COMMUN. MATH. SCI.
Vol. 12, No. 8, pp. 1409–1425

APPROXIMATING NEAR-GEODESIC NATURAL CUBIC SPLINES*

LYLE NOAKES†

Abstract. A method is given for calculating approximations to natural Riemannian cubic splines in symmetric spaces with computational effort comparable to what is needed for the classical case of a natural cubic spline in Euclidean space. Interpolation of $n+1$ points in the unit sphere S^m requires the solution of a sparse linear system of $4mn$ linear equations. For $n+1$ points in bi-invariant $SO(p)$ we have a sparse linear system of $2np(p-1)$ equations. Examples are given for the Euclidean sphere S^2 and for bi-invariant $SO(3)$ showing significant improvements over standard chart-based interpolants.

Key words. Lie group, Riemannian manifold, trajectory planning, mechanics, rigid body.

Mathematics Subject Classifications (2000). 70E17, 34E05, 49K99, 70E18, 53A99.

1. Introduction

Natural (Riemannian) cubic splines are extrema of a variational problem for interpolation of point data in a Riemannian m-manifold M, and turn out to be C^2 track-sums of solutions of a nonlinear ODE for curves called *Riemannian cubics* [15, 17, 18, 19, 16, 14, 2, 3, 4, 6, 20, 8]. Except when M is flat, Riemannian cubics are usually calculated as numerical solutions of initial value problems; calculating natural cubic splines as solutions of boundary and interior point problems is much more difficult. On the other hand, in the Euclidean case, Riemannian cubics are just cubic polynomial curves, and interpolation by natural cubic splines reduces to solving a tridiagonal linear system [5].

For interpolation of rigid body motion relative to a fixed centre, M is the rotation group $SO(3)$. There is no algorithm for calculating natural cubic splines in $SO(3)$ approaching the simplicity and convenience of the standard algorithm for the Euclidean case. Chart-based interpolants reduce locally to interpolants in Euclidean space, with only slightly more computational effort, but variational interpolants such as natural cubic splines are better behaved than chart-based interpolants. By definition, variational interpolants are critical for some quantity of interest (the mean square norm of covariant acceleration for the case of natural cubic splines) and depend only on the geometry of M and on the data. Chart-based interpolants are nonintrinsic, ungeometrical, unstable, and awkward (for instance when charts must be switched).

Our aim is to capture most of the good behaviour of natural cubic splines in $SO(3)$, with computational effort comparable to what is needed for the Euclidean case. To achieve this we first restrict attention to a circumstance that often occurs in practice: the data D should be near some geodesic γ (a numerical quantity $\mu_\gamma(D)$ should not be too large). The other tradeoff for computational simplicity is to allow suboptimal interpolants, namely the conditions for a natural cubic spline need only be *nearly* satisfied (to the extent that $\mu_\gamma(D)^2$ is small).

Our main result, Theorem 4.4, gives a suitable interpolator in terms of so-called *bi-Jacobi fields* along γ. Bi-Jacobi fields are generalizations of Jacobi fields, comprising a $4m$-dimensional real vector space. The bi-Jacobi fields are used in combination with a *quasiexponential map* to construct an interpolator I_δ with the property that

*Received: January 9, 2013; accepted (in revised form): October 9, 2013. Communicated by Richard Tsai.

†School of Mathematics and Statistics, The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia (Lyle.Noakes@uwa.edu.au).

$x_{\tilde{Y}(D)} = I_\delta(D)$ nearly satisfies the conditions for a natural cubic spline. Calculation of $x_{\tilde{Y}(D)}$ reduces to a problem in linear algebra, closely resembling the classical case of natural cubic splines in E^m.

For symmetric spaces, such as Euclidean unit spheres S^m, and rotation groups $SO(p)$ with a bi-invariant Riemannian metric, the bi-Jacobi fields are easily written down in terms of trigonometric functions. For $M = S^m$ this is done in Section 5 where an algorithm is given for calculating $x_{\tilde{Y}(D)}$. The following example, for $M = S^2$ with 3 interpolation points, illustrates the quality of $x_{\tilde{Y}(D)}$ and gives an idea of what is needed to calculate it. For full details we refer ahead to Section 5.

EXAMPLE 1.1. Consider the geodesic γ in $M = S^2$ given by $\gamma(t) = (\cos t, \sin t, 0)$ for $t \in [0, 2\pi/3]$. Our data consists of times $t_0 = 0$, $t_1 = \pi/6$, and $t_2 = 2\pi/3$ and corresponding interpolation points $x_0 = \gamma(0)$, $x_1 = (0.85529, 0.486954, 0.177074)$ (not especially near $\gamma(t_1) = (1/2, \sqrt{3}/2, 0)$), and $x_2 = \gamma(2\pi/3)$.

For each $t \in [0, 2\pi/3]$, the column spaces of the 3×8 matrices

$$B(t) = \begin{bmatrix} -\sin t & -t\sin t & -t^2\sin t & -t^3\sin t & 0 & 0 & 0 & 0 \\ \cos t & t\cos t & t^2\cos t & t^3\cos t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos t & \sin t & t\cos t & t\sin t \end{bmatrix},$$

$$C(t) = \begin{bmatrix} 0 & -\sin t & -2t\sin t & -3t^2\sin t & 0 & 0 & 0 & 0 \\ 0 & \cos t & 2t\cos t & 3t^2\cos t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sin t & \cos t & -t\sin t + \cos t & t\cos t + \sin t \end{bmatrix},$$

$$D(t) = \begin{bmatrix} 0 & 0 & -2\sin t & -6t\sin t & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\cos t & 6t\cos t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\cos t & -\sin t & -t\cos t - 2\sin t & -t\sin t + 2\cos t \end{bmatrix},$$

$$HB(t) = \begin{bmatrix} 0 & 0 & -2\sin t & -6t\sin t & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\cos t & 6t\cos t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2\sin t & 2\cos t \end{bmatrix}$$

are orthogonal to $\gamma(t)$, and the columns of B spans the space of bi-Jacobi fields along γ. The nearly cubic interpolant $x_{\tilde{Y}(D)} : [0, 2\pi/3] \to S^2$ is given by

$$x_{\tilde{Y}(D)}(t) = \frac{\gamma(t) + B(t)y^1}{\|\gamma(t) + B(t)y^1\|} \quad \text{for } 0 \le t \le \pi/6,$$

$$x_{\tilde{Y}(D)}(t) = \frac{\gamma(t) + B(t)y^2}{\|\gamma(t) + B(t)y^2\|} \quad \text{for } \pi/6 \le t \le 2\pi/3,$$

where $y^1, y^2 \in \mathbb{R}^8$ satisfy 5 homogeneous linear equations (5.1), (5.2), (5.3), (5.4), (5.5), the nonhomogeneous linear equation (5.6), and two further nonhomogeneous linear equations given by (5.7). These equations[1] in E^2 give 16 equations for the 16 real coordinates of y^1, y^2. Solving this linear system, we find

$$y^1 = (0, -0.0134274, 0, 0.00699674, 0, 0.0674479, 0.322409, 0),$$
$$y^2 = (0.00133915, -0.0211002, 0.0146539, -0.00233225,$$
$$-0.0129975, 0.382353, 0.0806023, -0.139607),$$

and $J(x_{\tilde{Y}(D)}) = 0.109762$.

[1] In Section 5 we take $\xi_i : TS^2_{x_i} \to E^2$ to be projection to the last two coordinates for $i = 0, 1, 2$. The ξ_i are linear isomorphisms because $(1, 0, 0)$ is not orthogonal to any of our x_i.

Figure 1.1 shows $x_{\tilde{Y}(D)}$ and the geodesic arc γ (blue) with the interpolation points $\gamma(0), x_1, \gamma(2\pi/3)$ (black), and $\gamma(\pi/6)$ (red). Figure 1.1 also shows another interpolant x_{ch} (green), constructed as follows using a coordinate chart ϕ and a classical natural cubic polynomial spline x_{E^2} in E^2.

Stereographic projection from $(0,0,1)$ defines a chart $\phi: S^2 - \{(0,0,1)\} \to E^2$. Then $x_{ch} := \phi^{-1} \circ x_{E^2} : [0, 2\pi/3] \to S^2$, where x_{E^2} is the natural Euclidean cubic spline in E^2 interpolating $\phi(\gamma(0))$, $\phi(x_1)$, $\phi(\gamma(2\pi/3))$ at times 0, $\pi/6$, and $2\pi/3$. In figure 1.1, x_{ch} appears more wavelike than $x_{\tilde{Y}(D)}$ and, accordingly, $J(x_{ch}) = 3.72049$ is some 36 times larger than $J(x_{\tilde{Y}(D)})$.

