IMA Journal of Numerical Analysis (2002) **22**, 359–390

Newton's method on Riemannian manifolds and a geometric model for the human spine

ROY L. ADLER†
Department of Mathematical Sciences, IBM T.J.Watson Research Center, Yorktown Heights, NY 10598, USA

JEAN-PIERRE DEDIEU‡
MIP, Départment de Mathématiques, Université Paul Sabatier, 31062 Toulouse Cedex 04, France

JOSEPH Y. MARGULIES§
Orthopaedic Spine Surgeon, 8 Usonia Road, Pleasantville, NY 10570, USA

AND

MARCO MARTENS¶ AND MIKE SHUB**
Department of Mathematical Sciences, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA

[Received on 6 March 2000; revised on 18 July 2001]

To study a geometric model of the human spine we are led to finding a constrained minimum of a real valued function defined on a product of special orthogonal groups. To take advantge of its Lie group structure we consider Newton's method on this manifold. Comparisons between measured spines and computed spines show the pertinence of this approach.

Keywords: Newton's method; Riemannian manifold; orthogonal group; human spine.

1. Introduction

We are motivated by a geometric model of the spine to study a certain optimization problem. Since the orientation of vertebrae can be specified by a frame of three orthogonal vectors in Euclidean three-space, we are led to finding a constrained minimum, or at least a local minimum, of a real-valued function ϕ defined on a product $\mathbf{SO}(3)^N$ of special orthogonal groups. The function in question will turn out to be quadratic and the constraint affine.

In order to take advantage of the Lie-group structure of $\mathbf{SO}(3)$, we preferred to treat this constrained optimization problem as one of finding the zeros of a gradient vector field

†Email: adler@watson.ibm.com
‡Email: dedieu@mip.ups-tlse.fr
§Email: yossi@jymargulies.com
¶Email: martens@us.ibm.com
** Email: mshub@us.ibm.com

on a sub-manifold of $\mathbf{SO}(3)^N$ rather than use the method of Lagrange multipliers. There is some evidence that the numerics of our approach may be slightly better conditioned, but more analysis is needed in order to justify this claim.

It is known from Morse theory that for a generic real-valued function defined on $\mathbf{SO}(3)^N$ there are at least 4^N zeros of its gradient vector field. This suggests that there may also be many zeros for a gradient vector field of a function on the intersection of $\mathbf{SO}(3)^N$ with the constraint manifold. Such zeros are merely extreme points, not necessarily local minima. However, the computations we have run involving actual clinical cases resulted in solutions sufficiently isolated from the other local extrema to indicate that our method is effective.

Newton's method for finding a zero of a vector field proceeds by a sequence of updates. We get a new candidate from an old one by first solving a linear approximation to a nonlinear problem. Since a solution by linearization naturally lies in the tangent space to the manifold, whereas an updated point should lie in the manifold itself, we then employ a transformation, which we call a *retraction*, to map a tangent vector to a nearby element of the manifold.

In Section 2 we introduce and motivate our model for the shape of the spine. We compare spines predicted by our model with measured ones for five patients.

In Section 3 we describe Newton's method for the problem: Find a zero for the map $F : \mathbf{M} \to \mathbf{V}$ where $\mathbf{M}$ is a differentiable manifold, $\mathbf{V}$ is a Euclidean vector space and a zero is a point $x \in \mathbf{M}$ such that $F(x) = 0$. In this section we introduce retractions and give some examples.

Newton methods make zeros of F correspond to fixed points of an iterative process N_F with good convergence properties. For overdetermined systems and least-square problems, convergence occurs in more limited circumstances. In general only local minima correspond to attractors, but not all local minima are attractors, see Dennis & Schnabel (1983); Dedieu & Shub (2000); Dedieu & Kim (2002).

In Section 4 we study these questions on manifolds and even extend what is known on vector spaces.

An alternative to solving least-square problems by the Newton–Gauss method is to use Newton's method to find zeros of the gradient vector field grad $f(x)$ where $f(x) = \|F(x)\|^2/2$. In Section 5 we study Newton's method for finding zeros of vector fields on manifolds. The Rayleigh Quotient Iteration, Lord Rayleigh (1899), for the eigenvalue problem arises naturally in this context.

In Section 6 we derive formulae for the Newton iteration for the gradient vector field associated with our model of the spine.

There is quite a bit of previous work on such questions: Shub (1986) defined Newton's method for the problem of finding the zeros of a vector field on a manifold and used retractions as we do here; Udriste (1994) studied Newton's method to find the zeros of a gradient vector field defined on a Riemannian manifold; Owren & Welfert (1996) defined Newton iteration for solving the equation $F(x) = 0$ where F is a map from a Lie group to its corresponding Lie algebra; Smith (1994) and Edelman *et al.* (2000) developed Newton and conjugate gradient algorithms on the Grassmann and Stiefel manifolds. In general these authors define Newton's method via the exponential map which 'projects' the tangent bundle onto the manifold. Shub (1993); Shub & Smale (1993a,b,c, 1996, 1994), see also, Blum *et al.* (1998); Malajovich (1994) and Dedieu & Shub (2000) introduce and study Newton's method on projective spaces and their products.

2. A geometric model for the spine

The spine is a complex organ subject to a multitude of unquantified processes—e.g. muscular, anatomical, neurological, metabolic, developmental—driving its three-dimensional configuration. The only clinically measurable data available are that of the vertebrae themselves obtained by radiography. Thus we address the question: Can the balanced erect three-dimensional configuration of the spine be predicted from the shape of the bones alone?

If yes, then there are important consequences. Treatment of spinal problems by surgical intervention has been on the increase, and there are now over 500 000 such operations a year in the USA alone. Of all spinal problems, scoliosis is the most challenging to deal with. For the most part, it is a three-dimensional enigmatic deformity. It afflicts around 25 adolescents out of 1000 enough to require medical attention. About three of these will progress, in a 10:1 ratio of female to male, to a degree severe enough to be considered for surgery. Approximately one of these will then actually undergo one of the most heroic surgical procedures known to medicine: namely, spinal fusion, whose primary goal is to prevent further deterioration and secondarily to correct the deformity as much as is still possible. An affirmative answer to the question could lead to better corrections.

In view of what can be clinically measured about the spine, our approach to modelling it is to simplify, in fact oversimplify, to such an extent that one may well question how anything meaningful can result. Yet using almost no knowledge of human anatomy and musculature along with the crude approximations we are about to make, it is astounding how close our computer simulations resemble actual spines.

We base our model on a single hypothesis which seems to be consistent with medical experience: namely,

> The erect stationary three-dimensional configuration of the human spine expresses nature's intent to *balance* the head level over the pelvis while *aligning* neighbouring vertebrae as closely as possible.

The only other things which are incorporated into the model, have to do with expressing this hypothesis mathematically and taking advantage of every possible simplification.

Figure 1 depicts a typical vertebra. Pointed out is the vertebral body which is the portion of the vertebra which most closely resembles something that can be approximated by a simple mathematical object—namely a cylinder, though a non-circular one. Vertebral bodies bear most of the load on the spine so, except for the sacrum, our spine will consist of only them. We shall assume the top and bottom surfaces of these elements to be planar figures with an orientation. In this study, we shall not include the seven cervical vertebrae of the neck. So from bottom to top our spinal elements are the sacrum $S1$, the five lumbar vertebral bodies $L5, \dots, L1$, and the twelve thoracic vertebral bodies $T12, \dots, T1$. In Fig. 3, these are also labelled $V_1 = S1$, $V_2 = L5, \dots, V_6 = L1$, $V_7 = T12, \dots, V_{18} = T1$. The orientation of bottom surface of V_k will be given by an orthogonal matrix m_k, the columns of which form a frame of orthogonal vectors (see Fig. 2). We adopt the following convention: the positive x-axis points to the right of the patient, the positive

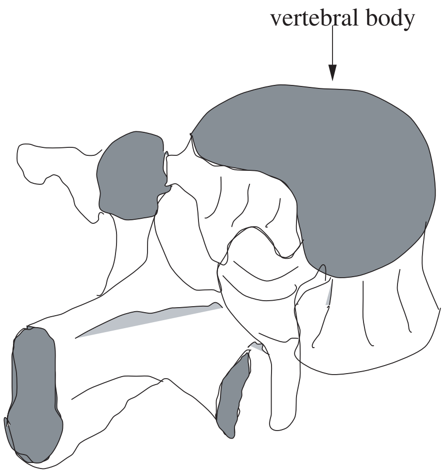

FIG. 1. Typical vertebra.

y-axis forward, and the positive z-axis up. For a normal vertebra the first column of m_k is a vector pointing mostly in the x-direction—i.e. it has positive first component and very small second and third ones. The second and third columns are vectors lying close to the (y, z)-plane—i.e. they have very small first components. The second column has a positive second component and the third column a positive third component. The distortion of the top surface of V_k relative to the bottom is given by another orthogonal matrix q_k. Thus the orientation of the top surface of V_k is $m_k q_k$.

Our specification of the spine will consist of orthogonal frames m_k, q_k, heights h_k of vertebral bodies, and thicknesses t_k of the discs between vertebrae for $k = 1, \dots, 18$. The other dimensions, side-to-side diameters a_k and front-to-back diameters b_k of vertebral bodies, are really only needed for drawing purposes. However, we did employ them for another role as we shall mention later.

In Fig. 3, the vector v_k is the third column of m_k multiplied by the scalar $h_k + t_{k-1}/2$, and the vector v'_k is the third column of $m_k q_k$ multiplied by the scalar $t_k/2$. The base of v_k is placed at the head of v'_{k-1} and the base of v'_k at the head of v_k. So mathematically for us the spine is given by v_k, v'_k, along with the other orientation information contained in m_k and q_k. When we come to draw the spine, we replace the artistic version of a vertebral body in Fig. 1 by the stylized one in Fig. 2. Doing so significantly speeds up computer graphics and yet captures what we take to be the essential features of these bodies. We locate the point representing the centre of our stylized sacrum at the origin of our coordinate system. We place the stylized vertebra V_k so that v_k, the base of which lies midway in the disc between V_k and V_{k-1}, passes through a point representing the centre of the bottom surface of V_k and ends at a similar point on the top surface. The base v'_k is also located at this point on the top surface and its head at a point midway in the disc between V_k and V_{k+1}. The head of v'_{18} lies midway between $T1$ and $C7$ the seventh cervical vertebra. That point, which is the sum $\sum_{k=1}^{18} v_k + v'_k$, has coordinates (δ_x, δ_y, z). According to the *balance* assumption of our hypothesis (δ_x, δ_y) is approximately $(0, 0)$, a fact consistent with clinical data.

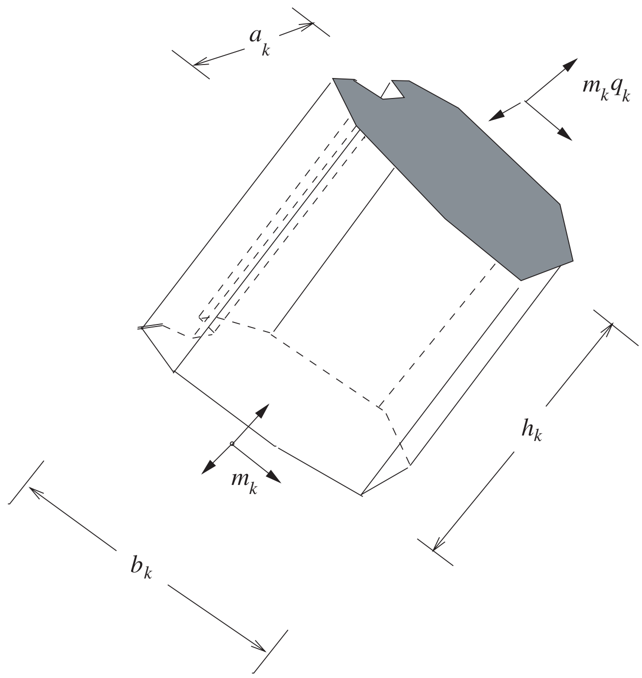

FIG. 2. Stylized vertebral body.

