Robot Sensor Calibration: Solving \( AX = XB \) on the Euclidean Group

Frank C. Park and Bryan J. Martin

Abstract—The equation \( AX = XB \) on the Euclidean group arises in the problem of calibrating wrist-mounted robotic sensors. In this article we derive, using methods of Lie theory, a closed-form exact solution that can be visualized geometrically, and a closed-form least squares solution when \( A \) and \( B \) are measured in the presence of noise.

I. INTRODUCTION

The equation \( AX = XB \) on the Euclidean group, where \( A \) and \( B \) are known and \( X \) is unknown, is fundamental in solving the problem of calibrating wrist-mounted robotic sensors. Typically the matrix \( A \) describes the position and orientation of the wrist frame relative to itself after some arbitrary movement, and \( B \) describes the position and orientation of the sensor (e.g., camera) frame relative to itself after the same movement. \( X \) then describes the position and orientation of the sensor frame relative to the wrist frame. Calibration involves performing several arbitrary movements of the robot arm and solving \( AX = XB \) for \( X \) to determine the precise location of the sensor. While solutions to this equation have been studied when \( A \) and \( B \) are general matrices (see, e.g., Gantmacher [5]), in robotic applications one is interested in solutions that belong to the Euclidean group.

Shiu and Ahmad [8] first motivate this equation in the context of sensor calibration and provide a closed-form solution and conditions for its uniqueness. Chou and Kamel [3] present a method for solving this equation using quaternions. In this paper we present both exact and least-squares solutions to this equation using methods of Lie group theory. The principal advantage of this approach, aside from its geometric appeal, is that there exists a set of canonical coordinates for the Euclidean group that leads to a particularly simple characterization of the solutions to \( AX = XB \). The solution can be expressed explicitly and also admits a simple geometric visualization.

Because \( AX = XB \) has a one-parameter family of solutions (as first shown by Shiu and Ahmad [8]), two pairs of \((A_i, B_i)\) satisfying certain constraints are required in order to obtain a unique solution. Unfortunately, for sensor calibration applications some noise is usually present in the measured values of \( A \) and \( B \), so that conditions for existence of a solution may not be satisfied. A more practical approach is to make several measurements \( \{(A_1, B_1), (A_2, B_2), \ldots, (A_k, B_k)\} \), and to find an \( X \) that minimizes the error criterion

\[
\eta = \sum_{i=1}^{k} d(A_iX, XB_i)
\]

where \( d(\cdot, \cdot) \) is some distance metric on the Euclidean group. Using the canonical coordinates for Lie groups the above minimization problem can be recast into a least-squares fitting problem that admits a simple and explicit solution. Specifically, given vectors \( x_1, x_2, \ldots, x_k \) and \( y_1, y_2, \ldots, y_k \) in Euclidean \( n\)-space, Nadas [6] provides explicit expressions for the orthogonal matrix \( \Theta \) and translation \( b \) that minimize

\[
\eta = \sum_{i=1}^{k} \|\Theta x_i + b - y_i\|^2
\]

The best values of \( \Theta \) and \( b \) turn out to depend only on the matrix \( M = \sum x_iy_i^t \). By applying the canonical coordinates and this result a "best-fit" solution to \( AX = XB \) can be obtained.

The paper is organized as follows. In Section II we examine the Lie group structure of the Euclidean group, and derive explicit formulas for the canonical coordinates. In Section III we derive closed-form exact solutions to the equation \( AX = XB \), and in Section IV we present a least-squares solution given a set of noisy measurements for \( A \) and \( B \).

II. THE EUCLIDEAN GROUP

For our purposes it is sufficient to think of \( SE(3) \), the Euclidean group of rigid-body motions, as consisting of matrices of the form

\[
\begin{pmatrix}
\Theta & b \\
0 & 1
\end{pmatrix}
\]

where \( \Theta \in SO(3) \) and \( b \in \mathbb{R}^3 \). Here \( SO(3) \) denotes the group of \( 3 \times 3 \) rotation matrices, \( SE(3) \) has the structure of both a differentiable manifold and an algebraic group, and is an example of a Lie group. Some well known examples of matrix Lie groups include \( Gl(n) \), the general linear group of \( n \times n \) nonsingular matrices, and \( Sl(n) \), the special linear group of \( n \times n \) nonsingular matrices with unit determinant.

