Averaging Rotations

Michael Taylor

1. Introduction

Suppose $A_1, \dots, A_N$ are rotation matrices on n-dimensional Euclidean space $\mathbb{R}^n$, i.e., $A_j \in SO(n)$. We want to consider some element of $SO(n)$ that represents an "average" of these elements A_j. There are a number of possible ways to define the notion of an average in this context. One approach has been to write $A_j = e^{Z_j}$ with Z_j a real, skew-symmetric $n \times n$ matrix (i.e., $Z_j \in \mathrm{skew}(n)$), and define an average as $e^{\overline{Z}}$, with $\overline{Z} = (Z_1 + \cdots + Z_N)/N$. This has ambiguities, arising from the fact that the exponential map $\exp : \mathrm{skew}(n) \to SO(n)$ is not one-to-one; one then needs to tackle the problem of finding the "best" candidates Z_j (logarithms of A_j) to produce the desired $e^{\overline{Z}}$; cf. [V]. Another drawback to this approach is mentioned in §4.

We discuss a different approach, considering the following minimization problem. Given $A_1, \dots, A_N \in SO(n)$, define

$$\psi(X) = \sum_{j=1}^{N} \|X - A_j\|^2, \tag{1.1}$$

and minimize this over $X \in SO(n)$. Here we take the Hilbert-Schmidt norm:

$$\|T\|^2 = \mathrm{Tr}\ T^* T. \tag{1.2}$$

It is clear that (1.1) has a minimum, though the minimizer might or might not be unique; more on that below. We mention parenthetically that if A_j, X belong to a vector space V, equipped with an inner product, and if (1.1) is minimized over all $X \in V$, then the unique minimizer is $X = (A_1 + \cdots + A_N)/N$. Of course, the current minimization problem is different, since we assume $A_j \in SO(n)$ and require $X \in SO(n)$.

If (1.1) has a minimum at $X \in SO(n)$, we say X is an R-average of $A_1, \dots, A_N$; if the minimizer X is unique, we say it is *the* R-average. The minimizer will be unique, except for exceptional choices of A_j.

The plan of this paper is the following. In §2 we establish some properties of the R-average. We show that it is determined by $\overline{A} = A_1 + \cdots + A_N$. How it is determined depends on whether the determinant of $\overline{A}$ is positive, negative, or zero. In §3 we give a number of examples of collections of elements of $SO(3)$ and compute their R-averages. In §4 we discuss the property of covariance. We show that the R-average has this property but the logarithmic average does not. In §5 we discuss how to implement the calculation of the R-average using Mathematica programs. We also discuss numerical calculation of the logarthmic average in some special cases, and compare the results with computations of the R-average. In §6 we discuss averages over other spaces, particularly the group $E(n)$ of rigid motions of Euclidean space $\mathbb{R}^n$, which is a semidirect product of $SO(n)$ and $\mathbb{R}^n$. We produce an average that is left-covariant and one that is right-covariant.

Finally, we attach three Mathematica notebooks, in which the numerical methods discussed in §5 are implemented.

2. The R-average

We tackle the problem of computing R-averages of sets of elements of $SO(n)$. To analyze (1.1), write

$$\begin{aligned} \|X - A_j\|^2 &= \operatorname{Tr}(X^* - A_j^*)(X - A_j) \\ &= \operatorname{Tr}(X^*X - X^*A_j - A_j^*X + A_j^*A_j) \\ &= 2n - 2\ \operatorname{Tr} A_j^*X, \end{aligned} \tag{2.1}$$

using $X^*X = A_j^*A_j = I$. Hence we have

$$\psi(X) = 2nN - 2\ \operatorname{Tr} \overline{A}^* X, \quad \overline{A} = A_1 + \cdots + A_N. \tag{2.2}$$

Thus the problem of minimizing (1.1) over $X \in SO(n)$ is equivalent to the problem:

$$\text{Maximize} \quad \operatorname{Tr} \overline{A}^* X \quad \text{over} \quad X \in SO(n). \tag{2.3}$$

We break the analysis into several cases:

CASE I. $\overline{A}$ is invertible.
Take the polar decomposition of $\overline{A}$:

$$\overline{A} = UP, \tag{2.4}$$

with U orthogonal and P positive definite. This polar decomposition is unique; in particular $P = (\overline{A}^* \overline{A})^{1/2}$. Then we are considering

$$\operatorname{Tr} \overline{A}^* X = \operatorname{Tr} PU^*X. \tag{2.5}$$

CASE IA. $\det \overline{A} > 0$.
In this case $U \in SO(n)$, so U^*X runs over $SO(n)$ in (2.5) as X runs over $SO(n)$, so the following result is useful.