Alternatively consider the natural Euclidean cubic polynomial spline $x_{E^3}: [0, 2\pi/3] \to E^3$ in E^3 interpolating $\gamma(0)$, x_1 and $\gamma(2\pi/3)$. The *normalised Euclidean spline* $x_{norm} : [0, T] \to S^2$ (not shown in figure 1.1) is defined by $x_{norm}(t) := x_{E^3}(t)/\|x_{E^3}(t)\|$. We find $J(x_{norm}) = 0.644315$, still 6 times larger than $J(x_{\tilde{Y}(D)})$. This is mainly due to significant differences in parameterizations of x_{norm} and $x_{\tilde{Y}(D)}$, whose plots are misleadingly similar.

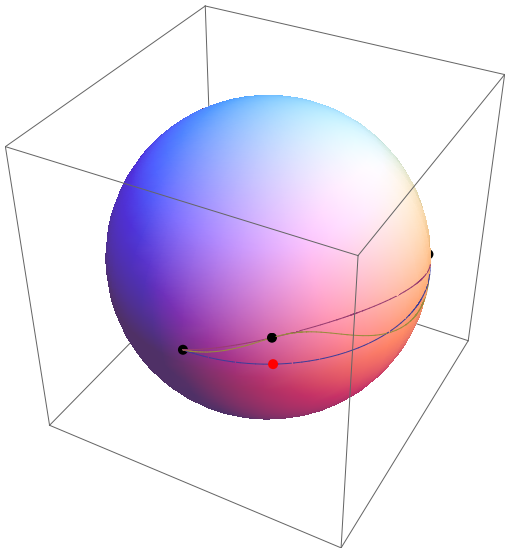

FIG. 1.1. *Geodesic, nearly cubic interpolant and chart-based interpolant (green) in Example 1.1.*

Such algorithms also apply when M is the double-cover S^3 of the bi-invariant rotation group $SO(3)$. However it is less complicated to deal with $SO(3)$ directly, without the covering; this avoids the use of unit quaternions for which 2^n optimization problems would need to be solved to find a natural cubic spline interpolating $n+1$ points in $SO(3)$. In Section 6 our algorithm for $SO(3)$ reduces to a sparse linear system of dimension $12n$ where $n+1$ is the number of data points. In Example 6.1, $n = 4$ and $x_{\tilde{Y}(D)}$ is seen to be significantly better than a standard chart-based interpolant.

In Section 2 we give a more detailed discussion of interpolation by curves in Riemannian manifolds, followed by the definition of natural cubic splines. Then Theorem 2.2 characterizes natural cubic splines as track-sums of solutions (Riemannian cubics) of a nonlinear 4th order ODE together with boundary and interior point conditions. The classical Euclidean case is reviewed (Example 2.2) and contrasted with the richer case of a curved space-form (Example 2.4). Example 2.5 compares Riemannian cu-

bics with Euler-Bernoulli elastica. Example 2.6 motivates the study of natural cubic splines in $SO(3)$. A precise definition is given of a nearly cubic spline, and then an outline of the rest of the paper.

2. Natural cubic and nearly cubic Riemannian splines

Let M be a C^∞ manifold of finite dimension m, with Riemannian metric $\langle\ ,\ \rangle$, associated Levi-Civita covariant derivative ∇, and Riemannian distance d. Let $\gamma\colon [0,T]\to M$ be a geodesic where $T>0$. A *data list* is an ordered $n+1$-tuple

$$D\ =\ ((t_0,x_0),(t_1,x_1),\ldots,(t_n,x_n)),$$

where $0=t_0<t_1<t_2<\ldots<t_n=T$, and $x_i\in M$ for $0\le i\le n$. Set $\tau:=(t_0,t_1,\ldots,t_n)$ and $\mu_\gamma(D):=\max\{d(x_i,\gamma(t_i)):i\le 0\le n\}$.

EXAMPLE 2.1. The *trivial data* is characterised by $\mu_\gamma(D_\tau)=0$, namely $D_\tau:=((t_0,\gamma(t_0)),(t_1,\gamma(t_1)),\ldots,(t_n,\gamma(t_n)))$.

A continuous map $x:[0,T]\to M$ is *D-feasible* when $x(t_i)=x_i$ for all $0\le i\le n$. Fix n and let $\mathcal{D}\subset([0,T]\times M)^{n+1}$ be a nonempty set of data lists, and let $I:\mathcal{D}\times[0,T]\to M$ be C^k where $k\ge 1$. Given $D\in\mathcal{D}$ define $x_D:[0,T]\to M$ by $x_D(t):=I(D,t)$.

DEFINITION 2.1. *I is an* interpolator *on $\mathcal{D}$ when x_D is D-feasible for all $D\in\mathcal{D}$. An interpolator I is* γ-consistent *when*

$$((t_0,\gamma(t_0)),(t_1,\gamma(t_1)),\ldots,(t_n,\gamma(t_n)))\in\mathcal{D}\implies x_{(t_0,\gamma(t_0)),(t_1,\gamma(t_1)),\ldots,(t_n,\gamma(t_n)))}=\gamma.$$

An interpolator I is linear *when $x_D(t)$ can be found from D by solving a system of linear equations.* Given $D\in\mathcal{D}$, define a functional J on C^2 D-feasible curves by $J(x):=\displaystyle\int_0^T\langle\nabla_t x^{(1)},\nabla_t x^{(1)}\rangle\,dt$. A *natural cubic spline* is a critical point of J.

THEOREM 2.2. *A C^2 D-feasible curve x is a natural cubic spline if and only if $\nabla_t x^{(1)}(0)=\mathbf{0}$, $\nabla_t x^{(1)}(T)=\mathbf{0}$, and*

$$\nabla_t^3 x^{(1)}+R(\nabla_t x^{(1)},x^{(1)})x^{(1)}=\mathbf{0} \tag{2.1}$$

for $t\ne t_i$ for any $0\le i\le n$.

A *Riemannian cubic* is a C^∞ curve satisfying equation (2.1). So a natural cubic spline is a C^2 track-sum of Riemannian cubics with vanishing covariant accelerations at $t_0=0$ and $t_n=T$.

EXAMPLE 2.2. In classical approximation theory [5] M is Euclidean m-space E^m and a natural cubic spline is the unique C^2 interpolant $x=x_D$ such that

- for all $i=1,2,\ldots,n$ the restriction of x to $[t_{i-1},t_i]$ is a cubic polynomial;
- $x^{(2)}(t_0)=\mathbf{0}=x^{(2)}(t_n)$.

These conditions reduce to a linear system for the coefficients of the cubic polynomials, where the coefficient matrix is tridiagonal. The widespread use of Euclidean natural cubic splines is due in part to ease of computation, which extends to cases where the Riemannian manifold M is flat.

When M is not flat it is much more difficult to compute natural cubic splines. To begin with, there are very few cases where Riemannian cubics are known explicitly in terms of standard functions.

EXAMPLE 2.3. Let $\gamma:[0,T]\to M$ be a geodesic in any Riemannian manifold M, and let $c:[0,S]\to[0,T]$ be a cubic polynomial. As is easily verified, $x:=\gamma\circ c:[0,S]\to M$ is a Riemannian cubic.

EXAMPLE 2.4. For $m>1$ Riemannian cubics in real projective m-space $\mathbb{R}P^m$ lift to Riemannian cubics in the Riemannian double-cover, which is the unit sphere S^m in E^{m+1}. Riemannian cubics in S^2 are also Riemannian cubics in S^3. For $m>3$ Riemannian cubics in S^m are Riemannian cubics in totally geodesic copies of S^3. So Riemannian cubics in S^m and $\mathbb{R}P^m$ reduce to Riemannian cubics in S^3.

In E^3 elastic curves are given by elliptic functions [23], whereas Riemannian cubics are just cubic polynomial curves. On the other hand, in bi-invariant $SO(3)\cong\mathbb{R}P^3$ elastic curves are also known in terms of elliptic functions [11, 22], but Riemannian cubics in bi-invariant $SO(3)$ are much less perfectly understood [16, 17]. The only explicitly nontrivial examples are the codimension-3 family of *null Riemannian cubics*, found by reducing to the completely integrable time-independent Schrödinger equation with quadratic potential [21, Chapter 4].