For a typical normal spine the sacrum is tilted counterclockwise about 40^o and $T1$ about 27^o. Thus the orthogonal matrices m_1 and m_{18} are approximately

$$m_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-40^o) & -\sin(-40^o) \\ 0 & -\sin(-40^o) & \cos(-40^o) \end{pmatrix} \tag{2.1}$$

and

$$m_{18} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-27^o) & -\sin(-27^o) \\ 0 & -\sin(-27^o) & \cos(-27^o) \end{pmatrix}. \tag{2.2}$$

The items $q_1, \ldots, q_{18}, m_1, m_{18}$, and (δ_x, δ_y) are data associated with a specific patient obtained from one type of radiography or another. In order to measure the total alignment discrepancies between spinal elements we form the function

$$\phi(m_2, \ldots, m_{17}) = \frac{1}{2} \sum_{k=1}^{17} w_k \, \|m_{k+1} - m_k q_k r_k\|^2 . \tag{2.3}$$

The coefficients w_k are non-negative quantities which allow for weighting individual alignment discrepancies differently as k varies. The r_k are orthogonal matrices which allow

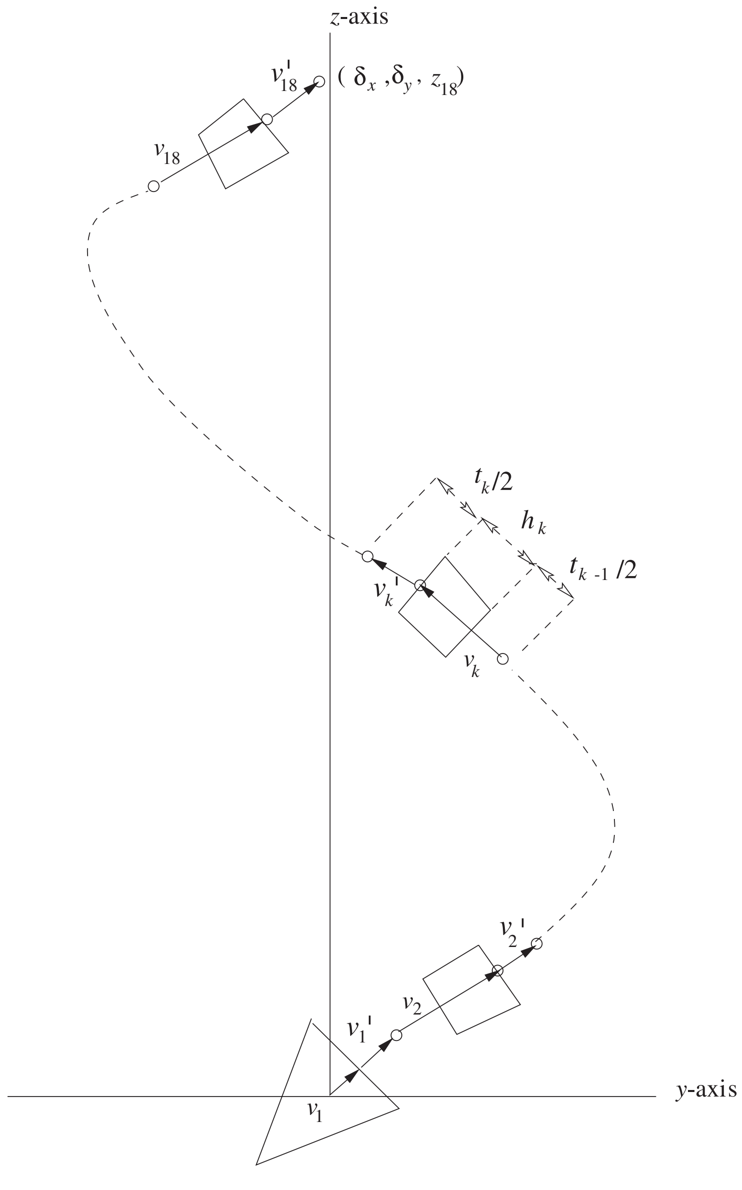

FIG. 3. Coordinatization of the spine.

for an offset with respect to which the discrepancy of alignment between V_k and V_{k+1} can be measured. The quantity in the right-hand-side of (2.3) has the appearance of potential energy à la Hooke's law. Potential energy according to that law is a quadratic function of displacements. Accordingly we call that quantity *spinal energy* and the quantities w_k *stiffness coefficients*. However, keep in mind that spinal energy has nothing to do with an actual physical energy. It is rather an ideal quantity which is meant to sum up all the unknown forces and processes shaping an erect spine and whose only virtue is in its use in empirically approximating spines.

To obtain the three-dimensional configuration of the spine, we seek $m_2, \ldots, m_{17}$ which are a consequence of our basic hypothesis. The fixed orientations of V_1, V_{18} given by m_1, m_{18} along with the position of the head of v'_{18} given by (δ_x, δ_y) are constraints in conflict with perfect alignments—i.e. with $\phi(m_2, \ldots, m_{17}) = 0$. So the hypothesis when translated into a mathematical statement becomes the optimization problem:

Find $m = (m_2, \ldots, m_{17}) \in \mathbf{SO}(3)^{16}$ at which $\phi(m)$ assumes a minimum (*)
subject to the additional constraint that

$$h_1(m) \equiv \sum_{k=1}^{18} [(h_k + \mathrm{t}_{k-1}/2) m_k(1,3) - (\mathrm{t}_k/2) m_k q_k(1,3)] - \delta_x = 0, \tag{2.4a}$$

$$h_2(m) \equiv \sum_{k=1}^{18} [(h_k + \mathrm{t}_{k-1}/2) m_k(2,3) - (\mathrm{t}_k/2) m_k q_k(2,3)] - \delta_y = 0. \tag{2.4b}$$

Since there is no spinal element V_0, we set $t_0 = 0$ in (2.4a) and (2.4b).

In (2.3), (2.4a) and (2.4b) there are two categories of parameters: namely, ones associated with a specific spine and universal ones, or at least ones applying to a large class of spines. Category one consists of $m_1, m_{18}; q_1, \ldots, q_{18}; t_1, \ldots, t_{18}; h_1, \ldots, h_{18}; \delta_x, \delta_y$ which are obtained from radiography of one form or another. Category two consists of the stiffness coefficients w_k and the offset matrices r_k. We chose these parameters by guesswork. A little experimentation indicated that the stiffness coefficients should increase with k: so for lack of anything better we chose $w_k = 1/t_k a_k b_k$. Furthermore, choosing $r_k = q_k$ gave the best results of anything we tried. Note that while these w_k and r_k depend on individual spines, the form of the dependency does not. Also note that these offsets square the influence of vertebral distortions on spinal energy. We have no explanation why our choice might be correct other than the fact that in some vague way the shape of a vertebra is a reflection of forces upon it.

We present results, Figs 4–8, for five patients which were drawn by means of Matlab graphics. Patient 1 has a normal spine, the others are scoliotic. Each figure compares two views of a measured spine with a computed one. The measured spines were determined from available x-rays by measuring angles of what appeared to be edges of surfaces and distances between edges. Estimation of category one parameters were made by standard analytic geometry calculations. Besides the two frames m_1 and m_{18} we obtained the other 16 orthogonal frames $m_2, \ldots, m_{17}$. Based on these 18 orthogonal frames we drew what we call measured spines in the figures below. Doubt is immediately cast as to how closely a measured spine represents reality. Problems arise with obtaining accurate three-dimensional surveys from the usual x-rays: just to name a few, a patient's posture is not held constant for the two views, location of the x-ray gun with respect to the patient and the film plate is not recorded; some vertebrae are not seen clearly in an x-ray, especially the sagittal one; it is almost impossible to get any usable twist data from x-rays; furthermore, what is be seen in any view is subject to interpretation. (δ_x, δ_y, z) can both be measured and calculated from $m_1, \ldots, m_{18}$, which is useful as a check on accuracy. This provided further evidence of the aforementioned difficulties. Nevertheless, visual comparison of measured spines with actual x-rays seems to indicate that our approximations are not so bad.

If we let

$$\Phi(m) = (\ldots, w_k(i, j)(m_{k+1}(i, j) - m_k q_k r_k(i, j)), \ldots), \tag{2.5}$$

then we can write

$$\phi = \frac{1}{2} \, \|\Phi\|^2 .$$

In 2.5 we have incorporated stiffness coefficients which are more general in that we have allowed them to vary not only with k but with also with i, j.

As is customary in searching for minima, instead of solving (*) we consider another problem:

$$\text{Find a zero of the vector field } \operatorname{grad}(\phi|\mathbf{SO}(3)^{16} \cap h^{-1}(0)), \tag{**}$$

where $h = (h_1, h_2)$. Computed spines were the result of numerical solutions of (**) obtained by using Matlab. The solutions were based on the Newton methods to which the rest of this paper is devoted. As one would expect, good starting points for Newton steps are the measured m. However, sometimes there are difficulties: for, example, the measured m of patient 5 was not in a Newton basin of attraction of the minimum but rather in one for a different extremum. We overcame this difficulty by first performing a few steps of a gradient method before applying the Newton method.

Despite the problem of oversimplification of the model, despite the difficulties with data acquisition, and despite the local discrepancies between computed spines and measured ones, global three-dimensional distortions have been captured quite well in all cases. We find this remarkable enough to deserve further study. In fact the results are so promising that the answer to our original question has a fighting chance of being yes. Finally, no serious attempt at tuning the parameters was made. Fitting w_k and r_k from a sample population of spines is also for the future.

3. Newton's Method on a Manifold

Newton's method is a classical numerical method to solve a system of nonlinear equations

$$F : \mathbf{E} \to \mathbf{V}$$

with $\mathbf{E}$ and $\mathbf{V}$ two Euclidean spaces of arbitrary dimensions. If $x \in \mathbf{E}$ is an approximation of a zero of this system then, Newton's method updates this approximation by linearizing the equation $F(y) = 0$ around x so that

$$F(x) + DF(x)(y - x) = 0.$$

When $DF(x)$ is an isomorphism we obtain the classical Newton iteration

$$y = N_F(x) = x - DF(x)^{-1}F(x).$$

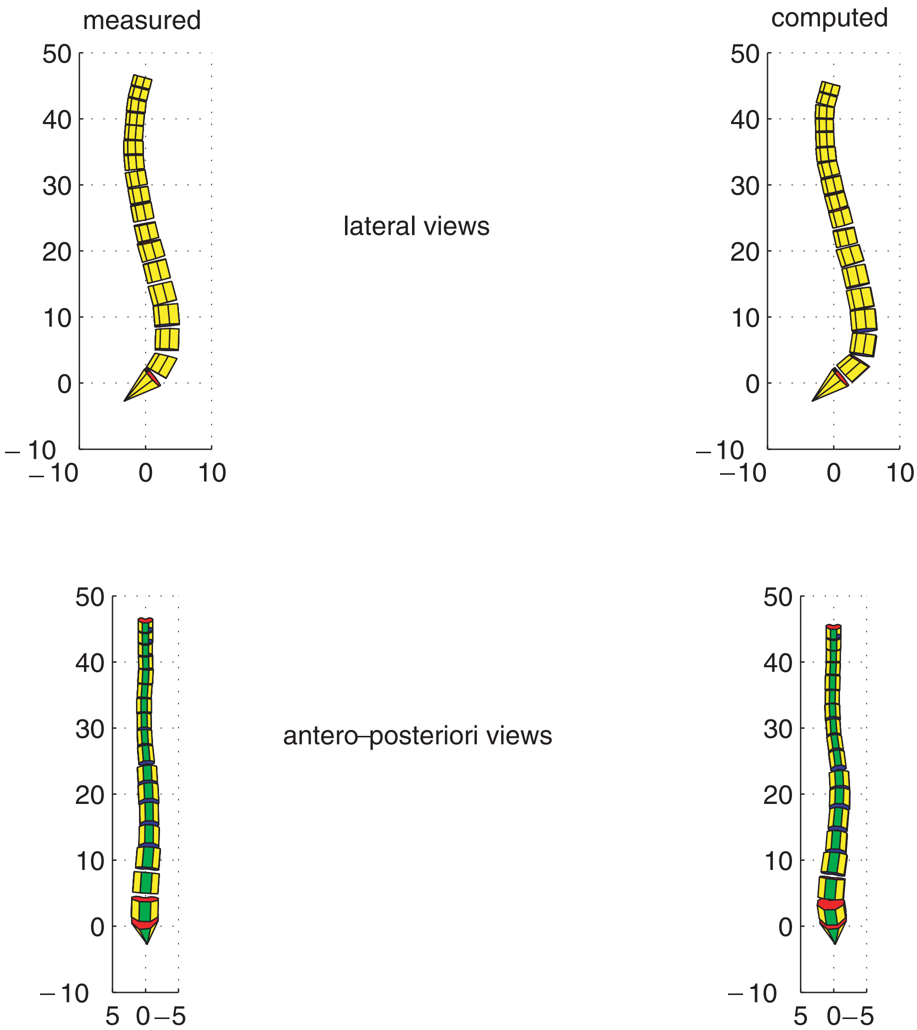

FIG. 4. Patient 1 (normal spine).