Lemma 2.1. *If P is positive-definite on $\mathbb{R}^n$ and $V \in SO(n)$, then*

$$\operatorname{Tr} PV \le \operatorname{Tr} P, \tag{2.6}$$

with identity if and only if $V = I$.

Proof. Let $v_1, \dots, v_n$ be an orthonormal basis of $\mathbb{R}^n$, consisting of eigenvectors of P, $Pv_j = \lambda_j v_j$, $\lambda_j > 0$. Then

$$\operatorname{Tr} PV = \sum_j (PVv_j, v_j) = \sum_j \lambda_j (Vv_j, v_j). \tag{2.7}$$

We have $(Vv_j, v_j) \le 1$, with equality if and only if $Vv_j = v_j$, given $V \in SO(n)$, and this proves the lemma.

Corollary 2.2. *In Case IA, the minimum for (1.1) over $X \in SO(n)$ is achieved at one point:*

$$X = U, \tag{2.8}$$

with $U \in SO(n)$ given by the polar decomposition (2.4).

Case IB. det $\overline{A} < 0$.
In this case U in (2.4) is orthogonal but det $U = -1$; we say $U \in O^-(n)$. Then U^*X runs over $O^-(n)$ in (2.5) as X runs over $SO(n)$, so the following result is useful.

Lemma 2.3. *If P is positive-definite on $\mathbb{R}^n$, with eigenvalues satisfying $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$, and $V \in O^-(n)$, then*

$$\text{Tr } PV \le \lambda_n + \cdots + \lambda_2 - \lambda_1. \tag{2.9}$$

If $\{v_j\}$ is an orthonormal basis such that $Pv_j = \lambda_j v_j$, the maximum is achieved when $Vv_j = v_j$ for $j \ge 2$ and $Vv_1 = -v_1$.

Proof. First suppose V is stationary for $\varphi : O^-(n) \to \mathbb{R}$, $\varphi(X) = \text{Tr } PX$. Differentiating $\varphi(Ve^{sZ})$ at $s = 0$, we deduce that Tr $PVZ = 0$ for all $Z \in \text{skew}(n)$, hence PV is symmetric, so $PV = V^*P$. Hence $PV^2 = V^*PV$. In particular Tr $PV^2 =$ Tr P, and $V^2 \in SO(n)$, so Lemma 2.1 implies $V^2 = I$. Hence $V^* = V$, so $PV = VP$.

The proof of Lemma 2.3 is now straightforward if P has only simple eigenvalues. In such a case $Vv_j = \pm v_j$, and the maximum of Tr PV is assumed only for V described in the lemma.

A fairly straightforward argument extends the treatment to the case where P has multiple eigenvalues. We describe the result. Suppose the distinct eigenvalues of P are $\lambda_1 = \mu_1 < \cdots < \mu_k = \lambda_n$. Let E_j be the μ_j-eigenspace. Then $V : E_j \to E_j$ for each j. For $j > 1$, $V|_{E_j}$ is the identity if (2.9) is maximized, and $V|_{E_1}$ must be a reflection across some hyperplane in E_1, so it has one eigenvalue equal to -1.

Corollary 2.4. *In Case IB, the minimum for (1.1) over $X \in SO(n)$ is achieved at*

$$X = UV, \tag{2.10}$$

where $U \in O^-(n)$ is given by the polar decomposition (2.4) and $V \in O^-(n)$ is the identity on E_j for all $j > 1$ and an orthogonal reflection on E_1. Thus the minimizer X is unique if dim $E_1 = 1$. If dim $E_1 = \mu > 1$, the set of minimizers for (1.1) is diffeomorphic to $\mathbb{RP}^{\mu-1}$.

Case II. $\overline{A}$ is not invertible.
We can still write

$$\overline{A} = UP. \tag{2.11}$$

This time $P = (\overline{A}^*\overline{A})^{1/2}$ is positive semi-definite, with null space $\mathcal{N}(P) = \mathcal{N}(\overline{A})$. The factor U is a uniquely defined orthogonal linear map from the range $\mathcal{R}(P)$ to $\mathcal{R}(\overline{A})$. We can extend U to provide an orthogonal linear map from $\mathcal{N}(P) = \mathcal{R}(P)^\perp$ to $\mathcal{R}(\overline{A})^\perp$. Several such choices can be made. Make one choice, and arrange that $U \in SO(n)$. Again we are considering a function of the form (2.5); the only difference is that now P is only positive semi-definite. Hence the following lemma is useful.