Associated with every Lie group is its Lie algebra. In general a Lie algebra is a vector space \( V \), together with a bilinear map \( [,] : V \times V \rightarrow V \) (called the Lie bracket) that satisfies, for every \( \eta, \mu, \nu \in V \), (i) \([\eta, \eta] = 0\), and (ii) \([\eta, \mu + \nu] + [\eta, \mu] + [\eta, \nu] = 0\). The Lie algebra of \( SO(3) \), denoted \( so(3) \), consists of the \( 3 \times 3 \) skew-symmetric matrices of the form

\[
\begin{pmatrix}
0 & -\omega_3 & \omega_2 \\
\omega_3 & 0 & -\omega_1 \\
-\omega_2 & \omega_1 & 0
\end{pmatrix}
\]

\( \omega = [\omega] \)

Observe from this definition that \( so(3) \) can be identified with \( \mathbb{R}^3 \). The Lie algebra of \( SE(3) \), denoted \( se(3) \), consists of the \( 4 \times 4 \) matrices of the form

\[
\begin{pmatrix}
\omega \times v & v \\
0 & 0
\end{pmatrix}
\]

where \( [\omega] \in so(3) \) and \( v \in \mathbb{R}^3 \). For both \( so(3) \) and \( se(3) \) (and for general matrix Lie algebras) the Lie bracket is given by the matrix commutator: \( [A, B] = AB - BA \).

A fundamental concept related to Lie groups is the exponential mapping. Given a matrix Lie group \( G \) and its corresponding matrix Lie algebra \( g \), the exponential mapping is the map \( \exp : g \rightarrow G \) defined by the matrix exponential: \( \exp A = I + A + \frac{1}{2!} A^2 + \cdots \) for \( A \in \mathbb{R}^3 \). Over some open set \( U \subseteq G \) containing \( 0 \) the mapping \( \exp : U \rightarrow G \) is a diffeomorphism.\footnote{A diffeomorphism is a differentiable 1-1 and onto mapping whose inverse is also differentiable.}

In the remainder of this section we derive explicit formulas for the canonical coordinates on \( SE(3) \) and its subgroup \( SO(3) \). One of

1042-296X/94$04.00 © 1994 IEEE
the classical results from screw theory is that every rigid motion can be decomposed into a rotation and translation that commute. Mathematically this is equivalent to the statement that the exponential is an onto mapping from se(3) to SE(3), i.e., for any $X \in \text{SE}(3)$ there exists some $x \in \text{se}(3)$ such that $exp x = X$.

The exponential mapping from $\text{so}(3)$ to $\text{SO}(3)$ is given by the following explicit formula:

**Lemma 1:** Given $[\omega] \in \text{so}(3)$, $\exp [\omega]$ is an element of $\text{SO}(3)$ given by
\[
\exp [\omega] = I + \frac{\sin \| \omega \|}{\| \omega \|} [\omega] + \left[1 - \cos \| \omega \| + \frac{\| \omega \|^2}{2} \right] [\omega]^2
\]
where $\| \omega \|^2 = w_1^2 + w_2^2 + w_3^2$.

*Proof:* That $\exp [\omega]$ is an element of $\text{SO}(3)$ is easily shown (see, e.g., Curtis [4]). The characteristic polynomial of $[\omega]$ is $s^3 + \| \omega \|^2 s$, and by the Cayley-Hamilton Theorem $[\omega]^3 = -\| \omega \|^2 [\omega]$. The formula follows by applying this identity to the series expansion for $\exp [\omega]$.

By this lemma $\text{SO}(3)$ can be visualized as a 3-dimensional solid ball of radius $\pi$ centered at the origin; a point $\omega$ in the ball represents a rotation (in the right hand sense) by an angle $\| \omega \|$ radians about the ray directed from the origin through $\omega$. From this interpretation it can be seen that antipodal points on the boundary represent the same rotation.

The exponential map provides local coordinates for the set of rotation matrices whose rotation angles are less than $\pi$. Over this set the inverse of the exponential map, or logarithm, is well-defined:

**Lemma 2:** Let $\Theta \in \text{SO}(3)$ such that $\text{Tr}(\Theta) \neq -1$. Then
\[
\log \Theta = \frac{\Theta - \Theta^T}{2 \sin \phi} (\Theta - \Theta^T),
\]
where $\phi$ satisfies $1 + 2 \cos \phi = \text{Tr}(\Theta)$, $|\phi| < \pi$, and $\| \log \Theta \|^2 = \phi^2$.