Null Riemannian cubics in $SO(3)$ also correspond to self-similar solutions $z:[0,S]\times[0,\infty)\to E^3$ of the Da Rios equations [9] for the localized induction approximation

$$\frac{\partial z}{\partial t}=\frac{\partial z}{\partial s}\times\frac{\partial^2 z}{\partial s^2} \tag{2.2}$$

modelling the motion of an isolated vortex filament in an incompresible inviscid fluid. Here $z(s,t)$ is the position at time t of a point on the filament with arc-length parameter $s\in[0,S]$. Self-similar solutions of (2.2) have the form $z(s,t)=t^{1/2}w(st^{-1/2})$, where $w:[0,\infty)\to E^3$ satisfies [9, Equation (20)], namely

$$2w''(u)\times w'(u)=uw'(u)-w(u). \tag{2.3}$$

Because $w'(u)$ is a unit vector, $\langle w'',w'\rangle=0$. Taking cross-products with $w'(u)$ on both sides of (2.3) gives $2w''(u)=w'(u)\times w(u)$. Define $V:[0,\infty)\to E^3$ by $V(t):=2^{-1/2}w(2^{1/2}t)$. Then

$$V^{(2)}(t)=V^{(1)}(t)\times V(t),$$

namely V is the left Lie reduction of a null Riemannian cubic [15, 19, 14].

Lack of closed-form expressions for Riemannian cubics is not the main barrier to calculation of Riemannianian natural cubic splines. After all, Riemannian cubics can be computed numerically as solutions of initial value problems for equation (2.1). By far the largest obstruction to computing natural cubic splines is the problem of satisfying the boundary and interior point conditions of Theorem 2.2 for Riemannian cubics. In theory this can be reduced by (multiple) shooting to a sequence of initial value problems [12], but this seems a daunting task.

A suggestion by Jerrold Marsden was to apply a finite-dimensional optimizer to a discretization of the variational problem for Riemannian cubic splines. Following this direct method, good approximations were achieved by Marin Kobilarov for M the unit 3-sphere S^3 in E^4 (private communication). Alternatively, Riemannian cubic splines in S^3 can be found using the indirect method of [7], achieving in seconds what once took hours. Using either approach, the computational effort for accurate approximations to Riemannian cubic splines is much greater than for the Euclidean

case, and 2^n natural cubic splines in S^3 are needed to calculate a natural cubic spline interpolating $n+1$ points in bi-invariant $SO(3)$.

Indeed, solving nonlinear boundary and interior point problems is a familiar difficulty in approximation theory.

EXAMPLE 2.5. The elastica in E^2 are completely known [23], but the use of elastic splines in E^2 is limited by the difficulty of satisfying boundary value and interior point conditions, as well as the condition that splines should be C^2 with vanishing covariant accelerations at endpoints. This is the principal reason for the popularity of polynomial splines as an alternative.

More than anything, a useful interpolator should be robust and easily calculable, and we are prepared to accept some small degree of suboptimality in order to achieve this. The present paper constructs a γ-consistent linear interpolator I_δ where, for near geodesic data D, the interpolator x_D *approximately* satisfies the conditions for a natural cubic spline. This would be applicable when data is sampled from trajectories of an object subject to small unknown external forces, namely *nearly geodesic* trajectories with respect to the polarization of kinetic energy [10].

EXAMPLE 2.6. At time t the configuration of a rigid body relative to its centre of mass is given by a point $y(t)$ in the rotation group $SO(3)$. Measurements $x_i := y(t_i)$ are made of the configurations at times $t_0<t_1<\ldots<t_n$. The best estimate $x:[t_0,t_n]\to SO(3)$ of y depends on what can be assumed about the dynamics of the body, and on what is considered best in an approximation.

If the body can be assumed to move nearly freely, with small unknown external torques, then $y:[t_0,t_n]\to SO(3)$ is near some geodesic γ with respect to a left-invariant Riemannian metric on $SO(3)$ (which might in principle be inferred from observations [13]). An interpolant x may be considered optimal when it minimises the mean-square norm of the applied external torque. Then, in particular, x is a natural Riemannian cubic spline.

DEFINITION 2.3. *A C^2 interpolator I on $\mathcal{D}$ is* nearly cubic *when, for any $D\in\mathcal{D}$, and $x=x_D$,*

$$\max\{\|\nabla_t x^{(1)}(0)\|_{x(0)},\ \|\nabla_t x^{(1)}(T)\|_{x(T)},\ \|\nabla_t^3 x^{(1)}+R(\nabla_t x^{(1)},x^{(1)})x^{(1)}\|_{x(t)} : t\neq t_i\}$$
$$\leq K\mu_\gamma(D)^2,$$

where i ranges from 1 to $n-1$, and $K>0$ depends on γ and on $t_0,t_1,\ldots,t_n$, but not on $x_0,x_1,\ldots,x_n$. Then x_D is called a nearly cubic spline.

Because natural cubic splines are difficult to compute, a nearly cubic spline $\hat{x}$ is acceptable as a first approximation. If $\mu_\gamma(D)$ is small then $J(\hat{x})$ should be very small. If I is linear then computation of $\hat{x}$ should be straightforward.

In Section 3 we introduce the notion of *bi-Jacobi fields* Y along a geodesic γ. Then, after some choice of quasi-exponential map, we obtain approximations x_Y to Riemannian cubics. In Section 4 track-sums of the x_Y are used to construct an interpolator I_δ. The main result Theorem 4.4 says that I_δ is C^2, γ-consistent, nearly cubic, and (crucially) linear. The case where $M=S^m$ is treated in detail in Section 5. In Section 6 we do the same for bi-invariant $SO(3)$, and a nearly cubic interpolant $x_{\tilde{Y}(D)}$ is compared with a chart-based interpolant using cubic polynomial splines: the former is markedly better behaved, much more nearly optimal and found in less than one hundredth of a second. After a brief summary, Section 7 lists topics for further study.

3. Near-geodesic Riemannian cubics and bi-Jacobi fields

Let $\gamma:[0,T]\to M$ be a geodesic in a Riemannian m-manifold M. For $t\in[0,T]$ define a linear endomorphism S_t of $TM_{\gamma(t)}$ by $S_t(v):=R_{\gamma(t)}(v,\gamma^{(1)}(t))\gamma^{(1)}(t)$, where R denotes Riemannian curvature. As is well-known S_t is self-adjoint with respect to the inner product $\langle\ ,\ \rangle_{\gamma(t)}$.

EXAMPLE 3.1. If M is locally-symmetric then, given γ, the eigenvalues of S_t are independent of t.

The *Jacobi operator* defined on C^∞ fields X along γ is given by

$$H(X)(t)\ :=\ \nabla_t^2X(t)+S_t(X(t)).$$

Elements of the kernel of H are *Jacobi fields* along γ, and comprise a $2m$-dimensional real vector space. More generally, consider the $4m$-dimensional space of *bi-Jacobi fields*.

DEFINITION 3.1. *A C^∞ field Y along γ is a* bi-Jacobi field *when* $H^2Y=\mathbf{0}$.

Our construction of curves in M from vector fields requires a generalized exponential map.

DEFINITION 3.2. *Let N be an open neighbourhood of the $\mathbf{0}$-section of TM. A C^∞ function $E:N\to M$, with all derivatives bounded, is said to be* quasiexponential *when, for any $z\in M$ and any $v\in TM_z$,*

$$E(z,\mathbf{0})=z\quad\textit{and}\quad\frac{d}{dh}E(z,hv)|_{h=0}=v.$$

Although the exponential map (with respect to any Riemannian metric) is quasi-exponential, more elementary constructions may also be used.

EXAMPLE 3.2. Let M be the unit m-sphere S^m in Euclidean $m+1$-space E^{m+1}. Define $E(z,w):=(z+w)/\|z+w\|$ where $\|\ \|$ is the Euclidean norm. Then E is quasi-exponential.

EXAMPLE 3.3. Let $M=SO(p)$ where $p\geq 2$. Given $(z,w)\in TSO(p)$, the $p\times p$ matrix $v:=z^{-1}w$ is skew-symmetric. Define $E(z,w):=z(\mathbf{1}+v/2)(\mathbf{1}-v/2)^{-1}$. Then E is quasiexponential.

REMARK 3.1. A local inverse of E is needed for calculation of F in the feasibility equation (4.3) for near cubic splines. So, because of our preference for rational functions over exponentials and logarithms, the quasiexponentials of examples 3.2 and 3.3 are used in examples 1.1 and 6.1 respectively.

