Frame Transformations in Computer Integrated Surgery: A Graphical Introduction

Benjamin D. Killeen and Russell H. Taylor

August 10, 2022

The notation in this tutorial aligns with the Computer Integrated Surgery course at Johns Hopkins University. Minimal linear algebra is assumed for the main text. Proofs assume familiarity with basic graph theory.

1 Introduction

Frame transformations are often the cause of confusing, frustratingly simple errors in computer integrated surgery. For beginners, developing the intuition to answer the question, "is this the A to B transform or the B to A transform?" can take some time, especially when the meaning of "A to B" may vary based on convention. In this tutorial, we will establish a consistent notation and vocabulary for talking about points, frames, and the transformations among them, which reflect the graph-like structure of any tracking setup. We also show how to translate from a problem statement, which may use informal language to describe these transformations, to a rigorous description of the problem, suitable for writing a computer program.

2 Frames and Points

First, let's define precisely what we mean by "frame," "point," and "transform." For this tutorial, we will constrain ourselves to points and frames in 3D.

Definition 2.1. A **frame** is a basis for numerical measurements of object locations, orientations, or poses. It can be thought of as a virtual object floating in physical space.

Definition 2.2. A **point** u, is a singular location in space. It can be measured relative to a frame A with a 3-vector:[1]

$$\mathbf{u}_A = \begin{bmatrix} u_{A,x} \\ u_{A,y} \\ u_{A,z} \end{bmatrix} \tag{1}$$

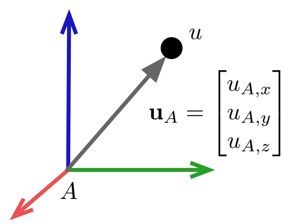

Figure 1: A point u as measured in frame A.

Definition 2.3. An **orientation** is measured by a rotation $\mathbf{R} \in \mathrm{SO}(3)$. There are several convenient ways to describe the rotation group SO(3) in a given frame A. We will use rotation matrices $\mathbf{R}_A \in \mathbb{R}^{3\times 3}$.

Together, a location and an orientation fully describe the degrees of freedom for a rigid body in 3D Euclidean space. The combination is often referred to as a "pose," and it is the same information needed to describe a second frame B. Thus, a "pose" can be thought of as a *measurement of a frame*, or a frame transformations.[2]

Definition 2.4. The "A from B" **frame transformation $\mathbf{F}_{AB}$** is a measurement of frame B's pose with respect to frame A. It consists of a rotation and a translation

$$\mathbf{F}_{AB} = [\mathbf{R}_{AB}, \mathbf{p}_{AB}] \tag{2}$$

such that for a given point u,

$$\mathbf{u}_A = \mathbf{F}_{AB}\mathbf{u}_B = \mathbf{R}_{AB}\mathbf{u}_B + \mathbf{p}_{AB} \tag{3}$$

as in Fig. 3. 3D frame transformations are elements of the special Euclidean group with $n = 3$: SE(3).

It is important to ensure consistency between notation like Equation 2, frame transform diagrams like Fig. 2, and the vocabulary we use to talk about them. Here, we use the notation $\mathbf{F}_{AB}$ for frame transformations because, mathematically, these are left-hand-side operators.[3]A point u measured in some frame

[1]For notational convenience, frame subscripts are often dropped in favor of $\mathbf{a}, \mathbf{b}$ corresponding to measurements of the same point u in frames A, B respectively. To avoid confusion, in this tutorial, we will always refer to a physical point with the same letter, using a frame subscript to specify which basis.

[2]The following terms are often used synonymously to refer to frame transformations: "pose," "frame transform," "rigid body transform," "transformation," or simply "transform."

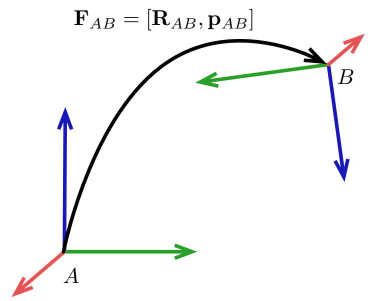

Figure 2: The "A from B" frame transformation $\mathbf{F}_{AB} = [\mathbf{R}_{AB}, \mathbf{p}_{AB}]$.

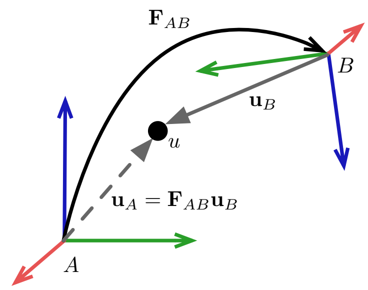

Figure 3: A frame transformation of a point u. If $\mathbf{u}_B$ is the measurement in frame B, and $\mathbf{F}_{AB}$ is known, then $\mathbf{u}_A = \mathbf{F}_{AB}\mathbf{u}_B$.

B is transformed as $\mathbf{F}_{AB}\mathbf{u}_B$, where the right-hand subscript of the transform matches the subscript of the point. Otherwise, the result is nonsensical (and will likely cause confusing behavior in your program!). Compositions of frame transforms behave similarly. In order for $\mathbf{F}_{AB}\mathbf{F}_{CD}$ to make sense, we must have $B = C$. In essence, adjacent frames must match when multiplying frame transformations and points, and the result maps between the outer frames:

$$\mathbf{F}_{AB}\mathbf{F}_{BC} = \mathbf{F}_{AC}. \tag{4}$$

We call $\mathbf{F}_{AB}$ the "A from B" transformation, because as an operator, $\mathbf{F}_{AB}$ takes in measurements of points in frame B and returns measurements in frame A. So why not call this the "B to A" transform? Or, as in quite a lot of source code, the `B2A` transform? The reason is that "A from B" reads from left to right in the same order as $\mathbf{F}_{AB}$, avoiding *so much confusion*. Saying "A from B" makes clear how the transformation operates: it takes a measurement