When **E** and **V** are two Euclidean spaces and when $DF(x)$ is not an isomorphism we choose its Moore–Penrose inverse $DF(x)^{\dagger}$ instead of its classical inverse:

$$y = N_F(x) = x - DF(x)^{\dagger} F(x).$$

We recall that the Moore–Penrose inverse (also called psuedo-inverse) of a linear operator

$$A : \mathbf{E} \to \mathbf{V}$$

is the minimum norm one sided 'inverse' of A. It is the composition $A^{\dagger} = B \circ \pi_{\mathrm{Im}\ A}$ of two maps where $\pi_{\mathrm{Im}\ A}$ is the orthogonal projection in **V** onto Im A and B is the inverse

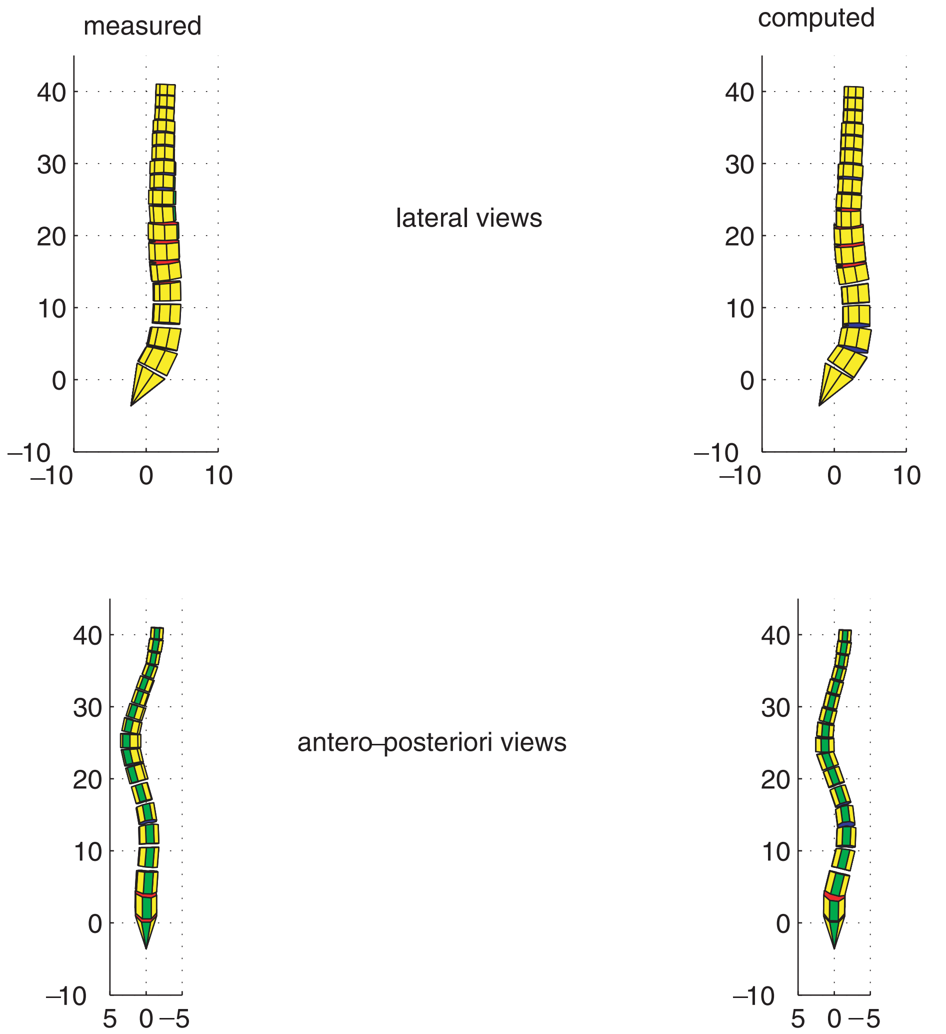

FIG. 5. Patient 2

of the restriction $A|_{(\text{Ker } A)^\perp} : (\text{Ker } A)^\perp \to \text{Im } A$. We have $A^\dagger = (A^*A)^{-1}A^*$ when A is injective, $A^\dagger = A^*(AA^*)^{-1}$ when A is surjective. Notice that $A^\dagger A = \pi_{(\text{Ker } A)^\perp}$ and $AA^\dagger = \pi_{\text{Im } A}$.

When the source space is a manifold $\mathbf{M}$ instead of the Euclidean space $\mathbf{E}$ the previous definition of Newton's method is meaningless: x and $N_F(x)$ should lie in $\mathbf{M}$ while the Moore–Penrose inverse of the derivative $DF(x)$ takes its values in the tangent space at x to $\mathbf{M}$:

$$DF(x)^\dagger : \mathbf{V} \to \mathbf{T_x M}.$$

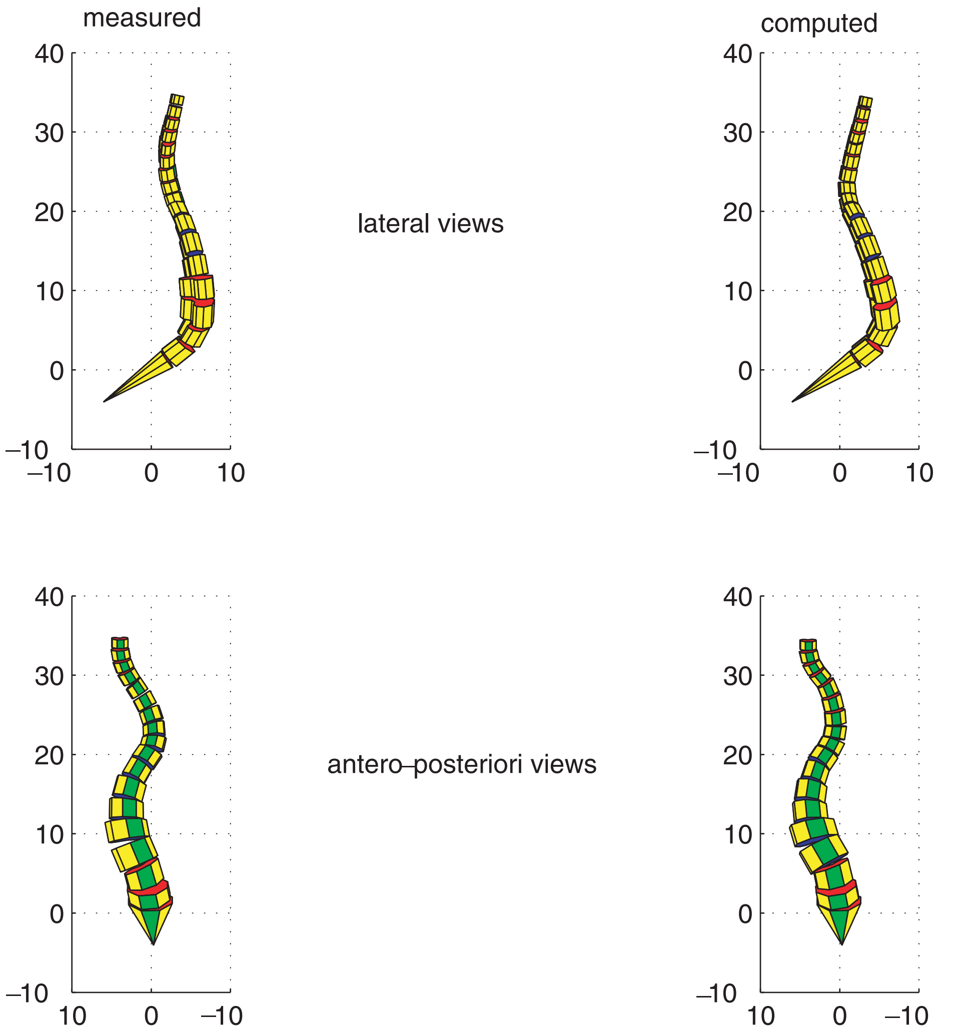

FIG. 6. Patient 3

For this reason we introduce a *retraction*

$$R : T\mathbf{M} \to \mathbf{M}$$

where $T\mathbf{M}$, the disjoint union of $T_x\mathbf{M}$ for $x \in \mathbf{M}$, is the tangent bundle of $\mathbf{M}$. See Hirsch (1976) or any other introductory book on differential topology for a discussion of manifolds and their tangent bundles. The tangent bundle to a manifold $\mathbf{M}$ is naturally a manifold itself and it contains $\mathbf{M}$ as a submanifold. A point $x \in \mathbf{M}$ is identified with the zero tangent vector 0_x of $T_x\mathbf{M}$. R is assumed to be a smooth map defined on a neighbourhood of $\mathbf{M}$ in $T\mathbf{M}$,

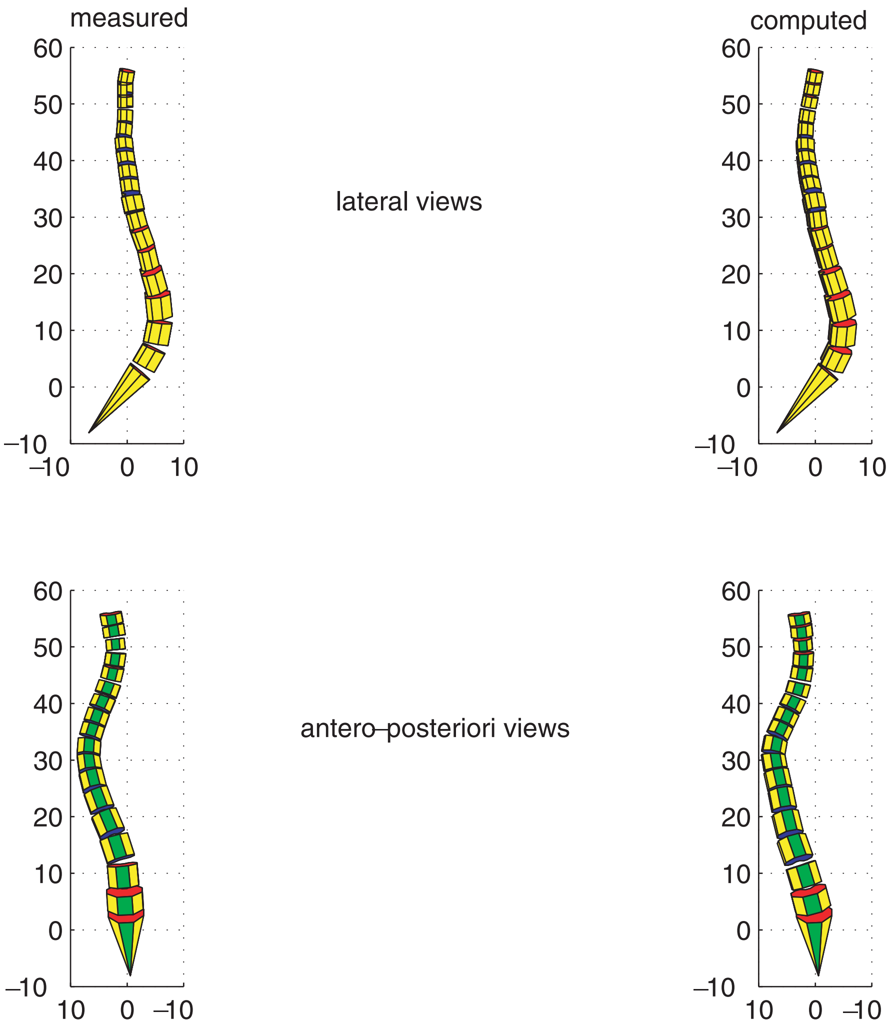

FIG. 7. Patient 4.

taking its values in the manifold $\mathbf{M}$ and satisfying the following properties. Let us denote by R_x the restriction of R to $T_x\mathbf{M}$, then

1. R_x is defined in an open ball $B_{r_x}(0_x) \subset T_x\mathbf{M}$ of radius r_x about 0_x.
2. $R_x(\dot{x}) = x$ if and only if $\dot{x} = 0_x$.
3. $DR_x(0_x) = \mathrm{id}_{T_x\mathbf{M}}$
   The tangent bundle of $T\mathbf{M}$ at a point $x \in \mathbf{M}$ may be identified with the product $T_x\mathbf{M} \times T_x\mathbf{M}$ of two copies of $T_x\mathbf{M}$. The first factor corresponds to vectors tangent to $\mathbf{M}$ at x and the second to vectors tangent to zero in the vector space $T_x\mathbf{M}$. Since

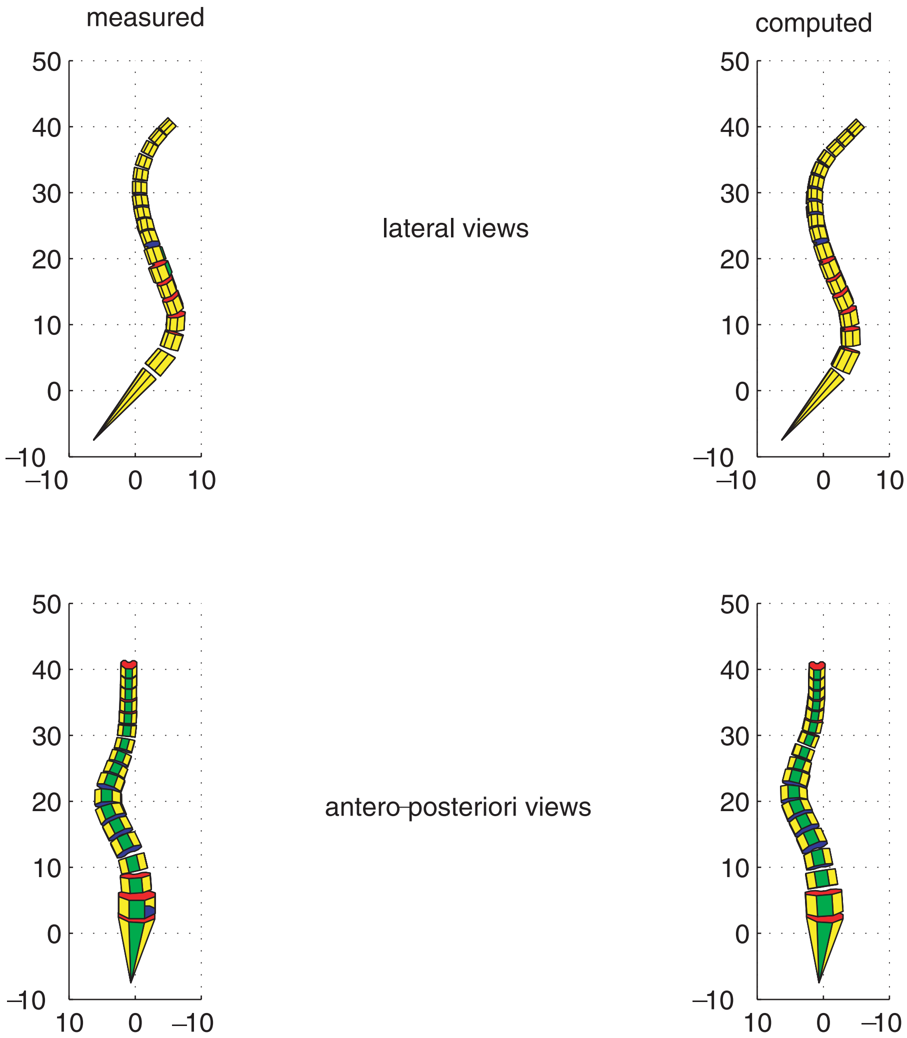

FIG. 8. Patient 5.

