“My own business always bores me to death; I prefer other people's.” Oscar Wilde
Before we get started...

- **Last Time:**
  - Algebraic Vectors (the algebraic counterpart of Geometric Vectors)
  - Understand what it takes to do a change of RF
  - Started discussion about time derivative of the orientation matrix $A$
    - We barely introduced the concept of angular velocity

- **Today:**
  - Finish introducing the concept of angular velocity of a rigid body
  - Talk about the number of generalized coordinates required to characterize the orientation of a rigid body

- HW1 returned, solutions posted at Learn@UW

- I’ll be out on Th Feb. 4
  - Justin and Makarand will present an overview of ADAMS
  - Today’s lecture is 20 minutes shorter
Angular Velocity: Getting There…

- Recall that $\mathbf{AA}^T = \mathbf{I}_3$. Take time derivative to get:

$$\dot{\mathbf{A}} \mathbf{A}^T + \mathbf{A} \dot{\mathbf{A}}^T = 0_{3 \times 3} \implies \dot{\mathbf{A}} \mathbf{A}^T = -\mathbf{A} \dot{\mathbf{A}}^T$$

- Notice the following:
  - The matrix $\dot{\mathbf{A}} \mathbf{A}^T$ is a $3 \times 3$ matrix
  - The matrix $\dot{\mathbf{A}} \mathbf{A}^T$ is skew-symmetric

- CONCLUSION: there must be a vector, $\omega$, whose cross product matrix is equal to the $3 \times 3$ skew symmetric matrix $\dot{\mathbf{A}} \mathbf{A}^T$:

$$\ddot{\omega} = \dot{\mathbf{A}} \mathbf{A}^T$$

- This vector $\omega$ is called the angular velocity of the L-RF with respect to the G-RF.
Problem Setup:
- You have one rigid body and two different L-RF rigidly attached to that body
  - Rigidly attached means that their relative orientation never change
  - Rigidly attached to the body = “welded” to the body ⇒ they move as the body moves

- Call the local references frames and orientation matrices L-RF$_1$, A$_1$, and L-RF$_2$, A$_2$

Question: what is the relationship between A$_1$ and A$_2$?

\[
A_1 = \begin{bmatrix} f_1 & g_1 & h_1 \end{bmatrix} \quad A_2 = \begin{bmatrix} f_2 & g_2 & h_2 \end{bmatrix}
\]
The important observation: since both \( \text{L-RF}_1 \) and \( \text{L-RF}_2 \) are “welded” to the rigid body, their relative attitude (orientation) doesn’t change in time.

Equivalent way of saying this:

\[
\begin{align*}
\textbf{f}_2(t) &= c_{11} \textbf{f}_1(t) + c_{21} \textbf{g}_1(t) + c_{31} \textbf{h}_1 \\
\textbf{g}_2(t) &= c_{12} \textbf{f}_1(t) + c_{22} \textbf{g}_1(t) + c_{32} \textbf{h}_1 \\
\textbf{h}_2(t) &= c_{13} \textbf{f}_1(t) + c_{23} \textbf{g}_1(t) + c_{33} \textbf{h}_1
\end{align*}
\]

\[
\begin{bmatrix}
\textbf{f}_2(t) & \textbf{g}_2(t) & \textbf{h}_2(t)
\end{bmatrix} =
\begin{bmatrix}
\textbf{f}_1(t) & \textbf{g}_1(t) & \textbf{h}_1(t)
\end{bmatrix}
\begin{bmatrix}
c_{11} & c_{12} & c_{13} \\
c_{21} & c_{22} & c_{23} \\
c_{31} & c_{32} & c_{33}
\end{bmatrix}
\]

\[
\textbf{A}_2(t) = \textbf{A}_1(t) \cdot \mathbf{C}
\]

The important point: \( \mathbf{C} \) is a constant matrix (since \( \text{L-RF}_1 \) and \( \text{L-RF}_2 \) are “welded” to the rigid body).
The Invariance Property of $\omega$

- Recall that we saw that the angular velocity was implicitly defined by the identity $\tilde{\omega} = \dot{\mathbf{A}} \mathbf{A}$. This somewhat suggests that $\omega$ is related to the matrix $\mathbf{A}$. What follows proves that this is not the case, instead, $\omega$ is an attribute of the rigid body the L-RF is attached to.