[3] In practice, you may see $\mathbf{F}_{AB}$ abbreviated to $\mathbf{F}_B$. This is acceptable for convenience, particularly when manipulating equations by hand, since $\mathbf{F}_{AB}$ is the measurement of frame B in frame A, just as $\mathbf{u}_A$ is the measurement of point u in A is $\mathbf{u}_A$. However, this introduces a possible ambiguity between the component of a transformation, such as $\mathbf{p}_{AB}$, with a measurement of a point $\mathbf{p}$ in a frame A: $\mathbf{p}_A$, and so in this tutorial we maintain both subscripts.

from frame B *to* a measurement in frame A, while maintaining consistency in our left-to-right ordering. Using this convention when naming variables, e.g. `A_from_B`, will make code easier to read, since the adjacency rule carries over, and more likely to be correct on the first try.

2.1 Frame Transformation Diagrams

Frame transformation diagrams are an informal, helpful tool for understanding the relationships in a given tracking setup. Fig. 2 is a simple diagram, where the point u is measured in frame B, and B is measured in frame A. The diagram uses arrows to signify these measurements, following the same left-to-right convention as the notation thus far. If $\mathbf{F}_{AB}$ is the "A from B" frame transform, then its arrow starts at frame A and ends at frame B, as in Fig. 2. This is often confusing, because the *arrow in the diagram* starts at A and ends at B, but $\mathbf{F}_{AB}$ maps measurements **to** frame A **from** frame B. For this reason, when talking about arrows, we will say the arrow "starts" and "ends," reserving "to" and "from" for the frame transformation as a mathematical operator. With this vocabulary, we eliminate a single point of confusion while maintaining the *many* advantages of drawing frame transform diagrams with arrow directions matching the left-to-right ordering in $\mathbf{F}_{AB}$ and "A from B."

One of these advantages has to do with constructing the frame transformation. If $\mathbf{o}$ is the point at the origin of B, and $\mathbf{o}_A$ is the measurement in frame A, the translational component of $\mathbf{F}_{AB}$ is simply $\mathbf{o}_A$. Thus the arrow in Fig. 4 representing $\mathbf{o}_A$ (left) starts at the same frame and ends in the same location as the arrow representing $\mathbf{F}_{AB}$ (right). Similarly, if $\mathbf{R}_A$ is the rotation describing frame B's orientation in frame A, then the rotational component of $\mathbf{F}_{AB}$ is simply $\mathbf{R}_A$.

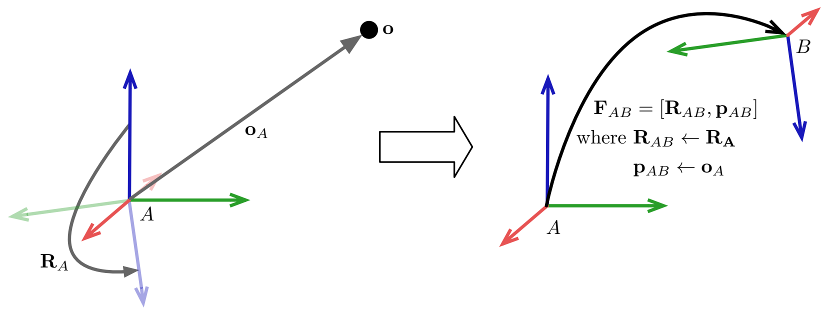

Figure 4: Constructing a frame transform.

Further advantages of drawing frame transformation diagrams in this manner become clear as tracking setups become more complicated. With a little formalization, these diagrams can help us understand measurements across a chain of frame transformations. In the following section, we will use the powerful tools of graph theory to outline the rules for dealing with frame transformations.

3 Reference Graphs

We will formalize frame transformation diagrams as "reference graphs" so that we can define the rules by which these objects operate, and how we can use them to recover the desired information.

Definition 3.1 (Reference Graphs)**.** A **reference graph** is a connected, directed graph $G = (V, E)$, representing measurements of points and frames in physical space, that satisfies the following:

- The nodes V are partitioned into frames $\mathcal{F}$ and points $\mathcal{U}$. That is, $V = \mathcal{F} \sqcup \mathcal{U}$, where $\sqcup$ is the disjoint union, and $\mathcal{F}, \mathcal{U}$ are both nonempty.
- Every point node $u \in \mathcal{U}$ has no outgoing edges, or out-degree 0. Precisely, $\nexists\, (v, w) \in E$ such that $v = u$.
- There is a measurement $\mathbf{u}_A$ of a point u in a frame A *iff* there is an edge $(A, u) \in E$ connecting the corresponding nodes.
- There is a measurement $\mathbf{F}_{AB}$ of the "A from B" frame transformation *iff* the edge (A, B) is in E.
- There is a one-to-one **measurement map** $\varphi : E \to \mathrm{SE}(3) \cup R^3$ that associates edges in $\mathcal{F} \times \mathcal{F}$ with frame transforms and edges in $\mathcal{F} \times \mathcal{U}$ to point measurements.