Fix a quasiexponential map E. Given a C^k field Y along γ where $k\geq 2$, define $x_Y:[0,T]\to M$ by $x_Y(t):=E(\gamma(t),Y(t))$. A C^k homotopy $\tilde{x}_Y:[0,1]\times[0,T]\to M$ from γ to x_Y is defined by $\tilde{x}_Y(h,t):=x_{hY}(t)$. Because E is quasiexponential,

$$\frac{\partial\tilde{x}_Y(h,t)}{\partial h}|_{h=0}=Y. \tag{3.1}$$

LEMMA 3.3. *Let Y be C^∞. Then $\nabla_h\nabla_t\tilde{x}_Y^{(1)}(h,t)|_{h=0}=HY(t)$.*

Proof. From the definition of Riemannian curvature

$$\nabla_h\nabla_t\tilde{x}_Y^{(1)}(h,t)=\nabla_t\nabla_h\tilde{x}_Y^{(1)}+R\Big(\frac{\partial\tilde{x}_Y}{\partial h},\tilde{x}_Y^{(1)}\Big)\tilde{x}_Y^{(1)}=\nabla_t^2\frac{\partial\tilde{x}_Y}{\partial h}+R\Big(\frac{\partial\tilde{x}_Y}{\partial h},\tilde{x}_Y^{(1)}\Big)\tilde{x}_Y^{(1)}.$$

Taking $h=0$ the right hand side becomes HY, by (3.1). □

DEFINITION 3.4. *For $k\geq 1$ set*

$$\|Y\|_{k,\infty} := \max\{\|Y(t)\|_{\gamma(t)}, \|\nabla_t^j Y(t)\|_{\gamma(t)} : 1\leq j\leq k,\ 0\leq t\leq T\}.$$

LEMMA 3.5. *For some $L_{ends}>0$ depending on E and γ, and for $t_*=0,T$, if Y is C^∞ with $HY(t_*)=\mathbf{0}$ then, for all $h\in[0,1]$,*

$$\|\nabla_t \tilde{x}_Y^{(1)}\|_{\tilde{x}_Y(h,t_*)} \ \leq\ L_{ends}\|Y\|_{2,\infty}^2.$$

Proof. It suffices to argue with $t_*=0$. Then we have $\tilde{x}_Y^{(1)}(t)=E_1(hY,h\nabla_t Y,t) := dE_{x_{hY(t)}}(\gamma^{(1)}(t),hY^{(1)}(t))$, where $E_1 : TN\oplus TN\times[0,T]\to TM$ is C^∞ with all derivatives bounded. Similarly $\nabla_t\tilde{x}_Y^{(1)}=E_2(hY,h\nabla_t Y,h\nabla_t^2 Y,t)$, where

$$E_2 : TN\oplus TN\oplus TN\times[0,T]\to TM$$

is C^∞ with all derivatives bounded. Then, for some $L_{ends}>0$ depending on E and γ, and all $(h,t)\in[0,1]\times[0,T]$,

$$\frac{1}{2}\|\nabla_h^2\nabla_t\tilde{x}_Y^{(1)}(h,t)\|_{\tilde{x}_Y(h,t)}\leq L_{ends}\|Y\|_{2,\infty}^2.$$

The result now follows from Taylor's Theorem and Lemma 3.3, since $HY(0)=\mathbf{0}$. □

LEMMA 3.6. *Let Y be C^∞. Then $\nabla_h(\nabla_t^3\tilde{x}_Y^{(1)}+R(\nabla_t\tilde{x}_Y^{(1)},\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)})|_{h=0}=H^2Y(t)$.*

Proof. The left hand side expands as

$$\begin{aligned}&(\nabla_h\nabla_t^3\tilde{x}_Y^{(1)}+\nabla_h(R)(\nabla_t\tilde{x}_Y^{(1)},\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)}+R(\nabla_h\nabla_t\tilde{x}_Y^{(1)},\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)}\\ &\quad+R(\nabla_t\tilde{x}_Y^{(1)},\nabla_h\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)}+R(\nabla_t\tilde{x}_Y^{(1)},\tilde{x}_Y^{(1)})\nabla_h\tilde{x}_Y^{(1)})|_{h=0}\\ &=(\nabla_h\nabla_t^3\tilde{x}_Y^{(1)}+R(HY,\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)})|_{h=0}=(\nabla_h\nabla_t^3\tilde{x}_Y^{(1)})|_{h=0}+S_t(HY)\end{aligned}$$

by Lemma 3.3, and because $\tilde{x}_Y(0,t)=\gamma(t)$ with γ a geodesic. By (3.1) and the definition of Riemannian curvature,

$$\begin{aligned}(\nabla_h\nabla_t^3\tilde{x}_Y^{(1)})|_{h=0}&=(\nabla_t\nabla_h\nabla_t^2\tilde{x}_Y^{(1)}+R(Y,\tilde{x}_Y^{(1)})\nabla_t^2\tilde{x}_Y^{(1)})|_{h=0}\\ &=(\nabla_t(\nabla_t\nabla_h\nabla_t\tilde{x}_Y^{(1)}+R(Y,\tilde{x}_Y^{(1)})\nabla_t\tilde{x}_Y^{(1)}))|_{h=0}.\end{aligned}$$

Because γ is a geodesic, the right hand side is

$$\begin{aligned}(\nabla_t^2\nabla_h\nabla_t\tilde{x}_Y^{(1)})|_{h=0}&=(\nabla_t^2(\nabla_t\nabla_h\tilde{x}_Y^{(1)}+R(Y,\tilde{x}^{(1)})\tilde{x}_Y^{(1)})|_{h=0}\\ &=\nabla_t^2(\nabla_t^2Y+S_t(Y))=\nabla_t^2HY.\end{aligned}$$

So $\nabla_h(\nabla_t^3\tilde{x}_Y^{(1)}+R(\nabla_t\tilde{x}_Y^{(1)},\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)})|_{h=0}=\nabla_t^2HY+S_tHY=H^2Y$, as stated.

REMARK 3.2. Lemma 3.6 may be compared with the second variation of J along a general Riemannian cubic, given by Theorem 2.4 of [4] "after many tedious manipulations". Lemma 3.6 of the present paper calculates the variation of the Euler-Lagrange equation for J along a very special kind of Riemannian cubic, namely a geodesic. So

(modulo boundary conditions and some additional effort) Lemma 3.6 should follow from [4]. Because geodesics are so special it is easier to prove Lemma 3.6 directly, but a connection with the more general results in [4] can be made as follows.

Formula (14) of [4] for $K(Y,x^{(1)})$ has 12 terms involving covariant derivatives and Riemannian curvature. For the special case where x is a geodesic, 7 terms vanish, leaving 5 whose sum agrees with the expansion of H^2Y. It seems that this simple expression for the case of a geodesic has not been noted previously.

LEMMA 3.7. *For some $L_{int}>0$ depending on E and γ, if Y is a bi-Jacobi field then, for all $h\in[0,1]$,*

$$\|\nabla_t^3\tilde{x}_Y^{(1)}+R(\nabla_t\tilde{x}_Y^{(1)},\tilde{x}_Y^{(1)})\tilde{x}_Y^{(1)}\|_{\tilde{x}_Y(h,t)}\le L_{int}\|Y\|_{3,\infty}^2.$$

The proof is similar to that of Lemma 3.5, with Lemma 3.6 replacing Lemma 3.3. □

4. Track-sums

Given bi-Jacobi fields $Y^1,Y^2,\dots,Y^n$ along γ, their *track-sum* $\tilde{Y}$ is defined by $\tilde{Y}(t):=Y^i(t)$ for $t\in[t_{i-1},t_i)$ and $\tilde{Y}(t_n):=Y^n(t_n)$. Then $\tilde{Y}$ is a field defined along γ, and is C^∞ except possibly at $t_1,t_2,\dots,t_{n-1}$.

DEFINITION 4.1. *The track-sum $\tilde{Y}$ is* infinitesimally natural cubic *when it is C^2 and $(HY^1)(t_0)=\mathbf{0}$ and $(HY^n)(t_n)=\mathbf{0}$.*

The next result follows from lemmas 3.5, 3.7.

LEMMA 4.2. *Let the track-sum $\tilde{Y}$ be infinitesimally natural cubic. Then $x_{\tilde{Y}}$ is C^2 and, on taking $x=x_{\tilde{Y}}$,*

$$\max\{\|\nabla_t x^{(1)}(0)\|_{x(0)},\ \|\nabla_t x^{(1)}(T)\|_{x(T)},\ \|\nabla_t^3x^{(1)}+R(\nabla_t x^{(1)},x^{(1)})x^{(1)}\|_{x(t)}:t\ne t_i\}$$
$$\le L\|\tilde{Y}\|_{3,\infty}^2,$$

where $L=\max\{L_{ends},L_{int}\}$.

