\,\mathbf{C_t}^{-1}\,(\mathbf{t}(\mathbf{g}_k) - \mathbf{t}_{\text{obs}}) \\ &+ (\mathbf{\Sigma}_k\,\mathbf{C_g})^t\,\mathbf{C}_\sigma^{-1}\,(\sigma(\mathbf{g}_k) - \sigma_{\text{obs}}) \\ &+ (\mathbf{A}_k\,\mathbf{C_g})^t\,\mathbf{C_a}^{-1}\,(\mathbf{a}(\mathbf{g}_k) - \mathbf{a}_{\text{obs}}) \\ &+ (\mathbf{\Pi}_k\,\mathbf{C_g})^t\,\mathbf{C}_\pi^{-1}\,(\pi(\mathbf{g}_k) - \pi_{\text{obs}}) \\ &+ (\mathbf{\Omega}_k\,\mathbf{C_g})^t\,\mathbf{C}_\omega^{-1}\,(\omega(\mathbf{g}_k) - \omega_{\text{obs}}) \quad , \end{aligned}} \tag{54}$$

where the linear operators $\mathbf{Z}_k$, $\mathbf{T}_k$, $\mathbf{\Sigma}_k$, $\mathbf{A}_k$, $\mathbf{\Pi}_k$, and $\mathbf{\Omega}_k$ have all been introduced above. The meaning of $(\cdot)^t$ in the above should be obvious. We say *transpose* operators, better than *dual* operators, because the difference between the two notions matters inside the theory of least-squares[9].

The Newton algorithm presented in appendix 6.4, is typically not used as such. One rather uses a 'preconditioned steepest descent algorithm',

$$\boxed{\mathbf{g}_{k+1} \; = \; \mathbf{g}_k - \mathbf{P}_k\,\gamma_k \quad ,} \tag{55}$$

where $\mathbf{P}_k$ is an ad-hoc positive definite operator, suitably chosen to produce a convergence of the algorithm as rapid as possible. Should one choose to use for $\mathbf{P}_k$ the Hessian of the misfit function, one would obtain the Newton algorithm.

Alternatively, one may choose to use a 'relaxation algorithm', where succesive 'jumps' are performed along the different directions defined by the different terms in equation 54.

One should keep in mind that to obtain a proper estimation of the posterior uncertainties in the metric, one needs the evaluation of the inverse of the Hessian operator (see appendix 6.4).

3.2 Transpose of a Linear Application

Let $\mathbf{L}$ be a linear operator mapping a linear space $\mathbb{A}$ into a linear space $\mathbb{B}$. Let $\mathbb{A}^*$ and $\mathbb{B}^*$ the respective dual spaces, and $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbb{A}}$ and $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbb{B}}$ the respective duality products. The transpose of $\mathbf{L}$ is the linear operator mapping $\mathbb{B}^*$ into $\mathbb{A}^*$ such that for any $\mathbf{a} \in \mathbb{A}$ and any $\mathbf{b}^* \in \mathbb{B}^*$,

$$\langle\,\mathbf{b}^*\,,\,\mathbf{L}\,\mathbf{a}\,\rangle_{\mathbb{B}} \; = \; \langle\,\mathbf{L}^t\,\mathbf{b}^*\,,\,\mathbf{a}\,\rangle_{\mathbb{A}} \quad . \tag{56}$$

For a good text on functional analysis, in particular on the transpose and adjoint of a linear operator, see Taylor and Lay (1980). Some of the results in the following sections are provided without demonstration: to check the proposed results, the reader should become familiar with the concepts proposed in that book. Let us only mention here two elementary results. The transpose of the linear operator defined through

$$y^{ij\ldots}{}_{k\ell\ldots} \; = \; A^{ij\ldots\mu\nu\ldots}{}_{k\ell\ldots\alpha\beta\ldots}\,x^{\alpha\beta\ldots}{}_{\mu\nu\ldots\ldots} \tag{57}$$

is the linear operator defined through

$$x_{\alpha\beta\ldots}{}^{\mu\nu\ldots\ldots} \; = \; A^{ij\ldots\mu\nu\ldots}{}_{k\ell\ldots\alpha\beta\ldots}\,y_{ij\ldots}{}^{k\ell\ldots} \tag{58}$$

[9] In fact, the dual operators (denoted with a 'star') ere respectively $\mathbf{Z}_k^* = \mathbf{C_g}\,\mathbf{Z}_k^t\,\mathbf{C_z}^{-1}$, $\mathbf{T}_k^* = \mathbf{C_g}\,\mathbf{T}_k^t\,\mathbf{C_t}^{-1}$, $\mathbf{\Sigma}_k^* = \mathbf{C_g}\,\mathbf{\Sigma}_k^t\,\mathbf{C}_\sigma^{-1}$, $\mathbf{A}_k^* = \mathbf{C_g}\,\mathbf{A}_k^t\,\mathbf{C_a}^{-1}$, $\mathbf{\Pi}_k^* = \mathbf{C_g}\,\mathbf{\Pi}_k^t\,\mathbf{C}_\pi^{-1}$, and $\mathbf{\Omega}_k^* = \mathbf{C_g}\,\mathbf{\Omega}_k^t\,\mathbf{C}_\omega^{-1}$. See Tarantola (2005) for details.

The transpose of the linear operator defined through

$$y^{\alpha\beta\dots}{}_{\gamma\mu\nu\dots\dots} \;=\; \nabla_\gamma\, x^{\alpha\beta\dots}{}_{\mu\nu\dots\dots} \tag{59}$$

is the linear operator defined through

$$x_{\alpha\beta\dots}{}^{\mu\nu\dots\dots} \;=\; -\nabla_\gamma\, y_{\alpha\beta\dots}{}^{\gamma\mu\nu\dots\dots} \tag{60}$$

There are typically some boundary conditions to be attached to a differential operator, what implies for the transpose operator a set of 'dual' boundary conditions, but we shall not enter into these 'details' in this preliminary version of the theory.

3.2.1 Einstein Equation

Considering the two duality products

$$\begin{aligned} \langle\, \delta\hat{\mathbf{t}}\,,\, \delta\mathbf{t}\,\rangle \;&=\; \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \delta\hat{t}^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\, \delta t_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \\ \langle\, \delta\hat{\mathbf{g}}\,,\, \delta\mathbf{g}\,\rangle \;&=\; \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \delta\hat{g}^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\, \delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \quad , \end{aligned} \tag{61}$$

the transpose operator $\mathbf{T}^t$ is defined by the condition that for any $\delta\hat{\mathbf{t}}$ and for any $\delta\mathbf{g}$, one must have

$$\langle\, \mathbf{T}^t \delta\hat{\mathbf{t}}\,,\, \delta\mathbf{g}\,\rangle \;=\; \langle\, \delta\hat{\mathbf{t}}\,,\, \mathbf{T}\,\delta\mathbf{g}\,\rangle \quad . \tag{62}$$

Using equation 12 and the expression of the duality products, this can be written

$$\begin{aligned} &\int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\;\; \left[\mathbf{T}^t\, \delta\hat{\mathbf{t}}\right]^{\alpha\beta}\;\; \delta g_{\alpha\beta} \;=\; \\ &\frac{1}{\chi} \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \delta\hat{t}^{\alpha\beta}\; \big(\, A_{\alpha\beta}{}^{\mu\nu,\rho\sigma}\; \nabla_{(\rho}\nabla_{\sigma)}\, \delta g_{\mu\nu} + B_{\alpha\beta}{}^{\mu\nu}\, \delta g_{\mu\nu}\,\big) \quad . \end{aligned} \tag{63}$$

To compute the direction of steepest descent we need to evaluate the term $\mathbf{C}_\mathbf{g}\, \mathbf{T}^t\, \delta\hat{\mathbf{t}}$, i.e., we need to evaluate

$$\begin{aligned} &\left[\mathbf{C}_\mathbf{g}\, \mathbf{T}^t\, \delta\hat{\mathbf{t}}\right]_{\gamma\delta} (\sigma^1,\sigma^2,\sigma^3,\sigma^4) \\ &= \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\;\; \left[\mathbf{T}^t\, \delta\hat{\mathbf{t}}\right]^{\alpha\beta} (\tau^1,\tau^2,\tau^3,\tau^4)\; C_{\alpha\beta\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4) \quad . \end{aligned} \tag{64}$$