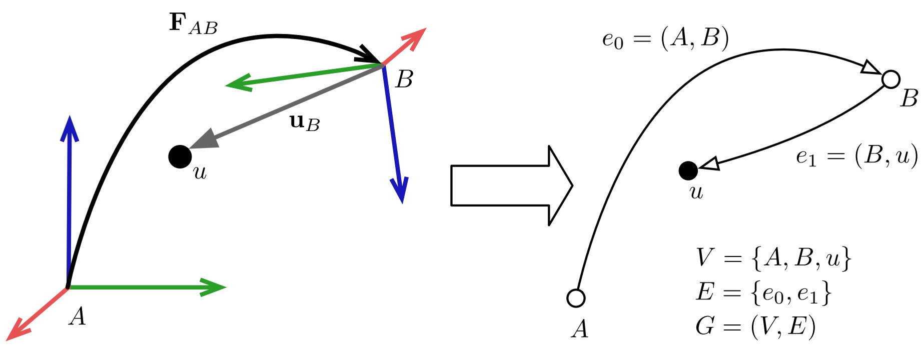

Figure 5: A simple frame transformation diagram (left) and its corresponding reference graph (right).

Don't confuse the frames and points themselves, which are nodes in the graph, with the *measurements* of those frames and points, which correspond to edges. Remember, a "frame" is just an abstract object floating in physical space, with 6 degrees of freedom, and points are similarly abstract locations in physical space with 3 degrees of freedom. We can create sets of these objects, but we cannot (in good faith) associate them *numbers* until we have measured them in a *separate* frame of reference. Hence, the frame or point in physical space is the node, and its measurement is really just a relationship (edge) with another node.

The measurement map φ is just the association of edges $e \in E$ with numerical measurements, which are elements of SE(3) or $\mathbb{R}^3$.

$$\varphi(e) = \begin{cases} \mathbf{F}_{AB} & \text{if } e = (A, B) \in \mathcal{F} \times \mathcal{F} \\ \mathbf{u}_A & \text{if } e = (A, u) \in \mathcal{F} \times \mathcal{U} \end{cases} \tag{5}$$

Based on the definition of G, this is a complete definition of φ. In general, the definition of reference graphs and φ can be extended to include other objects of interest, such as point clouds, triangle meshes, camera projections, and non-Euclidean manifolds, but here we will restrict ourselves to points and frame transforms in 3D Euclidean space.

The reference graph G has some useful properties, which help us determine whether a given object can be measured in a given frame, and how to obtain that measurement from the given quantities (assuming no error). In the following section, we will provide two *simple* theorems that define how one should deal with (1) inverse transformations and (2) kinematic chains, which are really just paths on the reference graph.

3.1 Inverse Transformations

Theorem 3.2 (Inverse Transformations)**.** *If $A, B \in \mathcal{F}$, $(A, B) \in E$, and $\varphi(e) = \mathbf{F}_{AB} = [\mathbf{R}_{AB}, \mathbf{p}_{AB}]$, then the reverse edge (B, A) is also in E, and*

$$\varphi((B, A)) = \mathbf{F}_{AB}^{-1} = [\mathbf{R}_{AB}^{-1}, -\mathbf{R}_{AB}^{-1}\mathbf{p}_{AB}]. \tag{6}$$

Proof. It suffices to prove that if $\mathbf{F}_{AB}$ is known, then $\mathbf{F}_{BA}$ is also known, and is as given.

Suppose

$$\mathbf{u}_A = \mathbf{F}_{AB}\mathbf{u}_B = \mathbf{R}_{AB}\mathbf{u}_B + \mathbf{p}_{AB}. \tag{7}$$

Then

$$\begin{aligned} \mathbf{u}_B &= \mathbf{R}_{AB}^{-1}(\mathbf{u}_A - \mathbf{p}_{AB}) && (8) \\ &= \mathbf{R}_{AB}^{-1}\mathbf{u}_A - \mathbf{R}_{AB}^{-1}\mathbf{p}_{AB} && (9) \\ &= [\mathbf{R}_{AB}^{-1}, -\mathbf{R}_{AB}^{-1}\mathbf{p}_{AB}]\mathbf{u}_A && (10) \\ &= \mathbf{F}_{AB}^{-1}\mathbf{u}_A && (11) \end{aligned}$$

This is precisely the definition of the "B from A" transformation, $\mathbf{F}_{BA}$. □

Thereom 3.2 provides a useful method for obtaining the desired transform, when only its inverse is known. Essentially, the inverse operation flips the subscripts of a transform, so $\mathbf{F}_{AB}^{-1}\mathbf{u}_A$ makes sense according to our earlier rule, because the right-hand frame of $\mathbf{F}_{AB}^{-1} = \mathbf{F}_{BA}$ is frame A, which matches $\mathbf{u}_A$.

Let's make this a little more concrete with a simple example, which can be done by hand. Remember, rotation matrices are orthonormal, which means $\mathbf{R}^{-1} = \mathbf{R}^T$.

Problem 3.3. Suppose a tracked patient frame B is measured relative to the tracker frame A to be at

$$\mathbf{F}_{AB} = [\mathbf{R}_{AB}, \mathbf{p}_{AB}] = \left[\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right]. \tag{12}$$

If the tip of a surgical instrument is measured at $\mathbf{u}_A = [4, 5, 6]^T$ in the tracker frame, what is the same point's measurement in the patient frame B? See Fig. 6.

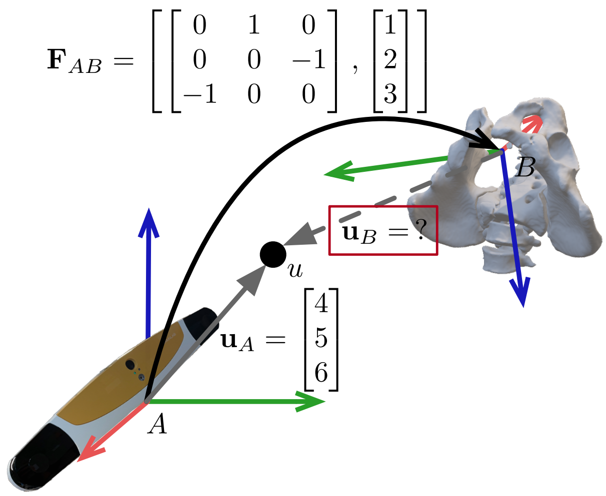

Figure 6: Find the measurement of the point u in the anatomical frame B, given $\mathbf{F}_{AB}$ and $\mathbf{u}_A$.