*Proof:* The result follows directly from Euler's Theorem, which states that for any $\Theta \in \text{SO}(3)$ there exists $Q \in \text{SO}(3)$ and $0 \leq \phi \leq 2\pi$ such that
\[
\Theta = Q \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} Q^T.
\]
Note that the eigenvalues of $Q$ are $1, e^{\pm i\phi}$; the condition $\text{Tr}(\Theta) \neq -1$ is therefore equivalent to $\phi \neq \pi$. The result now follows by an application of the general matrix identity $\Theta^A B^T \Theta^B = e^A B e^{-A}$.

The log formula determines the point in the solid ball of radius $\pi$ that corresponds to a particular rotation. $\log \Theta$ has two solutions when $\text{Tr}(\Theta) = -1$: denoting by $\vartheta$ the unit length eigenvector of $\Theta$ associated with the eigenvalue 1, a simple calculation shows that $\log \Theta = \pm \vartheta \pi$.

We now derive explicit formulas for the exponential and logarithm on SE(3).

**Lemma 3:** Let $[\omega] \in \text{so}(3)$ and $r \in \mathbb{R}^3$. Then
\[
\exp [\omega] r = \begin{bmatrix} \exp [\omega] \rvec \\ 0 \end{bmatrix}
\]
is an element of $\text{SE}(3)$, where $\exp [\omega]$ is as given in Lemma 1, and
\[
A = I + \frac{1 - \cos \| \omega \|}{\| \omega \|^2} [\omega] + \left[1 - \cos \| \omega \| + \frac{\| \omega \|^2}{2} \right] [\omega]^2
\]
where $\| \omega \|^2 = w_1^2 + w_2^2 + w_3^2$.

*Proof:* Writing out the series expansion for the exponential,
\[
\exp [\omega] r = \begin{bmatrix} \exp [\omega] \rvec \\ 0 \end{bmatrix} \sum_{n=0}^{\infty} (\frac{-\| \omega \|^{2n}}{n!}) r^n
\]
$A = \sum_{n=1}^{\infty} \frac{[\omega]^{n-1}}{n!}$ can also be written as \( \int_0^1 e^{\omega t} dt \), which simplifies to the above.

Like the SO(3) case, the exponential map provides a set of local coordinates for all of SE(3), except for those elements where $\text{Tr}(\Theta) = -1$. Over this set the logarithm is given by the following formula:

**Lemma 4:** Let $\Theta \in \text{SO}(3)$ such that $\text{Tr}(\Theta) \neq -1$, and let $b \in \mathbb{R}^3$. Then
\[
\log \begin{bmatrix} \Theta & b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} [\omega] & A^{-1} b \\ 0 & 0 \end{bmatrix}
\]
where $[\omega] = \log \Theta$, and
\[
A^{-1} = I - \frac{1}{2} \begin{bmatrix} [\omega] & 2 \sin \| \omega \| \end{bmatrix} \begin{bmatrix} [\omega]^T & -1 \end{bmatrix} + \frac{\| \omega \|^2}{2} \sin \| \omega \| [\omega]^T
\]

*Proof:* Since by the Cayley-Hamilton Theorem $[\omega]^3 = -\| \omega \|^2 [\omega]$, $A^{-1}$ is a quadratic matrix polynomial in $[\omega]$. An elementary calculation then establishes the formula.

III. SOLUTION TO $AX = XB$

The equation $AX = XB$ in matrix form is
\[
\begin{bmatrix} \Theta_A & b_A \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Theta_X & b_X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \Theta_B & b_B \\ 0 & 1 \end{bmatrix}
\]
This equation can be rewritten as the pair
\[
\Theta_A \Theta_X + b_A = \Theta_X \Theta_B + b_B
\]
(2)

*Proof:* Suppose that both $\Theta_A$ and $\Theta_B$ are not equal to $-1$, so that their logarithms are uniquely defined (as points in the open ball of radius $\pi$ in $\mathbb{R}^3$). Rewriting $AX = XB$ as $A = XB X^T$, it follows that $\log A$ must equal $\log (XB X^T)$. Let $\log A$ and $\log B$ be denoted by $[\alpha]$ and $[\beta]$, respectively. The preceding equality then implies that $[\alpha] = \log (XB X^T) = X[\beta] X^T$. Applying the easily established identity $[\alpha] [\beta]^T = [\Theta_A] [\Theta_B]$ for $\Theta_A \in \text{SO}(3)$ and $[\omega] \in \text{so}(3)$, it follows that $\alpha = X \beta$. However, since $X$ is orthogonal, a solution exists if and only if $[\alpha] = [\beta]$. The lemma also holds when $A$ and $B$ have trace $-1$, since then $\| \log A \| = \| \log B \| = \pi$.

