

Review of Interpolation in Special Orthogonal Groups [2]

Edward Samson

June 3, 2011

Contents

1	Abstract	3
2	Introduction	3
2.1	Groups	3
2.2	Lie Groups	4
2.3	Special Orthogonal Group and Operators	4
3	Interpolation	6
3.1	Classical approach to interpolation	6
4	Calculations	7
4.1	Results	8
5	Iterative Improvements	8
6	Improving the iteration	10
7	Conclusion	11

1 Abstract

The purpose of this paper is to discuss the results that T. Shingel produced in the paper, Interpolation in special orthogonal groups [2]. In physics, underlying symmetries in the $SO(3 + 1)$ group explain numerous properties of particles and systems of particles. The primary goal of Shingel's paper was to develop a technique capable of interpolating every element of the $SO(n)$ group. The developed method is iterative and quadratically convergent for the entire group, and using Padé approximation, becomes computationally inexpensive. Using the surjective relationship between the Lie group $SO(n)$ and the Lie algebra $\mathfrak{so}(n)$, we summarize the process Shingel uses to create an exponential representation of any $Q \in SO(n)$ and then the iterative method which allows us to generate $A_{r+1} \in \mathfrak{so}(n)$ from A_r and Q_r .

2 Introduction

Before we can begin to summarize Shingel's work, it is important to build up an understanding of some of the fundamental properties of Lie groups and Lie algebras. The following summary of group theory and Lie groups expands from Sephen Ellis' of the University of Washington [1].

2.1 Groups

Groups are useful tools for describing a number of different elements. In general we will focus on the specific $SO(n)$ group, with each element being a matrix. A group, G , is defined as a set of elements with an associated operator, \star . In addition, the operator must satisfy several properties, First off, the operator must be closed, such that

$$g_x \star g_y \in G \forall g_x, g_y \in G$$

Second, the operator must be associative, such that

$$(g_x \star g_y) \star g_z = g_x \star (g_y \star g_z) \forall g_x, g_y, g_z \in G$$

Along with these restrictions on the operator, there are two other restrictions on the elements in any group. For any valid group there must exist an identity

element and an inverse element for all elements.

$$\begin{aligned} 1 \star g_x &= g_x & \forall g_x \in G \\ \exists g_y : g_x \star g_y &= 1 & g_x, g_y \in G \end{aligned}$$

Groups can have a finite or infinite number of elements. Groups that have an infinite number of elements are often parametrized, where any element in the group can be written in terms of a finite number of parameters.

2.2 Lie Groups

Lie groups are a specific type of group where given a set of parameters \mathbf{t} , there exists a finite number of generators that are differentiable near the identity element I . A generator can be written as

$$X_n = \lim_{|\mathbf{t}| \rightarrow 0} \frac{\partial p}{\partial t_n}$$

The origin is unique for Lie groups because the only element in the group must be the identity element, I . Therefore, generators are the infinitesimal changes near the origin which give us all the elements of a Lie group. In addition, for every Lie group, there exists a complimentary Lie algebra. The Lie algebra is similar to a group as it is a collection of elements, but instead of an associated \star operator, there is an associated commutator, $[\cdot, \cdot]$. This commutator also has several useful properties. First it must take two elements of the algebra back into the algebra

$$[a_x, a_y] \in A \forall a_x, a_y \in A$$

Secondly it must equal exactly zero for any two similar elements. And finally it must satisfy the Jacobi identity.

$$[a_x, [a_y, a_z]] + [a_y, [a_z, a_x]] + [a_z, [a_x, a_y]] = 0$$

2.3 Special Orthogonal Group and Operators

Specifically important to physics, there is a special Lie group called the Special Orthogonal Group or $SO(n)$. n denotes the number of dimensions of the

group. The most relevant group is the $SO(3)$ which describes the underlying symmetries of particle physics. For the Lie group $SO(n)$ the corresponding Lie algebra is known as $\mathfrak{so}(n)$. Specifically $SO(n)$ is defined to be the set of all matrices such that

$$A^T A = A A^T = I \forall A \in SO(n)$$

In addition to the group, there are several important operators that we will define for $SO(n)$ and $\mathfrak{so}(n)$. Shingel uses these operators frequently to simplify the expression of complex matrix operations.

$$|A| = \max_{j \neq 0} \frac{|A_x|_2}{|x|_2} = \sqrt{(\text{Max eigenvalue of } A A^T)}$$

