A new algorithm for the computation of the group logarithm of diffeomorphisms

Matias Bossa and Salvador Olmos

GTC, I3A, University of Zaragoza, Spain, \{bossa, olmos\}@unizar.es

Abstract. There is an increasing interest on computing statistics of spatial transformations, in particular diffeomorphisms. In the Log-Euclidean framework proposed recently the group exponential and logarithm are essential operators to map elements from the tangent space to the manifold and vice versa. Currently, one of the main bottlenecks in the Log-Euclidean framework applied on diffeomorphisms is the large computation times required to estimate the logarithm. Up to now, the fastest approach to estimate the logarithm of diffeomorphisms is the Inverse Scaling and Squaring (ISS) method. This paper presents a new method for the estimation of the group logarithm of diffeomorphisms, based on a series in terms of the group exponential and the Baker-Campbell-Hausdorff formula. The proposed method was tested on 3D MRI brain images as well as on random diffeomorphisms. A performance comparison showed a significant improvement in accuracy-speed trade-off vs. the ISS method.

1 Introduction

Computational Anatomy is an emerging research field in which anatomy are characterized by means of large diffeomorphic deformation mappings of a given template [1]. The transformation is obtained by non-rigid registration, minimizing a cost function that includes an image matching term, and a regularization term that penalizes large and non-smooth deformations. Several approaches have been proposed in order to analyze the information contained in the transformation. Some methods consist in introducing a right-invariant Riemannian distance between diffeomorphisms, yielding methods with high computational load [2, 3]. Recently, an alternative framework was proposed [4] and consists in endowing the group of transformations with a Log-Euclidean metric. Although this metric is not translation invariant (with respect to the diffeomorphism composition), geodesics are identified with one-parameter subgroups, which can be obtained faster and more easily than the geodesics of a right-invariant Riemannian metric.

One-parameter subgroups of diffeomorphisms $\varphi_t(x)$ are obtained as solutions of the stationary Ordinary Differential Equation (ODE)

$$\frac{d\varphi_t(x)}{dt} = v \circ \varphi_t \equiv v(\varphi_t(x)).$$

\(1\)

* This work was funded by research grants TEC2006-13966-C03-02 from CICYT, Spain
A diffeomorphism \( \phi \equiv \phi(x) \equiv \varphi_1(x) \) is defined as the value of the flow \( \varphi_t \) at time one. Any velocity vector field can be written as a linear expansion \( v(x) = \sum_{i=1}^{D} v_i(x) e_i \), where \( \{e_i\}_{i=1}^{D} \) is an orthogonal basis of \( \mathbb{R}^D \). If the components \( v_i(x) \) are analytic then the solution of Eq. (1) is also analytic, and is given by the following formal power series (a.k.a. Gröbner’s Lie Series) [5]:

\[
\varphi_1(x) = e^{TV} x = \sum_{n=0}^{\infty} \frac{t^n}{n!} V^n x,
\]

where \( V \equiv \sum_{i=1}^{D} v_i(x) \frac{\partial}{\partial x_i} \) is a differential operator and \( V^n \) denotes the \( n \)-fold self-composition of \( V \).

The Log-Euclidean framework to compute statistics on diffeomorphisms consists in defining a distance between two diffeomorphisms \( \phi_1 \) and \( \phi_2 \) via a norm \( ||.|| \) on vector fields: \( dist(\phi_1, \phi_2) = ||v_1 - v_2|| \), where \( \phi_i = \exp(v_i) \). Assuming that such a \( v_i \) exists we call it the logarithm of \( \phi_i \), \( v_i = \log(\phi_i) \). This metric is equivalent to a bi-invariant Riemannian metric defined on the (abelian) group with the following composition rule: \( \phi_1 \circ \phi_2 = \exp(\log(\phi_1) + \log(\phi_2)) \). With such a group structure, the distances in the space of diffeomorphisms is computed as the Euclidean distance in the space of vector fields. This distance is inversion-invariant, i.e. \( dist(\phi_1, \phi_2) = dist(\phi_1^{-1}, \phi_2^{-1}) \) since \( \log(\phi_1^{-1}) = -\log(\phi_1) \) (in fact, it is invariant with respect to the exponentiation to any real power \( \neq 0 \)), and invariant with respect to the new group product, i.e. \( dist(\phi_1 \circ \phi_3, \phi_2 \circ \phi_3) = dist(\phi_1, \phi_2) \), but is not invariant under the standard composition, i.e. \( dist(\phi_1 \circ \phi_3, \phi_2 \circ \phi_3) = dist(\phi_1(\phi_3(x)), \phi_2(\phi_3(x))) \neq dist(\phi_1, \phi_2) \). Assuming that the logarithm and exponential can be (fast and accurately) computed, any standard statistical analysis can be performed directly on vector fields \( v_i \). This provides a simple way of computing statistics on transformations that avoids the problems of the small deformation frameworks, such as the likely occurrence of non-invertible mappings, and the ones of a right-invariant Riemannian framework, such as the intensive computation cost [6, 7].