Let $\{B_1,B_2,\dots,B_{4m}\}$ be a basis of bi-Jacobi fields along γ. A bi-Jacobi field Y^i along γ is By^i where

$$B:=\begin{bmatrix}B_1\ B_2\ \dots\ B_{4m}\end{bmatrix}\quad\text{and}\quad y^i:=\begin{bmatrix}y_1^i\\ y_2^i\\ \dots\\ y_{4m}^i\end{bmatrix}\in\mathbb{R}^{4m}.$$

So the condition that the track-sum $\tilde{Y}$ of $Y^1,Y^2,\dots,Y^n$ be infinitesimally natural cubic reads

$$\begin{matrix}B(t_i)y^i=B(t_i)y^{i+1},\\ \nabla_tB(t_i)y^i=\nabla_tB(t_i)y^{i+1},\\ \nabla_t^2B(t_i)y^i=\nabla_t^2B(t_i)y^{i+1},\end{matrix}\quad\text{where}\ \ 1\le i\le n-1,\quad\text{and}\quad\begin{matrix}HB(0)y^1=\mathbf{0},\\ HB(T)y^n=\mathbf{0},\end{matrix}\qquad(4.1)$$

comprising $(3n-1)m$ homogeneous linear equations in $4mn$ real unknowns y_j^i. Then $x_{\tilde{Y}}$ is D-feasible when

$$E(\gamma(t_i),B(t_i)y^i)=x_i\quad\text{for}\ \ 0\le i\le n,\qquad(4.2)$$

where $y^0:=y^1$. Given $\delta>0$ set $\mathcal{E}_\delta:=\{(t,y)\in[0,T]\times M:d(\gamma(t),y)<\delta\}$. Because E is quasiexponential, for some $\delta>0$ there is a C^∞ function $F:\mathcal{E}_\delta\to TM$ with the property that $F(t,\gamma(t))=(\gamma(t),\mathbf{0})$ and $E(\gamma(t),F(t,y))=y$.

EXAMPLE 4.1. Let E be defined as in Example 3.2 for the case where $M=S^m$. Then

$$F(t,y)=y/\langle\gamma(t),y\rangle-\gamma(t),$$

where $\langle\ ,\ \rangle$ is the Euclidean inner product.

EXAMPLE 4.2. Let E be defined as in Example 3.3 for the case where $M=SO(p)$. Then

$$F(t,y)=-2\gamma(t)(\mathbf{1}+\gamma(t)^{-1}y)^{-1}(\mathbf{1}-\gamma(t)^{-1}y).$$

For $D\in\mathcal{D}_\delta:=\{D=((t_0,x_0),(t_1,x_1),\dots,(t_n,x_n)):\mu_\gamma(D)<\delta\}$, (4.2) is equivalent to the linear equations

$$B(t_i)y^i=F(t_i,x_i) \tag{4.3}$$

for the y^i where $0\le i\le n$ and $y^0:=y^1$.

DEFINITION 4.3. *τ is* ncs-nonconjugate along γ *when for $D=D_\tau$ the linear system (4.1), (4.3) has rank $4mn$.*

A separate study will be made of ncs-nonconjugacy. Meanwhile suppose the condition is verified numerically for some choice of τ. By the implicit function theorem, for some small $\delta>0$, there is a unique C^∞ assignment

$$D\in\mathcal{D}_\delta\mapsto y(D):=(y^1,y^2,\dots,y^n)\in\mathbb{R}^{4mn}$$

with $y(D_{(t_0,t_1,\dots,t_n)})=\mathbf{0}$ such that (4.1), (4.3) hold for any $D\in\mathcal{D}_\delta$. For $D\in\mathcal{D}_\delta$ let $\tilde{Y}(D)$ be the infinitesimally natural cubic track-sum of bi-Jacobi fields given by $y(D)$.

THEOREM 4.4. *Define $I_\delta:\mathcal{D}_\delta\times[0,T]\to M$ by $I_\delta(D,t):=x_{\tilde{Y}(D)}(t)$. For $\delta>0$ sufficiently small, I_δ is C^2, γ-consistent, nearly cubic and linear.*

Proof. By (4.2) and because $y(D_{(t_0,t_1,\dots,t_n)})=\mathbf{0}$, the interpolator I_δ is γ-consistent. By (4.1) each $x_{\tilde{Y}(D)}$ is C^2 and so is I_δ because y is C^∞. Because y is C^1 it is Lipschitz for $\delta>0$ sufficiently small. So $\|\tilde{Y}(D)\|_{3,\infty}\le b\mu_\gamma(D)$ where $b>0$ depends on B. Then I_δ is nearly cubic by Lemma 4.2. Because $y(D)$ satisfies the $4mn$-dimensional linear system (4.1), (4.3) we have I_δ linear. □

So finding a C^2, γ-consistent, nearly cubic linear interpolator reduces to calculating the bi-Jacobi fields along γ. When M is a symmetric space, in particular for a Euclidean sphere S^m or for bi-invariant $SO(3)$, this is a simple task.

5. Euclidean spheres

For $m\ge2$ let M be the unit sphere S^m in Euclidean $m+1$-space E^{m+1}, with Riemannian metric $\langle\ ,\ \rangle$ given by the first fundamental form. The group of isometries of S^m is the orthogonal group $O(m+1)$. So after rotation any nonconstant geodesic $\gamma:[0,T]\to S^m$ can be written $\gamma(t)=(\cos\omega t,\sin\omega t,\mathbf{0})$ where $\mathbf{0}\in E^{m-1}$ and $\omega>0$. Any C^∞ field X along γ has the form

$$X(t)=f(t)\gamma^{(1)}(t)+X_\perp(t),$$

where $f:[0,T]\to\mathbb{R}$ and $X_\perp:[0,T]\to\{(0,0)\}\times E^{m-1}\cong E^{m-1}$.

The Levi-Civita covariant derivative of a vector field Y on S^m in the direction of a vector field X at $z\in S^m$ is given by $\nabla_X Y = dY_z(X(z)) + \langle X(z), Y(z)\rangle z$. For the Riemannian curvature we have $R(X,Y)Z = \langle Y,Z\rangle X - \langle X,Z\rangle Y$. So

$$(HX)(t) = f^{(2)}(t)\gamma^{(1)}(t) + \Big(\frac{d^2}{dt^2} + \omega^2\Big)X_\perp(t) \implies$$

$$(H^2X)(t) = f^{(4)}(t)x^{(1)}(t) + \Big(\frac{d^2}{dt^2} + \omega^2\Big)^2 X_\perp(t).$$

So X is a bi-Jacobi field when f is a cubic polynomial and $\Big(\frac{d^2}{dt^2} + \omega^2\Big)^2 X_\perp(t) = \mathbf{0}$. A basis of such fields is given by

$$B_1(t) = \gamma^{(1)}(t),\ B_2(t) = t\gamma^{(1)}(t),\ B_3(t) = t^2\gamma^{(1)}(t),\ B_4(t) = t^3\gamma^{(1)}(t),$$

$$B_{2+j}(t) = e_j\cos\omega t,\ B_{m+1+j} = e_j\sin\omega t,\ B_{2m+j}(t) = e_j t\cos\omega t,\ B_{3m+j-1} = e_j t\sin\omega t,$$

where $3\le j\le m+1$, and $\{e_1, e_2, \dots, e_{m+1}\}$ denotes the standard basis of E^{m+1}. Here B_1, B_2 are also Jacobi fields, as are the B_{2+j}, B_{m+1+j} for $3\le j\le m+1$. For $t\in[0,T]$ let $\pi(t): E^3 \to TS^m_{\gamma(t)}$ be projection orthogonal to $\gamma(t)$. Then

- $\nabla_t B_j(t) = \pi(t)\circ B_j^{(1)}(t)$,
- $\nabla_t^2 B_j = \pi(t)\circ(\nabla_t B_j)^{(1)}(t)$, and
- HB_j is $\nabla_t^2 B_j$ or $\nabla_t^2 B_j + \omega^2 B_j$ according as $1\le j\le 4$ or $5\le j\le 4m$.