This can be evaluated by just replacing $\delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)$ by $C_{\alpha\beta\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4)$ in equation 63. One gets

$$\begin{aligned} &\left[\mathbf{C}_\mathbf{g}\, \mathbf{T}^t\, \delta\hat{\mathbf{t}}\right]_{\gamma\delta} (\sigma^1,\sigma^2,\sigma^3,\sigma^4) \;=\; \\ &\frac{1}{\chi} \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \delta\hat{t}^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \\ &\Big(\; A_{\alpha\beta}{}^{\mu\nu,\rho\sigma}(\tau^1,\tau^2,\tau^3,\tau^4)\; \nabla_{(\rho}\nabla_{\sigma)}\, C_{\mu\nu\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4) \\ &+ B_{\alpha\beta}{}^{\mu\nu}(\tau^1,\tau^2,\tau^3,\tau^4)\; C_{\mu\nu\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4)\,\Big) \quad . \end{aligned} \tag{65}$$

3.2.2 Arrival Time Data

For each satellite trajectory, to any space-time metric field perturbation $\delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)$, expression 27 associates the scalar $\delta\sigma(\tau)$ defined along the trajectory. Let us here characterize the transpose, $\boldsymbol{\Sigma}^t$, of this operator. Considering the two duality products

$$\begin{aligned} \langle\, \delta\hat{\boldsymbol{\sigma}}\,,\, \delta\boldsymbol{\sigma}\,\rangle \;&=\; \int d\tau\; \delta\hat{\sigma}(\tau)\, \delta\sigma(\tau) \\ \langle\, \delta\hat{\mathbf{g}}\,,\, \delta\mathbf{g}\,\rangle \;&=\; \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \delta\hat{g}^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\, \delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \quad , \end{aligned} \tag{66}$$

the transpose operator is defined by the condition that for any $\delta\hat{\boldsymbol{\sigma}}$ and for any $\delta\mathbf{g}$, one must have

$$\langle\, \boldsymbol{\Sigma}^t \delta\hat{\boldsymbol{\sigma}}\,,\, \delta\mathbf{g}\,\rangle \;=\; \langle\, \delta\hat{\boldsymbol{\sigma}}\,,\, \boldsymbol{\Sigma}\,\delta\mathbf{g}\,\rangle \quad . \tag{67}$$

Using equation 27 and the expression of the duality products, this can be written

$$\begin{aligned} &\int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \left[\boldsymbol{\Sigma}^t\,\delta\hat{\boldsymbol{\sigma}}\right]^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\; \delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \\ &= \int d\tau\; \delta\hat{\sigma}(\tau)\; \left(- \frac{1/2}{g_{\mu\nu}(\tau^\kappa(\tau))\, u^\mu(\tau)\, \ell^\nu(\tau)} \int_{\lambda(\mathbf{g},\tau)} d\lambda\; \ell^\alpha(\lambda)\, \ell^\beta(\lambda)\, \delta g_{\alpha\beta}(\tau^\kappa(\lambda)) \right) \quad . \end{aligned} \tag{68}$$

To compute the direction of steepest descent we need to evaluate the term $\mathbf{C_g}\,\boldsymbol{\Sigma}^t\,\delta\hat{\boldsymbol{\sigma}}$, i.e., we need to evaluate

$$\begin{aligned} &\left[\mathbf{C_g}\,\boldsymbol{\Sigma}^t\,\delta\hat{\boldsymbol{\sigma}}\right]_{\gamma\delta}(\sigma^1,\sigma^2,\sigma^3,\sigma^4) \\ &= \int d\tau^1 \int d\tau^2 \int d\tau^3 \int d\tau^4\; \left[\boldsymbol{\Sigma}^t\,\delta\hat{\boldsymbol{\sigma}}\right]^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\; C_{\alpha\beta\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4) \quad . \end{aligned} \tag{69}$$

This can be evaluated by just replacing $\delta g_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)$ by $C_{\alpha\beta\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4)$ in equation 68. One gets

$$\begin{aligned} \left[\mathbf{C_g}\,\boldsymbol{\Sigma}^t\,\delta\hat{\boldsymbol{\sigma}}\right]_{\gamma\delta}(\sigma^1,\sigma^2,\sigma^3,\sigma^4) \;&=\; \int d\tau\; \delta\hat{\sigma}(\tau)\; \Bigg(- \frac{1/2}{g_{\mu\nu}(\tau^\kappa(\tau))\, u^\mu(\tau)\, \ell^\nu(\tau)} \\ \int_{\lambda(\mathbf{g},\tau)} d\lambda\; \ell^\alpha(\lambda)\, \ell^\beta(\lambda)\, C_{\alpha\beta\gamma}&(\tau^1(\lambda),\tau^2(\lambda),\tau^3(\lambda),\tau^4(\lambda),\sigma^1,\sigma^2,\sigma^3,\sigma^4) \Bigg) \quad . \end{aligned} \tag{70}$$

As the covariance function $C_{\alpha\beta\gamma\delta}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4)$ shall be a smooth function, we see that this equation 'spreads' the arrival time residuals into the region of the space-time that is around each satellite trajectory.

3.2.3 Accelerometer Data

It follows from equation 33 that the transpose operator $\mathbf{A}^t_{\mathbf{g}}$ is the linear operator that to any $\delta\hat{a}_\alpha$, defined at the points where the acceleration was measured, associates, at the same points, the values

$$\boxed{\;\delta\hat{g}^{\alpha\beta} \;=\; -\tfrac{1}{2}\left(g^{\alpha\nu}\, u^\beta\, u^\mu + g^{\beta\nu}\, u^\alpha\, u^\mu - g^{\mu\nu}\, u^\alpha\, u^\beta\right)\, \nabla_\mu\, \delta\hat{\pi}_\nu \quad .\;} \tag{71}$$

This operator appears in equations 142 and 147.

3.2.4 Gyroscope Data

The operator, $\mathbf{\Pi}^t_\mathbf{g}$, transpose of the operator $\mathbf{\Pi}_\mathbf{g}$ characterized in equation 41 is

$$\boxed{\delta\hat{g}^{\alpha\beta} \;=\; -\tfrac{1}{4}\left(g^{\alpha\nu}\,(u^\beta\,\sigma^\mu + u^\mu\,\sigma^\beta) + g^{\beta\nu}\,(u^\alpha\,\sigma^\mu + u^\mu\,\sigma^\alpha) - g^{\mu\nu}\,(u^\alpha\,\sigma^\beta + u^\beta\,\sigma^\alpha)\right)\,\nabla_\mu\,\delta\hat{\pi}_\nu \;\;.} \tag{72}$$

It appears in equations 142 and 147.

3.2.5 Gradiometer Data

The operator $\mathbf{\Omega}_\mathbf{g}$ was characterized in equations 48–50. The transpose operator, $\mathbf{\Omega}^t_\mathbf{g}$, associates to any $\delta\hat{\omega}_\alpha$ the $\delta\hat{g}^{\alpha\beta}$ given by

$$\begin{aligned}\delta\hat{g}^{\alpha\beta} \;&= v^\mu\,u^\nu\,(u^\alpha\,\nabla_{\mu\nu}\delta\hat{\omega}^\beta + u^\beta\,\nabla_{\mu\nu}\delta\hat{\omega}^\alpha) + (u^\alpha\,v^\beta + u^\beta\,v^\alpha)\,u^\mu\,\nabla_{\mu\nu}\delta\hat{\omega}^\nu \\ &\quad -u^\mu\,u^\nu\,(v^\alpha\,\nabla_{\mu\nu}\delta\hat{\omega}^\beta + v^\beta\,\nabla_{\mu\nu}\delta\hat{\omega}^\alpha - 2\,u^\alpha\,u^\beta\,v^\mu\,\nabla_{\mu\nu}\delta\hat{\omega}^\nu \quad .\end{aligned} \tag{73}$$

It appears in equations 142 and 147.

4 Discussion and Conclusion

We have been able to develop a consistent theory, fully relativistic, where the data brought by satellites emitting and receiving time signals is used to infer trajectories and the space-time metric. This constitutes both, a kind of ultimate gravimeter and a positioning system. Any observer with receiving capabilities shall know its own space-time trajectory “in real time”. These coordinates are not the usual ‘geographical’ coordinates plus a time, but are four times. The problem of attaching these four time coordinates to any terrestrial system of coordinates is just an attachment problem that should not interfere with the basic problem of defining an accurate reference system, and of knowing space-time trajectories into this system.

For more generality, we have considered the possibility that the satellites may have accelerometers, gradiometers, or gyroscopes. This is because the positioning problem and the problem of estimating the gravity field (i.e., the space-time metric) are coupled. If fact, all modern gravimetry satellite missions are coupled with GNSS satellites. Our theory applies, in particular, to the GOCE satellite mission (orbiting gradiometers). It also applies to the Gravity Probe B or the LISA[10] experiments, that could be analyzed using the concepts presented here.

The optimization algorithm proposed (Newton algorithm) is by no means the more economical to be used in the present context, and considerable effort is required to propose a practical algorithm, possibly using the ‘Kalman filter’ approach briefly mentioned in appendix 6.5. We are quite confident in our prediction that, some day, all positioning systems will be run using the basic principles exposed in this paper: the ever-increasing accuracy of time measurements with eventually force everyone to take relativity theory seriously —at last.—

5 Bibliography

Ashby, N., 2002, Relativity and the global positioning system, Physics Today, 55 (5), May 2002, pp. 41–47.