**Theorem 1:** Let $A, B$ be elements of $\text{SO}(3)$ whose traces are not equal to $-1$, and denote their logarithms by $[\alpha]$ and $[\beta]$, respectively. Suppose that $[\alpha] = [\beta]$, and define $\hat{\alpha} = \alpha/\|\alpha\|$, $\hat{\beta} = \beta/\|\beta\|$. If $X_p \in \text{SO}(3)$ is a particular solution to $AX = XB$, then
\[
X = X_p e^{[\beta]r} = e^{[\hat{\beta}]r} X_p^T, \quad 0 \leq t \leq 2\pi
\]
is also a solution.

*Proof:* The equation $AX = XB$ can be recast using canonical coordinates as $X \beta = \alpha$. Consider the two-parameter set
\[
X = X_p e^{[\hat{\beta}]r} e^{[\hat{\beta}]s}, \quad 0 \leq r, s \leq 2\pi
\]
Observe that $e^{[\hat{\beta}]r} e^{[\hat{\beta}]s} \alpha = \alpha$ for any $r$ and $s$, since any vector is invariant with respect to rotations about itself. Also, because $X_p$ is a particular solution it follows that $X_p \beta = \alpha$. $X$ given above therefore satisfies $AX = XB$. Now, from the general matrix identity $P e^{[\alpha]} P^{-1} = e^{[\hat{\beta}]r} X_p$, it can also be expressed as
\[
X = X_p e^{[\hat{\beta}]r} X_p^T X_p e^{[\hat{\beta}]s} = X_p e^{[\hat{\beta}]r} e^{[\hat{\beta}]s}
\]
But \( X^T \tilde{\alpha} = \beta \), and
\[
X = X e^{(\beta / \beta) \theta^T} = X e^{(\beta / \beta) \theta}
\]
where \( t \equiv r + s \). Similarly,
\[
X = e^{(\beta / \beta) e^{(\beta / \beta) \theta^T}} X \theta^T = e^{(\beta / \beta) e^{(\beta / \beta) \theta^T}} X \theta^T = e^{(\beta / \beta) \theta^T} X \theta^T
\]
To show that no other solutions exist, suppose \( Y \) also satisfies \( AX = XB \). After some manipulation,
\[
A = (X e^{(\beta / \beta) \theta^T} X \theta^T)^T A (X e^{(\beta / \beta) \theta^T} X \theta^T)
\]
Since \( \text{Tr}(A) \neq -1 \) the logarithm is uniquely defined; taking the logarithm of both sides,\( \alpha = (Y e^{(\beta / \beta) \theta^T} X \theta^T) \alpha \), where \( |\alpha| = \log A \).

The preceding equality implies that \( Y e^{(\beta / \beta) \theta^T} X \theta^T \) is a rotation about \( \tilde{\alpha} \), i.e., \( Y e^{(\beta / \beta) \theta^T} X \theta^T = e^{(\beta / \beta) \theta^T} \) for all \( \beta \), but recall that \( X e^{(\beta / \beta) \theta^T} = e^{(\beta / \beta) \theta^T} X_\theta^T \), so that \( Y = e^{(\beta / \beta) (\alpha + \beta) X \theta^T} \) is of the general form as claimed.

**Remark 1:** If \( \alpha, \beta \in \mathbb{R}^n \) such that \( |\alpha| = |\beta| \), one particular solution \( \Theta_\alpha = \theta \) is a rotation about the axis \( \omega = \beta \times \alpha \), i.e.,
\[
\Theta_\beta = e^{(\omega / \omega) \theta^T} X \theta^T,
\]
where \( \omega = \alpha (|\omega|, |\beta| / |\beta|) \), and \( \theta_\beta = \phi / |\omega| \), where \( \phi \) is the angle from \( \beta \) to \( \alpha \). This particular solution is valid as long as \( \alpha \) and \( \beta \) are not parallel. When \( \alpha = \beta, \Theta_\beta = I \) is a particular solution, and when \( \alpha = -\beta, \Theta_\beta = \alpha, \) any \( \Theta_\beta \) of the form \( e^{(\beta / \beta) \theta} \), where \( \|\beta\| = \pi \) and \( \alpha_\beta = 0 \), is a solution.