For $0\le i\le n$ choose linear isomorphisms $\xi_i : TS^m_{x_i} \to E^m$. Denote by $B(t)$, $C(t)$, $D(t)$, and $HB(t)$ the $(m+1)\times 4m$ matrices whose jth columns are $B_j(t)$, $\nabla_t B_j(t)$, $\nabla_t^2 B_j(t)$, and $HB_j(t)$ respectively, for $1\le j\le 4m$.

Given a data list D, there are $(3n-1)m$ homogeneous linear equations (4.1) for the $4mn$ coordinates y^i_j of the bi-Jacobi fields Y^i in Section 4. For the remaining $(n+1)m$ nonhomogeneous equations (4.3) we use the map F from Example 4.1, namely

$$F(t,x) = \frac{x}{\langle\gamma(t), x\rangle} - \gamma(t).$$

Taking the quasiexponential E from Example 3.2, we have for $t\in[t_{i-1}, t_i]$, $x_{\tilde{Y}(D)}(t) = \dfrac{\gamma(t) + B(t)y^i}{\|\gamma(t) + B(t)y^i\|}$, where $y^1, y^2, \dots, y^n \in \mathbb{R}^{4m}$ satisfy the linear system of equations corresponding to (4.1), (4.3):

$$\xi_i\circ B(t_i)(y^{i+1} - y^i) = \mathbf{0}, \quad \text{for } 1\le i\le n-1, \tag{5.1}$$
$$\xi_i\circ C(t_i)(y^{i+1} - y^i) = \mathbf{0}, \quad \text{for } 1\le i\le n-1, \tag{5.2}$$
$$\xi_i\circ D(t_i)(y^{i+1} - y^i) = \mathbf{0}, \quad \text{for } 1\le i\le n-1, \tag{5.3}$$
$$\xi_0\circ HB(0)y^1 = \mathbf{0}, \tag{5.4}$$
$$\xi_n\circ HB(T)y^n = \mathbf{0}, \tag{5.5}$$
$$\xi_0\circ B(0)y^1 = F(0, x_0), \tag{5.6}$$
$$\xi_i\circ B(t_i)y^i = F(t_i, x_i), \quad \text{for } 1\le i\le n. \tag{5.7}$$

So we have $(3n-1)m$ homogeneous linear equations and $(n+1)m$ nonhomogeneous linear equations for the $4mn$ real coordinates of the y^i where $1\le i\le n$.

The algorithm is illustrated in Example 1.1 of Section 1 for a case where $m = n = 2$ and $\omega = 1$.

6. Bi-invariant $SO(3)$

Considering Euclidean 3-space E^3 as a Lie algebra with respect to the cross-product, the adjoint action defines a Lie isomorphism $\mathrm{ad}: E^3 \to so(3)$. Corresponding under ad to the Euclidean inner product on E^3 is an ad-invariant inner product on $so(3)$, which extends by left multiplication to a bi-invariant Riemannian metric $\langle\ ,\ \rangle$ on $M = SO(3)$. For a nonconstant geodesic $\gamma: [0,T] \to SO(3)$ satisfying $\gamma(0) = \mathbf{1}$, we write $\gamma^{(1)}(t) = \omega E_1(t)$ where $E_1(t) \in TSO(3)_{\gamma(t)}$ is a unit vector and $\omega > 0$.

Extend E_1 to an orthonormal basis $\{E_1, E_2, E_3\}$ of fields translated parallely along γ, and ordered so that $[E_1(0), E_2(0)] = E_3(0)$. Any vector field defined along γ has the form $X(t) = \sum_{k=1}^3 x_k(t) E_k(t)$, where $x_k: [0,T] \to \mathbb{R}$ for $1 \le k \le 3$. Then $\nabla_t X = \sum_{k=1}^3 x_k^{(1)} E_k$. So E_1 is in the kernel of S_t, and E_2, E_3 are also eigenvectors of S_t, with eigenvalues $\omega^2/4$. So

$$HX = \sum_{k=1}^3 x_k^{(2)} E_k + \frac{\omega^2}{4}(x_2 E_2 + x_3 E_3) \implies$$

$$H^2 X = \sum_{k=1}^3 x_k^{(4)} E_k + \frac{\omega^2}{2}(x_2^{(2)} E_2 + x_3^{(2)} E_3) + \frac{\omega^4}{16}(x_2 E_2 + x_3 E_3).$$

So X is a bi-Jacobi field when both x_1 is a cubic polynomial and x_2, x_3 satisfy $x^{(4)} + \frac{\omega^2}{2} x^{(2)} + \frac{\omega^4}{16} x = 0$. A basis of bi-Jacobi fields along γ is given by $B_1(t) = E_1$, $B_2(t) = tE_1$, $B_7(t) = t^2 E_1$, $B_8(t) = t^3 E_1$, and $B_9(t) = tB_3(t)$, $B_{10}(t) = tB_4(t)$, $B_{11}(t) = tB_5(t)$, $B_{12}(t) = tB_6(t)$ with

$$B_3(t) = E_2 \cos\frac{\omega t}{2} + E_3 \sin\frac{\omega t}{2}, \qquad B_4(t) = -E_2 \sin\frac{\omega t}{2} + E_3 \cos\frac{\omega t}{2},$$
$$B_5(t) = E_2 \sin\frac{\omega t}{2} + E_3 \cos\frac{\omega t}{2}, \qquad B_6(t) = -E_2 \cos\frac{\omega t}{2} + E_3 \sin\frac{\omega t}{2}.$$

Here B_1, B_2, B_3, B_4, B_5, and B_6 are also Jacobi fields. Define a linear isomorphism $E(t): \mathbb{R}^3 \to TSO(3)_{\gamma(t)}$ by $E(t)(v) := \begin{bmatrix} E_1(t) & E_2(t) & E_3(t) \end{bmatrix} v$, $v \in \mathbb{R}^3$, and set $D := \begin{bmatrix} 0&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}$.

Then $B(t) = E(t)\tilde{B}(t)$, $\nabla_t B(t) = E(t)\tilde{B}^{(1)}(t)$, $\nabla_t^2 B(t) = E(t)\tilde{B}^{(2)}(t)$, and

$$HB(t) = E(t)(\tilde{B}^{(2)}(t) + \frac{\omega^2}{4} D\tilde{B}(t)), \qquad \text{where}$$

$$\tilde{B}(t) := \begin{bmatrix} 1 & t & 0 & 0 & 0 & 0 & t^2 & t^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos\frac{\omega t}{2} & -\sin\frac{\omega t}{2} & \sin\frac{\omega t}{2} & -\cos\frac{\omega t}{2} & 0 & 0 & t\cos\frac{\omega t}{2} & -t\sin\frac{\omega t}{2} & t\sin\frac{\omega t}{2} & -t\cos\frac{\omega t}{2} \\ 0 & 0 & \sin\frac{\omega t}{2} & \cos\frac{\omega t}{2} & \cos\frac{\omega t}{2} & \sin\frac{\omega t}{2} & 0 & 0 & t\sin\frac{\omega t}{2} & t\cos\frac{\omega t}{2} & t\cos\frac{\omega t}{2} & t\sin\frac{\omega t}{2} \end{bmatrix}.$$

Equations (4.1), (4.3) for the $y^k \in \mathbb{R}^{12}$ where $1 \le k \le n$ read as

$$\begin{array}{l} \tilde{B}(t_i) y^i = \tilde{B}(t_i) y^{i+1}, \\ \tilde{B}^{(1)}(t_i) y^i = \tilde{B}^{(1)}(t_i) y^{i+1}, \\ \tilde{B}^{(2)}(t_i) y^i = \tilde{B}^{(2)}(t_i) y^{i+1}, \end{array} \quad \text{where } 1 \le i \le n-1, \quad \begin{array}{l} (\tilde{B}^{(2)}(0) + \frac{\omega^2}{4} D\tilde{B}(0)) y^1 = \mathbf{0}, \\ (\tilde{B}^{(2)}(T) + \frac{\omega^2}{4} D\tilde{B}(T)) y^n = \mathbf{0} \end{array} \tag{6.1}$$

and (still with $y^0 := y^1$)

$$E(t_i)\tilde{B}(t_i)y^i = -2\gamma(t_i)(\mathbf{1}+\gamma(t_i)^{-1}x(t_i))^{-1}(\mathbf{1}-\gamma(t_i)^{-1}x(t_i)) \text{ for } 0 \le i \le n \qquad (6.2)$$

using the quasiexponential function from Example 3.3.