[10]Laser Interferometer Space Antenna.

Blagojević, M., J. Garecki, F.W. Hehl, and Yu.N. Obukhov, 2001, Real null coframes in general relativity and GPS type coordinates, arXiv:gr-qc/0110078v3.
Ciarlet, P.G., 1982, Introduction à l'analyse numérique matricielle et à l'optimisation, Masson, Paris.
Červený, V., 2002, Fermat's variational principle for anisotropic inhomogeneous media, Stud. geophys. geod., vol. 46, pp. 567–588. (`http://sw3d.mff.cuni.cz`).
Ciufolini, I., and Pavlis, E.C., 2004, A confirmation of the general relativistic prediction of the Lense-Thirring effect, Nature, Vol. 431, pp. 958–960.
Coll, B., 1985, Coordenadas luz en relatividad, Proc. Spanish Relativity Meeting, ERE-85, Publ. Servei Publ. ETSEIB, Barcelona, p 29–39. English translation (Light coordinates in relativity) at `http://syrte.obspm.fr/~coll/Papers/CoordinateSystems/LightCoordinates.pdf`
Coll, B., and Morales, J.A., 1992, 199 causal classes of space-time frames, International Journal of Theoretical Physics, Vol. 31, No. 6, pp. 1045–1062.
Coll, B., 2000, Elements for a theory of relativistic coordinate systems, formal and Physical aspects, ERES 2000, Valladolid.
Coll, B., 2001a, Elements for a Theory of Relativistic Coordinate Systems; Formal and Physical Aspects, Proc. XXIII Spanish Relativity Meeting ERE2000, World Scientific.
Coll, B., 2001b, Physical Relativistic Frames, JSR 2001, ed. N. Capitaine, Pub. Observatoire de Paris, pp. 169–174.
Coll, B., 2002, A principal positioning system for the Earth, JSR 2002, eds. N. Capitaine and M. Stavinschi, Pub. Observatoire de Paris, pp. 34–38.
Coll, B. and Tarantola, A., 2003, Galactic positioning system; physical relativistic coordinates for the Solar system and its surroundings, eds. A. Finkelstein and N. Capitaine, JSR 2003, Pub. St. Petersbourg Observatory, pp. 333-334.
Gordon, W., 1923, Ann. Phys. (Leipzig), 72, p. 421.
Grewal, M.S., Weill, L.R., and Andrews, A.P., 2001, Global positioning systems, inertial navigation, and integration, John Wiley & Sons.
Hernández-Pastora, J.L., Martín, J., and Ruiz, E., 2001, On gyroscope precession, arXiv:gr-qc/0009062.
Klimeš, L., 2002, Second-order and higher-order perturbations of travel time in isotropic and anisotropic media, Stud. geophys. geod., vol. 46, pp. 213–248. (`http://sw3d.mff.cuni.cz`).
Papapetrou, A., 1951, Spinning test-particles in general relativity (I), Proceedings of the Royal Society of London. A209, pp. 248–258
Pham Mau Quan, 1957, Arch. Rat. Mech. and Anal., 1, p. 54.
Rovelli, C., 2001, arXiv:gr-qc/0110003v2.
Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation, SIAM.
Taylor, A.E., and Lay, D.C., 1980, Introduction to functional analysis, Wiley.
Weinberg, S., 1972, Gravitation and Cosmology, Wiley.

6 Appendixes

6.1 Perturbation of Einstein's Tensor

Note that, as for any matrix $\mathbf{a}$, $(\mathbf{a}+\delta\mathbf{a})^{-1} = \mathbf{a}^{-1} - \mathbf{a}^{-1}\,\delta\mathbf{a}\,\mathbf{a}^{-1} + \cdots$, when imposing to the metric the perturbation

$$g_{\alpha\beta} \;\mapsto\; g_{\alpha\beta} + \delta g_{\alpha\beta} \quad , \tag{74}$$

the contravariant components have the perturbation

$$g^{\alpha\beta} \;\mapsto\; g^{\alpha\beta} - g^{\alpha\gamma}\,\delta g_{\gamma\delta}\,g^{\delta\beta} + \cdots \quad . \tag{75}$$

When introducing the perturbations 74–75 in the expressions 5, one obtains, keeping only first order terms in $\delta g_{\alpha\beta}$, the perturbation $\delta E_{\alpha\beta}$ of the Einstein tensor.

We consider the perturbation of the metric and the perturbation of the connection.

$$g_{\alpha\beta} \;\to\; g'_{\alpha\beta} \;=\; g_{\alpha\beta} + \delta g_{\alpha\beta} \qquad ; \qquad \Gamma^{\alpha}{}_{\beta\gamma} \;\to\; \Gamma'^{\alpha}{}_{\beta\gamma} \;=\; \Gamma^{\alpha}{}_{\beta\gamma} + \delta\Gamma^{\alpha}{}_{\beta\gamma} \quad . \tag{76}$$

In this appendix, and in order to make the expressions more compact, let us denote

$$\delta g_{\alpha\beta} \;=\; h_{\alpha\beta} \qquad ; \qquad \delta\Gamma^{\alpha}{}_{\beta\gamma} \;=\; \Omega^{\alpha}{}_{\beta\gamma} \quad . \tag{77}$$

We will use the unperturbed metric to raise and lower indices. For instance, we will write $h^{\alpha}{}_{\beta} \equiv g^{\alpha\gamma}h_{\gamma\beta}$.

By requiring that both, the unperturbed and the perturbed connection, to be metric, $\nabla g = \nabla' g' = 0$, and symmetric, $\Omega^{\alpha}{}_{[\beta\gamma]} = 0$, we get:

$$\left.\begin{array}{r} \nabla'_{\gamma}g'_{\alpha\beta} = \nabla_{\gamma}h_{\alpha\beta} - \Omega^{\delta}{}_{\alpha\gamma}g'_{\delta\beta} - \Omega^{\delta}{}_{\beta\gamma}g'_{\alpha\delta} = 0 \\ \Omega^{\alpha}{}_{[\beta\gamma]} = 0 \end{array}\right\} \;\Rightarrow\; g'_{\alpha\delta}\Omega^{\delta}{}_{\beta\gamma} = \frac{1}{2}\left(\nabla_{\gamma}h_{\alpha\beta} + \nabla_{\beta}h_{\alpha\gamma} - \nabla_{\alpha}h_{\beta\gamma}\right). \tag{78}$$

Then, $\Omega^{\mu}{}_{\beta\gamma}$ is obtained by contracting this expression with the inverse of the perturbed metric $g'^{\mu\alpha}$. But the first order is given by the contraction with the unperturbed one:

$$\Omega^{\alpha}{}_{\beta\gamma} = g^{\alpha\delta}\Omega_{\delta\beta\gamma} \qquad \text{where} \qquad \Omega_{\alpha\beta\gamma} = \frac{1}{2}\left(\nabla_{\gamma}h_{\alpha\beta} + \nabla_{\beta}h_{\alpha\gamma} - \nabla_{\alpha}h_{\beta\gamma}\right). \tag{79}$$

The two Riemann tensors are related by

$$R'^{\alpha}{}_{\beta\gamma\delta} = R^{\alpha}{}_{\beta\gamma\delta} + 2\nabla_{[\gamma}\Omega^{\alpha}{}_{\delta]\beta} + 2\Omega^{\alpha}{}_{\mu[\gamma}\Omega^{\mu}{}_{\delta]\beta}\,. \tag{80}$$

Thus, the first order perturbation of the Riemann is given by

$$\delta R^{\alpha}{}_{\beta\gamma\delta} = 2\nabla_{[\gamma}\Omega^{\alpha}{}_{\delta]\beta}\,, \tag{81}$$

from which we get the first order perturbation of the Ricci tensor:

$$\delta R_{\alpha\beta} = \delta R^{\delta}{}_{\alpha\delta\beta} = 2\nabla_{[\delta}\Omega^{\delta}{}_{\beta]\alpha} = \nabla_{[\delta}\nabla_{\beta]}h^{\delta}{}_{\alpha} + \frac{1}{2}\left(\nabla_{\delta}\nabla_{\alpha}h^{\delta}{}_{\beta} - \nabla_{\beta}\nabla_{\alpha}h^{\delta}{}_{\delta}\right) - \frac{1}{2}\left(\nabla_{\delta}\nabla^{\delta}h_{\alpha\beta} + \nabla_{\beta}\nabla^{\delta}h_{\alpha\delta}\right). \tag{82}$$