Regarding to the computation of the exponential, it was recently proposed to extend the well known Scaling and Squaring (SS) method for computing the matrix exponential to diffeomorphisms [4]. This method basically consist in squaring (self-composing) recursively \( N \) times \( x + v/2^N \approx \exp(v/2^N) = \exp(v)^{-2^N} \). In a recent study [8], we presented a detailed performance comparison of several methods to compute the group exponential of diffeomorphisms, including the SS method, the forward Euler method and the direct application of the Lie series (2). The SS method achieved the best speed-accuracy trade-off, though two main drawbacks were found: first, the transformation must be computed in the whole domain, contrary to the forward Euler method and the Lie series expansion, that can be computed at a single point; and secondly, there exists an intrinsic lower bound in the accuracy due to the interpolation scheme and the finite size of the sampling grid. Despite of this lower bound, the SS method seems to be fast and accurate enough for most medical image analysis applications.

Regarding to the group logarithm of diffeomorphisms, it was proposed to apply the Inverse Scaling and Squaring (ISS) method [4], based on the following
approximation $v \approx 2^N(\exp(v)^2 - x)$, where the square root of $\phi$ must be recursively estimated $N$ times. The ISS method is much slower (about 100 times) than the SS method, as the computation of the square root involves an energy functional minimization. In the cases where a diffeomorphism can be written as a composition of two exponentials, $\phi = \exp(v_1) \circ \exp(v_2)$, the logarithm can be estimated with the Baker-Campbell-Hausdorff (BCH) formula, which is a series in terms of the Lie Bracket. In [7] it was tested the BCH formula applied to diffeomorphisms and it was shown that it provides similar accuracy than the ISS method, but with a much lower computational time. In a general case, where the diffeomorphism is neither an exponential of a known vector field, nor a composition of known exponentials, the ISS method seems to be the only available method for estimating the logarithm.

In this work, we propose a new method of computing the logarithm of arbitrary diffeomorphisms in any dimension based on a series involving the group exponential and the BCH formula.

2 A series for the logarithm of diffeomorphisms

The Lie series of the diffeomorphism exponential in Eq. (2) is a generalization of the Taylor expansion of the scalar exponential. However, to our knowledge, the Taylor expansion of the scalar logarithm cannot be generalized to the logarithm of diffeomorphisms in the same way. In fact, there exist diffeomorphisms (even infinitely close to the identity) that cannot be written as the exponential of any vector field in the tangent space [9], i.e. the exponential $\phi = \exp(v)$ is not a local diffeomorphism at $v = 0$, therefore a Lie series for the logarithm cannot exist. Nevertheless, we will talk about the logarithm $v$ of a diffeomorphism $\phi$, and define it as the vector field $v$ whose exponential is closer to $\phi$.