LEMMA 6.1.

$$E(t) = \gamma(t)E(0)\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\frac{\omega t}{2} & \sin\frac{\omega t}{2} \\ 0 & -\sin\frac{\omega t}{2} & \cos\frac{\omega t}{2} \end{bmatrix}.$$

Proof. For $p = 1,2,3$ denote by $\tilde{E}_p(t)$ the left Lie reduction $\gamma(t)^{-1}E_p(t)$ of E_p along γ. Then $\tilde{E}_p : [0,T] \to so(3)$, $\tilde{E}_p(0) = E_p(0)$, and $\tilde{E}_1(t) = E_1(0)$. The left Lie reduction of $\nabla_t E_p$ is

$$\mathbf{0} = \tilde{E}_p^{(1)}(t) + \frac{\omega}{2}\mathrm{ad}(E_1(0))\tilde{E}_p(t) \implies$$

$$\tilde{E}_2 = E_2(0)\cos\frac{\omega t}{2} - E_3(0)\sin\frac{\omega t}{2},\ \tilde{E}_3 = E_2(0)\sin\frac{\omega t}{2} + E_3(0)\cos\frac{\omega t}{2}.$$

□

From Lemma 6.1 we find $E(t)\tilde{B}(t) = \gamma(t)C(t)$, where

$$C(t) := E(0)\begin{bmatrix} 1 & t & 0 & 0 & 0 & 0 & t^2 & t^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \sin\omega t & -\cos\omega t & 0 & 0 & t & 0 & t\sin\omega t & -t\cos\omega t \\ 0 & 0 & 0 & 1 & \cos\omega t & \sin\omega t & 0 & 0 & 0 & t & t\cos\omega t & t\sin\omega t \end{bmatrix}. \qquad (6.3)$$

Then equation (6.2) is equivalent to

$$C(t_i)y^i = -2(\mathbf{1}+\gamma(t_i)^{-1}x(t_i))^{-1}(\mathbf{1}-\gamma(t_i)^{-1}x(t_i)) \text{ for } 0 \le i \le n. \qquad (6.4)$$

Once the linear system (6.1), (6.4) is solved for the y^k we have, for $t \in [t_{i-1}, t_i]$ where $1 \le i \le n$,

$$\begin{aligned} x_{\tilde{Y}(D)}(t) &= \gamma(t)(\mathbf{1}+\gamma(t)^{-1}E(t)\tilde{B}(t)y^i/2)(\mathbf{1}-\gamma(t)^{-1}E(t)\tilde{B}(t)y^i/2)^{-1} \\ &= \gamma(t)(\mathbf{1}+C(t)y^i/2)(\mathbf{1}-C(t)y^i/2)^{-1}. \end{aligned} \qquad (6.5)$$

EXAMPLE 6.1. Let $\gamma : [0,T] \to SO(3)$ be the geodesic given by $\gamma(t) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega t & -\sin\omega t \\ 0 & \sin\omega t & \cos\omega t \end{bmatrix}$, where $T = 2.0$ and $\omega = 0.8$. Then

$$\gamma^{(1)}(0) = \omega\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \omega\ \mathrm{ad}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \implies E_1(t) = -\begin{bmatrix} 0 & 0 & 0 \\ 0 & \sin\omega t & \cos\omega t \\ 0 & -\cos\omega t & \sin\omega t \end{bmatrix}$$

and $\gamma^{(1)}(t)$ has length ω.

With $E_2(0) := \mathrm{ad}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$, $E_3(0) := \mathrm{ad}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ we have

$$E(0)\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}.$$

Taking $n=4$, set $t_0=0$, $t_1=0.206811$, $t_2=0.43208$, $t_3=0.89426$, $t_4=2$, and the x_i as

$$\begin{bmatrix} .996136 & -.0725445 & .0494948 \\ .0701445 & .996351 & .0486186 \\ -.0528412 & -.0449589 & .99759 \end{bmatrix}, \quad \begin{bmatrix} .999253 & -.0184091 & -.0339881 \\ .00628044 & .944945 & -.32717 \\ .0381398 & .326712 & .944354 \end{bmatrix},$$

$$\begin{bmatrix} .996511 & -.0803521 & -.0225835 \\ .0365764 & .663613 & -.747181 \\ .0750243 & .743748 & .664236 \end{bmatrix},$$

$$\begin{bmatrix} .998571 & .0501545 & -.0184616 \\ -.0158622 & -.0517393 & -.998535 \\ -.0510362 & .9974 & -.0508698 \end{bmatrix}, \quad \begin{bmatrix} .998105 & .0308313 & -.0532549 \\ .0531121 & -.868682 & .492515 \\ -.0310767 & -.49441 & -.868673 \end{bmatrix}$$

respectively. With this choice of data D, figure 6.1 shows the curves in S^2 obtained by applying $\gamma(t)$ (red) and $x_{\tilde{Y}(D)}(t)$ (black) to a reference point $p_r=(0,1,0)$. The $x_i p_r \in S^2$ are labelled i where $0 \le i \le 4$, and the data is neither far from the geodesic nor especially close.

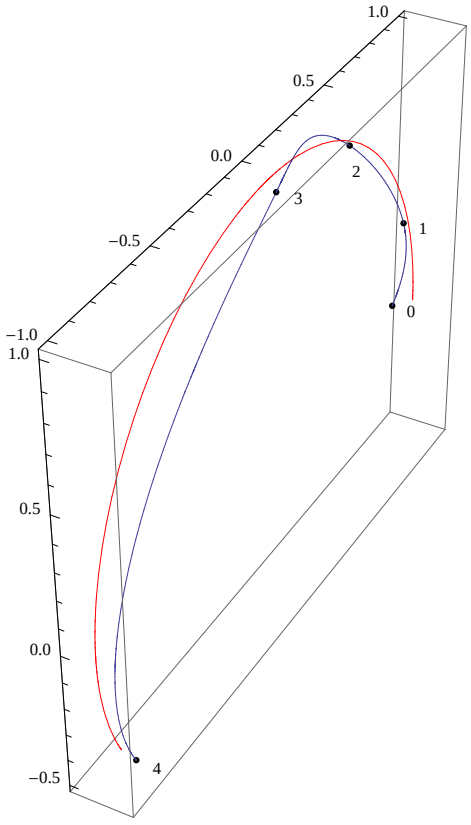

FIG. 6.1. *Geodesic, nearly cubic interpolant in Example 6.1.*

Chart based methods can also be used to interpolate in $SO(3)$. Suppose given a diffeomorphism $\phi : U \to V$ where U and V are open subsets of $SO(3)$ and E^3, with $x_0, x_1, x_2, x_3, x_4 \in U$. Then the natural Euclidean cubic spline $x_E : [t_0, t_4] \to E^3$ for the data

$$(t_0, \phi(x_0)),\ (t_1, \phi(x_1)),\ (t_2, \phi(x_2)),\ (t_3, \phi(x_3)),\ (t_4, \phi(x_4))$$

is unique and easily calculated in 0.006 seconds on a 1.7GHz MacBook Air running Mathematica (no attempt was made to exploit sparseness of the linear system determining $\tilde{Y}(D)$). If $x_E([t_0, t_4]) \subset V$ then $x_{ch} := \phi^{-1} \circ x_E : [t_0, t_4] \to SO(3)$ is D-feasible. Define

$$\psi : V \subset E^3 \cong so(3) \to SO(3)$$

by $\psi(v) := (\mathbf{1} + v/2)(\mathbf{1} - v/2)^{-1}$, where $v \in so(3)$ and V is chosen so that ψ is one-to-one. Then set $\phi := \psi^{-1}$ with $U := \psi(V)$. Figure 6.2 shows the curves in S^2 obtained by applying $x_{\tilde{Y}(D)}(t)$ (black) and $x_{ch}(t)$ (green) to p_r.