**Remark 2:** So far, we have assumed that both \( A \) and \( B \) do not have trace equal to \(-1\). If \( A \) and \( B \) both have trace \(-1\), then log \( A = \pi i \alpha \) and log \( B = \pm \pi \beta \), where \( \alpha \) and \( \beta \) are unit-length eigenvectors of \( A \) and \( B \) associated with the eigenvalue \( 1 \). A particular solution to \( AX = XB \) can then be obtained as above.

### A. Finding a Unique Solution on SO(3)

Since the equation \( AX = XB \) on SO(3) has a one-parameter family of solutions, a minimum of two equations is required to determine a unique solution. Assuming that a solution exists, we now state conditions on \( A \) and \( B \) under which the solution is unique.

**Theorem 2:** Let \((A_1, B_1)\) and \((A_2, B_2)\) be pairs of elements of SO(3) such that \( \text{Tr}(A_1) \neq -1, \text{Tr}(B_1) \neq -1 \), and \( \| \log A_1 \| = \| \log B_1 \| \), \( i = 1, 2 \). Then the solution \( X \) to the set of equations \( AX = XB \), \( i = 1, 2 \), will be unique if only if \( \log A_1 \times \log A_2 \neq 0 \) and \( \log B_1 \times \log B_2 \neq 0 \).

**Proof:** From the hypotheses the logarithm is well-defined on these pairs of elements, denote by \([\alpha_i] \) and \([\beta_i] \) the logarithms of \( A_i \) and \( B_i \), respectively. We prove the forward direction first. Suppose that \( \log A_1 \times \log B_2 \neq 0 \), or equivalently, that \( \beta_1 = e^{i \delta} \) for some scalar constant \( \delta \). Now, if \( X \) is a solution, i.e., \( X \beta = \alpha_1 \beta_1 \), then so is \( X e^{i \delta} \) for \( i = 1, 2 \), so that the solution will not be unique. A parallel argument shows that \( \log A_1 \times \log A_2 \neq 0 \) in order for a unique solution to exist. To prove the reverse direction, observe that the unique solution is simply
\[
X = AB^{-1}
\]
where \( A \) and \( B \) are matrices whose columns are the vectors \( \alpha_1 \times \alpha_2, \alpha_1 \times \beta_2, \beta_1 \times \beta_2 \), respectively.

A physical interpretation of the above requirements on \( A \) and \( B \) is that their axes of rotation not be parallel. As noted in the proof, the unique solution to the equation in which the hypotheses above are satisfied is \( X = AB^{-1} \).

### B. Solution on SE(3)

We now consider the equation \( AX = XB \) on SE(3), which can be decomposed into equations (2) and (3) as before. Equation (3) can also be written
\[
(A \Theta_Y - I) b_X = \Theta_Y b_B - b_A
\]
We showed earlier that in order to obtain a unique solution to (2), two pairs of \((A_i, B_i)\) whose rotational parts satisfy certain conditions are required. Since the equation \( AX = XB \) arises in a physical setting where a solution is known to exist, as long as no noise is present in the measurements of \((A_1, B_1)\) and \((A_2, B_2)\) a solution is guaranteed to exist. Suppose \((A_1, B_1)\) and \((A_2, B_2)\) are such measurements whose rotational parts also satisfy the conditions of Theorem 2; a unique solution \( \Theta_X \) to (2) can then be found. By substituting this value of \( \Theta_X \) into
\[
\begin{pmatrix}
\Theta_{A_1} - I \\
\Theta_{A_2} - I
\end{pmatrix}
b_X =
\begin{pmatrix}
\Theta_{B_1} b_B - b_A \\
\Theta_{B_2} b_B - b_A
\end{pmatrix}
\]
a unique solution for \( b_X \) can be obtained. Clearly from physical considerations a solution must exist. To see that it is unique observe that, from Euler's theorem, the orthogonal matrix \( \Theta_A \) can be factored as in (1) where, if the eigenvalues are \( \{ \cos \phi_i \pm i \sin \phi_i, 1 \} \), and their corresponding eigenvectors are \( \{ x_i, \pm iy_i, z_i \} \), such that \( \| x_i \| = \| y_i \| = \| z_i \| = 1 \), then the columns of \( Q \) are given by \( x_i, y_i \), and \( z_i \), respectively. Observe also that \( z_i \) is the unit-axis of rotation for \( \Theta_A \), i.e., \( \Theta_{z_i} = \log \Theta_A / \log \Theta_{\alpha_A} \| \Theta_{\alpha_A} \| \). After some manipulation the row space of \( \Theta_A - I \) is known to be spanned \( \{ x_i, y_i \} \). Now by hypothesis \( z_1 \times z_2 \neq 0 \), and since \( \{ x_i, y_i, z_i \} \) forms an orthogonal matrix, it follows that \( \text{span} \{ x_i, y_i \} \cup \text{span} \{ x_2, y_2 \} = \mathbb{R}^3 \). Hence, the matrix in (6) is of rank 3, and the solution is unique.