Splitting into symmetric and intisymmetric parts of the two covariant derivatives

$$\begin{aligned} \delta R_{\alpha\beta} = {} & \nabla_{[\delta}\nabla_{\beta]}h^{\delta}{}_{\alpha} + \tfrac{1}{2}\,\nabla_{[\delta}\nabla_{\alpha]}h^{\delta}{}_{\beta} + \tfrac{1}{2}\,\nabla_{[\beta}\nabla_{\delta]}h^{\delta}{}_{\alpha} \\ & + \tfrac{1}{2}\,\nabla_{(\delta}\nabla_{\alpha)}h^{\delta}{}_{\beta} - \tfrac{1}{2}\,\nabla_{(\beta}\nabla_{\alpha)}h^{\delta}{}_{\delta} - \tfrac{1}{2}\,\nabla_{\delta}\nabla^{\delta}h_{\alpha\beta} + \tfrac{1}{2}\,\nabla_{(\beta}\nabla_{\delta)}h^{\delta}{}_{\alpha}\,. \end{aligned} \tag{83}$$

Substituting now the identity $2\nabla_{[\delta}\nabla_{\alpha]}h^{\delta}{}_{\beta} = R^{\delta}{}_{\mu\delta\alpha}\,h^{\mu}{}_{\beta} - R^{\mu}{}_{\beta\delta\alpha}\,h^{\delta}{}_{\mu}$ and introducing the notation

$$H_{\alpha\beta,\gamma\delta} \;\equiv\; \nabla_{(\gamma}\nabla_{\delta)}h_{\alpha\beta} \quad , \tag{84}$$

$$2\,\delta R_{\alpha\beta} = R_{\mu(\alpha}h^{\mu}{}_{\beta)} + R^{\mu}{}_{(\alpha\beta)\delta}h^{\delta}{}_{\mu} + H^{\delta}{}_{\beta,\alpha\delta} + H^{\delta}{}_{\alpha,\beta\delta} - H^{\delta}{}_{\delta,\alpha\beta} - H_{\alpha\beta,}{}^{\delta}{}_{\delta}\,. \tag{85}$$

In order to obtain the perturbation of the Ricci scalar we also need the perturbation of the contravariant metric:

$$g^{\alpha\gamma}\,g_{\gamma\beta} = \delta^{\alpha}_{\beta} \;\Rightarrow\; \delta g^{\alpha\beta} = -g^{\alpha\gamma}\,h_{\gamma\delta}\,g^{\delta\beta} = -h^{\alpha\beta}\,. \tag{86}$$

Thus, the perturbation of the Ricci scalar is

$$\delta R = \delta g^{\alpha\beta}\, R_{\alpha\beta} + g^{\alpha\beta}\, \delta R_{\alpha\beta} = -R^{\alpha\beta}\, h_{\alpha\beta} + H^{\alpha\beta}{}_{,\alpha\beta} - H^{\alpha}{}_{\alpha,}{}^{\beta}{}_{\beta} \,. \tag{87}$$

Finally we obtain the first order perturbation of the Einstein tensor,

$$\begin{aligned} \delta E_{\alpha\beta} &= \delta R_{\alpha\beta} - \tfrac{1}{2}\,(\delta R\, g_{\alpha\beta} + R\, \delta g_{\alpha\beta}) \\ &= \tfrac{1}{2}\,(R^{\delta}{}_{(\alpha\beta)}{}^{\mu}\, h_{\delta\mu} + R^{\mu}{}_{(\alpha}\, h_{\beta)\mu} + R^{\gamma\delta}\, h_{\gamma\delta}\, g_{\alpha\beta} - R\, h_{\alpha\beta}) \\ &\quad + \tfrac{1}{2}\,(H^{\delta}{}_{\beta,\alpha\delta} + H^{\delta}{}_{\alpha,\beta\delta} - H^{\delta}{}_{\delta,\alpha\beta} - H_{\alpha\beta,}{}^{\delta}{}_{\delta} + H^{\gamma}{}_{\gamma,}{}^{\delta}{}_{\delta} g_{\alpha\beta} - H^{\gamma\delta}{}_{,\gamma\delta}\, g_{\alpha\beta}) \end{aligned} \tag{88}$$

This result can be rewriten as

$$\delta E_{\alpha\beta} \;=\; A_{\alpha\beta}{}^{\gamma\delta,\rho\sigma}\, H_{\gamma\delta,\rho\sigma} + B_{\alpha\beta}{}^{\gamma\delta}\, h_{\gamma\delta} \tag{89}$$

with

$$\begin{aligned} A_{\alpha\beta}{}^{\gamma\delta,\rho\sigma} &= 2\, g^{(\gamma|(\sigma}\, \delta^{\rho)}_{(\alpha}, \delta^{|\delta)}_{\beta)} - \tfrac{1}{2}\, g^{\gamma\delta}\, \delta^{\rho}_{(\alpha}\, \delta^{\sigma}_{\beta)} - \tfrac{1}{2}\, g^{\rho\sigma}\, \delta^{\gamma}_{(\alpha}\, \delta^{\delta}_{\beta)} + \tfrac{1}{2}\, g^{\gamma\delta}\, g^{\rho\sigma}\, g_{\alpha\beta} - \tfrac{1}{2}\, g^{\gamma(\rho}\, g^{\sigma)\delta}\, g_{\alpha\beta} \\ B_{\alpha\beta}{}^{\gamma\delta} &= \tfrac{1}{2}\,(R^{\gamma}{}_{(\alpha\beta)}{}^{\delta} + R^{(\gamma}{}_{(\alpha}\, \delta^{\delta)}_{\beta)} + R^{\gamma\delta}\, g_{\alpha\beta} - R\, \delta^{\gamma}_{(\alpha}\, \delta^{\delta}_{\beta)}) \,. \end{aligned} \tag{90}$$

Note that using he definition 84, equation 89 can be written, explicitly,

$$\delta E_{\alpha\beta} = A_{\alpha\beta}{}^{\gamma\delta,\rho\sigma}\, \nabla_{(\rho}\nabla_{\sigma)} h_{\gamma\delta} + B_{\alpha\beta}{}^{\gamma\delta}\, h_{\gamma\delta} \quad . \tag{91}$$

Observe that, by construction, the two tensors $\mathbf{A}$ and $\mathbf{B}$ are symmetrics in each pair of indices,

$$A_{\alpha\beta}{}^{\gamma\delta,\mu\nu} = A_{(\alpha\beta)}{}^{(\gamma\delta),(\mu\nu)} \qquad \text{and} \qquad B_{\alpha\beta}{}^{\gamma\delta} = B_{(\alpha\beta)}{}^{(\gamma\delta)} \,. \tag{92}$$

In addition, it results that $A_{\alpha\beta}{}^{\gamma\delta,\mu\nu}$ is symmetric respect to the interchange of the two contravariant pairs:

$$A_{\alpha\beta}{}^{\gamma\delta,\mu\nu} = A_{\alpha\beta}{}^{\mu\nu,\gamma\delta} \,. \tag{93}$$

This implies that not all the information in $H_{\alpha\beta,\gamma\delta}$ contrubutes to $\delta E_{\alpha\beta}$. In fact, we can express the term $A_{\alpha\beta}{}^{\gamma\delta,\mu\nu} H_{\alpha\beta,\gamma\delta}$ in an interesting form. Let us define $J_{\alpha\beta\gamma\delta} \equiv 2H_{[\delta|[\alpha,\beta]|\gamma]}$. This tensor contains less information than $H_{[\gamma|[\alpha,\beta]|\delta]}$, and has the same symmetries as a Riemann:

$$J_{\alpha\beta\gamma\delta} = J_{[\alpha\beta][\gamma\delta]} = J_{\gamma\delta\alpha\beta} \qquad \text{and} \qquad J_{\alpha[\beta\gamma\delta]} = 0 \,. \tag{94}$$

We can then take the traces of this tensor (obtaining a Ricci-like tensor and scalar): $J_{\alpha\beta} \equiv J^{\gamma}{}_{\alpha\gamma\beta}$ and $\beta \equiv J^{\alpha}{}_{\alpha}$. Then, it is easy to check that the contribution of $H_{\alpha\beta,\gamma\delta}$ is only the Einstein-like tensor of $J_{\alpha\beta\gamma\delta}$:

$$A_{\alpha\beta}{}^{\gamma\delta,\mu\nu} H_{\gamma\delta,\mu\nu} \;=\; J_{\alpha\beta} - \frac{1}{2}\, J\, g_{\alpha\beta} \tag{95}$$

In contrast, observe that $B_{\alpha\beta}{}^{\gamma\delta}$ contains all the information of the Riemann tensor.