Lemma 2.5. *If P is positive semi-definite on $\mathbb{R}^n$ and $V \in SO(n)$, then*

$$\operatorname{Tr} PV \le \operatorname{Tr} P, \tag{2.12}$$

with equality if and only if $V|_{\mathcal{R}(P)} = I$.

The proof is a simple analogue of the proof of Lemma 2.1.

Corollary 2.6. *In Case II, the minimum for (1.1) over $X \in SO(n)$ is achieved at*

$$X = UV, \tag{2.13}$$

where $U \in SO(n)$ is as described above for (2.11) and V is any element of $SO(n)$ such that $V = I$ on $\mathcal{R}(P)$. If dim $\mathcal{N}(\overline{A}) = 1$, then necessarily $V = I$ on $\mathbb{R}^n$ and X is unique. If dim $\mathcal{N}(\overline{A}) = \mu > 1$ then the set of minimizers for (1.1) is diffeomorphic to $SO(\mu)$.

Remark. While the analysis above includes cases for which the R-average is not unique, we emphasize that this situation has probability zero in the set of random collections $\{A_1, \dots, A_N\} \subset SO(n)$.

Remark. L. Vicci brought to our attention the paper [CJK], which contains a brief treatment of averaging a *pair* of rotations A_1, and A_2, in a fashion like that in Case IA. We note that, when $A_j \in SO(n)$, $A_1 + A_2$ has determinant ≥ 0, so Case IB does not arise when considering only two rotations.

3. Examples

We illustrate the results of §2 with some examples. Let Q be the unit cube in $\mathbb{R}^3$, centered at the origin, with edges parallel to the coordinate axes. Let G be the group of rotations of $\mathbb{R}^3$ preserving Q. It is known that G is a group of order 24, isomorphic to the symmetric group S_4. Furthermore, G is generated by rotations R_{xy}, R_{yz}, and R_{zx}, where R_{xy} is counterclockwise rotation by 90° in the xy-plane, etc. We will select various subsets of G.

Example 1. Let $A_1 = R_{xy},\ A_2 = R_{yz},\ A_3 = R_{zx}$. Then

$$\overline{A} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}, \quad \det \overline{A} = 4. \tag{3.1}$$

A calculation gives $\overline{A} = UP$ with

$$P = \frac{1}{3}\begin{pmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{pmatrix}, \quad U = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}. \tag{3.2}$$

This element $U \in SO(3)$ is the unique minimizer for (1.1), so U is the R-average of A_1, A_2, A_3.

Example 2. Let $A_1, \dots, A_{24}$ enumerate all the elements of G. Well known group properties imply

$$\overline{A} = A_1 + \cdots + A_{24} = 0. \tag{3.3}$$

In this case, the function $\psi : SO(3) \to \mathbb{R}$ defined by (1.1) is constant, and hence achieves its minimum at each point of $SO(3)$. Thus the R-average of this set of rotations is completely arbitrary, an expression that this set of rotations is evenly distributed in $SO(3)$.

Example 3. Let $A_1, \dots, A_{23}$ enumerate all the elements of G except the identity. Then

$$\overline{A} = A_1 + \cdots + A_{23} = -I, \quad \det \overline{A} = -1.$$

We hence have $\overline{A} = UP$ with $P = I$, $U = -I$. Corollary 2.4 applies and we see that the minimum for (1.1) over $X \in SO(3)$ is achieved precisely when

$$X = -R, \tag{3.4}$$

where R is an arbitrary reflection across some 2D plane in $\mathbb{R}^3$. This set of minimizers is diffeomorphic to $\mathbb{RP}^2$.