### IV. A Least-squares Solution

The previous sections assumed that in determining a unique solution to \( AX = XB \) no noise was present in the measured values for \( A \) and \( B \). Unfortunately this assumption is physically unrealistic; a more practical approach is to find some type of "best-fit" solution from a set of noisy measurements \( \{ (A_1, B_1), \ldots, (A_k, B_k) \} \), i.e., to find \( X \in \text{SE}(3) \) that minimizes an error criterion of the form
\[
\eta = \sum_{i=1}^{k} d(A_i, X, X B_i)
\]
where \( d(\cdot, \cdot) \) is some suitably defined distance metric on \( \text{SE}(3) \).

Many choices for the metric \( d(\cdot, \cdot) \) exist; for example, in Park et al. [7] the following metric is used for mechanism design: if \( A = (\Theta_A, b_A) \) and \( B = (\Theta_B, b_B) \) are elements of \( \text{SE}(3) \), then
\[
d^2(A, B) = \| \log \Theta_A \Theta_B^T \|^2 + \| b_B - b_A \|^2
\]
where \( \| \cdot \| \) denotes the standard Euclidean norm in \( \mathbb{R}^3 \). Here the distance between \( A \) and \( B \) is defined to be the length of the minimal geodesic on \( \text{SE}(3) \) connecting \( A \) and \( B \), measured with respect to a certain physically meaningful left-invariant Riemannian metric on \( \text{SE}(3) \) (see [7] for details). The metric is left-invariant in the sense that \( d(A, B) = d(TA, TB) \) for any \( T \in \text{SE}(3) \). This is clearly important for mechanism design, since the choice of base frame should not affect the outcome of the design procedure. However, for sensor calibration applications the choice of base frame has no effect on the calibration equations. More significantly, by applying this metric to our problem we are then faced with a difficult nonlinear least-squares minimization problem.

One of the advantages of the Lie group approach is that the canonical coordinates enable us to formulate the problem as one of linear least-squares fitting, in which the solution assumes a particularly simple form. Recall first that the equation \( \Theta_A \Theta_X = \Theta_A \Theta_B \) can be recast via the logarithmic mapping as \( \Theta_X \beta = \alpha \), where...
\( \alpha \) and \( \beta \) are the logarithms of \( \Theta_A \) and \( \Theta_B \), respectively. In [6] Nádas shows that if \( x_1, x_2, \cdots, x_P \) and \( y_1, y_2, \cdots, y_P \) are given vectors in Euclidean \( n \)-space, then the translation \( b \) and orthogonal matrix \( \Theta \) that minimize

\[
\eta = \sum_{i=1}^{p} \| \Theta x_i + b - y_i \|^2
\]

can be expressed explicitly. It is not difficult to see that the given data enter \( \eta \) only through the matrices \( X = \sum x_i x_i^T, M = \sum x_i y_i^T \), and the centroids \( \bar{x} = (x_1 + x_2 + \cdots + x_P) / p, \bar{y} = (y_1 + y_2 + \cdots + y_P) / p \). In fact the best values of \( \Theta \) and \( b \) do not even depend on \( X \) and are simply

\[
\Theta = (M^T M)^{-1/2} M^T
\]

and

\[
b = \bar{y} - \Theta \bar{x}
\]

where the square root is the symmetric, positive definite square root (see, e.g., Gantmacher [5]). For a given \( \Theta \) the choice of \( b \) is unique. The choice of \( \Theta \) is unique if \( M^T M \) is nonsingular and has no repeated eigenvalues.