$R(x) = x$ for $x \in \mathbf{M}$, it follows that $DR(x)$ is also $\mathrm{id}_{T_x\mathbf{M}}$ when restricted to the first factor of $T_x\mathbf{M} \times T_x\mathbf{M}$.

4. If we write $\frac{\partial R(x)}{\partial x}$ to denote the restriction of $DR(x)$ to the first factor of $T_x\mathbf{M} \times T_x\mathbf{M}$ and $\frac{\partial R(x)}{\partial \dot{x}}$ to denote the restriction to the second factor (i.e. $DR_x(x)$), then

$$DR(x)(u, v) = \frac{\partial R(x)}{\partial x}u + \frac{\partial R(x)}{\partial \dot{x}}v = u + v$$

for $x \in \mathbf{M}$ and for $(u, v) \in T_x\mathbf{M} \times T_x\mathbf{M}$.

We suppose here that $\mathbf{M}$ has a Riemannian metric, that is for each $x \in \mathbf{M}$, $T_x\mathbf{M}$ is equipped with an inner product $\langle \dot{x}, \dot{y} \rangle_x$ which varies smoothly with x. So the Moore–Penrose inverses of linear operators defined on $T_x\mathbf{M}$ are computed with respect to this inner product structure and vary smoothly where the rank of DF is constant.

DEFINITION 1 For

$$F : \mathbf{M} \to \mathbf{V}$$

and $x \in \mathbf{M}$ the Newton operator $N_F(x)$ is defined by

$$N_F(x) = R_x \left(-DF(x)^\dagger F(x) \right)$$

so that $N_F : \mathbf{M} \to \mathbf{M}$.

If $-DF(x)^\dagger F(x)$ is not in B_{r_x}, then $N_F(x)$ is not defined.

Let us check that this definition is consistent with the usual one when $\mathbf{M} = \mathbf{E}$, a Euclidean space. The tangent space $T_x\mathbf{E}$ may be identified with $\mathbf{E}$ itself. We define the retraction R by $R_x(\dot{x}) = x + \dot{x}$ and we obtain $N_F(x) = x - DF(x)^\dagger F(x)$.

Let us now study three examples.

EXAMPLE 1 **The unit sphere $\mathbf{S}^{n-1} \subset \mathbb{R}^n$.**

The unit sphere is a compact submanifold in $\mathbb{R}^n$ and its dimension is $n-1$. The tangent space to the unit sphere at x is identified to the hyperplane in $\mathbb{R}^n$ orthogonal to the vector x:

$$T_x\mathbf{S}^{n-1} = x^\perp = \{\dot{x} \in \mathbb{R}^n \ : \ \langle \dot{x}, x \rangle \ = \ 0\}.$$

This tangent space is equipped with the inner product structure induced by $\mathbb{R}^n$. We define a retraction by

$$R_x(\dot{x}) = \frac{x + \dot{x}}{\|x + \dot{x}\|}$$

for any $\dot{x} \in T_x\mathbf{S}^{n-1}$. Its inverse is defined for any $y \in \mathbf{S}^{n-1}$ such that $\langle y, x \rangle > 0$:

$$R_x^{-1}(y) = \frac{y}{\langle y, x \rangle} - x.$$

In this context, for $F : \mathbf{S}^{n-1} \to \mathbf{V}$, Newton's map is defined by

$$N_F(x) = \frac{x - DF(x)^\dagger F(x)}{\|x - DF(x)^\dagger F(x)\|}.$$

EXAMPLE 2 **The special orthogonal group $\mathbf{SO}(n) \subset \mathcal{M}(n, n)$.**

The special orthogonal group is a compact submanifold in the space of n by n real matrices with determinant +1. Its dimension is equal to $n(n-1)/2$. The tangent space at the identity matrix id is equal to $\mathcal{A}(n, n)$, the space of n by n antisymmetric matrices. More generally, the tangent space at $u \in \mathbf{SO}(n)$ is given by

$$T_u\mathbf{SO}(n) = u\mathcal{A}(n, n).$$

This tangent space is equipped with the usual scalar product

$$\langle \dot{a}, \dot{b} \rangle = \mathrm{T}race(\dot{b}^* \dot{a}).$$

A retraction is given by the exponential of a matrix

$$R_u(\dot{u}) = u \exp(u^{-1}\dot{u})$$

defined for any $u \in \mathbf{SO}(n)$ and $\dot{u} \in T_u\mathbf{SO}(n)$. Here $\exp(a)$ is the exponential of the matrix a defined by $\exp(a) = \sum_{k=0}^{\infty} a^k/k!$. Let us also consider the logarithm $\log(\mathrm{id} + x) = \sum_{k=1}^{\infty}(-1)^{k+1}x^k/k$ defined for any matrix satisfying $\|x\| < 1$. The inverse of the retraction R_u is defined for any orthogonal matrix $v \in \mathbf{SO}(n)$ with $\|v - u\| < 1$:

$$R_u^{-1}(v) = u \log(u^{-1}v).$$

Using this retraction, Newton's map associated with a function $F : \mathbf{SO}(n) \to \mathbf{V}$ is given by

$$N_F(u) = u \exp\left(-u^{-1}DF(u)^{\dagger}F(u)\right).$$

Here is another retraction defined on $\mathbf{SO}(n)$. For an antisymmetric matrix a, let

$$Q_{\mathrm{id}}(a) = (\mathrm{id} + a)((\mathrm{id} + a^*)(\mathrm{id} + a))^{-1/2} = (\mathrm{id} + a)(\mathrm{id} - a^2)^{-1/2}.$$

Then $Q_{\mathrm{id}}(a)$ is the closest orthogonal matrix to $\mathrm{id}+a$. It is the orthogonal matrix appearing in the polar decomposition of $(\mathrm{id} + a)$ (Dedieu, 1996; Fan & Hoffman, 1955). For any $u \in \mathbf{SO}(n)$ and $\dot{u} \in T_u\mathbf{SO}(n)$ let

$$Q_u(\dot{u}) = u Q_{\mathrm{id}}(u^{-1}\dot{u}).$$

The retraction Q is defined by Q_u for each orthogonal matrix u.

A third retraction Orth is given by defining Orth_u for each orthogonal matrix u as follows. For an antisymmetric matrix a, let $\mathrm{Orth}_{\mathrm{id}}(a) = \prod_{1 \leqslant i < j \leqslant n} R_{ij}(a_{ij})$ where the order of multiplication is any fixed order, say lexicographic, and where $R_{ij}(a_{ij})$ is the rotation in the ij-plane by an angle a_{ij}. The entries of $R_{ij}(a_{ij})$ are as follows:

$$\begin{aligned} ij - \text{entry} &= -\sin(a_{ij}) \\ ji - \text{entry} &= \sin(a_{ij}) \\ ii - \text{entry} &= \cos(a_{ij}) \\ jj - \text{entry} &= \cos(a_{ij}) \\ kk - \text{entry} &= 1, \quad k \neq i, j \\ \text{other entries} &= 0 \end{aligned}$$

When $n = 3$, $\mathrm{Orth}_{\mathrm{id}}$ is injective for $-\pi/2 < a_{ij} < \pi/2$.

For any $u \in \mathbf{SO}(n)$ and $\dot{u} \in T_u\mathbf{SO}(n)$ let

$$\mathrm{Orth}_u(\dot{u}) = u\,\mathrm{Orth}_{\mathrm{id}}(u^{-1}\dot{u}).$$

In addition a family of retractions on $\mathbf{SO}(n)$ is given by $u\,\mathrm{Cay}(u^{-1}\dot{u})$ where Cay denotes a Cayley transform (Iserles, 2001).

These and other retractions defined on Lie groups have been extensively used in the literature on numerical integration on manifolds (see Iserles, 2001; Iserles *et al.*, 2001; Celledoni & Iserles, 2001; Dieci & Van Vleck, 1999; Owren & Marthinsen, 1999; Munthe-Kaas *et al.*, 2001).

EXAMPLE 3 $\mathbf{M} = \mathbf{SO}(3)^N \subset \mathcal{M}(3,3)^N$.

In this example, which has already been introduced in the first section, $\mathbf{M}$ is a compact submanifold in $\mathcal{M}(3,3)^N$ and its dimension is $3N$. The tangent space at $m = (m_1, \dots, m_N) \in \mathbf{M}$ is given by

$$T_m\mathbf{M} = \prod_{i=1}^{N} T_{m_i}\mathbf{SO}(3)$$

so that $\dot m = (\dot m_1, \dots, \dot m_N) \in T_m\mathbf{M}$ if and only if $\dot m = (m_1a_1, \dots, m_Na_N)$ with $a_1, \dots, a_N \in \mathcal{A}(3,3)$. Using the exponential of a matrix we can define a retraction on the product structure:

$$R_m(\dot m) = (m_1 \exp({m_1}^{-1}\dot m_1), \dots, m_N \exp({m_N}^{-1}\dot m_N)).$$

Similarly, using the previously defined Orth and Q we can define additional retractions for $\mathbf{M} = \mathbf{SO}(3)^N$ which we again denote by Orth and by the Q.

We remark that exp, $\mathrm{Orth}_{\mathrm{id}}$ and Q_{id} are all easy to compute for three by three antisymmetric matrices, and therefore, so are R, Q and Orth for $\mathbf{M} = \mathbf{SO}(3)^N$.

4. What Newton's method does

In the classical case, when $\mathbf{M} = \mathbf{E}$ is a Euclidean space and when the derivative $DF(x)$ is an isomorphism, fixed points for Newton's map, $N_F(x) = x$, correspond to zeros for F: $F(x) = 0$. It is also the case when the system of equations $F(x) = 0$ is underdetermined and when $DF(x)$ is onto. Newton's method for underdetermined systems of equations is studied in Ben-Israel (1966); Allgower & Georg (1990); Shub & Smale (1996); Beyn (1993); Dedieu & Kim (2002). We will return to the underdetermined case shortly.

When the system $F(x) = 0$ is overdetermined and when the derivative $DF(x)$ is injective, these fixed points correspond to the the least-square solutions of the system $F(x) = 0$. By least-square solution we mean here a zero of the derivative of the residue function: $Df(x) = 0$ with

$$f(x) = \tfrac{1}{2}\|F(x)\|^2.$$

Newton's method, in this context, was introduced by Gauss (1809), to solve the nonlinear least-square problem. For this reason it is called the Newton–Gauss method. Convergence properties of the Newton–Gauss method are studied by Ben-Israel (1966); Dennis & Schnabel (1983); Dedieu & Shub (2000); Dedieu & Kim (2002).

When, instead of a Euclidean space $\mathbf{E}$ we consider a manifold $\mathbf{M}$, we have a similar characterization of the fixed points.

PROPOSITION 1 When the derivative $DF(x)$ is injective, fixed points for N_F correspond to least-square solutions for F: $N_F(x) = x$ if and only if $Df(x) = 0$ with $f(x) = \frac{1}{2}\|F(x)\|^2$. When $DF(x)$ is onto then $N_F(x) = x$ if and only if $F(x) = 0$.

Proof. Using $R_x(\dot x) = x$ if and only if $\dot x = 0$, a fixed point for N_F is given by $DF(x)^\dagger F(x) = 0$ or equivalently $DF(x)^* F(x) = 0$. On the other hand, for any $x \in \mathbf{M}$ and $\dot x \in T_x\mathbf{M}$ we have

$$Df(x)\dot x = \langle DF(x)\dot x, F(x)\rangle_{T_x\mathbf{M}}$$

so that, when $x \in \mathbf{M}$ and $\dot{x} \in T_x\mathbf{M}$

$$Df(x)\dot{x} = \langle \dot{x}, DF(x)^* F(x) \rangle_{T_x\mathbf{M}}.$$

This proves the correspondence between fixed points and least-square solutions. When $DF(x)$ is onto, its adjoint $DF(x)^*$ is injective so that $DF(x)^* F(x) = 0$ if and only if $F(x) = 0$: x is a zero for F. □

The next proposition is a technical result which will be used later. We recall that when the derivative of a real-valued function g on a manifold $\mathbf{M}$ is zero at a point x then the second derivative of the function $D^2 g(x)$ makes sense (see Milnor, 1963, Section 2). It is a symmetric bilinear map from $T_x\mathbf{M} \times T_x\mathbf{M}$ to $\mathbb{R}$. Thus, there is a symmetric linear map Hess $g(x)$, called the *Hessian* of g at x, of $T_x\mathbf{M}$ into $T_x\mathbf{M}$ defined by

$$D^2 g(x)(\dot{x}, \dot{y}) = \langle \dot{y}, \text{Hess } g(x)(\dot{x}) \rangle,$$

for $\dot{x}, \dot{y} \in T_x\mathbf{M}$. More generally, the fibre component of the derivative of a section σ of $T_x\mathbf{M}$ makes sense at a zero x of σ and maps $T_x\mathbf{M}$ to $T_x\mathbf{M}$. By abuse of notation we call this derivative D.