Example 4. Let $A_1, \dots, A_{21}$ enumerate all the elements of G except the three rotations R_{xy}, R_{yz}, and R_{zx} considered in Example 1. Hence (by (3.1) and (3.3)),

$$\overline{A} = A_1 + \cdots + A_{21} = -\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}, \quad \det \overline{A} = -4. \tag{3.5}$$

(Anti)parallel to (3.2), we have $\overline{A} = UP$ with P as in (3.2) and

$$U = -\frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix} \in O^-(3). \tag{3.6}$$

We note that the eigenvalues of P are $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 2$; in particular

$$P(e_1 + e_2 + e_3) = e_1 + e_2 + e_3, \tag{3.7}$$

where $\{e_j\}$ is the stamdard orthonormal basis of $\mathbb{R}^3$. Corollary 2.4 applies and we see that (1.1) has the unique minimizer,

$$X = UV, \tag{3.8}$$

where U is as in (2.4) and $V \in O^-(3)$ has the property

$$V(e_1 + e_2 + e_3) = -(e_1 + e_2 + e_3), \tag{3.9}$$

with $V = I$ on the orthogonal complement of the span of this vector. In other words,

$$V = \frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}, \tag{3.10}$$

and hence the R-average in this case is

$$X=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}. \tag{3.11}$$

4. Covariance

We mention a covariance property of the R-average, and contrast it with the type of average defined in terms of logarithms, described in the introduction. Namely, let $A_1,\dots,A_N\in SO(n)$, take $U\in SO(n)$, and set $B_j=UA_j\in SO(n)$. Then X is an R-average of $A_1,\dots,A_N$ if and only if UX is an R-average of $B_1,\dots,B_N$. This is an obvious consequence of the definition in terms of (1.1). Similarly the R-average is covariant with respect to $A_j\mapsto A_jU$ for $U\in SO(n)$, and with respect to $A_j\mapsto VA_jW$, for $V,W\in O^-(n)$.

We show that the logarithmic method of averaging does not always possess this property. Suppose $n\geq 3$ and $A_j\in SO(n)$ are very close to the identity, so we have uniquely

$$A_j=e^{Z_j}, \tag{4.1}$$

with Z_j close to 0. Given $Y\in \text{skew}(n)$ and $s\in\mathbb{R}$ close to 0, consider

$$A_j(s)=e^{sY}e^{Z_j}. \tag{4.2}$$

We can write

$$A_j(s)=e^{\mathcal{C}(sY,Z_j)}, \tag{4.3}$$

where $\mathcal{C}(sY,Z_j)$ is given by the Campbell-Hausdorff formula, whose general form is

$$\begin{aligned}\mathcal{C}(Y,Z)=Y+Z+\frac{1}{2}[Y,Z]&+\frac{1}{12}[[Y,Z],Z]\\&-\frac{1}{12}[[Y,Z],Y]+\cdots\end{aligned} \tag{4.4}$$

The logarithmic method of averaging two such rotations $A_1(s),A_2(s)$ gives

$$B(s)=e^{[\mathcal{C}(sY,Z_1)+\mathcal{C}(sY,Z_2)]/2}. \tag{4.5}$$

If the covariance property held we would have $B(s)=e^{sY}B(0)$, or

$$B(s)=e^{\mathcal{C}(sY,(Z_1+Z_2)/2)}. \tag{4.6}$$

Since these elements range over sets where the exponential map is a diffeomorphism, this implies

$$\mathcal{C}\big(sY,(Z_1+Z_2)/2\big)=\frac{\mathcal{C}(sY,Z_1)+\mathcal{C}(sY,Z_2)}{2}. \tag{4.7}$$

If this were true for all (small) Y, Z_1, Z_2, it would imply that $\mathcal{C}(Y,Z)$ is affine-linear in Z. However, for such a non-commutative Lie algebra as skew(n), $n \geq 3$, this is clearly not the case (as (4.4) shows), so covariance fails for the logarithmic method of averaging rotations.

The fact that the covariance property holds for the R-average, as defined here, seems to give this notion of average an advantage over a logarithmic approach to avaraging rotations.

5. Numerical work

It is a simple matter to write Mathematica functions to compute the R-average of a generic collection of A_j in $SO(3)$. One needs to implement the polar decomposition, which can be done in terms of the singular value decomposition. Indeed, if the singular value decomposition of $\overline{A}$ (with nonzero determinant) is given by

$$\overline{A} = u_1^* D v_1, \tag{5.1}$$

with D diagonal, having positive entries that decrease as you move down the diagonal, then the polar decomposition $\overline{A} = UP$ (as in (2.4)) holds, with

$$U = u_1^* v_1, \quad P = v_1^* D v_1. \tag{5.2}$$

In Case IA (when $\det \overline{A} > 0$), the R-average is simply given by U. See the attached Mathematica notebook, "Ravg.nb," where the function "Matpolar" is defined. There it is applied to Example 1 in §3, to compute U in (3.2), given $\overline{A}$ in (3.1). Next it is applied to

$$\overline{A} = A_1 + A_2, \quad A_j = e^{Z_j}, \tag{5.3}$$

where

$$Z_1 = \frac{1}{2}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z_2 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{5.4}$$

giving Ravg(A_1, A_2) as MatrixForm[Ravg].