This result can be directly applied to our problem as follows. Suppose we have \( k \) sets of measurements \( \{ (A_1, B_1), (A_2, B_2), \ldots, (A_k, B_k) \} \). We first find the \( \Theta_X \) that minimizes

\[
\eta_1 = \sum_{i=1}^{k} \| (\Theta_X \beta_i - \alpha_i) \|^2
\]

where \( \alpha_i = \log \Theta_{A_i}, \beta_i = \log \Theta_{B_i} \). With this value of \( \Theta_X \) the next step is to find the \( b_X \) that minimizes

\[
\eta_2 = \sum_{i=1}^{k} \| (\Theta_X - I) b_X - \Theta_X b_{A_i} + b_{A_i} \|^2
\]

The optimal value of \( \Theta_X \) is given by \( \Theta_X = (M^T M)^{-1/2} M^T \), where the \( M \) matrix is now

\[
M = \sum_{i=1}^{k} \beta_i \alpha_i^T
\]

Since \( M^T M \) will in general be nonsingular, the optimal \( \Theta_X \) will also be unique. The value of \( b_X \) that minimizes \( \eta_2 \) (with the \( \Theta_X \) just obtained) is then the standard least-squares solution \( b_X = (C^T C)^{-1} C^T d \), where

\[
C = \begin{bmatrix}
I - \Theta_A \\
I - \Theta_B \\
\vdots \\
I - \Theta_A
\end{bmatrix}, \quad d = \begin{bmatrix}
b_{A_1} - \Theta_X b_{A_1} \\
b_{A_2} - \Theta_X b_{A_2} \\
\vdots \\
b_{A_k} - \Theta_X b_{A_k}
\end{bmatrix}
\]

In this manner a simple, computationally efficient least-squares solution to the equation \( AX = XB \) can be obtained from a set of measured values of \( A \) and \( B \). By finding \( \Theta_X \) first, the solutions generated are independent of the choice of length scale for physical space; that is, the least-squares solution will be the same regardless of whether \( b_{A_i} \) and \( b_{B_i} \) are expressed in inches, meters, etc. This would in general not be the case if the optimal \( \Theta_X \) and \( b_X \) were found simultaneously.

\section{Example}

We now present simulation examples for finding an exact solution given noiseless measurements, and a least-squares solution obtained from a set of noisy measurements. Suppose the measured values of \( (A_1, B_1) \) and \( (A_2, B_2) \) are (from an example in [8])

\[
A_1 = \begin{bmatrix}
-0.989992 & -0.141120 & 0.000000 & 0 \\
0.141120 & -0.989992 & 0.000000 & 0 \\
0.000000 & 0.000000 & 1.000000 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
B_1 = \begin{bmatrix}
-0.989992 & -0.138307 & 0.028036 & -26.9559 \\
0.138307 & -0.911449 & 0.387470 & -96.1332 \\
-0.028036 & 0.387470 & 0.921456 & 19.4872
\end{bmatrix}
\]

\[
A_2 = \begin{bmatrix}
0.070737 & 0.000000 & 0.997453 & -400.0000 \\
0.000000 & 1.000000 & 0.000000 & 0 \\
-0.997453 & 0.000000 & 0.070737 & 400.0000
\end{bmatrix}
\]

\[
B_2 = \begin{bmatrix}
0.070737 & 0.198172 & 0.997612 & -309.543
\end{bmatrix}
\]

These matrices satisfy the requirements of Theorem 2 for the existence of a unique solution. Denote by \( \Theta_{A_i} \) and \( \Theta_{B_i} \) the SO(3) components of \( A_i \) and \( B_i \), respectively, and let \( \alpha_i = \log \Theta_{A_i}, \beta_i = \log \Theta_{B_i} \). If the solution \( X \) is denoted \( (\Theta_X, b_X) \), then recall that

\[
\Theta_X = AB^{-1}
\]

where \( A \) and \( B \) are \( 3 \times 3 \) matrices with columns \( \{ \alpha_1, \alpha_2, \alpha_1 \times \alpha_2 \} \) and \( \{ \beta_1, \beta_2, \beta_1 \times \beta_2 \} \), respectively, \( b_X \) is then the solution of the linear equations (6). From the logarithm formula on SO(3),

\[
\alpha_1 = (0, 0, 3)^T, \beta_1 = (0, 0.596, 2.9402)^T, \alpha_2 = (0, 1.5, 0)^T,
\]

and \( \beta_2 = (0, 1.4701, -0.2985)^T \), so that the solution is

\[
X = \begin{bmatrix}
1.000000 & 0.000000 & 0.000000 & 10 \\
0.000000 & 0.98067 & -0.198669 & 50 \\
0.000000 & 0.198669 & 0.98067 & 100
\end{bmatrix}
\]