PROPOSITION 2 If $x \in \mathbf{M}$ is a least-square zero of F and if $DF(x)$ is injective then for any $\dot{x} \in T_x\mathbf{M}$

$$\text{Hess } f(x)\dot{x} = D(DF(x)^* F(x))\dot{x} + DF(x)^* DF(x)\dot{x}$$

and

$$\begin{aligned} DN_F(x)(\dot{x}) &= -(DF(x)^* DF(x))^{-1} D(DF(x)^* F(x))\dot{x} \\ &= \dot{x} - (DF(x)^* DF(x))^{-1} \text{Hess } f(x)\dot{x}. \end{aligned}$$

Proof. The first formula is obtained by straightforward differentiation starting from the formula given in the proof of Proposition 1:

$$Df(x) = DF(x)^* F(x).$$

Since $DF(x)$ is an injective linear operator its Moore–Penrose inverse is given by

$$DF(x)^\dagger = (DF(x)^* DF(x))^{-1} DF(x)^*.$$

The derivative of Newton's map is given by

$$DN_F(x)\dot{x} = \frac{\partial R(x,0)}{\partial x}\dot{x} + \frac{\partial R(x,0)}{\partial \dot{x}} D(-(DF(x)^* DF(x))^{-1} DF(x)^* F(x))\dot{x}.$$

We now use the equalities $Df(x) = DF(x)^* F(x) = 0$ to obtain

$$DN_F(x)\dot{x} = \frac{\partial R(x,0)}{\partial x}\dot{x} - \frac{\partial R(x,0)}{\partial \dot{x}}(\dot{x} + (DF(x)^* DF(x))^{-1} D(DF(x)^* F(x))\dot{x})$$

which gives the first expression for $DN_F(x)\dot{x}$. The second expression follows by substituting the value for Hess $f(u)\dot{n}$ given in the first part of the proposition. □

For a smooth function $N : \mathbf{M} \to \mathbf{M}$ and $x \in \mathbf{M}$, we say that x is an attractive fixed point for N when $N(x) = x$ and when the eigenvalues of the derivative $DN(x) : T_x\mathbf{M} \to T_x\mathbf{M}$ have modulus less than Proposition 1. In such a case the sequence $x_k = N^{(k)}(x_0)$ converges geometrically to x for any starting point x_0 taken in a neighborhood of x.

PROPOSITION 3 If $x \in \mathbf{M}$ is a least-square zero of F and if $DF(x)$ is injective then

1. The eigenvalues of $DN_F(x)$ are real,
2. When x is a local minimum for the residue function f then, these eigenvalues are less than or equal to 1,
3. When x is a local maximum for the residue function f then, these eigenvalues are greater than or equal to 1,
4. When x is an attractive fixed point for N_F then x is a strict local minimum for the residue function.

Proof. From Proposition 2 we get

$$\text{Hess } f(x)\dot{x} = D(DF(x)^* F(x))\dot{x} + DF(x)^* DF(x)\dot{x}$$

and

$$DN_F(x)(\dot{x}) = -(DF(x)^* DF(x))^{-1} D(DF(x)^* F(x))\dot{x}$$

or in other words Hess $f(x) = a + b$ and $DN_F(x) = -b^{-1}a$. We also notice that the operator $b = DF(x)^* DF(x)$ is positive definite and thus possesses a positive definite square root. Given two $n \times n$ matrices x and y it is a well known fact that the eigenvalues of xy and yx are the same. Thus, the eigenvalues of $DN_F(x) = -b^{-1}a$ and $-b^{-1/2}ab^{-1/2}$ are the same. We let Spec (x) denote the set of eigenvalues of x so that

$$\text{Spec}\,(b^{-1/2}\,\text{Hess } f(x)b^{-1/2}) = \text{Spec}\,(b^{-1/2}(a+b)b^{-1/2}) =$$

$$\text{Spec}\,(\text{id} - (-b^{-1/2}ab^{-1/2})) = 1 - \text{Spec } DN_F(x).$$

Since Hess $f(x)$ is real symmetric, the operator $b^{-1/2}$ Hess $f(x)b^{-1/2}$ is also real symmetric so that its eigenvalues are real. This proves the first assertion. Using the same inequality between the spectra, when x is a local minimum, Hess $f(x)$ is positive semidefinite, see Hestenes (1975, Chapter 3, Theorem 2.2). Thus $b^{-1/2}D^2 f_{\mathbf{M}}(x)b^{-1/2}$ is also positive semidefinite and consequently Spec $DN_F(x) \leqslant 1$. The third assertion is proved similarly. Now, if x is an attractive fixed point then $-1 <$ Spec $DN_F(x) < 1$ so that Hess $f(x)$ is positive definite and x is a strict local minimum, see Hestenes (1975, Chapter 3, Theorem 3.1). □

REMARK It is important to notice the following facts.

1. The fixed points for Newton-Gauss method correspond to the zeros of the derivative of the residue function.
2. Generically, if we are able to compute a fixed point x for N_F it is necessarily attractive: its eigenvalues have a modulus less than 1. In that case x is necessarily a strict local minimum for the residue function, not only a stationary point.
3. Local maxima are repelling points for Newton-Gauss.

4. When $\dim(\mathbf{M}) = \dim(\mathbf{V})$ and the DF is invertible, then the fixed points of N_F are indeed zeros of F and the convergence is quadratic.

REMARK In the Euclidean case, i.e. when $\mathbf{M} = \mathbf{E}$ is a Euclidean space, let x^* be a fixed point for the Gauss-Newton method. If the eigenvalues of $DN_F(x)$ are contained in the open interval $]-n, 1[$ and not all of them > -1 then, according to Remark 7, x^* is a hyperbolic non-attractive fixed point and thus cannot be computed via the Gauss-Newton algorithm. However, we may stabilize x^* by using the iterative method $P_n[N_F]$ instead, which is defined as follows. Let $P_n(x)$ be the polynomial obtained by multiplying out

$$\frac{x(x+1)\dots(x+n-1)}{n!}.$$

Then $P_n[N_F](x)$ is defined by replacing x^j for $j > 0$ by the j-fold composition of N_F with itself evaluated at x.

For example,

$$P_2[N_F](x) = (N_F(x) + N_F \circ N_F(x))/2.$$

To prove the stability of x^* it is sufficient to notice the following facts:

1. $P_n(1) = 1$ from which it can be seen that x^* is also a fixed point for $P_n[N_F]$;
2. $D(P_n[N_F])(x^*) = P_n(DN_F(x^*))$ and $\operatorname{Spec}(D(P_n[N_F])(x^*)) = P_n(\operatorname{Spec}(DN_F(x^*)))$;
3. $P_n(]-n, 1[) \subset]-1, 1[$.

Thus all the eigenvalues of $D(P_n[N_F])(x^*)$ are contained in the interval $]-1, 1[$ and x^* is an attractive fixed point for the map $P_n[N_F]$. Note that $P_n[N_F]$ is computed using n iterates of N_F. If in following a homotopy a contractive fixed point becomes unstable, it may be stablized using already computed data.

We now return to Newton's method in the underdetermined case. We have already seen in Proposition 1 that the zeros of F correspond to fixed points of N_F. If we assume that $D(F)$ is onto at the zeros of F then the implicit function theorem proves that the set of zeros, $\mathbf{N} = F^{-1}(0)$, is a submanifold of $\mathbf{M}$ of codimension equal to the dimension of $\mathbf{V}$. As the manifold $\mathbf{N}$ is pointwise fixed by N_F it follows that the derivative of N_F is the identity when restricted to the the tangent space of $\mathbf{N}$.

PROPOSITION 4 If $F : \mathbf{M} \to \mathbf{V}$ and $DF(x)$ is onto at any zero of F then

1. N_F is defined in a neighborhood of $\mathbf{N} = F^{-1}(0)$.
2. $DN_F|T\mathbf{N}$, the restriction of DN_F to $T\mathbf{N}$, is the identity map of $T\mathbf{N}$.
3. $DN_F|T\mathbf{N}^\perp$ is identically zero, where $T\mathbf{N}^\perp$is the orthogonal complement of $T\mathbf{N}$ in $T\mathbf{M}$.
4. There is a neighborhood $\mathbf{U}$ of $\mathbf{N}$ in $\mathbf{M}$ such that the kth iterate of N_F applied to x, $N_F^k(x)$, is defined for all $x \in \mathbf{U}$ and all positive integers k. Moreover, $N_F^k(x)$ converges to a point in $\mathbf{N}$. The set of points in $\mathbf{U}$ which converge to $v \in \mathbf{N}$ under the iteration N_F^k forms a disc tangent to $T_v\mathbf{N}^\perp$ at v. This disc has dimension equal to the dimension of $\mathbf{V}$, varies smoothly with v and is mapped into itself by N_F^k.

5. For $x \in \mathbf{U}$ let $q_{\mathbf{N}}(x) = \lim N_F^k(x)$. Then the mapping $q_{\mathbf{N}} : \mathbf{U} \to \mathbf{N}$ defined by $x \to q_{\mathbf{N}}(x)$ is smooth. The derivative $Dq_{\mathbf{N}}(x)$ is the identity on $T\mathbf{N}$ and zero on $T\mathbf{N}^{\perp}$.

Proof. The proof of this proposition in the case that M is a Euclidean space is contained in Beyn (1993) and the references there. We indicate a proof here item by item.

1. Since DF is onto $\mathbf{N}$, it is onto a neighborhood of $\mathbf{N}$, so N_F is defined on this neighborhood.
2. We have already seen that DF is the identity on $T\mathbf{N}$.
3. Since $DF(x)$ is onto its Moore–Penrose inverse is given by

$$DF(x)^{\dagger} = DF(x)^*(DF(x)DF(x)^*)^{-1}.$$

   The derivative of Newton's map is given by

$$\begin{aligned} DN_F(x)\dot{x} =& \frac{\partial R(x,0)}{\partial x}\dot{x} + \frac{\partial R(x,0)}{\partial \dot{x}} D(-DF(x)^* \\ &\times (DF(x)^*DF(x))^{-1})F(x))\dot{x} + DF(x)^{\dagger}DF(x)\dot{x}. \end{aligned}$$

   We now use the equality $F(x) = 0$ and the fact that $DF(x)^{\dagger}DF(x)$ is orthogonal projection on $T_x\mathbf{N}^{\perp}$ to see that $DN_F(v)$ is orthogonal projection on $T_v\mathbf{N}$ for any $v \in \mathbf{N}$. Now the rest of the assertions follow from Beyn (1993) or stable manifold theory and graph transform techniques for the derivative statements as in Hirsch *et al.* (1977) see also Shub & Smale (1996).
4. The smoothness of $q_{\mathbf{N}}$ follows from the previous item.

□

EXAMPLE 4 **A constrained retraction**
Proposition 4 allows us to define a retraction for the submanifold $\mathbf{N}$ of $\mathbf{M}$ in terms of one for $\mathbf{M}$ when $\mathbf{N} = F^{-1}(0)$ as in proposition 4. Recall that $R_{\mathbf{M}}$ maps a neighborhood of the zero section in $T\mathbf{M}$ to $\mathbf{M}$. $R_{\mathbf{M}}$ does not necessarily map a neighborhood of $T\mathbf{N}$ to $\mathbf{N}$. In order to return to the constraint manifold $\mathbf{N}$ we employ the map $q_{\mathbf{N}}$ and define $R_{\mathbf{N}} = q_{\mathbf{N}}R_{\mathbf{M}}$.

It follows from Proposition 4 that $R_{\mathbf{N}}$ is a retraction. We may apply our results with Newton's method defined by $R_{\mathbf{N}}$ and conclude the quadratic convergence of Newton's method to find zeros of constrained problems. In practical terms $q_{\mathbf{N}}(x)$ may not be able to be computed because it involves an infinite limit. It is replaced by a finite iterate $N_F^k(x)$. This is the approach we take to solve the constrained minimization problem of Sections 2 and 3. It is common practice in optimization theory to first solve the tangential problem and then to satisfy the constraints.

5. Newton's method for vector fields

By a vector field on manifold $\mathbf{M}$ we mean a smooth section $X : \mathbf{M} \to T\mathbf{M}$ i.e. a smooth map X which assigns to each $x \in \mathbf{M}$ a tangent vector $X(x) \in T_x\mathbf{M}$. We are interested in using Newton's method to find zeros of X, i.e. points $x \in \mathbf{M}$ such that $X(x) = 0_x$ the zero vector in $T_x\mathbf{M}$.

We shall be concerned with $X = \text{grad}\,(\phi)$ where $\phi : \mathbf{M} \to \mathbb{R}$ is a smooth real-valued function. So the zeros of X are the critical points of ϕ.