There we also examine

$$A_1(s) = e^{sY} A_1, \quad A_2(s) = e^{sY} A_2, \tag{5.5}$$

with

$$Y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad s = \frac{1}{2}. \tag{5.6}$$

The last two matrices printed in that notebook are

$$\text{Ravg}(A_1(s), A_2(s)) \quad \text{and} \quad e^{sY}\,\text{Ravg}(A_1, A_2), \tag{5.7}$$

and it is seen that they agree to all 6 decimal places displayed.

In Case IB, where $\det \overline{A} < 0$, one has generically for the polar decomposition (2.4) that the smallest eigenvalue of P is simple. (In fact, generically all the eigenvalues of P are simple.) When this property holds, then the R-average computed by (2.10) is given by

$$X = u_1^* \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix} v_1, \tag{5.8}$$

in case $n = 3$, where u_1 and v_1 arise in the singular value decomposition via (5.1). See the attached Mathematica notebook, "Ravg2.nb" where this is implemented. The example discussed in Example 4 of §3 is treated here.

Finally, the Mathematica notebook "Lavg.nb" is attached. There we compute the logarithmic average $\mathrm{Lavg}(A_1, A_2)$, for $A_j = e^{Z_j}$, with Z_j given by (5.4), using the formula

$$\mathrm{Lavg}(A_1, A_2) = e^{(Z_1+Z_2)/2}. \tag{5.9}$$

We also compute $\mathrm{Lavg}(A_1(s), A_2(s))$, for $A_j(s)$ as in (5.5)–(5.6), and we compare

$$\mathrm{Lavg}\big(A_1(s), A_2(s)\big) \quad \text{vs} \quad e^{sY}\,\mathrm{Lavg}(A_1, A_2), \tag{5.10}$$

in the last two matrices given in this notebook. We note that they are close, but they differ in the third digit after the decimal point. Furthermore, this differance is real, not the product of numerical error. It is not surprising that the difference is small, since it crops up first in the cubic term in the expansion of the Campbell-Hausdorff formula. Note also that $\mathrm{Ravg}(A_1, A_2)$ and $\mathrm{Lavg}(A_1, A_2)$ differ only in the third decimal in this example.

We comment on the numerical cost of computing $\mathrm{Ravg}(S)$ and $\mathrm{Lavg}(S)$ when S is a set of k elements of $SO(3)$. To compute $\mathrm{Ravg}(S)$ we need $k-1$ matrix sums, one determinant calculation, one use of the singular value decomposition, and one or two matrix products. To compute $\mathrm{Lavg}(S)$, we need to compute k matrix logs, then $k-1$ matrix sums, then a matrix exponential. Apparently $\mathrm{Ravg}(S)$ requires fewer numerical operations.

6. Averaging other objects

The notion of R-average in §1 can be extended to a rather general context. In fact, let (X, d) be a metric space, having the following property:

$$\begin{aligned} &\text{For each } \ p \in X, \ R \in (0, \infty), \text{the set} \\ &\{x \in X : d(x, p) \le R\} \ \text{ is compact.} \end{aligned} \tag{6.1}$$

In particular, (6.1) holds if X is compact, but it also holds in many other cases. Now, given a collection $p_1, \dots, p_N \in X$, define

$$\psi(x) = \sum_{j=1}^{N} d(x, p_j)^2, \quad \psi : X \to \mathbb{R}^+. \tag{6.2}$$

The hypothesis (6.1) guarantees that ψ assumes a minimum on X. We denote by

$$(X,d)\mathrm{Avg}(p_1,\dots,p_N)$$

the set on which this minimum is achieved. This set might consist of one point, or it might be larger, as we have already seen for the example $X = SO(n)$ in §§1–3.

Specializing, we can take $X = G$, a Lie group. If $d : G \times G \to \mathbb{R}^+$ is a metric on G with the property that

$$\forall\ R \in (0,\infty),\ \{g \in G : d(e,g) \le R\}\ \text{ is compact,} \tag{6.3}$$

where $e \in G$ is the identity element, and if d is either left-invariant or right-invariant, then the hypothesis (6.1) holds. Also, parallel to the observation made about the covariance of the R-average in §4, we have the following result.