Solution 3.4. Our goal is to find $\mathbf{u}_B$. Since we know $\mathbf{u}_A$, we need to find the frame transformation $\mathbf{F}_{BA}$ so that

$$\mathbf{u}_B = \mathbf{F}_{BA}\mathbf{u}_A. \tag{13}$$

$\mathbf{F}_{BA}$ is not given. Looking at the reference graph in Fig. 7 (left), we know from Theorem 3.2 the edge (B, A) is in the graph, however and its corresponding frame transform is given by $\mathbf{F}_{AB}^{-1}$.

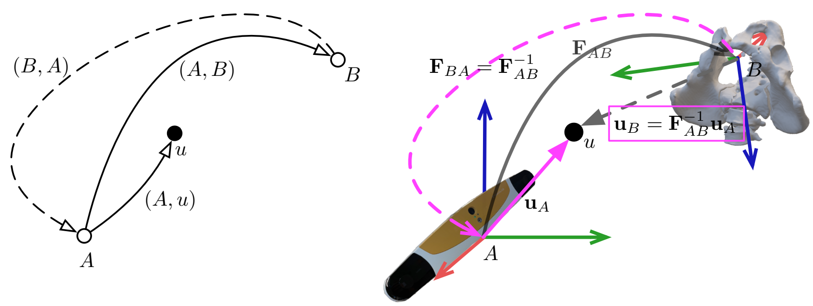

Figure 7: The reference graph and solution for the problem.

$$\begin{aligned}
\mathbf{u}_B = \mathbf{F}_{BA}\mathbf{u}_A &= \mathbf{F}_{AB}^{-1}\mathbf{u}_A && (14)\\
&= [\mathbf{R}_{AB}^{-1}, -\mathbf{R}_{AB}^{-1}\mathbf{p}_{AB}]\mathbf{u}_A && (15)\\
&= \mathbf{R}_{AB}^{-1}\mathbf{u}_A - \mathbf{R}_{AB}^{-1}\mathbf{p}_{AB} && (16)\\
&= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} && (17)\\
&= \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} && (18)\\
&= \begin{bmatrix} -6 \\ 4 \\ -5 \end{bmatrix} - \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix} && (19)\\
&= \begin{bmatrix} -3 \\ 3 \\ -3 \end{bmatrix} && (20)
\end{aligned}$$

In this example, we used the concept of reference graphs in a very minor way, namely to have a vocabulary for how the inverse frame transformation operates on points. However, you might notice that the end result $\mathbf{u}_B = \mathbf{F}_{AB}^{-1}\mathbf{u}_A$ corresponds to a series of connected edges on the reference graph, namely $\{\,(B, A), (A, u)\,\}$. That is, to find the measurement $\mathbf{u}_B$, we have found a *directed* *path* starting at the node B and ending at u, then multiplied the corresponding mathematical objects in the same order. This is no coincidence, and it leads us to the primary advantage of using reference graphs to think about coordinate frame transformations.

3.2 Kinematic Chains

Theorem 3.5 (Kinematic Paths of Points)**.** *Given a reference graph $G = (\mathcal{F} \sqcup \mathcal{U}, E)$, the measurement $\mathbf{u}_A$ of a point $u \in \mathcal{U}$ in a frame $A \in \mathcal{F}$ can be determined if G contains a directed path $P_{A \to u} = \{e_0, e_1, \ldots, e_{n-1}\}$ starting at A and ending at u. Moreover,*

$$\mathbf{u}_A = \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-1}). \tag{21}$$

Put another way, if

$$P_{A \to u} = \{(A, A_1), (A_1, A_2), \ldots, (A_{n-2}, A_{n-1}), (A_{n-1}, u)\} \tag{22}$$

then

$$\mathbf{u}_A = \mathbf{F}_{AA_1}\mathbf{F}_{A_1A_2}\cdots\mathbf{F}_{A_{n-2}A_{n-1}}\mathbf{u}_{A_{n-1}}. \tag{23}$$

Thus we have a simple rule for obtaining the measurement of a given point in any frame: find a path starting at the frame and ending at the given point and compose the transforms along the way.

The proof of Theorem 3.5 is included here for completeness, but the reader may skip ahead to Section 3.2.1 without too much concern. It hinges on the notion of equivalent reference graphs. These are reference graphs which differ in their measurements but ultimately describe the same set of points in physical space. As a simple example, given any reference graph G, I can create an equivalent reference graph G' by eliminating an edge $\{(A, u)\}$ and still preserve all measurements by adding a "dummy" frame B along with new edges $(A, B), (B, u)$, and define a new measurement map φ' such that $\varphi'((A, B)) = \mathbf{I}, \varphi'((B, u)) = \varphi((A, u))$, and $\varphi' = \varphi$ otherwise (see Fig. **??**).

Formally, let $\sim$ denote this equivalence relation on reference graphs, which describe the same points in physical space.

Definition 3.6. If $G = (\mathcal{F} \sqcup \mathcal{U}, E), G' = (\mathcal{F}' \sqcup \mathcal{U}', E')$ are reference graphs, then $G \sim G'$ *iff* $\mathcal{U} = \mathcal{U}'$ and $\mathcal{F} \cup \mathcal{F}' \neq \emptyset$.