In order to define Newton's method for vector fields, we resort to an object studied in differential geometry: namely, the covariant derivative of vector fields. The covariant derivative of a vector field X defines a linear map $\nabla X(x) : T_x\mathbf{M} \to T_x\mathbf{M}$ for any $x \in \mathbf{M}$.

REMARK Note that if $X = \text{grad}\ \phi$ for ϕ a real-valued function on $\mathbf{M}$ and $X(x) = 0$, then $\nabla X(x) = \text{Hess}\ \phi(x)$ the Hessian of ϕ at x.

For submanifolds $\mathbf{M}$ of a Euclidean space $\mathbf{E}$, $\mathbf{M} \subset \mathbf{E}$, an induced covariant derivative is easy to define. Recall that $T\mathbf{M} \subset \mathbf{M} \times \mathbf{E}$. Thus we may consider $X : \mathbf{M} \to \mathbf{M} \times \mathbf{E}$. Let $\pi_2 : \mathbf{M} \times \mathbf{E} \to \mathbf{E}$ be the projection on the second factor and $\pi_{T_x\mathbf{M}} : \mathbf{E} \to T_x\mathbf{M}$ be the orthogonal projection from $\mathbf{E}$ to $T_x\mathbf{M}$. Then for $x \in \mathbf{M}$,

$$\nabla X(x) = \pi_{T_x\mathbf{M}} D(\pi_2 X)(x)$$

defines the covariant derivative of X. When dealing with manifolds which are naturally submanifolds of Euclidean spaces we will mean the induced covariant derivative unless otherwise specified.

DEFINITION 2 We define the Newton iteration for the vector field X by

$$N_X(x) = R_x(-\nabla X(x)^{-1} X(x))$$

as long as $\nabla X(x)$ is invertible and $-\nabla X(x)^{-1} X(x)$ is contained is the domain of R_x.

When $\mathbf{M}$ is a submanifold of Euclidean space and the connection the induced connection

$$\begin{aligned} N_X(x) &= R(x, -\nabla X(x)^{-1}(\pi_2 X)(x)) \\ &= R(x, -\nabla X(x)^{-1}\pi_{T_x\mathbf{M}}(\pi_2 X)(x)) \end{aligned}$$

as long as $\nabla X(x)$ is invertible and $-\nabla X(x)^{-1}(\pi_2 X)(x)$ is contained is the domain of R_x.

Newton's method has the usual property of quadratic convergence for simple zeros of vector fields (Shub, 1986).

PROPOSITION 5 If $x \in \mathbf{M}$ is a fixed point for $N_X(x)$ then $X(x) = 0_x$ and $DN_X(x) = 0$.

Proof. For $N_X(x)$ to be defined $\nabla X(x)$ must be invertible. Since $R_x(\dot{x}) = x$ if and only if $\dot{x} = 0_x$ we have $X(x) = 0_x$. In computing the derivative of the expression in above definition for $N_X(x)$ using the chain and product rules, several terms vanish since they are evaluated at zero. A short computation shows that

$$DN_X(x) = \text{id}_{T_x\mathbf{M}} - \nabla X(x)^{-1} \nabla X(x).$$

□

EXAMPLE 5 The Rayleigh Quotient Iteration, introduced by Lord Rayleigh a century ago (Lord Rayleigh, 1899), for the eigenvalue problem may now be seen in this context, as in

Shub (1986). Let a be an $n \times n$ real matrix. We consider the vector field X defined on the unit sphere $\mathbf{S}^{n-1}$

$$X(x) = ax - \langle ax, x\rangle x.$$

Then the zeros of X are unit eigenvectors of a. Let us denote $\rho(x) = \langle ax, x\rangle$ so that $X(x) = ax - \rho(x)x$. In R^n we have

$$DX(x)v = av - \langle av, x\rangle x - \langle ax, v\rangle x - \rho(x)v.$$

To compute $\nabla X(x)v$ project $DX(x)v$ onto the tangent space of $\mathbf{S}^{n-1}$ at x: i.e. substract the projection of $DX(x)v$ on x from $DX(x)v$:

$$\langle DX(x)v, x\rangle = \langle av, x\rangle - \langle av, x\rangle\langle x, x\rangle - \langle ax, v\rangle\langle x, x\rangle - \rho(x)\langle x, v\rangle.$$

Since $\langle x, x\rangle = 1$ and $\langle x, v\rangle = 0$ for v tangent to $\mathbf{S}^{n-1}$ at x we obtain

$$\langle DX(x)v, x\rangle = -\langle ax, v\rangle$$

and

$$\nabla X(x)v = av - \langle av, x\rangle x - \rho(x)v.$$

Now solving

$$\nabla X(x)w = X(x)$$

gives

$$aw - \langle aw, x\rangle x - \rho(x)w = ax - \rho(x)x$$

or

$$(a - \rho(x)\mathrm{id})w = (a - \rho(x)\mathrm{id})x + \langle aw, x\rangle x.$$

Thus, if we can solve for w and $a - \rho(x)\mathrm{id}$ is invertible,

$$w = x + (a - \rho(x)\mathrm{id})^{-1}\langle aw, x\rangle x.$$

Newton's iteration is then

$$N_X(x) = \frac{x - w}{\|x - w\|} = \frac{-(a - \rho(x)\mathrm{id})^{-1}\langle aw, x\rangle x}{\|(a - \rho(x)\mathrm{id})^{-1}\langle aw, x\rangle x\|} = \pm\frac{-(a - \rho(x)\mathrm{id})^{-1}x}{\|(a - \rho(x)\mathrm{id})^{-1}x\|}$$

which is precisely the classical Rayleigh Quotient Iteration.

6. Formulae for a Newton method for the spinal optimization problem

NOTATION The space $\mathbf{SO}(3)$ of orthogonal matrices with $det = 1$ is naturally contained in the nine-dimensional Euclidean space $\mathcal{M}(3, 3)$ of all 3×3 real matrices with inner product structure defined by $\langle m, n\rangle = \mathrm{T}race(n^*m)$ where $m, n \in \mathcal{M}(3, 3)$ and n^* denotes the transpose of n. If $m = (m_{i,j})$ and $n = (n_{i,j})$, then $\langle a, b\rangle = \sum_{i,j} m_{i,j}n_{i,j}$. Thus $\mathcal{M}(3, 3)$, $\langle , \rangle$ may be identified with $\mathbb{R}^9$ and the standard inner product, the entries of the matrices corresponding to the coordinates of the vectors in $\mathbb{R}^9$. Similarly, $\mathbf{SO}(3)^N$ is contained in the $9N$-dimensional space $\mathcal{M}(3, 3)^N$ with the 'product' inner product,

$\langle m, n\rangle = \sum_{i=1}^{N}\langle m_i, n_i\rangle$ where $m = (m_1, \dots, m_N)$ and $n = (n_1, \dots, n_N)$. From the inner product we get a norm $\|m\| = \sqrt{\langle m, m\rangle}$. We shall be concerned with the following maps:

$$
\begin{aligned}
&A : \mathcal{M}(3,3)^N \to \mathcal{M}(3,3)^{N+1},\\
&\Phi : \mathcal{M}(3,3)^N \to \mathcal{M}(3,3)^{N+1},\\
&H : \mathcal{M}(3,3)^N \to \mathbb{R}^2,\\
&h : \mathcal{M}(3,3)^N \to \mathbb{R}^2,
\end{aligned}
$$

where A is linear and injective, Φ affine given by $\Phi(m) = A(m) + c$, $c \in \mathcal{M}(3,3)^{N+1}$, H linear and onto, h affine given by $h(m) = H(m) - d,\ d \in \mathbb{R}^2$.

In our particular application, Φ is the affine map of (2.5) extended to all 3×3 matrices, and c is its linear part. H is the affine map also extended whose components are given by (2.4a) and (2.4b) and d is the vector whose components are δ_x, δ_y.

Let $\mathbf{M} = \mathbf{SO}(3)^N \subset \mathcal{M}(3,3)^N$ and $\mathbf{N} = \mathbf{SO}(3)^N \cap h^{-1}(0)$. By Sard's theorem almost every $d \in \mathbb{R}^2$ is a regular value of $H|\mathbf{M}$. For such d, $\mathbf{N}$ is a smooth sub-manifold of $\mathbf{M}$. Thus we shall assume throughout the remainder of this section that d is a regular value and $\mathbf{N}$ is nonempty. Thus Proposition 4 provides the map $q_{\mathbf{N}}$. Since $\mathbf{M}$ itself is a subspace of the Euclidean space $\mathcal{M}(3,3)^N$, we have induced connections on $\mathbf{M}$ and $\mathbf{N}$, $\nabla_{\mathbf{M}}$ and $\nabla_{\mathbf{N}}$ respectively. We let $R_{\mathbf{M}}$ be one of the retractions defined earlier on $\mathbf{SO}(3)^N$. Thus, as in Example 4 we have a retraction defined on $\mathbf{N}$, $R_{\mathbf{N}} = q_{\mathbf{N}} R_{\mathbf{M}}|T\mathbf{N}$. Finally, let

$$
\begin{aligned}
\phi &= \tfrac{1}{2}\|\Phi\|^2,\\
E &= \operatorname{grad} \phi,\\
X &= \operatorname{grad} (\phi|\mathbf{N}).
\end{aligned}
$$

Our optimization problem is to minimize $\phi|\mathbf{N}$. Our main tool is to apply Newton's method to find the zeros of X. Recall that Newton's method to find the zero's of X is given by $N_X : \mathbf{N} \to \mathbf{N}$ where

$$
\begin{aligned}
N_X(m) &= R_{\mathbf{N},m}(-\nabla_{\mathbf{N}} X(m)^{-1} X(m))\\
&= q_{\mathbf{N}} R_{\mathbf{M}}(-\nabla_{\mathbf{N}} X(m)^{-1} X(m))
\end{aligned}
$$

and $q_{\mathbf{N}}$ is defined by the Newton iteration

$$N_h(m) = R_{\mathbf{M},m}(-Dh(m)^{\dagger} h(m)).$$

Let $N_{\mathbf{M}} : \mathbf{N} \to \mathbf{M}$ be defined by

$$N_{\mathbf{M}} = R_{\mathbf{M}}(-\nabla_{\mathbf{N}} X(m)^{-1} X(m))$$

so that we have

$$N_X = q_{\mathbf{N}} N_{\mathbf{M}}.$$

It is the goal of this section to compute Matlab convenient formulae for $N_{\mathbf{M}}$ and N_h. As explained in Example 4, our numerical results are then achieved by approximating N_X by several iterates of N_h applied to $N_X(m)$.

We define the operations left-multiplication $\mathfrak{L}_m$ and right-multiplication $\mathfrak{R}_m$ on $\mathcal{M}(3,3)^N$, by

$$\mathfrak{L}_m(a_1, \ldots, a_N) = (m_1 a_1, \ldots, m_N a_N),$$
$$\mathfrak{R}_m(a_1, \ldots, a_N) = (a_1 m_1, \ldots, a_N m_N).$$

Let l_m denote the restriction of $\mathfrak{L}_m$ to $\mathcal{A}(3,3)^N$.

Let $\mathcal{S}(3,3)$ denote the 3×3 symmetric matrices and recall that $\mathcal{A}(3,3)$ denotes the 3×3 antisymmetric ones. We define the maps sym : $\mathcal{M}(3,3)^N \to \mathcal{S}(3,3)^N$ and asym : $\mathcal{M}(3,3)^N \to \mathcal{A}(3,3)^N$ by

$$\mathrm{sym}(m) = (m + m^*)/2,$$
$$\mathrm{asym}(m) = (m - m^*)/2,$$

where $m^* = (m_1^*, \ldots, m_N^*)$.

Also let

$$P_m = H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*.$$

We shall need a group invariance property of the retraction R, which for the remainder of this section we assume to be any one of the retractions R, Q or Orth of Example 2. This property is stated in the following easily verifiable proposition.

PROPOSITION 6 Let R be any one of the three retractions R, Q or Orth of Example 2. For any $m, n \in \mathbf{M}$ and $u \in T_n\mathbf{M}$,

$$\mathfrak{L}_m R_n(u) = R_{\mathfrak{L}_m(n)}(\mathfrak{L}_m(u)).$$

THEOREM 1 A formula for the iteration N_h on $\mathbf{M}$ is given by $N_h(m) = \mathfrak{L}_m R_{\mathrm{id}}(\dot{a})$ where $\dot{a} \in T_{\mathrm{id}}\mathbf{M} = \mathcal{A}(3,3)^N$ satisfies

$$\dot{a} = -l_m^* H^*(H l_m l_m^* H^*)^{-1} h(m).$$

THEOREM 2 A formula for $N_{\mathbf{M}}(m)$ for the gradient vector field $X(m) = \mathrm{grad}\ \phi|\mathbf{N}$ on $\mathbf{N}$ is given by

$$N_{\mathbf{M}}(m) = \mathfrak{L}_m R_{\mathrm{id}}(\dot{a})$$

where $\dot{a} \in T_{\mathrm{id}}\mathbf{M} = \mathcal{A}(3,3)^N$ satisfies

$$\begin{aligned} &- l_m^*(I - P_m)A^*\Phi(m) \\ &= (I - l_m^* P_m l_m)\left(l_m^* A^* A l_m - \frac{\mathfrak{L}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))} + \mathfrak{R}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))}}{2}\right)\dot{a} \end{aligned}$$

provided that $\nabla_{\mathbf{N}} X(m)$ is invertible at m.