The basic idea is that given an initial guess $v_0$ for $v$ (being $v$ the ‘true’ logarithm of $\phi$), $\exp(-v_0)$ is close to $\phi^{-1}$, therefore $\exp(-v_0) \circ \phi$ is close to the identity and can be approximated by $\exp(-v_0) \circ \phi \equiv \exp(\delta v_0) \approx x + \delta v_0$. Then $\delta v_0 \equiv \delta v_0 \equiv \exp(-v_0) \circ \phi - x$ and $v_0$ can be corrected with $\delta v_0$ in order to get a better estimation of $v$:

$$
\phi \equiv \exp(v) = \exp(v_0) \circ (\exp(-v_0) \circ \phi)
= \exp(v_0) \circ \exp(\delta v_0)
\approx \exp(v_0) \circ \exp(\tilde{\delta} v_0)
$$

Recalling that the set of diffeomorphisms is a noncommutative group, $v$ can be approximated by the BCH formula [7]: $v = v_0 + \delta v_0 + 1/2[v_0, \delta v_0] + \cdots \approx v_0 + \delta v_0 + 1/2[v_0, \delta v_0] + \cdots$, where $[v, w] \equiv \sum_i w_i \partial_i v - v_i \partial_i w$ is the Lie bracket.

Finally, we will show that the sequence $v_i = v_{i-1} + \delta v_{i-1} + 1/2[v_{i-1}, \delta v_{i-1}] + \cdots$, with $\delta v_{i-1} = \exp(-v_{i-1}) \circ \phi - x$, quickly converges to $v$. Before going to the more general case of diffeomorphisms, a convergence analysis is presented for the scalar case.
Proposition 1. Let be \( f = e^v, \ v \in \mathbb{R} \), and let \( v_n(f) \) be defined by
\[
v_0 = 0 \\
v_n = v_{n-1} + f e^{-v_{n-1}} - 1
\] (3)
then the sequence \( v_n \) converges to \( \lim_{n \to \infty} v_n(f) = v \) and the error in the \( n \)-th term, \( \delta_n = v - v_n \), decreases with \( n \) as
\[
\delta_n \propto O(\|f - 1\|^{2^n})
\] (4)
Proof. Replacing \( f = e^v \) in (3) and expanding the exponential in its power series we get
\[
v_n = v_{n-1} + e^{v - v_{n-1}} - 1 = v_{n-1} + e^{\delta_{n-1}} - 1 \\
= v_{n-1} + 1 + v - v_{n-1} + \sum_{k=2}^{\infty} \frac{\delta_{n-1}^k}{k!} - 1 \\
-\delta_n = \sum_{k=2}^{\infty} \frac{\delta_{n-1}^k}{k!} \propto O(\|\delta_{n-1}\|^2)
\] (5)
Recalling that \( \delta_1 = v - v_1 = v - (f - 1) \), and expanding \( v \) in its power series \( v = \log(f) = \sum_{k=1}^{\infty} \frac{(f-1)^k}{k} (-1)^{k+1} \) we have \( \delta_1 = (f - 1) - 1/2(f - 1)^2 + O(\|f - 1\|^3) - (f - 1) \), and with (5) we get (4). \(\square\)
In fact, the reader can check that the expansion of \( v_n \) in power series of \( f \) is
\[
v_1 = f - 1 \\
v_2 = (f - 1) - \frac{(f - 1)^2}{2} + \frac{(f - 1)^3}{3} - \frac{(f - 1)^4}{8} + \frac{(f - 1)^5}{30} + O(\|f - 1\|^6) \\
v_3 = \sum_{k=1}^{7} \frac{(f - 1)^k}{k} (-1)^{k+1} - \frac{15}{128}(f - 1)^8 + \frac{13}{144}(f - 1)^9 + O((f - 1)^{10}) \\
\vdots \\
v_n = \sum_{k=1}^{2^n-1} \frac{(f - 1)^k}{k} (-1)^{k+1} + O((f - 1)^{2^n})
\]
Note that the first \( 2^k - 1 \) terms of the Taylor expansion of the \( k \)-th element of the sequence are equal to the Taylor expansion of the logarithm.
Of course it is not practical to compute the logarithm of a scalar number as the limit of a sequence where an exponential must be computed for each term. However, in the case of diffeomorphisms there is no Taylor expansion (or an alternative method except for the ISS) available for the logarithm, and the
\footnote{Or equivalently the series \( v_n = \sum_{i=0}^{n-1} (g^n(f) - 1) \), where \( g(f) = e^{1-f} f \) and \( g^n(f) \) is the \( n \)-fold self-composition of \( g(f) \), i.e. \( g^0(f) = f \), \( g^1(f) = g(f) \) and \( g^n(f) = g(g^{n-1}(f)) \).}
exponential is not very expensive to compute for the usual numerical accuracy required in medical image analysis.