That is, every point measured in G is also measured in G' at the same location, and the graphs share at least one frame in common. We shall now prove Theorem 3.5, using induction on the path length n, by constructing an equivalent graph where the path $P_{A \to u}$ is contracted to a single edge (A, u). In this proof, we will use the somewhat cumbersome notation for a measurement in one reference graph $[\mathbf{u}_A]_G$ to show that it is equivalent to the corresponding measurement in another, but this notation will be unnecessary once the equivalence is shown.

Proof. In the $n = 1$ case, $P_{A\to u} = \{(A, u)\}$, $\varphi((A, u)) = \mathbf{u}_A$ by definition.

In the $n = 2$ case, G contains $P_{A\to u} = \{(A, A_1), (A_1, u)\}$. We claim there is an equivalent reference graph G' formed by removing the (A, A_1) edge from the path:

$$G' = (V, (E\backslash\{(A_1, u)\}) \cup \{(A, u)\} \tag{24}$$

We claim $G' \sim G$ if φ' is defined as

$$\varphi'(e) = \begin{cases} \varphi((A, A_1))\varphi((A_1, u)) & \text{if } e = (A, u) \\ \varphi(e) & \text{otherwise} \end{cases} \tag{25}$$

This fact follows from the definition of the frame transformation, since the measurement $\mathbf{u}_A$ is the same in both graphs:

$$\begin{aligned} [\mathbf{u}_A]_G &= \varphi'((A, u)) && (26)\\ &= \varphi((A, A_1))\varphi((A_1, u)) && (27)\\ &= \mathbf{F}_{AA_1}\mathbf{u}_{A_1} && (28)\\ &= [\mathbf{u}_A]_{G'} && (29) \end{aligned}$$

by the definigion of $\mathbf{F}_{AA_1}$. Thus, the measurement of the point is the same in both graphs, and $G' \sim G$.

In the $n+1$ case, assume the theorem holds for the nth case. Then $P_{A\to u} = \{e_0, e_1, \ldots, e_{n-1}, e_n\}$. Since every node in $\mathcal{U}$ has out-degree 0, every edge $e \in P_{A\to u}$ must start at a node in $\mathcal{F}$. By the definition of a path, then,

$$e_i \in \mathcal{F} \times \mathcal{F} \; \forall \, i \in 0, 1, \ldots, n - 1. \tag{30}$$

That is, every edge in the path ends at a frame, except the last. Denote $e_{n-1} = (A_{n-1}, A_n)$, $e_n = (A_n, u)$. Then,

$$\begin{aligned} P_{A\to u} = \{e_0, e_1, \ldots, e_{n-2}, (A_{n-1}, A_n), (A_n, u)\} && (31)\\ && (32) \end{aligned}$$

Following the same graph operation in the $n = 2$ case, contract the path by removing the (A_{n-1}, A_n) edge, and constructing φ' similarly. Then we have an equivalent reference graph G' with an n-path starting at A and ending at u. From the inductive hypothesis,

$$\begin{aligned} [\mathbf{u}_A]_{G'} &= \varphi'(e_0)\varphi'(e_1)\cdots\varphi'(e_{n-2})\varphi'((A_{n-1}, u)) && (33)\\ &= \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-2})\mathbf{u}_{A_{n-1}} && (34)\\ &= \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-2})\mathbf{F}_{A_{n-1}A_n}\mathbf{u}_{A_n} && (35)\\ &= \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-2})\varphi((A_{n-1}, A_n))\varphi((A_n, u)) && (36)\\ &= [\mathbf{u}_A]_G && (37) \end{aligned}$$

which is the $n + 1$ claim. □

Thus, we have established the usual relationship, namely that kinematic chains ending at points are formed by composing the frame transformations along the path in the reference graph. A similar fact holds for measurements of frames.

Corollary 3.6.1 (Kinematic Paths on Frames)**.** *Given a reference graph $G = (\mathcal{F} \sqcup \mathcal{U}, E)$, the measurement $\mathbf{F}_{AB}$ of a frame $f \in \mathcal{U}$ in a frame $A \in \mathcal{F}$ can be determined if G contains a directed path $P_{A\to B} = \{e_0, e_1, \ldots, e_{n-1}\}$ starting at A and ending at B. Moreover,*

$$\mathbf{F}_{AB} = \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-1}). \tag{38}$$

Proof. It suffices to prove that any given point measured in the graph G' formed by contracting $P_{A\to B}$ to a single edge (A, B) with

$$\varphi'((A, B)) = \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-1}) \tag{39}$$

is the same point measured in G. This clearly holds for any point not measured by G in B. Then consider an arbitrary point u with $(B, u) \in E$. From Theorem 3.5, we have the desired claim. □

In order to avoid the construction of endless reference graphs which are equivalent but reduced in the manner above (and therefore less reflective of the original measurements), it is convenient to extend the definition of the measurement map φ to cover paths on the graph in a recursive manner.

$$\varphi(x) = \begin{cases} \mathbf{F}_{AB} & \text{if } x = (A, B) \in \mathcal{F} \times \mathcal{F} \\ \mathbf{u}_A & \text{if } x = (A, u) \in \mathcal{F} \times \mathcal{U} \\ \varphi(e_0)\varphi(e_1)\cdots\varphi(e_{n-1}) & \text{if } x = \{e_0, e_1, \ldots, e_{n-1}\} \text{ is a path on } G \end{cases} \tag{40}$$

3.2.1 Frame Composition Formula

Although we have shown that the composition of frames results in the desired measurement, we have not given a formula for composing frames unless they are attached to points. We will do so now for the composition of two frames, from which the composition of arbitrary numbers of frames can be derived.