Proposition 6.1. *Suppose d is a metric on a Lie group satisfying (6.3). If d is left-invariant then, for each $h \in G,\ g_j \in G$,*

$$(G,d)Avg(hg_1,\dots,hg_N) = h \cdot (G,d)Avg(g_1,\dots,g_N), \tag{6.4}$$

while if d is right-inveriant then

$$(G,d)Avg(g_1h,\dots,g_Nh) = (G,d)Avg(g_1,\dots,g_N) \cdot h. \tag{6.5}$$

In §§1–5 we have considered $G = SO(3)$, with the bi-invariant metric given by

$$d_0(g,h)^2 = \mathrm{Tr}\big((g^* - h^*)(g-h)\big). \tag{6.6}$$

Here we consider another example, namely the Euclidean group $E(n)$ of (orientation-preserving) rigid motions of $\mathbb{R}^n$. Set-theoretically $E(n)$ is equivalent to a Cartesian product of $SO(n)$ and $\mathbb{R}^n$. An element g of $E(n)$ can be written uniquely as

$$g = U\tau_v, \quad U \in SO(n),\ v \in \mathbb{R}^n, \tag{6.7}$$

where $\tau_v(x) = x + v$, so $g(x) = U(x+v)$. It can also be written uniquely as

$$g = \tau_w U, \quad U \in SO(n),\ w \in \mathbb{R}^n, \tag{6.8}$$

so $g(x) = Ux + w$. The two representations are related by the identity $w = Uv$.

We can define metrics d_ℓ and d_r on $E(n)$ by

$$d_\ell(\tau_{v_1}U_1, \tau_{v_2}U_2)^2 = |v_1 - v_2|^2 + d_0(U_1,U_2)^2 \tag{6.9}$$

and

$$d_r(U_1\tau_{v_1}, U_2\tau_{v_2})^2 = |v_1 - v_2|^2 + d_0(U_1,U_2)^2. \tag{6.10}$$

It is readily verified that d_ℓ is left-invariant and d_r is right-invariant, and both satisfy (6.1). Suppose

$$g_j = U_j\tau_{v_j} = \tau_{w_j}U_j \tag{6.11}$$

is a collection of elements of $E(n)$. One finds that

(6.12)
$$(E(n), d_\ell)\mathrm{Avg}(g_1, \ldots, g_N) = \tau_{\overline{w}} \cdot \mathrm{Ravg}(U_1, \ldots, U_N), \quad \overline{w} = \frac{1}{N}(w_1 + \cdots + w_N),$$

and

$$(E(n), d_r)\mathrm{Avg}(g_1, \ldots, g_N) = \mathrm{Ravg}(U_1, \ldots, U_N) \cdot \tau_{\overline{v}}, \quad \overline{v} = \frac{1}{N}(v_1 + \cdots + v_N). \tag{6.13}$$

We see that $(E(n), d_\ell)\mathrm{Avg}$ has the left-covariance property (6.4) and $(E(n), d_r)\mathrm{Avg}$ has the right-covariance property (6.5). Of these two averages, one can show that the first is not right-covariant and the second is not left-covariant. For example, with g_j given by (6.11), we have

$$(E(n), d_\ell)\mathrm{Avg}(g_1\tau_y S, g_2\tau_y S) = \tau_{(w_1+w_2)/2+(U_1+U_2)y/2}\,\mathrm{Ravg}(U_1 S, U_2 S), \tag{6.14}$$

while

$$(E(n), d_\ell)\mathrm{Avg}(g_1, g_2)\tau_y S = \tau_{(w_1+w_2)/2+\mathrm{Ravg}(U_1,U_2)y}\,\mathrm{Ravg}(U_1, U_2)S. \tag{6.15}$$

These two are equal if and only if

$$\mathrm{Ravg}(U_1, U_2)y = \frac{1}{2}(U_1 + U_2)y, \tag{6.16}$$

an identity that does not generally hold.

References

[CJK] W. Curtis, A. Janin, and K. Zikan, A Note on averaging rotations, SIGGRAPH ’95 course notes #9, 0-7803-1363-1/93, IEEE.
[V] L. Vicci, Averages of rotations and orientations in 3-space. Preprint, 1999.