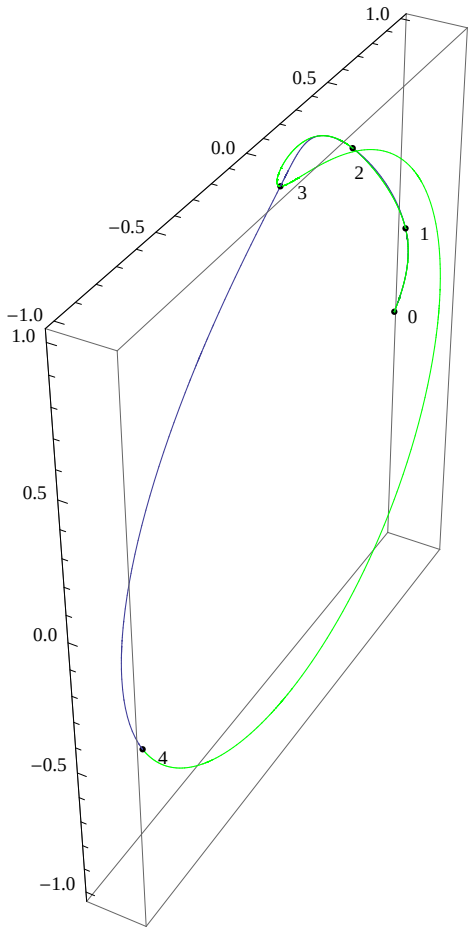

FIG. 6.2. *Nearly cubic and chart-based interpolants in Example 6.1.*

In figure 6.2 there is not much difference between the nearly cubic interpolant $x_{\tilde{Y}(D)}$ and the chart-based x_{ch} on the intervals $[t_{i-1},t_i]$ for $i=1,2$, but the difference is noticeable for $i=3$. When $i=4$ we find that x_{ch} takes a roundabout detour, as is not uncommon for chart-based interpolants. The nearly cubic interpolant $x_{\tilde{Y}(D)}$ is intrinsic and better behaved.

Because $x_{\tilde{Y}(D)}$ is nearly cubic we expect $J(x_{\tilde{Y}(D)})<J(x_{ch})$, and indeed $J(x_{\tilde{Y}(D)})=8.49251<<J(x_{ch})=343.664$. To see how much of the relatively poor performance of x_{ch} is due to the detour on $[t_3,t_4]$ we tabulate the contributions to J for both curves from each of the subintervals $I=[t_{i-1},t_i]$ for $1\le i\le 4$.

I	$J(x_{\tilde{Y}(D)}\|I)$	$J(x_{ch}\|I)$
$[t_0,t_1]$	2.67332	2.05117
$[t_1,t_2]$	2.63957	4.03758
$[t_2,t_3]$	2.63603	23.5679
$[t_3,t_4]$	0.543595	314.007

The detour when $i=4$ is very costly for x_{ch}, but $x_{\tilde{Y}(D)}$ also outperforms x_{ch} when $i=2,3$. To a smaller extent x_{ch} outperforms $x_{\tilde{Y}(D)}$ when $i=1$. On balance, even excluding $[t_3,t_4]$, the nearly cubic interpolant $x_{\tilde{Y}(D)}$ outperforms x_{ch} in terms of J. Note that J measures cumulative squared accelerations in $SO(3)$, including distortions due to parameterizations, whereas figure 6.2 depicts only the unparameterized images in S^2 of curves in $SO(3)$.

Figure 6.2 certainly depicts an extreme state of affairs. For natural splines (as opposed to splines with prescribed end velocities) there is no condition on the end velocities, and so it is unsurprising that the end velocities for the representations of $x_{\tilde{Y}(D)}$ and x_{ch} appear different at t_4, but it is remarkable that they appear so different (nearly opposite). This is related to the surprising detour made by x_{ch} on the interval $[t_3,t_4]$, and it may be useful to make some heuristic remarks, explaining how the detour came about.

The key reason is that x_0 and x_4 are nearly as far away as it is possible to be in $SO(3)$, because $T\omega\approx\pi/2$. Therefore, and because the chart-based interpolant has little regard for geometry, there was a near even chance of going the wrong way round $SO(3)=\mathbb{R}P^3$, as happened on our first attempt. This illustrates dramatically the advantages of a geometrically defined interpolant such as $x_{\tilde{Y}(D)}$. If our data had been chosen randomly there would still have been a noticeable advantage with $J(x_{\tilde{Y}(D)})<J(x_{ch})$, but not so dramatic (maybe no detour for x_{ch}), and if the data had been very localised the advantage would have been correspondingly small. Generally speaking, depending on the application, the pathology observed in figure 6.2 for chart-based interpolants should not be unusual.

7. Conclusion

A method is given for calculating approximations to natural Riemannian cubic splines in symmetric spaces such as S^m and $SO(p)$, with computational effort comparable to what is needed for the classical case of a natural cubic spline in Euclidean space. Our method offers significant improvements over standard chart-based interpolants.

Further work is needed to study the condition that τ is ncs-nonconjugate, and take account of sparseness of the linear system for calculating $\tilde{Y}(D)$. Whereas our nearly cubic splines approximately satisfy the necessary conditions for a natural cubic spline, there is also scope for investigation of the suboptimal value achieved for J. Finally there is a need to find I_δ for left-invariant Riemannian metrics on $SO(3)$, rather than

just the bi-invariant case of Section 6. This reduces to computation of bi-Jacobi fields for left-invariant Riemannian metrics, which is the subject of ongoing work with Tudor Ratiu.

Acknowledgment. I am grateful to the reviewer for a thoughtful reading and helpful suggestions which have improved the presentation.

REFERENCES

[1] A.H. Barr, B. Currin, S. Gabriel, and J.F. Hughes, *Smooth interpolation of orientations with angular velocity constraints using quaternions*, ACM SIGGRAPH Computer Graphics, 26(2), 313–320, 1992.
[2] A.M. Bloch, J. Baillieul, P. Crouch, and J. Marsden, *Nonholonomic Mechanics and Control*, Inter. Appl. Math., Springer, 2003.
[3] M. Camarinha, F. Silva Leite, and P. Crouch, *On the geometry of Riemannian cubic polynomials*, Diff. Geom. Appl., 15(2), 107–135, 2001.
[4] P. Crouch and F. Silva Leite, *The dynamic interpolation problem: On Riemannian manifolds, Lie groups, and symmetric spaces*, J. Dynam. Control Systems, 1(2), 177–202, 1995.
[5] C. De Boor, *A Practical Guide to Splines*, Appl. Math. Sci., Springer-Verlag New York, 27, 2001.
[6] R. Giambo, F. Giannoni, and P. Piccione, *An analytical theory for Riemannian cubic polynomials*, IMA J. Math. Cont. Inform., 19, 445–460, 2002.
[7] R. Ramamoorthi and A. Barr, *Fast construction of accurate quaternion splines*, in Proc. of SIGGRAPH'97, Los Angeles, 287–292, 1997.
[8] S. Gabriel and J. Kajiya, *Spline interpolation in curved space*, in SIGGRAPH 85 course notes for State of the Art in Image Synthesis, 11, 1985.
[9] S. Gutiérrez, J. Rivas, and L. Vega, *Formation of singularities and self-similar vortex motion under the localized induction approximation*, Commun. Part. Diff. Eqs., 28, 927–968, 2003.
[10] D. Holm, T. Schmah, and C. Stoica, *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*, Oxford Texts in Applied and Engineering Mathematics, 2009.
[11] V. Jurdjevic, *Non-Euclidean elastica*, Amer. J. Math., 117, 93–124, 1995.
[12] H.B. Keller, *Numerical Methods for Two-Point Boundary-Value Problems*, Waltham Mass. Blaisdell, 1968.
[13] W. Lawton and L. Noakes, *Computing the inertia matrix of a rigid body*, J. Math. Phys., 42(4), 1655–1665, 2001.
[14] L. Noakes, G. Heinzinger, and B. Paden, *Cubic splines on curved spaces*, IMA J. Math. Cont. Inform., 6, 465–473, 1989.
[15] L. Noakes, *Null cubics and Lie quadratics*, J. Math. Phys., 44(3), 1436–1448, 2003.
[16] L. Noakes, *Non-null Lie quadratics in E^3*, J. Math. Phys., 45(11), 4334–4351, 2004.
[17] L. Noakes, *Duality and Riemannian cubics*, Adv. Comput. Math., 25, 195–209, 2006.
[18] L. Noakes, *Lax constraints in semisimple Lie groups*, Quart. J. Math., 57, 527–538, 2006.
[19] L. Noakes, *Asymptotics of null Lie quadratics in E^3*, SIAM J. Appl. Dynamical Systems, 7(2), 437–460, 2008.
[20] L. Noakes and T. Popiel, *Geometry for robot path planning*, Robotica, 25, 691–701, 2007.
[21] M. Pauley, *Cubics, Curvature and Asymptotics*, PhD Thesis, University of Western Australia, 2011.
[22] T. Popiel and L. Noakes, *Elastica in $SO(3)$*, J. Australian Math. Soc., 83, 105–125, 2007.
[23] D.A. Singer, *Lectures on elastic curves and rods*, Curvature and Variational Modeling in Physics and Biophysics, AIP Conf. Proc., 1002, 3–32, 2008.