Corollary 3.6.2 (Frame Transformation Composition)**.** *The composition of two frame transformations $\mathbf{F}_{AB}, \mathbf{F}_{BC}$ is*

$$\mathbf{F}_{AC} = \mathbf{F}_{AB}\mathbf{F}_{BC} = [\mathbf{R}_{AB}\mathbf{R}_{BC}, \mathbf{R}_{AB}\mathbf{p}_{BC} + \mathbf{p}_{AB}]. \tag{41}$$

Proof. Let $\mathbf{u}_C$ be the measurement of some point in frame C. Then according to the kinematic path,

$$\begin{aligned}
\mathbf{u}_A &= \mathbf{F}_{AB}\mathbf{F}_{BC}\mathbf{u}_C && (42)\\
&= \mathbf{F}_{AB}(\mathbf{R}_{BC}\mathbf{u}_C + \mathbf{p}_{BC}) && (43)\\
&= \mathbf{R}_{AB}(\mathbf{R}_{BC}\mathbf{u}_C + \mathbf{p}_{BC}) + \mathbf{p}_{AB} && (44)\\
&= (\mathbf{R}_{AB}\mathbf{R}_{BC})\mathbf{u}_C + (\mathbf{R}_{AB}\mathbf{p}_{BC} + \mathbf{p}_{AB}) && (45)\\
&= [\mathbf{R}_{AB}\mathbf{R}_{BC}, \mathbf{R}_{AB}\mathbf{p}_{BC} + \mathbf{p}_{AB}]\mathbf{u}_C && (46)
\end{aligned}$$

$\square$

It is worth taking a moment to confirm that this composition formula makes sense with respect to the points being manipulated. Recall that $\mathbf{p}_{BC}$ is equivalent to the measurement of the origin of frame C as measured in frame B. Then $\mathbf{R}_{AB}\mathbf{p}_{BC} + \mathbf{p}_{AB} = \mathbf{F}_{AB}\mathbf{p}_{BC}$ is just the origin of C measured in frame A, as we would expect.

4 Summary

To summarize, the first approach to solving tracker problems in computer integrated surgery is as follows:

1. Identify the points and frames in the problem as nodes on a reference graph.

2. Identify the known measurements as edges on the graph.

3. Remember that inverse edges of known transformations are also on the graph.

4. Identify the desired measurement as the path starting at the *frame they are in* ending at the *object being measured*.

5. Compose the desired transformations and evaluate.

The transformations in question may be more complex than those discussed here, such as projections or nonlinear maps, but those details only affect the mathematical evaluation in the final step. Of course, the hardest part in any problem is not algebraic manipulation but rather mapping from real-world descriptions and constraints to diagrams, reference graphs, and equations. Thus, in the following example, we present the problem at a very high level, using none of the vocabulary we have developed so far, and encourage the reader to follow steps 1 - 5 on their own, including setting up the frames and drawing the initial reference graph. Keep in mind, the pose of "B with respect to A" is really the "A from B" transform $\mathbf{F}_{AB}$. The phrase "relative to" is also often used to mean the same thing, so that "tool pose relative to tracker" really means the "tracker from tool" transformation. Translating from common vernacular used in many problems to our more consistent vocabulary for frame transformations is an important first step. It is easier to remember that the arrow always starts at the frame of reference being provided, in these cases.

In the example, we give the problem statement, then a schematic diagram clarifying the problem statement, then an overview of the frames and points involved, then a reference graph describing the setup, and finally a solution. At each step, we encourage the reader to attempt the next step on their own, understanding that the hardest part may be translating the problem statement, which requires understanding some medical jargon, to frames and transforms. This is an essential skill, and the best way to acquire it is practicing problems much like this one.

Problem 4.1. Internal pelvic trauma fixation involves the insertion of a rigid metal rod, called a K-wire, into the pelvic bone along a specific trajectory. It is essential to identify the correct trajectory so as to avoid "cortical breach," where the K-wire punches through the outer cortical bone and enters soft tissue, leading to further complications. For this example, assume the fracture is along the superior pubic ramus and is already shortened (aligned).

The operating room is equipped with a calibrated X-ray imaging device, called a C-arm, with infrared reflective markers on its gantry. The calibration provides the pose of the C-arm's camera frame—a frame with its origin at the center of the X-ray tube. There is a Polaris optical tracker positioned such that it measures the pose of both the C-arm and a reference marker fixed to the patient table.

The surgeon is wearing a HoloLens 2, an "augmented" or "mixed reality" headset, which is capable of displaying holograms to the surgeon. The HoloLens tracks the surgeon's head movements using onboard sensors, relative to an estimated "world" frame. The HoloLens is also capable of tracking infrared reflective markers, relative to the headset frame, but the markers must be in the field of view. Since the surgeon is generally looking down, only the marker fixed to the patient table can be reliably tracked by the HoloLens.

There is a machine intelligent system which is capable of automatically determining the safe corridor for K-wire insertion based on X-ray images. The input to this system is the X-ray image, and the output is a start- and end-point for the safe trajectory in the camera frame.

Finally, suppose we want to support the safe insertion of the K-wire by displaying a hologram for the surgeon to align the K-wire with. The hologram is a virtual object in the Hololens "world," with its own coordinate frame such that the +Z-axis should aligns with the safe trajectory (going into the body), and the origin should be placed at the trajectory startpoint.

Virtual objects are controlled by setting their rotation and translation relative to the HoloLens "world." What should these be set to, in terms of the measurements provided, so as to guide the surgical alignment?

Solution 4.2. First, let's talk about the hardware in this setup. In Fig. 8, the X-ray imaging device is a Siemens C-arm with intrared markers attached to the detector side of the gantry. The locations of these markers have already been encoded relative to some frame, call it M. When the Polaris tracker, the stereo camera object in the top right of the figure, measures the pose of the C-arm, it measuring the $\mathbf{F}_{TM}$ frame transform, where T is the tracker frame. It doesn't matter precisely how M is defined, because the calibration of the C-arm provides the "C from M" transform, where C is the "camera" frame centered at the X-ray source. Let u and v be the start- and endpoints of a suitable trajectory, so $\mathbf{u}_C$ and $\mathbf{v}_C$ are estimated by the machine learning system in the camera frame.