\textit{Diffeomorphism logarithm.} Let’s assume that a diffeomorphisms \( \phi \) can be written as \( \phi = \exp(v) \), for some \( v \), in the sense of the formal power series (2). And let’s also assume that, for a given vector field \( \delta v_n \) close enough to 0, the BCH formula can be applied to compute \( v_{n+1} = \log(\exp(v_n) \circ \exp(\delta v_n)) \).

\[
v_{n+1} = v_n + \delta v_n + 1/2 \left[ v_n, \delta v_n \right] + 1/12 \left[ v_n, \left[ v_n, \delta v_n \right] \right] + 1/12 \left[ v_n, \delta v_n \right], \delta v_n \right] + \\
+ 1/48 \left[ v_n, \left[ v_n, \delta v_n \right] \right], \delta v_n \right] + 1/48 \left[ v_n, \left[ v_n, \delta v_n \right] \right], \delta v_n \right] \right) + O(\|\delta v_n\|)^3
\]

where \( [v, w] = \sum_i (v_i \partial w / \partial x_i - v_i \partial v / \partial x_i) \) is the Lie bracket, then the following proposition can be stated:

\textbf{Proposition 2.} The sequence

\[
v_0 = 0 \\
v_n = v_{n-1} + \delta v_{n-1} + 1/2 \left[ v_{n-1}, \delta v_{n-1} \right] + \cdots
\]

with \( \delta v_{n-1} = \exp(-v_{n-1}) \circ \phi - x \), converges to \( v \) with error

\[
\delta_n \equiv \log (\exp(v) \circ \exp(-v_n)) \propto O(\|\phi - x\|^2).
\]

\textbf{Proof.} Eq. (6) is equivalent to

\[
\exp(v_n) = \exp(v_{n-1}) \circ \exp(\delta v_{n-1}) \\
\exp(v_n) = \exp(v_{n-1}) \circ \exp(\exp(-v_{n-1}) \circ \exp(v) - x)
\]

where we used \( \phi = \exp(v) \). Now, multiplying on the right by \( \phi^{-1} = \exp(-v) \) and expanding \( \exp(\exp(-v_{n-1}) \circ \exp(v) - x) \) in its power series we have

\[
\exp(v_{n-1}) \circ \exp(-v) = \exp(v_{n-1}) \circ \exp(-v_{n-1}) \circ \exp(v) - x \circ \exp(-v) \\
\exp(-\delta_n) = \exp(v_{n-1}) \circ \left( x + \left( \exp(-v_{n-1}) \circ \exp(v) - x \right) + \\
+ \sum_{k=2}^{\infty} \frac{\left( \exp(-v_{n-1}) \circ \exp(v) - x \right)^k}{k!} \right) \circ \exp(-v)
\]

\[
= x + \exp(v_{n-1}) \circ \left( \exp(-v_{n-1}) \circ \exp(v) - x \right)^2 \circ \exp(-v) / 2 \\
+ \sum_{k=3}^{\infty} \frac{\exp(v_{n-1}) \circ \left( \exp(-v_{n-1}) \circ \exp(v) - x \right)^k \circ \exp(-v)}{k!}
\]

It is not difficult to see that the last term of r.h.s. is of order \( O(\delta_{n-1}^3) \) and \( (\exp(-v_{n-1}) \circ \exp(v) - x)^2 = (\exp(-v_{n-1}) \circ \exp(v) - x) \circ (\exp(-v_{n-1}) \circ \exp(v) - x) = \exp(-v_{n-1}) \circ \exp(v) \circ \exp(-v_{n-1}) \circ \exp(v) - 2 \exp(-v_{n-1}) \circ \exp(v) + x \), and left multiplying by \( \exp(v_{n-1}) \) and right multiplying by \( \exp(-v) \) gives \( \exp(v) \circ
\[
\exp(-v_{n-1}) - 2x + \exp(v_{n-1}) \circ \exp(-v) = \exp(\delta_{n-1}) - 2x + \exp(-\delta_{n-1}) = \delta_{n-1}^2 + \mathcal{O}(\delta_{n-1}^4), \text{ therefore} \\
\exp(-\delta_n) = x + \delta_{n-1}^2/2 + \mathcal{O}(\delta_{n-1}^3) \\
x - \delta_n + \mathcal{O}(\delta_n^2) = x + \delta_{n-1}^2/2 + \mathcal{O}(\delta_{n-1}^3) \\
\delta_n \propto \mathcal{O}(\delta_{n-1}^2) \quad (8)
\]