For exp and log we can define them by their Taylor series.

$$\exp : A \in \mathfrak{so}(n) \rightarrow \exp(A) \in SO(n)$$

$$\exp(A) = \sum_0^{\infty} \frac{A^k}{k!}$$

$$\log : A \in SO(n) \rightarrow \log(A) \in \mathfrak{so}(n)$$

$$\log A = \sum_1^{\infty} (-1)^{k+1} \frac{(A - \mathbf{1})^k}{k}, \forall |A - \mathbf{1}| < 1$$

Similarly, Shingel defines the following operators to refer to conjugation and commutation.

$$\text{Ad}_Q E = Q E Q^{-1} \forall Q \in SO(n), E \in \mathfrak{so}(n)$$

$$\text{ad}_A E = [A, E] \forall A, E \in \mathfrak{so}(n)$$

And finally, the useful dexp operator. The purpose of this operator is beyond the scope of this paper.

$$\text{dexp}_A = \frac{e^u - 1}{u} \Big|_{u=\text{ad}_A} \quad \forall A \in \mathfrak{so}(n)$$

3 Interpolation

The main problem presented in the paper was a way to find a function that smoothly interpolates every element of $SO(n)$. Similar to how we can find generators for any infinite group, we want to find a function that maps an interval to the group. The goal of this paper is to find a suitable method to find an interpolation function of $SO(n)$. Given that $Q_1, Q_2, \dots, Q_m \in SO(n)$, is an ordered set of elements and $t_1 < t_2 < \dots < t_m \in [0, 1]$, we want to find $F : [0, 1] \rightarrow SO(n)$ such that

$$F(t_k) = Q_k \forall k = 1, \dots, m$$

To accomplish this, Shingel will use the uniqueness of the surjective exp map that takes $\mathfrak{so}(n) \rightarrow SO(n)$. The useful property of $\mathfrak{so}(n)$ is that it can be embedded in a euclidean space, which are easy to work in. However, finding a proper interpolation is not as simple as using the exponential map and interpolating the Lie algebra. Special considerations have to be taken into account about what correct branch of log must be used along with other constraints. Most importantly, Shingel's method provides a global interpolation, meaning every element in $SO(n)$ is interpolated. Without this method, small errors could only be achieved on sub-groups of $SO(n)$.

3.1 Classical approach to interpolation

The quickest, and easiest approach to this solution, as Shingel points out, does not necessarily lead to the correct answer. For example we can use matrix factorization on, $\gamma(t)v$ where $v \in S^2$, S^2 is the projection of a two dimensional sphere, and $\gamma(\mathbf{t})$ is a parametrized curve. If we factorize the matrix representation of γ into $\gamma(t) = Q(t)R(t)$ then we can interpolate just $Q(t)$ or $R(t)$, however the interpolation of either of $Q(t)v$ or $R(t)v$ leads to large errors particularly near the equator of the sphere. As the figure below shows the difference between the actual points and the spiral curve interpolation.

The large error accumulated by this simple interpolation comes from the fact that at times, we must cross the boundray of a slit domain for log. Shingel's method fixes this error by finding a way to smoothly transition to the correct branch always.

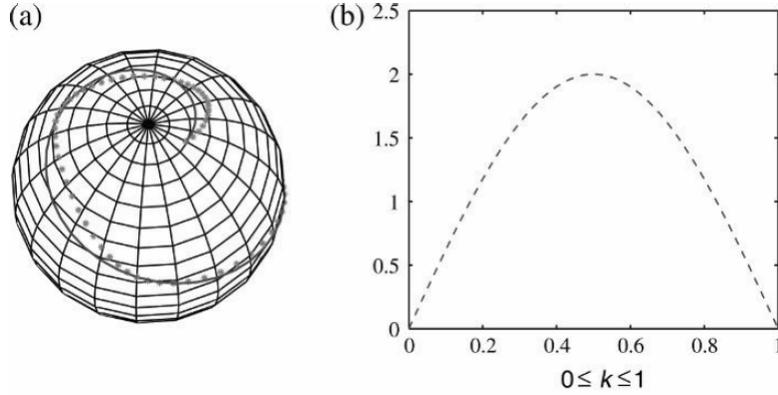


Figure 1: (a) A graph of the spiral curve with dots at every point, and the interpolated curve, the solid line. (b) Graph of the error on the vertical axis against the parameter k [2].