We accomplish the proofs of these theorems by a series of lemmas.

If V_1 is a vector subspace of the V with an inner product we let π_{V_1} denote orthogonal projection of V on V_1.

In the next lemma we put together some of the components which follow more or less by definition and are required for the computation.

LEMMA 1

1. For $m \in \mathcal{M}(3,3)^N$, $E(m) = A^*\Phi(m)$.
2. For $m \in \mathbf{N}$, $X(m) = \pi_{T_m\mathbf{N}} E(m)$.
3. For $m \in \mathbf{N}$, $\nabla_{\mathbf{N}} X(m) = \pi_{T_m\mathbf{N}} DX(m)$.
4. For $m \in \mathbf{N}$, $T_m\mathbf{N} = \mathrm{Ker}\, H \cap T_m\mathbf{M}$.
5. For $m \in \mathbf{M}$, $Dh(m) = H|T_m\mathbf{M}$.

Proof. The first item is the only one which requires any proof at all,

$$\phi(m) = \frac{1}{2}\|\Phi(m)\|^2 = \frac{1}{2}\langle A(m) + c, A(M) + c\rangle$$

and so

$$E(m) = \mathrm{grad}\ \phi(m) = A^*(A(m) + c) = A^*\Phi(m).$$

□

Now we identify some operators and projections associated with the inclusion $\mathbf{M} = \mathbf{SO}(3)^N \subset \mathcal{M}(3,3)^N$.

LEMMA 2

1. asym : $\mathcal{M}(3,3)^N \to \mathcal{A}(3,3)^N$ is the orthogonal projection. Hence, asym $= i^*$, where $i : \mathcal{A}(3,3)^N \to \mathcal{M}(3,3)^N$ is the inclusion.
2. For $m \in \mathcal{M}(3,3)^N$, $\mathfrak{L}_m^* = \mathfrak{L}_{m^*}$.
3. For $m \in \mathcal{M}(3,3)^N$, $l_m^* = \mathrm{asym} \circ (\mathfrak{L}_{m^*})$.
4. For $m \in \mathbf{M}$,
   (a) $\mathfrak{L}_m$ is an orthogonal transformation from $\mathcal{M}(3,3)^N$ to itself;
   (b) $l_m^* l_m = \mathrm{id}_{\mathcal{A}(3,3)^N}$;
   (c) $l_m l_m^*$ is orthogonal projection from $\mathcal{M}(3,3)^N$ to $T_m\mathbf{M}$.

Proof. We only deal with the case $N = 1$, the general case is the same.

1. $\mathcal{A}(3,3)$ and $\mathcal{S}(3,3)$ are orthogonal: For if a is antisymmetric and b is symmetric, then

   $$\mathrm{T}race(b^*a) = \mathrm{T}race(a^*b) = -\mathrm{T}race(ab) = -\mathrm{T}race(ba) = -\mathrm{T}race(b^*a).$$

   Now asym is the identity on $\mathcal{A}(3,3)$ and zero on $\mathcal{S}(3,3)$ which are orthogonal and whose direct sum is all of $\mathcal{M}(3,3)$. So we are done.
2. $\mathrm{T}race((mb)^*a) = \mathrm{T}race(b^*(m^*a))$. So $\langle a, \mathfrak{L}_m b\rangle = \langle \mathfrak{L}_{m^*}a, b\rangle$ and $\mathfrak{L}_{m^*} = \mathfrak{L}_m^*$.
3. $l_m = (\mathfrak{L}_m \circ i)$, so $l_m^* = i^* \circ \mathfrak{L}_m^* = \mathrm{asym} \circ (\mathfrak{L}_{m^*})$.
4. For $m \in \mathbf{M}$,
   (a) By 2), $\mathfrak{L}_m^* = \mathfrak{L}_{m^*}$ so $\mathfrak{L}_m^*\mathfrak{L}_m = \mathfrak{L}_{m^*m} = \mathrm{id}_{\mathcal{M}(3,3)^N}$ and $\mathfrak{L}_m$ is orthogonal.
   (b) $l_m^* l_m = (\mathfrak{L}_m \circ i)^*(\mathfrak{L}_m \circ i) = \mathrm{asym} \circ \mathfrak{L}_{m^*}\mathfrak{L}_m \circ i = \mathrm{asym} \circ \mathfrak{L}_{m^*m} \circ i = \mathrm{asym} \circ i = \mathrm{id}_{\mathcal{A}(3,3)}$.
   (c) Similarly, $l_m l_m^* = \mathfrak{L}_m\, \mathrm{asym}\, \mathfrak{L}_{m^*}$ which by parts (1), (2) and (4)(a) is orthogonal projection on its image, $T_m\mathbf{M}$.

□

We may now prove theorem 1.

Proof.

$$N_h(m) = R_{\mathbf{M},m}(-Dh(m)^{\dagger}h(m)) = \mathfrak{L}_m R_{\mathrm{id}}(-l_m^* Dh(m)^{\dagger}h(m)).$$

Now

$$Dh(m)^{\dagger} = (H|T_m\mathbf{M})^{\dagger} = l_m l_m^* H^*(H l_m l_m^* H^*)^{-1}.$$

Substituting the last expression for $Dh(m)^{\dagger}$ in the equation above using that $l_m^* l_m$ is the identity on $\mathcal{A}(3,3)^N$ we are done. $\square$

In the next two lemmas we compute the orthogonal projection $\pi_{T_m\mathbf{N}}$ from $\mathcal{M}(3,3)^N$ to $T_m\mathbf{N}$ for $m \in \mathbf{N}$.

LEMMA 3 Let $H : E \to V$ be a surjective linear map of finite-dimensional vector spaces with inner product and $W \subset E$ a subspace.

1. $\pi_{\mathrm{Ker}(H)} = I - H^*(HH^*)^{-1}H$
2. $\pi_W \pi_{\mathrm{Ker}(H\pi_W)} = \pi_{\mathrm{Ker}(H)\cap W}$.

Proof. 1. If $v \in \mathrm{Ker}(H)$ then $(I - H^*(HH^*)^{-1}H)(v) = v$. If $v = H^*(w) \in \mathrm{Image}(H^*)$ then $(I - H^*(HH^*)^{-1}H)(v) = (I - H^*(HH^*)^{-1}H)H^*(w) = 0$.
2. $\mathrm{Ker}(H\pi_W) = {\pi_W}^{-1}(\mathrm{Ker}(H)) = \mathrm{Ker}(H) \cap W \oplus W^{\perp}$ where the direct sum is orthogonal direct sum. The item follows.

$\square$

Recall that $P_m = H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*$.

LEMMA 4 For $m \in \mathbf{N}$,

$$\begin{aligned}\pi_{T_m\mathbf{N}} &= \pi_{T_m\mathbf{M}}\pi_{\mathrm{Ker}(H\pi_{T_m\mathbf{M}})}\\ &= l_m l_m^*(I - H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*)\\ &= \pi_{T_m\mathbf{M}}(I - P_m).\end{aligned}$$

Proof. The first equality follows from item (4) of Lemmas 1 and 3. The second equality follows from substituting $l_m l_m^*$ for $\pi_{T_m\mathbf{M}}$ according to Lemma 2 item (4)(c) and applying item (1) of Lemma 3 to $H l_m l_m^*$. The third equality follows from the definition of P_m. $\square$

It is now a simple matter to identify $X = \mathrm{grad}\ \phi|\mathbf{N}$ defined on $\mathbf{N}$.

LEMMA 5 Let $X = \mathrm{grad}\ \phi_{\mathbf{N}}$ and $m \in \mathbf{N}$ then

$$\begin{aligned}X(m) &= \pi_{T_m\mathbf{N}}E(m)\\ &= \pi_{T_m\mathbf{M}}(I - P_m)E(m)\\ &= l_m l_m^*(I - H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*)A^*\Phi(m).\end{aligned}$$

Proof. $X(m)$ is the projection of $E(m)$ into $T_m\mathbf{N}$ as in item (2) of Lemma 1. Now Lemma 4 and item (1) of Lemma 1 finish the proof. $\square$

We now prove some lemmas which help in the computation of $\nabla_{\mathbf{N}} X$. First we remark on some properties of P_m.

LEMMA 6 For $m \in \mathbf{M}$

1. $P_m \pi_{T_m \mathbf{M}} = P_m$
2. $(I - P_m) H^* = 0$
3. $0 = \pi_{T_m \mathbf{M}}(I - P_m) P_m = \pi_{T_m \mathbf{N}} P_m$
4. $\pi_{T_m \mathbf{N}} \pi_{T_m \mathbf{M}} = \pi_{T_m \mathbf{N}}$
5. $\pi_{T_m \mathbf{N}} \pi_{T_m \mathbf{M}} D(P_m) = 0$ where $D(P_m)$ is the derivative of P_m.

Proof.

1. $l_m l_m^* l_m l_m^* = l_m l_m^*$, now use the fact that $\pi_{T_m \mathbf{M}} = l_m l_m^*$ and the definition of P_m.
2. $$\begin{aligned} (I - P_m) H^* &= (I - H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*) H^* \\ &= H^* - H^* \\ &= 0. \end{aligned}$$
3. $(I - P_m) P_m = (I - P_m) H^* (H l_m l_m^* H^*)^{-1} H l_m l_m^*$ which is already 0 by the previous item. The second equality follows from Lemma 4.
4. $$\begin{aligned} \pi_{T_m \mathbf{N}} \pi_{T_m \mathbf{M}} &= \pi_{T_m \mathbf{M}}(I - P_m) \pi_{T_m \mathbf{M}} \\ &= \pi_{T_m \mathbf{M}}(I - P_m) \\ &= \pi_{T_m \mathbf{N}}. \end{aligned}$$
   The first equality follows from Lemma 4, the second from item (1) of this lemma and the fact that $\pi_{T_m \mathbf{M}}$ is a projection and the third equality from Lemma 4 again.
5. $$\begin{aligned} \pi_{T_m \mathbf{N}} \pi_{T_m \mathbf{M}} D(P_m) &= \pi_{T_m \mathbf{N}} D(P_m) \\ &= \pi_{T_m \mathbf{M}}(I - P_m) D(P_m) \\ &= \pi_{T_m \mathbf{M}}(I - P_m) H^* D((H l_m l_m^* H^*)^{-1} H l_m l_m^*) \\ &= 0. \end{aligned}$$
   The first equality follows from the item above, the second from Lemma 4, the third from the definition of P_m and the last from item (2) of this lemma.

□

We are ready for an intermediate result on $\nabla_{\mathbf{N}} X$.

PROPOSITION 7 For $m \in \mathbf{N}$,

$$\nabla_{\mathbf{N}} X(m)(\dot{m}) = \pi_{T_m \mathbf{N}}(D(\pi_{T_m \mathbf{M}})(m)(\dot{m})(I - P_m) E(m) + D(E)(m)\dot{m})$$

Proof. From Lemma 5, the definition of $\nabla_{\mathbf{N}}$ and Lemma 6, item (5).

$$\begin{aligned}\nabla_{\mathbf{N}} X(m) &= \pi_{T_m\mathbf{N}} D(\pi_{T_m\mathbf{N}} E(m))\\ &= \pi_{T_m\mathbf{N}} D(\pi_{T_m\mathbf{M}}(I - P_m)E(m))\\ &= \pi_{T_m\mathbf{N}}(D(\pi_{T_m\mathbf{M}})(m)(\dot m)(I - P_m)E(m) + \pi_{T_m\mathbf{M}} D(-P_m)(m)(\dot m)E(m)\\ &\quad + \pi_{T_m\mathbf{M}}(I - P_m)D(E)(m)(\dot m))\\ &= \pi_{T_m\mathbf{N}}(D(\pi_{T_m\mathbf{M}})(m)(\dot m)(I - P_m)E(m) + \pi_{T_m\mathbf{M}}(I - P_m)D(E)(m)(\dot m)).\end{aligned}$$

□

We now prove two lemmas which will help put $D(\pi_{T_m\mathbf{M}})(m)(\dot m)(I - P_m)E(m)$ in more convenient form.

LEMMA 7 $D(\pi_{T_m\mathbf{M}})(m)(\dot m) = (l_{\dot m} l_m^* + l_m l_{\dot m}^*)$.

Proof. Differentiate $\pi_{T_m\mathbf{M}} = l_m l_m^*$. □

LEMMA 8 For $m \in \mathbf{M}$, $\dot a \in \mathcal{A}(3,3)^N$, $\dot m = m\dot a$ and $n \in \mathcal{M}(3,3)^N$,

1. $(l_{\dot m} l_m^* + l_m l_{\dot m}^*)(n) = -m\left(\frac{\dot a n^* m + n^* m \dot a}{2}\right)$
2. $l_m^*(l_{\dot m} l_m^* + l_m l_{\dot m}^*)(n) = -\left(\frac{\mathfrak{R}_{\mathrm{sym}(m^*n))} + \mathfrak{L}_{\mathrm{sym}(m^*n)}}{2}\right)\dot a$

Proof. Recall from Lemma 2 that For $m \in \mathcal{M}(3,3)^N$, $l_m^* = \mathrm{asym} \circ (\mathfrak{L}_{m^*})$. Thus,

1.
$$\begin{aligned}(l_{\dot m} l_m^* + l_m l_{\dot m}^*)(n) &= m\dot a \frac{(m^* n - n * m)}{2} + m\left(\frac{\dot a^* m^* n - n^* m \dot a}{2}\right)\\ &= -m\left(\frac{\dot a n^* m + n^* m \dot a}{2}\right).\end{aligned}$$
using that $\dot a^* = -\dot a$ since $\dot a \in \mathcal{A}(3,3)^N$.
2. From item 1 it follows that
$$\begin{aligned}l_m^*(l_{\dot m} l_m^* + l_m l_{\dot m}^*)(n) &= -\,\mathrm{asym}\left(\frac{\dot a n^* m + n^* m \dot a}{2}\right)\\ &= -\left(\frac{\mathfrak{R}_{\mathrm{sym}(m^*n)} + \mathfrak{L}_{\mathrm{sym}(m^*n)}}{2}\right)\dot a.\end{aligned}$$

□

We now restate Proposition 7 with the help of the previous lemmas.