We now find the least-squares solution \( \hat{X} \) given \( k \) noisy measurements of \( A_i, B_i, i = 1, 2, \cdots, k \). We first determine the rotation matrix \( \hat{\Theta} \) from the \( k \) pairs \( \{ (\Theta_{A_i}, \Theta_{B_i}) \} \). Define \( \alpha_i = \log \Theta_{A_i} \) and \( \beta_i = \log \Theta_{B_i} \). Now recall that \( \Theta = (M^T M)^{-1/2} M^T \), where

\[
M = \sum \beta_i \alpha_i^T
\]

is diagonalizable as \( Q \Lambda Q^{-1} \), in which case \( (M^T M)^{-1/2} = Q \Lambda^{-1/2} Q^{-1} \), where

\[
\Lambda^{-1/2} = \begin{bmatrix}
1 / \sqrt{\lambda_1} & 1 / \sqrt{\lambda_2} & 1 / \sqrt{\lambda_3}
\end{bmatrix}
\]

and each square root is positive. The translation component \( b_X \) can then be found by a standard least-squares procedure.

Note that a minimum of three \( (A_i, B_i) \) pairs are needed in order for \( M \) to be nonsingular (with only two data pairs the formula for the exact solution can be used to obtain a least-squares estimate). For the simulation let the true solution \( X \) be that given earlier in (9). We first generate a set of \( k \) uncorrupted pairs of \( (A_i, B_i) \) satisfying \( A_i X = X B_i \); the translation components of the \( A_i \) are chosen to range in magnitude up to \( 400 \sqrt{3} \), and \( \alpha_i \) are distributed uniformly over the solid ball of radius \( \pi \).

To each \((A_i, B_i)\) pair we now add noise as follows. A zero-mean, independent, uniformly-distributed random variable \( \epsilon \) is added to each component of \( \alpha_i \) and \( \beta_i \), where \( -\frac{1}{100} \leq \epsilon \leq \frac{1}{100} \). Similarly, a zero-mean, independent, uniformly-distributed random variable \( \xi \) is
added to each of the translation components $b_A$ and $b_B$, where $-5 \leq \xi \leq 5$. Fig. 1 plots the error $\eta$ as a function of $k$, where

$$\eta = \sum_{i=1}^{k} \| \hat{X} - X_i \|^2$$

The plot represents the average of ten independent simulations, each run with 100 measurements. Note that the error decreases rapidly as $k$ approaches 10. The estimated $\hat{X}$ for a single simulation, after 100 measurements, is

$$\hat{X} = \begin{bmatrix} 1.0000 & -0.0012 & 0.0019 & 11.0696 \\ 0.0015 & 0.9803 & -0.1974 & 49.5175 \\ -0.0017 & 0.1974 & 0.9803 & 100.2557 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

Fig. 2 is an averaged plot of $\eta(k)$, but this time the noise values added to the $(A_i, B_i)$ are increased: $-\frac{7}{20} \leq \eta \leq \frac{5}{20}$, and $-10 \leq \xi \leq 10$. As expected, in both cases the least-squares error decreases asymptotically with the number of measurements.

VI. CONCLUSIONS

One of the advantages of using Lie theoretic methods is that the conditions for existence and uniqueness of solutions to $AX = XB$ can be stated in a compact, elegant way, and the solution can be expressed in closed form (see (4) and (6)). Because noise is inevitable in actual calibration measurements, a closed-form least-squares solution based on the canonical coordinates for SO(3) has also been presented. While from a mathematical perspective it may be more preferable to minimize a geometrically-defined (i.e., coordinate-invariant) measure like (7) and (8), in practice such measures lead to difficult nonlinear optimization problems. The proposed engineering solution is not only simple and computationally efficient, but also eliminates the dependence on choice of length scale for physical space.

REFERENCES