4 Calculations

Shingel suggests to solve the interpolation problem, to first construct a function that maps $L_A : \mathfrak{so}(n) \rightarrow \mathfrak{so}$ in order to simplify the problem.

$$e^{\frac{-t}{2}A} e^{t(A+E)} e^{\frac{-t}{2}A} = e^{L_A t E} + R$$

Where $A, E \in \mathfrak{SO}$ and R is a remainder term which depends on $|A|$ and $|E|$. After solving the equation using a substitution and the Baker-Campbell-Hausdorff formula, the general solution of the form $Y(t) = e^{\Omega(t)}$ where

$$Y(t) = e^{\frac{-t}{2}A} e^{t(A+E)} e^{\frac{-t}{2}A}$$

From the original equation, Shingel re-writes it in terms of a differential equation, with $Y(t)$ and two other function $U(t)$ and $W(t)$ where

$$U(t) = \frac{1}{2} e^{-t \text{ad}_{\frac{A}{2}}}$$

$$W(t) = -\frac{1}{2} e^{t \text{ad}_{\frac{A}{2}}}$$

In this case $\text{ad}_E A$ refers to the commutator $[E, A]$. Drawing from solutions from several other sources, Shingel gets a solution in the form

$$\Omega(t) = \sum_1^{\infty} \Omega_k(t)$$

Each Ω_k can be written in terms of $U(t)$ and $W(t)$ and therefore the series becomes

$$\Omega_k = \sum_0^{\infty} \frac{t^{2k}}{(2k)!} \text{ad}_{\frac{1}{2}A}^{2k} E$$

From this formula now Shingel goes on to describe the first few terms in the series. The key is to realize we can generate an expression for the exponential representation of any element using the $e^{\frac{-t}{2}A} Q e^{\frac{-t}{2}A} |_{t=1}$ form.

4.1 Results

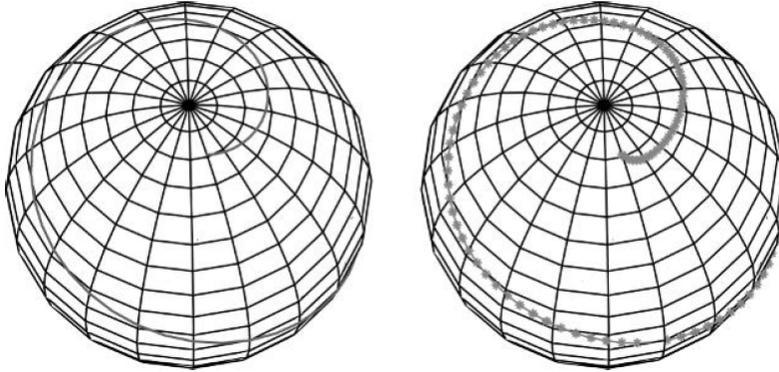


Figure 2: (a) Graph showing the comparison of the interpolated curve after a few iterations (left) and the points from which it was generated (right). The curve lies identically over the top of the points. [2].

Shingel's method utilizes the unique properties that relate the Lie algebra to the Lie group. We know the exponential map, $\exp : \mathfrak{so}(n) \rightarrow SO(n)$ is surjective, therefore.

$$\forall X \in SO(n) \exists Y \in \mathfrak{so}(n) \text{ s.t. } \exp(Y) = X$$

This method simplifies the interpolation problem since we can simply interpolate $\mathfrak{so}(n)$ which exists in euclidean space unlike $SO(n)$. Therefore, we can say $Q_r = e^{A_r}$, and to solve for A_r given a Q_r we need to use the logarithm. This causes a problem however, since we are forced to work on a slit domain. For now, it remains easiest to excise the entire negative real axis from our domain, so -1 is not the eigenvalue of any possible $Q \in SO(n)$. However, even if we fix the "correct" branch of the logarithm and ensure that every matrix has eigenvalues that work, the interpolated curve will only give a good local approximation. This is the key problem Shingel's paper tries to resolve. To get an accurate approximation for the entire $SO(n)$ group that works on a global scale, Shingel develops an iterative method which gives us a bound on the error.