6.2 Arrival Time Data

We need the linear operator $\mathbf{\Sigma}$ that is tangent to the forward operator $\boldsymbol{\sigma}$ at some $\mathbf{g}_0$. Formally,

$$\boldsymbol{\sigma}(\mathbf{g} + \delta\mathbf{g}) \;=\; \boldsymbol{\sigma}(\mathbf{g}) + \mathbf{\Sigma}\, \delta\mathbf{g} + O(\delta\mathbf{g})^2 \quad . \tag{96}$$

It is easy to understand the meaning of $\mathbf{\Sigma}$. While $\boldsymbol{\sigma}$ associates to any metric $\mathbf{g}$ some arrival times σ^i, the operator $\mathbf{\Sigma}$ associates to every metric perturbation $\delta\mathbf{g}$ (around $\mathbf{g}$) the perturbation $\delta\sigma^i$ of arrival times. Let us compute these perturbations.

6.2.1 Hamiltonian Formulation of Finsler Geometry

The Finsler space is a generalization of the Riemann space. This generalization is appropriate for the description of the propagation of light and many other waves.

Proper time τ in the Finsler space satisfies the stationary Hamilton-Jacobi equation

$$H(x^\kappa, \tau_{,\mu}) \;=\; \text{const.} \quad , \tag{97}$$

where $H(x^\kappa, p_\mu)$ is the Hamiltonian. The geodesics can then be described by the Hamilton equations

$$\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \;=\; \frac{\partial H}{\partial p_\alpha} \quad , \tag{98}$$

$$\frac{\mathrm{d}p_\alpha}{\mathrm{d}\lambda} \;=\; -\frac{\partial H}{\partial x^\alpha} \quad , \tag{99}$$

with initial conditions

$$x^\alpha(\lambda_0) \;=\; x_0^\alpha \qquad ; \qquad p_\alpha(\lambda_0) \;=\; \tau_{,\alpha}(x_0^\mu) \quad . \tag{100}$$

Then

$$\tau_{,\alpha}[x^\mu(\lambda)] \;=\; p_\alpha(\lambda) \tag{101}$$

along the geodesics. Parameter λ along a geodesic is determined by the form of the Hamiltonian and by initial conditions (equation 100) for the geodesic. Proper time τ along the geodesic is then given by

$$\tau(\lambda) \;=\; \tau(\lambda_0) \;=\; \int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; p_\alpha \frac{\partial H}{\partial p_\alpha} \quad , \tag{102}$$

which follows from equations 98 and 101. Note that equal geodesics may be generated by various Hamiltonians. For example, Hamiltonian $\tilde{H}(x^\kappa, p_\mu) = F[H(x^\kappa, p_\mu)]$, where $F(x)$ is an arbitrary function with a non–vanishing finite derivative at x equal to the right–hand side of equation 97, yields equal geodesics as Hamiltonian $H(x^\kappa, p_\mu)$. The Hamiltonian is often chosen as a homogeneous function of degree N in p_α. Especially, homogeneous Hamiltonians of degrees $N = 2$, $N = 1$ or $N = -1$ are frequently used.

If the Hamiltonian is chosen as a homogeneous function of degree $N = 2$ in p_α, and is properly normalized, then

$$g^{\alpha\beta}(x^\kappa, p_\mu) \;=\; \frac{\partial^2 H}{\partial p_\alpha \partial p_\beta}(x^\kappa, p_\mu) \tag{103}$$

is the contravariant Finslerian metric tensor. If metric tensor in equation 103 is independent of p_μ,

$$g^{\alpha\beta}(x^\kappa, p_\mu) \;=\; g^{\alpha\beta}(x^\kappa) \quad , \tag{104}$$

the Finsler space reduces to the *Riemann space*.

On the other hand, if we know the contravariant metric tensor, we may construct a homogeneous Hamiltonian of degree N in p_α as

$$H(x^\kappa, p_\mu) \;=\; \frac{1}{N}[p_\alpha g^{\alpha\beta}(x^\kappa, p_\mu) p_\beta]^{\frac{N}{2}} \quad . \tag{105}$$

Whereas degree N may be arbitrary for spatial or time-like geodesics, $N \neq 2$ should be avoided for zero-length geodesics in order to keep the right–hand sides of Hamilton equations 98 and 99 finite and non-vanishing identically.

For homogeneous Hamiltonians (equation 105), equation 102 reads

$$\tau(\lambda) \;=\; \tau(\lambda_0) + \int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; [p_\alpha g^{\alpha\beta}(x^\kappa, p_\mu) p_\beta]^{\frac{N}{2}} \quad , \tag{106}$$

and equation 98 yields

$$\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} g_{\alpha\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \;=\; [p_\alpha g^{\alpha\beta} p_\beta]^{N-1} \quad . \tag{107}$$

Considering equation 107, equation 106, can be expressed in the form

$$\tau(\lambda) \;=\; \tau(\lambda_0) + \int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; \left[\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} g_{\alpha\beta}(x^\kappa, p_\mu) \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\right]^{\frac{N}{2(N-1)}} \quad . \tag{108}$$

In the Hamiltonian formulation, the Finsler geometry is no more complex than the Riemann geometry.

6.2.2 Perturbation of Proper Time

The first-order perturbation of proper time (equation 102) is (Klimeš, 2002, eq. 25)

$$\delta\tau(\lambda) \;=\; \delta\tau(\lambda_0) - \int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; \delta H \quad . \tag{109}$$

If we wish to perform perturbations with respet to the components of the metric tensor along zero–length space–time geodesics, homogeneous Hamiltonians (equation 105) should be of degree $N = 2$ to avoid zero or infinite perturbations δH of the Hamiltonian.

One alternative to the present Hamiltomian formulation, would be to use a Lagrangian formulation of the first degree, this leading to the usual Fermat's integral. There are four reasons why the formulation here presented is better:

- Perturbations of a homogeneous Lagrangian of degree N with respect to the components of the metric tensor are zero for $N > 2$ and infinite for $N < 2$, which results in singularities in the computation;
- Hamilton equations beak down for $N \neq 2$, which would prevent us from using efficient tools of Hamiltonian formulation;
- Dual Legendre transform between homogeneous Hamiltonian and Lagrangian of the first degree is not possible (Červený, 2002), which also holds for spatial and time–like geodesics;
- The integral is generally complex-valued for indefinite metric tensors.

In the following, we shall thus consider an homogeneous Hamiltonian (equation 105) of degree $N = 2$,

$$H(x^\kappa, p_\mu) \;=\; \frac{1}{2} p_\alpha \, g^{\alpha\beta}(x^\kappa, p_\mu) \, p_\beta \quad . \tag{110}$$

Equation 98 then reads

$$\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \;=\; g^{\alpha\beta}(x^\kappa, p_\mu) p_\beta \quad , \tag{111}$$

and equation 109, with $\delta\tau(\lambda_0) = 0$, reads

$$\delta\tau(\lambda) \; = \; -\frac{1}{2}\int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; p_\alpha \delta g^{\alpha\beta} p_\beta \quad . \tag{112}$$

Inserting $\delta g^{\alpha\beta} = -g^{\alpha\kappa}\delta g_{\kappa\mu} g^{\mu\beta}$, we obtain

$$\delta\tau(\lambda) \; = \; \frac{1}{2}\int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; p_\alpha g^{\alpha\kappa} \delta g_{\kappa\mu} g^{\mu\beta} p_\beta \quad . \tag{113}$$

Inserting equation 111 into equation 113, we arrive at

$$\delta\tau(\lambda) \; = \; \frac{1}{2}\int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \delta g_{\alpha\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \quad . \tag{114}$$

6.2.3 Perturbation of arrival time

Assume the trajectory

$$x^i \; = \; y^i(\sigma) \tag{115}$$

parametrized by proper time σ along it (in general, σ may represent an arbitrary parameter along the trajectory). A light signal emitted at the given point will hit the given trajectory at proper time $\sigma = \sigma^0$. Assume now that the space–time metric is perturbed from g_{ij} to $g_{ij} + \delta g_{ij}$.The light signal will now hit the trajectory at proper time $\sigma^0 + \delta\sigma$. We shall now derive the first order relation between $\delta\sigma$ and delta g_{ij}.

The space-time wavefront may be expressed in the form

$$\tau(x^\alpha) \; = \; 0 \quad , \tag{116}$$

where $\tau(x^\alpha)$ is measured along the geodesic from the given point to point x^α. The geodesics can be calculated by Hamiltonian ray tracing from the given point.