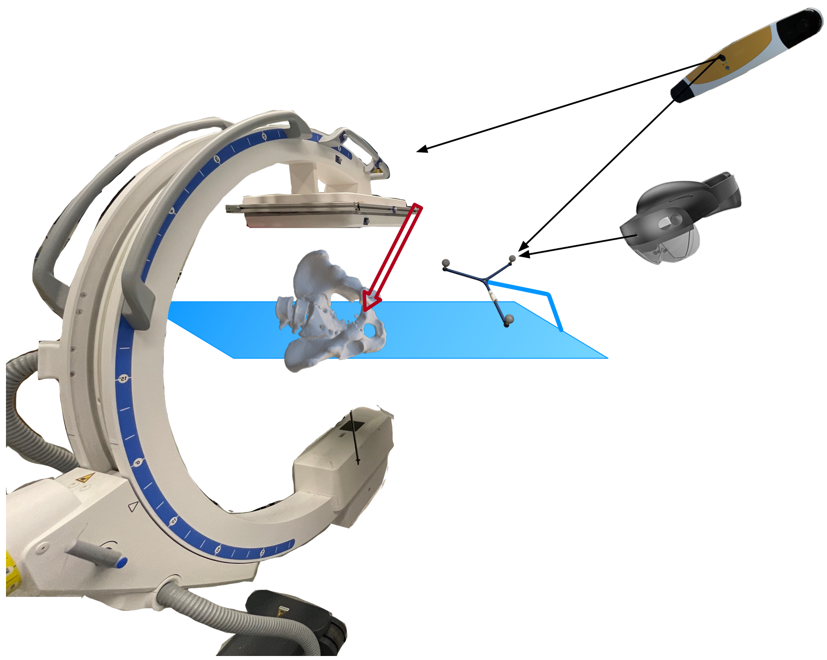

Figure 8: An overview of the hardware in the problem. The C-arm is equipped with infrared reflective markers which can be tracked, and a stable reference marker is fixed to the patient table (blue). The Polaris tracker (orange) tracks both the C-arm and reference marker, while the HoloLens only tracks the reference marker (since the surgeon is often looking down). The goal of the problem is to align a holographic indicator, such as the red arrow, with a desired trajectory with respect to the pelvis.

The tracker also measures the pose of a reference marker frame, call it R, which is simultaneously tracked by the HoloLens. The reason to include this marker is to provide a kinematic chain between the HoloLens "world" frame and the camera-centered coordinates measured by the C-arm. Denote the frame of the headset as H and the HoloLens world frame as W. Finally, denote the frame containing the hologram as A (for arrow), and assume that the hologram is aligned with the Z-axis.

The next step is to translate the above diagram into a more formal frame transformation diagram, with the frames discussed. (Try first before continuing to the next page.)

Fig. 9 captures the essential frames in the problem, as well as the raw measurements that exist between them. Already we can see the kinematic path that allows us to relate the virtual world of the HoloLens to the real world of the patient.

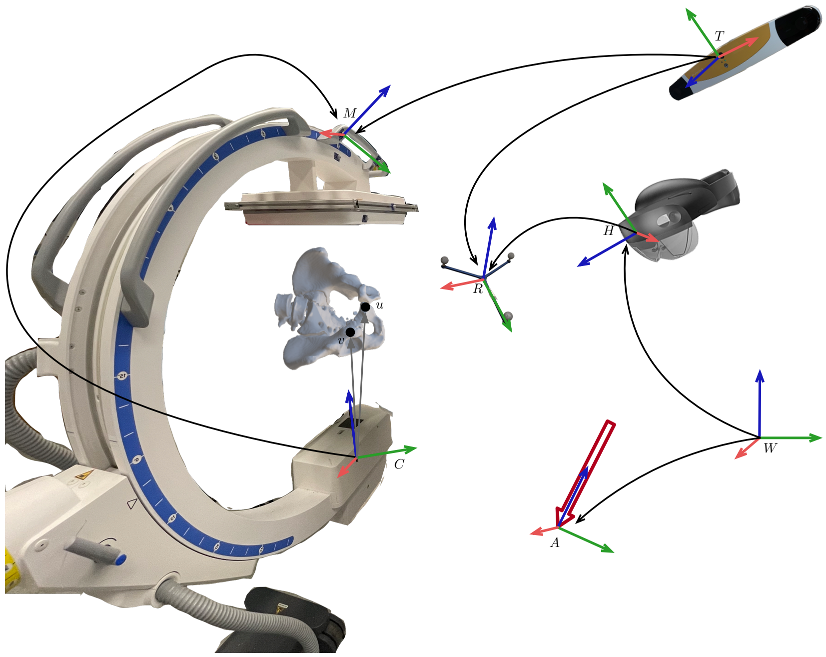

Figure 9: A frame transform diagram showing the initial measurements in the problem.

Often, Fig. 9 is sufficient to see the transformations that need to be composed, especially once the translation to reference graphs is clear. To be explicit, Fig. 10 shows the complete reference graph.

Now that we understand all the frames in the problem and the measurements between them, the hard part is over. All that remains is to define $\mathbf{F}_{WA} = [\mathbf{R}_{WA}, \mathbf{p}_{WA}$ such that the origin of A is at u, and the $-Z$ axis points in the same direction as the $u \rightarrow v$ vector.