5 Iterative Improvements

One key problem to address is the issue of branch cuts for the logarithm. We are looking to find a way to move away from the principle branch of the log and transition smoothly to the "correct" branch of the log when necessary. Start by assuming that for $Q_r, Q_{r+1} \in SO(n)$ are sufficiently close, such that

$$|Q_r^{-1}Q_{r+1} - I| < \epsilon \forall \epsilon > 0$$

Then we can work in the principle branch of the log and solve for Q_r in $Q_r = e^{A_r}$. As an initial value, WLOG we can set $Q_1 = I$ and $A_1 = O$. Then define E_{r+1} to satisfy $A_{r+1} = A_r + E_{r+1}$. Similarly we want the norm of E to be bounded and small. Shingel shows that E_{r+1} can be approximated by

$$\tilde{E}_{r+1} = T_{A_r} \log \left(e^{\frac{-1}{2}A_r} Q_{r+1} e^{\frac{-1}{2}A_r} \right)$$

Here $T_{A_r} = \frac{\frac{1}{2}\text{ad}_{A_r}}{\sinh(\frac{1}{2}\text{ad}_{A_r})}$. This T_{A_r} operator is defined by the values of

$f(z) = \frac{\frac{z}{2}}{\sinh \frac{z}{2}}$ for any eigenvalue of ad_{A_r} . Two important things to notice are the use of the exponential representation of Q and $f(z)$. The role of $f(z)$ will become more clear later in this section. This function has poles at $x = 2\pi ik$ with $k \in \mathbb{Z} \setminus 0$. If we do not allow x to be any of these singular points, then we can bound T_{A_r} and then get a bound on E_{r+1} . Shingel proves

that this bound is

$$\tilde{E}_{r+1} = E_{r+1} + \mathcal{O}(|T_{A_r}||E_{r+1}|^2)$$

Now, after some more algebra we can establish an error bound for the difference of our approximation and the actual value of E .

$$|e^{A_r+E_{r+1}} - e^{A_r+\tilde{E}_{r+1}}| \leq C|\tilde{E}_{r+1}|^2$$

Since we know $e^{\frac{-1}{2}A}$ is an orthogonal matrix, we can express

$$|e^{A_r+E_{r+1}} - e^{A_r+\tilde{E}_{r+1}}| = |e^{\frac{-1}{2}A_r}e^{A_r+E_{r+1}}e^{\frac{-1}{2}A_r} - e^{\frac{-1}{2}A_r}e^{A_r+\tilde{E}_{r+1}}e^{\frac{-1}{2}A_r}|$$

If we take the norm \tilde{E}_{r+1} to be small enough and work on the principle branch of the log

$$e^{(\frac{-1}{2}A_r+\tilde{E}_{r+1})}Q_{r+1}e^{(\frac{-1}{2}A_r+\tilde{E}_{r+1})}$$

This shows that now we can have an iterative algorithm, $A_r + \tilde{E}_{r+1}$, thus decreasing the error of the interpolation with every successive iteration. If we perform the same steps except denote each iteration with script $[n]$, Shingel shows that the iteration is quadratically convergent.

$$A_{r+1}^{[1]} = A_r$$

$$A_{r+1}^{[k+1]} = A_{r+1}^{[k]} + E_{r+1}^{[k]}$$

Starting with the same calculations as above, we will drop the $r+1$ subscript as Shingel does, and let A, E be any corresponding elements.

$$\begin{aligned} e^{\frac{-1}{2}(A^{[1]}+E^{[1]})}Qe^{\frac{-1}{2}(A^{[1]}+E^{[1]})} &= e^{-\frac{1}{2}(A^{[1]}+E^{[1]})}e^{A^{[1]}+E^{[1]}+E-E^{[1]}}e^{-\frac{1}{2}(A^{[1]}+E^{[1]})} \\ &= \exp\left(\frac{\sinh\frac{1}{2}\text{ad}_{A^{[1]}+E^{[1]}}}{\frac{1}{2}\text{ad}_{A^{[1]}+E^{[1]}}}(E-E^{[1]})\right) + \mathcal{O}(|E-E^{[1]}|^2) \end{aligned}$$

From our earlier calculation this gives us the change in E with each iteration

$$E^{[2]} = E - E^{[1]} + \mathcal{O}(|T_A^{[2]}||E-E^{[1]}|^2)$$

From there, Shingel proceeds by induction to prove for any $k \geq 2$

$$E^{[k]} = E - \sum_1^{k-1} E^{[j]} + \mathcal{O}\left(|T_A^{[k]}||E - \sum_1^{k-1} E^{[j]}|^2\right)$$