Recalling (7) the initial error \( \delta_1 \equiv \log(\exp(v) \circ \exp(-v_1)) \), where \( v_1 = \delta_0 = \phi - x = \sum_{k=1}^{\infty} \frac{v^k}{k!} \), and \( v \) commutes with \( v^k \) for all \( k \), therefore \( \delta_1 = v - \sum_{k=1}^{\infty} \frac{v^k}{k!} \propto \mathcal{O}(v^2) \propto \mathcal{O}((\phi - x)^2) \). Together with (8) we get (7).

In the estimation of the error (7) it was assumed that an infinite number of terms in the BCH formula was used. It can be argued that when a finite number \( N_{BCH} \) of terms is used, \( \delta_n \propto \mathcal{O}(\|\phi - x\|^{N_{BCH}+1}) \), as far as \( 2^n > N_{BCH} \). Therefore, in practice, \( N_{BCH} \) will limit the accuracy of the estimation.

3 Implementation details

The algorithm is initialized with \( v_1 = \phi - x \) and then updated following (6), where only 1 or 2 terms of the BCH formula are used. The computation of the Lie Bracket \([ \cdot \ , \cdot \ ]\) involves first order partial derivatives with respect to the spatial coordinates \( x_i \) that was implemented as centered finite differences after Gaussian filtering. The filtering is required because the noise in \( \tilde{\delta}v \) is quickly magnified after successive derivations. The filter width can be estimated using the following rule [8]: \( \nu_{\phi} \leq \nu_v \exp(\max(\|dv/dx\|)) \), where \( \nu_{\phi} \) (\( \nu_v \)) is the cut-off frequency of \( \phi \) (\( v \)). In our implementation there were still some isolated points in \( \delta v_k \) where the second derivative blown up, and a median filter was applied to these points. The exponential followed by a composition \( \exp(-v_k) \circ \phi \) present in \( \delta v_k \) was not computed with the SS method because, as explained in [8], both the composition and the SS methods introduce errors due to interpolation and the finite grid size. Instead, an integration scheme such as the Forward Euler method, starting at the locations defined by \( \phi(x_i) \), being \( x_i \) the grid points, is much more accurate.

The gradient descent method required to compute the square roots in the ISS method was based in a simpler gradient than in [4], in particular avoiding the estimation of the inverse diffeomorphism. This implementation provided a faster and more accurate convergence. It might be possible that the original proposal could provide more stable results for large diffeomorphisms.

4 Results

Firstly, a 60x60x60 smoothed random vector field \( v \) was exponentiated with the forward Euler method (step size 1/500) providing a diffeomorphism \( \phi \). We computed the logarithm \( \tilde{v} = \log(\phi) \) using (6) \( (N_{BCH} = 0, 1 \text{ and } 2) \), and the
ISS method. Accuracy was assessed by velocity vector field error \( e = (v - \tilde{v}) \) and its corresponding diffeomorphism error \( E = (\phi - \exp(\tilde{v})) \). Computations were performed using a 1.83GHz Core 2 Duo processor within a 2GB memory standard computer running Matlab 7.2 under Linux. Linear interpolation was implemented as C source mex files. Computation time was assessed with ‘cputime’ Matlab function. Figure 1 illustrates the accuracy-speed trade-off and a slice of the corresponding deformed grid. Each estimation method is described by two curves: a dashed/solid line corresponding to error in \( v \) and \( \phi \) respectively. Note the large amplitude of the deformation. Figure 2 shows a zoom detail of the error distributions and vector fields.