Proper time σ at the point of intersection of the trajectory with the space-time wavefront then satisfies equation

$$\tau(y^\alpha(\sigma)) \; = \; 0 \quad . \tag{117}$$

Perturbation of this equation yields

$$\delta\tau(y^\beta(\sigma)) + \tau_{,\alpha}(y^\beta(\sigma)) \, \frac{\mathrm{d}y^\alpha}{\mathrm{d}\sigma}(\sigma) \, \delta\sigma \; = \; 0 \quad . \tag{118}$$

Then

$$\delta\sigma \; = \; -\frac{\delta\tau(y^\beta(\sigma))}{\tau_{,\alpha}(y^\beta(\sigma)) \, \frac{\mathrm{d}y^\alpha}{\mathrm{d}\sigma}(\sigma)} \quad . \tag{119}$$

Inserting p_α from equation 111 for $\tau_{,\alpha}$ and equation 114 for $\delta\tau(y^\beta(\sigma))$, equation 119 can be expressed in the form

$$\delta\sigma \; = \; -\frac{1}{2}\left[\frac{\mathrm{d}y^\beta}{\mathrm{d}\sigma} \, g_{\alpha\beta} \, \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\right]^{-1} \int_{\lambda_0}^{\lambda} \mathrm{d}\lambda \; \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \delta g_{\alpha\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \quad . \tag{120}$$

6.3 A Priori Information on the Metric

Let $\mathbf{g}_{\text{prior}}$ some reference space-time metric (for instance the Minkowski or the Schwarzschild metric), and let $\mathbf{g}$ be the actual metric. In the simple (and a little bit simplistic) approach proposed here, it is assumed that the difference

$$\mathbf{g} - \mathbf{g}_{\text{prior}} \tag{121}$$

is small, and is assumed to be a random realization of a Gaussian random field with zero mean and prescribed covariance. Because in the light coordinates used here it is the contravariant metric that has some simple properties, the difference in equation 121 is taken using the contravariant components.

To obtain a reasonable model of covariance operator for the metric, we could perform a thought experiment. We imagine a large number of metric fields, all of the form

$$\{g^{\alpha\beta}\} = \begin{pmatrix} 0 & g^{12} & g^{13} & g^{14} \\ g^{12} & 0 & g^{23} & g^{24} \\ g^{13} & g^{23} & 0 & g^{34} \\ g^{14} & g^{24} & g^{34} & 0 \end{pmatrix} \tag{122}$$

at every point, all smoothly varying over space-time, and with the quantities $g^{\alpha\beta}$ randomly varying around the values corresponding to the reference metric, with prescribed, simple probability distributions (independent, to start with). The we could evaluate the covariance of such a 'random field' using the direct definition of covariance:

$$\begin{aligned} & C^{\alpha\beta\mu\nu}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4) = \\ & \overline{\left(g^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) - \overline{g^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)}\right)\ \left(g^{\mu\nu}(\sigma^1,\sigma^2,\sigma^3,\sigma^4) - \overline{g^{\mu\nu}(\sigma^1,\sigma^2,\sigma^3,\sigma^4)}\right)}\ , \end{aligned} \tag{123}$$

where $\overline{x}$ means the mean value of x. The mean metric $\overline{g^{\alpha\beta}}$ would be the reference metric.

Another option is to try to insert more constraints that we know are satisfied by the metric. For instance, Pozo (2005) shows that the metric has necessarily the form

$$\begin{pmatrix} 0 & g^{12} & g^{13} & g^{14} \\ g^{12} & 0 & g^{23} & g^{24} \\ g^{13} & g^{23} & 0 & g^{34} \\ g^{14} & g^{24} & g^{34} & 0 \end{pmatrix} = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix} \begin{pmatrix} 0 & A & B & 1 \\ A & 0 & 1 & B \\ B & 1 & 0 & A \\ 1 & B & A & 0 \end{pmatrix} \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix} \quad , \tag{124}$$

where the constants $\{a,b,c,d\}$ are positive, and the constants $\{A,B$ should satisfy the constraint that a triangle exists in the Euclidean plane whose sides have the lengths $\{A,B,1\}$. One could perhaps use the six quantities $\{a,b,c,d,A,B\}$ as basic quantities, and assume a Gaussian distribution for some simple functions of them.

We do not explore yet this possibility. Also, it is very likely that the basic variable to be used in the optimization problem is not the metric $g^{\alpha\beta}$, but the logarithmic metric. This point is, for the time being, not examined.

We don't try to be more specific at this point, we simply assume that some covariance function

$$C^{\alpha\beta\mu\nu}(\tau^1,\tau^2,\tau^3,\tau^4,\sigma^1,\sigma^2,\sigma^3,\sigma^4) \tag{125}$$

is chosen. The inverse $\mathbf{W} = \mathbf{C}^{-1}$ of the covariance operator (a distribution) has the kernel

$$W_{\alpha\beta\mu\nu}(\tau^1,\tau^2,\tau^3,\tau^4;\sigma^1,\sigma^2,\sigma^3,\sigma^4) \quad . \tag{126}$$

By definition (formally)

$$\int dv(\rho^1,\rho^2,\rho^3,\rho^4)\; W_{\alpha\beta\rho\sigma}(\tau^1,\tau^2,\tau^3,\tau^4;\rho^1,\rho^2,\rho^3,\rho^4)\times$$
$$\times\, C^{\rho\sigma\mu\nu}(\rho^1,\rho^2,\rho^3,\rho^4;\sigma^1,\sigma^2,\sigma^3,\sigma^4) \;=$$
$$=\; \delta_\alpha^\mu\,\delta_\beta^\nu\,\delta(\tau^1-\sigma^1)\,\delta(\tau^2-\sigma^2)\,\delta(\tau^3-\sigma^3)\,\delta(\tau^4-\sigma^4) \quad , \tag{127}$$

where

$$dv(\rho^1,\rho^2,\rho^3,\rho^4) \;=\; \sqrt{-\det \mathbf{g}_{\rm prior}(\rho^1,\rho^2,\rho^3,\rho^4)}\; d\rho^1\, d\rho^2\, d\rho^3\, d\rho^4 \quad . \tag{128}$$

The operators $\mathbf{C}(\mathbf{g})$ and $\mathbf{W}(\mathbf{g})$ being symmetric and positive definite, define a bijection between $\mathcal{G}$, the space of metric field perturbations and its dual, $\mathcal{G}^*$. We shall write

$$\delta\hat{\mathbf{g}} \;=\; \mathbf{W}\,\delta\mathbf{g} \qquad ; \qquad \delta\mathbf{g} \;=\; \mathbf{C}\,\delta\hat{\mathbf{g}} \quad . \tag{129}$$

Explicitly,

$$\delta\hat{g}_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \;=\; \int dv(\rho^1,\rho^2,\rho^3,\rho^4)$$
$$W_{\alpha\beta\mu\nu}(\tau^1,\tau^2,\tau^3,\tau^4;\sigma^1,\sigma^2,\sigma^3,\sigma^4)\;\delta g^{\mu\nu}(\sigma^1,\sigma^2,\sigma^3,\sigma^4) \tag{130}$$

and

$$\delta g^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \;=\; \int dv(\rho^1,\rho^2,\rho^3,\rho^4)$$
$$C^{\alpha\beta\mu\nu}(\tau^1,\tau^2,\tau^3,\tau^4;\sigma^1,\sigma^2,\sigma^3,\sigma^4)\;\delta\hat{g}_{\mu\nu}(\sigma^1,\sigma^2,\sigma^3,\sigma^4) \quad . \tag{131}$$

The duality product of a dual metric field perturbation $\delta\hat{\mathbf{g}}$ by a metric field perturbation $\delta\boldsymbol{\gamma}$ is defined as

$$\langle\, \delta\hat{\mathbf{g}}\,,\, \delta\boldsymbol{\gamma}\,\rangle \;=\; \int dv(\tau^1,\tau^2,\tau^3,\tau^4)\;\delta\hat{g}_{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4)\;\delta\gamma^{\alpha\beta}(\tau^1,\tau^2,\tau^3,\tau^4) \quad , \tag{132}$$

the scalar product of two metric field perturbations is

$$\delta\mathbf{g}_1\cdot\delta\mathbf{g}_2 \;=\; \langle\, \mathbf{W}\,\delta\mathbf{g}_1\,,\, \delta\mathbf{g}_2\,\rangle \quad , \tag{133}$$

and the norm of a metric field perturbation is

$$\|\,\delta\mathbf{g}\,\|_{\mathbf{C}_{\rm g}} \;=\; \sqrt{\delta\mathbf{g}\cdot\delta\mathbf{g}} \quad . \tag{134}$$

Denoting by $\mathbf{g}_{\rm prior}$ the a priori metric and by $\mathbf{g}$ our estimation of the actual metric field, we are later going to impose that the squared norm

$$2\,S_{\rm g}(\mathbf{g}) \;=\; \|\,\mathbf{g}-\mathbf{g}_{\rm prior}\,\|^2_{\mathbf{C}_{\rm g}} \tag{135}$$

is small.