Then with a simple substitution, we get back the formula for the difference between iterations

$$A_{r+1} - A_{r+1}^{[k+1]} = \mathcal{O} \left(|T_{A_{r+1}^{[k]}}| |A_{r+1} - A_{r+1}^{[k]}|^2 \right)$$

Now we have a algorithm and iterative method which converges quadratically. So long as we stay away from the singularities $x = 2\pi i k$ $k \in \mathbb{Z}$, then the interpolation function is valid and we have the desired convergence.

6 Improving the iteration

Implementation of the iterative method mentioned relies on computationally expensive eigenvalue-eigenvector decomposition of ad_{A_r} . To decrease the cost of this calculation Shingel suggests using the reduced commutator matrix and the diagonal Padé approximation. Here we will expand on Shingel's paper and give a more general definition of Pade's approximation. Padé's approximation states, for a given R , $R(z) = \frac{f(z)}{g(z)}$ where $f(z), g(z) \in \mathcal{O}(D)$ and D is a domain, then

$$R(z) = \frac{\sum_1^m a_k z^k}{\sum_1^n b_j z^j}$$

where m, n correspond to the degree of the polynomial representation of $f(z)$ and $g(z)$ [3]. In the case where $m = n$ Padé's approximation is referred to as a diagonal approximation. Shingel uses Padé's approximation for $f(z) = \frac{\frac{z}{2}}{\sinh \frac{z}{2}}$, letting $\frac{N(C_{A^{[k]}})}{P(C_{A^{[k]}})}$ be the diagonal approximation of $f(C_{A^{[k]}})$. C_A is the reduced commutator matrix. The commutator matrix has the special property that given a mapping v such that $v : \mathfrak{so}(n) \rightarrow \mathbb{R}^m$, and $m = \frac{1}{2}n(n-1)$

$$v(\text{ad}_A B) = C_A v(B) \forall B \in \mathfrak{so}(n)$$

If we apply this to $B^{[k]} = \log \left(e^{\frac{-1}{2}A^{[k]}} Q e^{\frac{-1}{2}A^{[k]}} \right)$, and apply Padé's approximation to $f(C_A^{[k]})$ we get

$$P(C_A^{[k]})v(E^{[k]}) = N(C_{A^{[k]}})v(B^{[k]})$$

After some algebraic simplifications, Shingel concludes, as a simplification to the iterative method, for any $Q \in SO(3)$ and $\theta : 1 + 2 \cos \theta = \text{Trace}(Q)$

$$\log Q = \frac{\theta}{2 \sin \theta} (Q - Q^T)$$

and

$$T_A = I + \left(\frac{1}{\theta \sin \theta} - \frac{1}{\theta^2} \right) \text{ad}_A^2$$

7 Conclusion

Shingel demonstrated how to generate a numerical method to find $A_{r+1} \in \mathfrak{so}(n)$, given two elements $Q_r, Q_{r+1} \in SO(n)$ which were within a neighbourhood of each other. Using the surjective nature of \exp , we concluded that there exists a numerical method to find A_{r+1} because of the exponential relation between $\mathfrak{so}(n)$ and $SO(n)$. Find the exponential representation of elements in $SO(n)$ we then found a map that embedded $\mathfrak{so}(n)$ in a euclidean m dimensional space, which then allowed for easy interpolation of the elements. This new method allowed for the interpolation of $SO(n)$ elements, and generated a numerical answer with an error that was quadratically convergent with each successive iteration. It also avoided issues of choosing the correct branch of the log and was a global interpolation for all elements of $SO(n)$. After expanding on Shingel's explanation of Padé's method, we showed also that several simplifications could be made to make the iterations less computationally expensive. Based from just a few simple properties, Shingel was able to derive a numerical interpolation of $SO(n)$ and $\mathfrak{so}(n)$.

References

- [1] ELLIS, S. D. Chapter 10: Introduction to group theory. Class notes from Physics 226.
- [2] SHINGEL, T. Interpolation in special orthogonal groups. *IMA Journal of Numerical Analysis* (2009).
- [3] VAJTA, M. Some remarks on pade approximations. In *TEMPUS-INTCOM Symposium*.