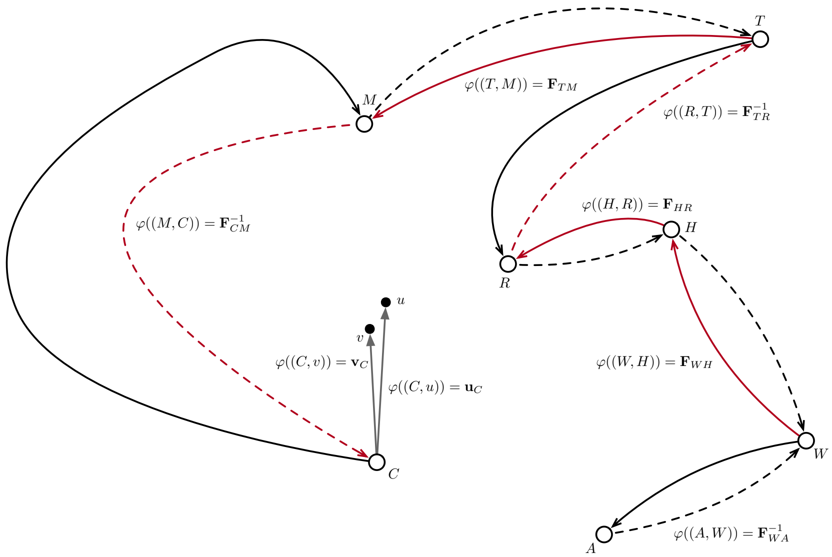

Figure 10: The complete reference graph for the problem, with the desired path in red.

The first step is straightforward using Theorem 3.5. We have

$$
\begin{aligned}
\mathbf{p}_{WA} &\leftarrow \mathbf{u}_W && (47)\\
&= \varphi(P_{W\to u}) && (48)\\
&= \mathbf{F}_{WH}\mathbf{F}_{HR}\mathbf{F}_{TR}^{-1}\mathbf{F}_{TM}\mathbf{F}_{CM}^{-1}\mathbf{u}_C && (49)\\
&= [\mathbf{R}_{WH}, \mathbf{p}_{WH}][\mathbf{R}_{HR}, \mathbf{p}_{HR}][\mathbf{R}_{TR}^{-1}, -\mathbf{R}_{TR}^{-1}][\mathbf{R}_{TM}, \mathbf{p}_{TM}][\mathbf{R}_{CM}^{-1}, -\mathbf{R}_{CM}^{-1}]\mathbf{u}_C && (50)\\
&= \Big[\mathbf{R}_{WH}\mathbf{R}_{HR}\mathbf{R}_{TR}^{-1}\mathbf{R}_{TM}\mathbf{R}_{CM}^{-1}, && (51)\\
&\qquad - \mathbf{R}_{WH}\mathbf{R}_{HR}\mathbf{R}_{TR}^{-1}\mathbf{R}_{TM}\mathbf{p}_{CM} && (52)\\
&\qquad + \mathbf{R}_{WH}\mathbf{R}_{HR}\mathbf{R}_{TR}^{-1}\mathbf{p}_{TM} && (53)\\
&\qquad - \mathbf{R}_{WH}\mathbf{R}_{HR}\mathbf{R}_{TR}^{-1}\mathbf{p}_{TR} && (54)\\
&\qquad - \mathbf{R}_{WH}\mathbf{p}_{HR} && (55)\\
&\qquad + \mathbf{p}_{WH}\Big]\mathbf{u}_C && (56)\\
&\equiv \mathbf{F}_{WC}\mathbf{u}_C && (57)\\
&= \mathbf{R}_{WC}\mathbf{u}_C + \mathbf{p}_{WC} && (58)
\end{aligned}
$$

Similarly, we note $\mathbf{v}_W = \mathbf{F}_{WC}\mathbf{u}_C$.

Next, all that remains is to ensure the hologram's rotation aligns with $u \to v$. This can be done by answering the question, "what is the rotation that aligns $[0, 0, 1]^T$ with $\mathbf{u}_W - \mathbf{v}_W$?" This question can be answered using Rodrigues's formula and the vector-angle formulation of rotations, which we will state but not explore in much detail. Briefly, Rodrigues's formula $\text{Rot}(\mathbf{w}, \theta)$ takes the vector $\mathbf{w}$ about which the rotation occurs and the angle θ, and returns the corresponding rotation matrix. Here, $\mathbf{w}$ must be perpendicular to both vectors, which we can obtain with the cross product:

$$\mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times (\mathbf{u}_W - \mathbf{v}_W) \tag{59}$$

Note that Rodriguez's formula does not take into account the magnitude of this vector, only its direction.

The angle θ is given by

$$\cos\theta = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \frac{\mathbf{u}_W - \mathbf{v}_W}{||\mathbf{u}_W - \mathbf{v}_W||} \tag{60}$$

Thus

$$\mathbf{R}_{WA} \leftarrow \text{Rot}(\mathbf{w}, \theta) \tag{61}$$

$$= \mathbf{I} + (\sin\theta)\,\mathbf{sk}(\hat{\mathbf{w}}) + (1 - \cos\theta)\,\mathbf{sk}(\hat{\mathbf{w}})^2 \tag{62}$$

where $\mathbf{sk}(\hat{\mathbf{w}})$ is the skew matrix of $\hat{\mathbf{w}}$.

5 Conclusion

In this tutorial, you learned how to take a description of a tracking system, such as those used in surgical navigation, and extract desired information. As you may have figured out, this is a pretty simple process once you have a consistent notation and vocabulary for discussing transformations, although the algebra can become tedious if done by hand. Fortunately software packages like tf2 in ROS or pytransform3d in Python/Numpy allow users to manage the frame transform graph and request transforms between frames automatically.

The next step, which we will discuss in future tutorials, is understanding how the potential *error* in each measurement, from the point itself to every transformation in between, affects the end result.