![Figure 1](image1.png)

**Fig. 1.** Left: Error vs. CPU time in the estimation of the logarithm corresponding to a random simulation. Solid/dashed lines correspond to \( E \) and \( e \) respectively. The horizontal lines correspond to the small deformation approximation: \( v(x) \approx \phi(x) - x \). Right and bottom: Illustration of the deformation grid. Fig. 2 will show the error distribution inside the red square.

Regarding to the accuracy in the estimation of the logarithm \( v \), which is actually our target, the ISS method only provided a midway accuracy between small deformation approximation and the proposed method for \( N_{BCH} = 1, 2 \). However, the corresponding diffeomorphism had similar accuracy for all methods. Regarding computation time, the proposed method with \( N_{BCH} = 1 \) was about 10 times faster than ISS method. From Fig. 2 it can be seen that the error was not due to outliers but in spatially correlated regions and far from the boundary.

A second set of experiments were performed on 3D MRI brain data sets. Two \( 181 \times 217 \times 181 \) brain images with isotropic 1mm resolution were randomly selected from LPBA40 database from LONI UCLA [10]. Two non-rigid registra-
tion methods were used: a diffeomorphic non-rigid registration [11] that provided a vector field $v$ as outcome; and Elastix [12] which is a registration method that provides a deformation field parameterized with B-Splines. In the later there is no warranty of the existence of $v$.

Left panels in Figures 3 and 4 show the error vs. computation time for the case of diffeomorphic and Elastix registration, respectively. In figure 4 only errors in $\phi$ are available. Additionally, a representative axial slice of the source image and the corresponding deformed grid is shown at right panel in both figures.

**Fig. 3.** Left: Error vs. CPU time in the estimation of the logarithm corresponding to a diffeomorphism computed with stationary LDDMM. Solid/dashed lines correspond to $E$ and $e$ respectively. The horizontal lines correspond to the small deformation approximation. Right: Illustration of the deformation grid superimposed on the brain image.
It is worthy to note that the error $e$ curve of the ISS method in Figures 1 and 3 are very different from the curve shown in [4]. We hypothesized that this behaviour could be explained by the large amplitude of the deformations. In order to verify this possibility the same experiment was performed on the same vector field $v$ divided by a factor of 10. Left panel of Figure 5 shows the error curves and right panel shows a detail of the deformed grid and the corresponding vector field. For this particular case of very small deformations, the ISS method was much more accurate than the logarithm series. Now the shape of the error curve was similar to the one in [4], with smaller error values. Note that all the error values, even for the small deformation approximation, are negligible for medical image analysis applications. When deformations are so small, $v \approx \phi(x) - x$ is accurate enough for standard statistical analysis.

![Figure 4](image1.png)

**Fig. 4.** Left: Error $E$ vs. CPU time in the estimation of the logarithm for a transformation computed with Elastix. The horizontal line corresponds to the small deformation approximation. Right: Illustration of deformation grids superimposed on the brain image.

![Figure 5](image2.png)

**Fig. 5.** Left: Idem Fig. 3 for $v/10$. Right: Illustration of the deformation grid and the corresponding velocity vector field.
In our opinion, the accuracy of the ISS method for large diffeomorphisms was limited by the fact that the right way to interpolate diffeomorphisms is unknown. Interpolation of diffeomorphisms is performed in the squaring (self-composition) operation. The composition of diffeomorphisms using a kernel interpolation scheme can provide non-diffeomorphic mappings. In contrast, velocity vector fields belong to a linear vector space, therefore they can be summed or interpolated without leaving the space.

5 Conclusion

We presented a new algorithm for the estimation of the group logarithm of arbitrary diffeomorphisms based on a series in terms of the Lie bracket and the group exponential. This method provided a much better accuracy-speed trade-off than the ISS method to estimate the vector field $v$ defining a diffeomorphism. In particular, at least one term of the BCH formula was essential for the series to provide a significant improvement vs. the ISS method.

Once a fast algorithm to compute the logarithm is available, statistics of the spatial transformations mapping image instances to a given atlas can be easily computed by means of standard multivariate statistics on the tangent space assuming the Log-Euclidean framework. This will be the topic of future studies.

References