6.4 Newton Algorithm

While in section 3.1 we have examined the simple steepest descent algorithm, let us now develop the quasi-Newton method. To obtain the actual algorithm, one may use the formulas developed in Tarantola (2005). The resulting iterative algorithm can be written

$$\mathbf{g}_{k+1} \;=\; \mathbf{g}_k - \mathbf{H}_k^{-1}\,\gamma_k \quad , \tag{136}$$

where the 'Hessian operator' $\mathbf{H}_k$ is

$$\begin{aligned} \mathbf{H}_k \;=\; & \mathbf{I} + (\mathbf{Z}_k\,\mathbf{C_g})^t\,\mathbf{C_z}^{-1}\,\mathbf{Z}_k + (\mathbf{T}_k\,\mathbf{C_g})^t\,\mathbf{C_t}^{-1}\,\mathbf{T}_k + (\mathbf{\Sigma}_k\,\mathbf{C_g})^t\,\mathbf{C}_\sigma^{-1}\,\mathbf{\Sigma}_k \\ & + (\mathbf{A}_k\,\mathbf{C_g})^t\,\mathbf{C_a}^{-1}\,\mathbf{A}_k + (\mathbf{\Pi}_k\,\mathbf{C_g})^t\,\mathbf{C}_\pi^{-1}\,\mathbf{\Pi}_k + (\mathbf{\Omega}_k\,\mathbf{C_g})^t\,\mathbf{C}_\omega^{-1}\,\mathbf{\Omega}_k \quad , \end{aligned} \tag{137}$$

the 'gradient vector' is

$$\begin{aligned} \gamma_k \;=\; (\mathbf{g}_k - \mathbf{g}_{\text{prior}}) & + (\mathbf{Z}_k\,\mathbf{C_g})^t\,\mathbf{C_z}^{-1}\,(\mathbf{z}(\mathbf{g}_k) - \mathbf{1}) \\ & + (\mathbf{T}_k\,\mathbf{C_g})^t\,\mathbf{C_t}^{-1}\,(\mathbf{t}(\mathbf{g}_k) - \mathbf{t}_{\text{obs}}) \\ & + (\mathbf{\Sigma}_k\,\mathbf{C_g})^t\,\mathbf{C}_\sigma^{-1}\,(\sigma(\mathbf{g}_k) - \sigma_{\text{obs}}) \\ & + (\mathbf{A}_k\,\mathbf{C_g})^t\,\mathbf{C_a}^{-1}\,(\mathbf{a}(\mathbf{g}_k) - \mathbf{a}_{\text{obs}}) \\ & + (\mathbf{\Pi}_k\,\mathbf{C_g})^t\,\mathbf{C}_\pi^{-1}\,(\pi(\mathbf{g}_k) - \pi_{\text{obs}}) \\ & + (\mathbf{\Omega}_k\,\mathbf{C_g})^t\,\mathbf{C}_\omega^{-1}\,(\omega(\mathbf{g}_k) - \omega_{\text{obs}}) \quad , \end{aligned} \tag{138}$$

where the linear operators $\mathbf{Z}_k$, $\mathbf{T}_k$, $\mathbf{\Sigma}_k$, $\mathbf{A}_k$, $\mathbf{\Pi}_k$, and $\mathbf{\Omega}_k$, are the Fréchet derivatives (tangent linear applications) of the operators $\mathbf{z}(\mathbf{g})$, $\mathbf{t}(\mathbf{g})$, $\sigma(\mathbf{g})$, $\mathbf{a}(\mathbf{g})$, $\pi(\mathbf{g})$, and $\omega(\mathbf{g})$, introduced in equations 9, **??**, **??**, 30, 37, and 45, all the operators evaluated for $\mathbf{g} = \mathbf{g}_k$, and where $\mathbf{L}^t$ denotes the transpose of a linear operator $\mathbf{L}$. We say *transpose* operators, better than *dual* operators, because the difference between the two notions matters inside the theory of least-squares.

All the linear operators just introduced are evaluated in section 3.2. But before going into these details, some comments on the iterative algorithm are needed.

The quasi-Newton algorithm 136 can be initialized at an arbitrary point (i.e., at any metric field) $\mathbf{g}_0$. If working in the vicinity of an ordinary planet, the present problem will only be mildly nonlinear, and the convergence point will be independent of the initial point. The simplest choice, of course, is

$$\mathbf{g}_0 \;=\; \mathbf{g}_{\text{prior}} \quad . \tag{139}$$

Before entering on the problem of how many iterations must the done in practice, let us take the strict mathematical point of view that, in principle, an infinite number of iterations should be performed. The optimal estimate of the space-time metric would then be

$$\widetilde{\mathbf{g}} \;=\; \mathbf{g}_\infty \quad . \tag{140}$$

The least-squared method not only provides an optimal solution, it also provides a mean of estimating the uncertainties on this solution. It can be shown (Tarantola, 2005) that these uncertainties are those represented by the covariance operator

$$\boxed{\;\widetilde{\mathbf{C}}_\mathbf{g} \;=\; \mathbf{H}_\infty^{-1}\,\mathbf{C_g} \quad .\;} \tag{141}$$

Crudely speaking, we started with the a priori metric $\mathbf{g}_{\text{prior}}$, with uncertainties represented by the covariance operator $\mathbf{C}_{\mathbf{g}}$, and we end up with the a posteriori metric $\widetilde{\mathbf{g}}$, with uncertainties represented by the covariance operator $\widetilde{\mathbf{C}}_{\mathbf{g}}$.

The practical experience we have with the quasi-Newton algorithm for travel-time fitting problems suggests that the algorithm should converge to the proper solution (with sufficient accuracy) in a few iterations (less than 10). Then, for all practical purposes, we can replace ∞ by 10 in the two equations 140–141.

An important practical consideration is the following. The Hessian operator (equation 137) shall be completely characterized below, and the different covariance operators shall be directly given. But the algorithm in equations 136–138 contains the inverse of these linear operators. It is a very basic result of numerical analysis (Ciarlet, 1982) that the numerical resolution of a linear system may be dramatically more economical than the numerical evaluation of the inverse of a linear operator. Therefore, we need to rewrite the quasi-Newton algorithm replacing every occurrence of the inverse of an operator by the associated resolution of a linear system.

Let us start by the evaluation of the gradient vector γ_k. Expression 138 can be rewritten

$$\boxed{\begin{aligned} \gamma_k \;=\; & \delta\mathbf{g}_k + (\mathbf{Z}_k\,\mathbf{C}_{\mathbf{g}})^t\,\delta\mathbf{z}^*_k + (\mathbf{T}_k\,\mathbf{C}_{\mathbf{g}})^t\,\delta\mathbf{t}^*_k + (\boldsymbol{\Sigma}_k\,\mathbf{C}_{\mathbf{g}})^t\,\delta\boldsymbol{\sigma}^*_k \\ & + (\mathbf{A}_k\,\mathbf{C}_{\mathbf{g}})^t\,\delta\mathbf{a}^*_k + (\boldsymbol{\Pi}_k\,\mathbf{C}_{\mathbf{g}})^t\,\delta\boldsymbol{\pi}^*_k + (\boldsymbol{\Omega}_k\,\mathbf{C}_{\mathbf{g}})^t\,\delta\boldsymbol{\omega}^*_k \quad , \end{aligned}} \tag{142}$$

where

$$\delta\mathbf{g}_k \;=\; \mathbf{g}_k - \mathbf{g}_{\text{prior}} \quad , \tag{143}$$

and where the vectors $\delta\mathbf{z}^*_k$, $\delta\mathbf{t}^*_k$, $\delta\boldsymbol{\sigma}^*_k$ $\delta\mathbf{a}^*_k$, $\delta\boldsymbol{\pi}^*_k$, and $\delta\boldsymbol{\omega}^*_k$, are the respective solutions of the linear systems

$$\boxed{\begin{aligned} \mathbf{C}_{\mathbf{z}}\,\delta\mathbf{z}^*_k \;&=\; \mathbf{z}(\mathbf{g}_k) - \mathbf{1} \\ \mathbf{C}_{\mathbf{t}}\,\delta\mathbf{t}^*_k \;&=\; \mathbf{t}(\mathbf{g}_k) - \mathbf{t}_{\text{obs}} \\ \mathbf{C}_{\sigma}\,\delta\boldsymbol{\sigma}^*_k \;&=\; \boldsymbol{\sigma}(\mathbf{g}_k) - \boldsymbol{\sigma}_{\text{obs}} \\ \mathbf{C}_{\mathbf{a}}\,\delta\mathbf{a}^*_k \;&=\; \mathbf{a}(\mathbf{g}_k) - \mathbf{a}_{\text{obs}} \\ \mathbf{C}_{\pi}\,\delta\boldsymbol{\pi}^*_k \;&=\; \boldsymbol{\pi}(\mathbf{g}_k) - \boldsymbol{\pi}_{\text{obs}} \\ \mathbf{C}_{\omega}\,\delta\boldsymbol{\omega}^*_k \;&=\; \boldsymbol{\omega}(\mathbf{g}_k) - \boldsymbol{\omega}_{\text{obs}} \quad . \end{aligned}} \tag{144}$$