PROPOSITION 8 For $m \in \mathbf{N}$, and $\dot m = m\dot a = l_m \dot a$,

$$\begin{aligned}&l_m^* \nabla_{\mathbf{N}} X(m)(\dot m)\\ &= (I - l_m^* P_m l_m)\left(l_m^* A^* A l_m - \frac{\mathfrak{L}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))} + \mathfrak{R}_{\mathrm{sym}(m^*(I-P_m l_m l_m^*)A^*\Phi(m))}}{2}\right)(\dot a).\end{aligned}$$

Proof.

$$l_m^* \nabla_{\mathbf{N}} X(m)(\dot{m}) = l_m^* \pi_{T_m\mathbf{N}}(D(\pi_{T_m\mathbf{M}})(m)(\dot{m})(I - P_m)E(m) + D(E)(m)\dot{m}).$$

Now

$$\begin{aligned} l_m^* \pi_{T_m\mathbf{N}} &= l_m^* l_m l_m^* (I - H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*) \\ &\quad \text{by lemma 4} \\ &= l_m^* (I - H^*(H l_m l_m^* H^*)^{-1} H l_m l_m^*) \\ &= (I - l_m^* H^*(H l_m l_m^* H^*)^{-1} H l_m) l_m^* \end{aligned}$$

and

$$\begin{aligned} &l_m^*(D(\pi_{T_m\mathbf{M}})(m)(\dot{m})(I - P_m)E(m) + D(E)(m)\dot{m}) \\ &= \left(l_m^* A^* A l_m - \frac{\mathfrak{L}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))} + \mathfrak{R}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))}}{2} \right)(\dot{a}) \end{aligned}$$

using the definition of P_m, that $E(m) = A^*F$, $D(E) = A^*A$, $\dot{m} = l_m\dot{a}$ and the two lemmas. □

We are now ready to prove Theorem 2.

Proof.

$$\begin{aligned} N_{\mathbf{M}}(m) &= R_{\mathbf{M}}(-\nabla_{\mathbf{N}} X(m)^{-1} X(m)) \\ &= \mathfrak{L}_m R_{\mathrm{id}}(-l_m^* \nabla_{\mathbf{N}} X(m)^{-1} X(m)) \\ &= \mathfrak{L}_m R_{\mathrm{id}}(\dot{a}) \end{aligned}$$

where

$$-l_m^* X(m) = l_m^* \nabla_{\mathbf{N}} X(m) l_m(\dot{a}).$$

Substituting the expression for $X(m)$ from Lemma 5 in the left-hand side of this equation and using that $l_m^* l_m$ is the identity gives the left-hand side of the second equation in the theorem. So it remains to identify the right-hand sides. Proposition 8 finishes the proof. □

The linear equations of Theorems 1 and 2 are solved in $\mathcal{A}(3,3)^N$ which is the Lie algebra of the group $\mathbf{SO}(3)^N$. Whereas $\mathcal{A}(3,3)^N$ is a $3N$-dimensional space we have been considering it as a subspace of the $9N$-dimensional space $\mathcal{M}(3,3)^N$. To take advantage of the lower dimension we recompute the N_h and $N_{\mathbf{M}}$ with respect to a natural three-dimensional representation of $\mathcal{A}(3,3)$.

Define Ant : $\mathbb{R}^3 \to \mathcal{A}(3,3)$ by

$$\mathrm{Ant}\,(\alpha, \beta, \gamma) = \begin{pmatrix} 0 & -\alpha & -\beta \\ \alpha & 0 & -\gamma \\ \beta & \gamma & 0 \end{pmatrix}.$$

We also use Ant to represent the map Ant : $(\mathbb{R}^3)^N \to \mathcal{A}(3,3)^N$ which is Ant for each $\mathbb{R}^3$. So Ant is an isomorphism. We also let $Ł_m = l_m \mathrm{Ant}$. If we put the usual inner product on $(\mathbb{R}^3)^N$, we have the following lemma.

LEMMA 9

$$\mathrm{Ant}^* = 2\mathrm{Ant}^{-1},$$

and

$$L_m L_m^* = 2 l_m l_m^*.$$

Proof. The proof follows easily from the observation that

$$\mathrm{T}race(\mathrm{Ant}(\alpha_2, \beta_2, \gamma_2)^* \mathrm{Ant}(\alpha_1, \beta_1, \gamma_1)) = 2(\alpha_1\alpha_2 + \beta_1\beta_2 + \alpha_3\beta_3).$$

□

Now we restate Theorems 1 an 2 in terms of $(\mathbb{R}^3)^N$.

THEOREM 3 A formula for the iteration N_h on $\mathbf{M}$ is given by $N_h(m) = \mathfrak{L}_m R_{\mathrm{id}}(\mathrm{Ant}(v))$ where $v \in (\mathbb{R}^3)^N$ satisfies

$$v = -L_m^* H^*(H L_m L_m^* H^*)^{-1} h(m).$$

For the next theorem we replace l_m in the definition of P_m by L_m so that now

$$P_m = H^*(H L_m L_m^* H^*)^{-1} H L_m L_m^*.$$

THEOREM 4 A formula for $N_{\mathbf{M}}(m)$ for the gradient vector field $X(m) = \mathrm{grad}\ \phi|\mathbf{N}$ on $\mathbf{N}$ is given by

$$N_{\mathbf{M}}(m) = \mathfrak{L}_m R_{\mathrm{id}}(v)$$

where $v \in (\mathbb{R}^3)^N$ satisfies

$$\begin{aligned} &- L_m^*(I - P_m)A^* F(m) \\ &= (I - L_m^* P_m L_m)(L_m^* A^* A L_m - \mathrm{Ant}^{-1}(\mathfrak{L}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))} \\ &\qquad + \mathfrak{R}_{\mathrm{sym}(m^*(I-P_m)A^*\Phi(m))})\mathrm{Ant})(v) \end{aligned}$$

provided that $\nabla_{\mathbf{N}} X(m)$ is invertible at m.

The proofs of both theorems are achieved by simple substitutions.

REFERENCES

ALLGOWER, E. & GEORG, K. (1990) *Numerical Continuation Methods*. Berlin: Springer.

BEN-ISRAEL, A. (1966) A Newton–Raphson method for the solution of systems of equations. *J. Math. Anal. Appl.*, **15**, 243–252.

BEYN, W.-J. (1993) On smoothness and invariance properties of the Gauss–Newton method. *Numer. Funct. Anal. Optimiz.*, **14**, 503–514.

BLUM, L., CUCKER, F., SHUB, M. & SMALE, S. (1998) *Complexity and Real Computation*. Berlin: Springer.

CELLEDONI, E. & ISERLES, A. (2001) Methods for the approximation of the matrix exponential in a Lie-algebraic setting. *IMA J. Numer. Anal.*, **21**, 463–488.

DEDIEU, J.-P. (1996) Approximate solutions of numerical problems, condition number analysis and condition number theorem. *The Mathematics of Numerical Analysis*, Lectures in Applied Mathematics 23. (J. Renegar, M. Shub & S. Smale, eds). American Mathematical Society.

DEDIEU, J.-P. & KIM, M.-H. (2002) Newton's method for analytic systems of equations with constant rank derivatives. *J. Complexity*, **18**, 187–209.

DEDIEU, J.-P. & SHUB, M. (2000) Newton's method for overdetermined systems of equations. *Mathematics of Computation*, **69**, 1099–1115.

DEDIEU, J.-P. & SHUB, M. (2000) Multihomogeneous Newton's method. *Mathematics of Computation*, **69**, 1071–1098.

DENNIS, J. & SCHNABEL, R. (1983) *Numerical methods for Unconstrained Optimization and Nonlinear Equation*. Englewood Cliffs, NJ: Prentice-Hall.

DIECI, L. & VAN VLECK, E. S. (1999) Computation of orthonormal factors for fundamental solution matrices. *Numer. Math.*, **83**, 599–620.

EDELMAN, A., ARIAS, T. & SMITH, S. (2000) The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, **21**, 1401–1418.

FAN, K. & HOFFMAN, A. (1955) Some metric inequalities in the space of matrices. *Proc. Am. Math. Soc.*, **6**, 111–116.

GAUSS, K. F. (1809) Theoria Motus Corporum Coelestian. *Werke*, **7**, 240–254.

HELGASON, K. (1979) *Differential Geometry, Lie Groups, and Symmetric Spaces*. New York: Academic.

HESTENES, M. (1975) *Optimization Theory*. New York: Wiley.

HIRSCH, M. W. (1976) *Differential Topology*. Berlin: Springer.

HIRSCH, M. W., PUGH, C. C. & SHUB, M. (1977) *Invariant Manifolds*, Lecture Notes in Mathematics 583. Berlin: Springer.

ISERLES, A. (2001) On Cayley-transform methods for the discretization of Lie-group equations. *Foundations of Computational Mathematics*, **1**, 129–160.

ISERLES, A., MUNTHE-KAAS, H. Z., NØRSET, S. P. & ZANNA, A. (2001) Lie-group methods. *Acta Numerica*, **9**, 215–365.

MALAJOVICH, G. (1994) On generalized Newton algorithms. *Theoretical Computer Science*, **133**, 65–84.

MILNOR, J. (1963) *Morse Theory*. Princeton, NJ: Princeton University Press.

MUNTHE-KAAS, H. Z., QUISPEL, G. R. W. & ZANNA, A. (2001) General polar decompositions on Lie groups with involutive automorphisms. *Foundations of Computational Mathematics*, **1**, 297–324.

ORTEGA, J. & RHEINBOLDT, V. (1968) *Numerical Solutions of Nonlinear Problems*. Philadelphia: SIAM.

OSTROWSKI, A. (1976) *Solutions of Equations in Euclidean and Banach Spaces*. New York: Academic.

OWREN, B. & WELFERT, B. (1996) The Newton iteration on Lie groups. *BIT*, **40**, 121–145.

OWREN, B. & MARTHINSEN, A. (1999) *Integration methods based on canonical coordinates of the second kind*, Technical Report Numerics 5/1999. Trondheim, Norway: Norwegian University of Science and Technology.

LORD RAYLEIGH, J. W. STRUTT (1899) On the calculation of the frequency of vibration of a system in its gravest mode, with examples from hydrodynamics. *The Philosophical Magazine*, **47**, 556–572.

SHUB, M. (1986) Some remarks on dynamical systems and numerical analysis. *Dynamical Systems and Partial Differential Equations: Proceedings of VII ELAM* (L. Lara-Carrero & J. Lewowicz, eds). Caracas: Equinoccio, Universidad Simon Bolivar, pp. 69–92.

SHUB, M. (1993) Some remarks on Bezout's theorem and complexity. *From Topology to Computation: Proceedings of the Smalefest* (J. E. Marsden, M. W. Hirsch & M. Shub, eds). Berlin: Springer, pp. 443–455.

SHUB, M. & SMALE, S. (1993a) Complexity of Bézout's theorem I: geometric aspects. *J. Am. Math. Soc.*, **6**, 459–501.

SHUB, M. & SMALE, S. (1993b) Complexity of Bézout's theorem II: volumes and probabilities. *Computational Algebraic Geometry*, Progress in Mathematics 109. (F. Eyssette & A. Galligo, eds). Berlin: Birkhäuser, pp. 267–285.

SHUB, M. & SMALE, S. (1993c) Complexity of Bézout's theorem III: condition number and packing. *J. Complexity*, **9**, 4–14.

SHUB, M. & SMALE, S. (1996) Complexity of Bezout's theorem IV: probability of success, extensions. *SIAM J. Numer. Anal.*, **33**, 128–148.

SHUB, M. & SMALE, S. (1994) Complexity of Bézout's theorem V: polynomial time. *Theoretical Computer Science*, **133**, 141–164.

SMALE, S. (1985) On the efficiency of algorithms of analysis. *Bull. AMS*, **13**, 87–121.

SMALE, S. (1986) Algorithms for solving equation. *Proceedings of the International Congress of Mathematicians* AMS, pp. 172–195.

SMALE, S. (1986) Newton's method estimates from data at one point. *The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics*. (R. Ewing, K. Gross & C. Martin, eds). Berlin: Springer.

SMITH, S. (1994) Optimization techniques on Riemannian manifolds. *Fields Institute Communications 3*. AMS, pp. 113–146.

UDRISTE, C. (1994) *Convex Functions and Optimization Methods on Riemannian Manifolds*. Dordrecht: Kluwer.