Once the gradient vector γ_k is evaluated, one can turn to the iterative step (equation 136). It can be written

$$\boxed{\mathbf{g}_{k+1} \;=\; \mathbf{g}_k - \Delta\mathbf{g}_k \quad ,} \tag{145}$$

where $\Delta\mathbf{g}_k$ is the solution of the linear system

$$\mathbf{H}_k\,\Delta\mathbf{g}_k \;=\; \gamma_k \quad . \tag{146}$$

Using the expression 137 for the operator $\mathbf{H}_k$ we can equivalently say that $\Delta\mathbf{g}_k$ is the solution of the linear system

$$\boxed{\begin{aligned} & \Delta\mathbf{g}_k + (\mathbf{Z}_k\,\mathbf{C}_{\mathbf{g}})^t\,\Delta\mathbf{z}^*_k + (\mathbf{T}_k\,\mathbf{C}_{\mathbf{g}})^t\,\Delta\mathbf{t}^*_k + (\boldsymbol{\Sigma}_k\,\mathbf{C}_{\mathbf{g}})^t\,\Delta\boldsymbol{\sigma}^*_k \\ & \quad + (\mathbf{A}_k\,\mathbf{C}_{\mathbf{g}})^t\,\Delta\mathbf{a}^*_k + (\boldsymbol{\Pi}_k\,\mathbf{C}_{\mathbf{g}})^t\,\Delta\boldsymbol{\pi}^*_k + (\boldsymbol{\Omega}_k\,\mathbf{C}_{\mathbf{g}})^t\,\Delta\boldsymbol{\omega}^*_k \;=\; \gamma_k \quad , \end{aligned}} \tag{147}$$

where, when introducing the vectors

$$\boxed{\begin{aligned} \Delta\mathbf{z}_k &= \mathbf{Z}_k\,\Delta\mathbf{g}_k \\ \Delta\mathbf{t}_k &= \mathbf{T}_k\,\Delta\mathbf{g}_k \\ \Delta\boldsymbol{\sigma}_k &= \boldsymbol{\Sigma}_k\,\Delta\mathbf{g}_k \\ \Delta\mathbf{a}_k &= \mathbf{A}_k\,\Delta\mathbf{g}_k \\ \Delta\boldsymbol{\pi}_k &= \boldsymbol{\Pi}_k\,\Delta\mathbf{g}_k \\ \Delta\boldsymbol{\omega}_k &= \boldsymbol{\Omega}_k\,\Delta\mathbf{g}_k \quad , \end{aligned}} \tag{148}$$

the vectors $\Delta\mathbf{z}_k^*$, $\Delta\mathbf{t}_k^*$, $\Delta\boldsymbol{\sigma}_k^*$, $\Delta\mathbf{a}_k^*$, $\Delta\boldsymbol{\pi}_k^*$, and $\Delta\boldsymbol{\omega}_k^*$, are the respective solutions of the linear systems

$$\boxed{\begin{aligned} \mathbf{C_z}\,\Delta\mathbf{z}_k^* &= \Delta\mathbf{z}_k \\ \mathbf{C_t}\,\Delta\mathbf{t}_k^* &= \Delta\mathbf{t}_k \\ \mathbf{C}_{\sigma}\,\Delta\boldsymbol{\sigma}_k^* &= \Delta\boldsymbol{\sigma}_k \\ \mathbf{C_a}\,\Delta\mathbf{a}_k^* &= \Delta\mathbf{a}_k \\ \mathbf{C}_{\pi}\,\Delta\boldsymbol{\pi}_k^* &= \Delta\boldsymbol{\pi}_k \\ \mathbf{C}_{\omega}\,\Delta\boldsymbol{\omega}_k^* &= \Delta\boldsymbol{\omega}_k \quad . \end{aligned}} \tag{149}$$

In equations 142 and 147 one needs to evaluate vectors whose generic form is

$$\mathbf{b} = (\mathbf{L}\,\mathbf{C})^t\,\mathbf{a} \quad . \tag{150}$$

the vector $\mathbf{a}$ being known. They involve the transpose of an operator. To evaluate these vectors one must resort to the very definition of transpose operator. By definition, the operator $(\mathbf{L}\,\mathbf{C})^t$ is the transpose of the linear operator $(\mathbf{L}\,\mathbf{C})$ if, and only if, for any $\mathbf{a}^*$ and any $\mathbf{a}$,

$$\langle\, \mathbf{a}^*\,,\,(\mathbf{L}\,\mathbf{C})^t\,\mathbf{a}\,\rangle = \langle\,(\mathbf{L}\,\mathbf{C})\,\mathbf{a}^*\,,\,\mathbf{a}\,\rangle \quad . \tag{151}$$

As the linear tangent operators are characterized (for all the nonlinear applications appearing above), we know how to write the right-hand side of this equation explicitly. As the vector $\mathbf{a}$ is known, the condition that the obtained expression must hold for any vector $\mathbf{a}^*$ gives an explicit expression for $\mathbf{b} = (\mathbf{L}\,\mathbf{C})^t\,\mathbf{a}$. Appendixes 3.2.1 and 3.2.2 provide two examples of this kind of evaluations.

6.5 Kalman Filter

Assume that some *linear model* allows to make a *preliminary* prediction of the state of the system at time k in terms of the state of the system at time $k-1$ (we retain here the notations in Grewal et al. (2001)):

$$x_k^- = \Phi_k\,x_{k-1}^+ \quad . \tag{152}$$

If the uncertainties we had on x_{k-1}^+ are represented by the covariance matrix P_{k-1}^+ and if the prediction by the linear model Φ_k has uncertainties described by the covariance matrix Q_{k-1} , the uncertainty we have on x_k^- is represented by the covariance matrix

$$P_k^- = \Phi_k\,P_{k-1}^+\,\Phi_k^t + Q_{k-1} \quad . \tag{153}$$

So we have the prior value x_k^- with uncertainties described by the prior covariance matrix P_k^- . To pass from the preliminary estimate x_k^- to the actual estimate x_k^+ we now use some observed data

z_k that is assumed to be related to x_k^+ via a linear relation $z_k \approx H_k\, x_k^+$, with uncertainties described by the covariance matrix R_k. The standard theory of linear least-squares then provides the posterior estimate as

$$x_k^+ \;=\; x_k^- + P_k^-\; H_k^t\; (\, H_k\, P_k^-\; H_k^t + R_k\,)^{-1}\; (\, z_k - H_k\; x_k^-\,) \quad , \tag{154}$$

that has uncertainties represented by the posterior covariance matrix

$$P_k^+ \;=\; P_k^- - P_k^-\; H_k^t\; (\, H_k\, P_k^-\; H_k^t + R_k\,)^{-1}\; H_k\; P_k^- \quad . \tag{155}$$

Then, if at each time step we input Φ_k, Q_{k-1}, z_k, H_k, and R_k, equations 152–155 allow to have a continuous estimation of the state of the system, x_k^+, together with an estimation of the uncertainties, P_k^+.

The reader may recognize that equations 154–155 are identical to the standard equations of linear least-squares theory (see equations 3.37 and 3.38 in Tarantola (2005)). The matrix

$$K_k \;=\; P_k^-\; H_k^t\; (\, H_k\, P_k^-\; H_k^t + R_k\,)^{-1} \quad , \tag{156}$$

that appears in the two equations 154–155, is called the 'Kalman gain matrix'.

Example. As a simple example, consider, in non-relativistic physics, the trajectory of a mass that has been equipped with some sensors. We can choose to represent the state of the system at any time by a 9-dimensional vector x, that contains the three values of the position, the three values of the velocity and the three values of the acceleration. Assume that, as a result of the previous iteration, at some moment we have the estimation x_{k-1}^+ with uncertainties P_{k-1}^+. Equation 152 may simply correspond to the use of the velocity to extrapolate the position one step in time, to use the acceleration to extrapolate the velocity, and to keep the acceleration unchanged. This perfectly characterizes the matrix Φ_k. Equation 153 then is used to update the estimation of uncertainty, where we can take for Q_{k-1} something as simple as a zero matrix excepted for the three diagonal elements associated to the acceleration, where a small variance will take into account that our extrapolation of acceleration is uncertain. The data z_k may consist in the output of some sensors, like accelerometers or data from a satellite positioning system. The relation $z_k = H_k\, x_k^-$ would correspond to the theoretical calculation of the data z_k given the state x_k^-. This is not a linear relation, and the theory should be developed to directly account for this, but if the time steps are small enough, we can always linearize the theory, this then defining the matrix H_k. Denoting now by z_k the actual output of the sensors, and by R_k the experimental uncertainties, equation 154 is used to obtain our second estimation of the state of the system, equation 155 providing the associated uncertainties.