- First, assume that there are two different angular velocities: $\omega_1$, which goes along with L-RF$_1$, and $\omega_2$, which goes along with L-RF$_2$, where the two L-RFs are rigidly attached to the same body.

- Then, since $\mathbf{A}_2 = \mathbf{A}_1 \mathbf{C}$, we have $\dot{\mathbf{A}}_2 = \dot{\mathbf{A}}_1 \mathbf{C}$, which implies that

$$\tilde{\omega}_2 \mathbf{A}_2 = \tilde{\omega}_1 \mathbf{A}_1 \mathbf{C}$$

- Since $\mathbf{A}_2 = \mathbf{A}_1 \mathbf{C}$, we get that

$$\tilde{\omega}_2 = \tilde{\omega}_1$$

- In other words, the angular velocity is an attribute of the body, not of the L-RF rigidly attached to it.
Angular Velocity: On Its Representation in the L-RF

- Assume you have a L-RF attached a body
- Assume that the angular velocity is $\omega$
- Question: what is its representation in the L-RF?

\[
\tilde{\omega} = \dot{\mathbf{A}} \mathbf{A}^T \quad \text{and} \quad \tilde{\omega} = \mathbf{A}^T \tilde{\omega} \mathbf{A}
\]

\[
\tilde{\omega} = \mathbf{A}^T \dot{\mathbf{A}}
\]

- Therefore, we have that

\[
\tilde{\omega} = \dot{\mathbf{A}} \mathbf{A}^T \quad \text{and} \quad \tilde{\omega} = \mathbf{A}^T \dot{\mathbf{A}}
\]

- Note that this also yields two ways of representing the time derivative of $\mathbf{A}$:

\[
\dot{\mathbf{A}} = \tilde{\omega} \mathbf{A} \quad \text{and} \quad \dot{\mathbf{A}} = \mathbf{A} \tilde{\omega}
\]
The Second Time Derivative of $\mathbf{A}$

- Straight forward application of the definition of the first time derivative of $\mathbf{A}$ combined with the chain rule of differentiation

- Using the angular velocity and its derivative expressed in the G-RF:
  \[ \ddot{\mathbf{A}} = \dot{\omega} \mathbf{A} + \ddot{\omega} \dot{\mathbf{A}} = \ddot{\omega} \mathbf{A} + \dddot{\omega} \mathbf{A} = (\ddot{\omega} + \dddot{\omega}) \mathbf{A} \]

- Using the angular velocity and its derivative expressed in the L-RF:
  \[ \ddot{\mathbf{A}} = \mathbf{A} \dot{\ddot{\omega}} + \dot{\mathbf{A}} \ddot{\omega} = \mathbf{A} \dddot{\omega} + \mathbf{A} \dddot{\omega} = \mathbf{A} (\dddot{\omega} + \dddot{\omega}) \]
The Implicit Function Theorem provides the guarantee that a relation can locally be turned into a function

- What do I mean by ‘relation’?
  - Here is an example:

\[ x^2 + y^2 - 8 = 0 \]  

(1)

- What do I mean by ‘function’?
  - Here is the function that goes with the above example:

\[ y(x) = \sqrt{8 - x^2} \]

- What do I mean by ‘locally’?
  - The meaning of ‘locally’ is the fact that if I take \( x = 3 \), relation (1) ceases to define a function anymore. When \( x = 2 \), in a neighborhood of this value things are good, but there is no guarantee that you can make a global assumption about the nature of the function \( y(x) \) that comes out of a relation.
There is one more important thing to be considered in relation to the “locality” aspect

- Note that both $x=2, y=2$ and $x=2, y=-2$ verify the relation (1).
- However, $x=2, y=2$ forces the relation to lead to this function:
  \[ y(x) = \sqrt{8 - x^2} \]

- Yet $x=2, y=-2$ forces the relation to lead to this different function,
  \[ y(x) = -\sqrt{8 - x^2} \]

Conclusion: the values $x_0, y_0$ around which you seek the function that comes out of a relation play a role in defining the expression of that function.
Let \( u : \mathbb{R}^{n+m} \to \mathbb{R}^m \), and for convenience we will use two letters to denote the entries of any element of \( \mathbb{R}^{n+m} \) such as in \((x_1, \ldots, x_n, y_1, \ldots, y_m)^T = (x, y) \equiv z\).

Note that by setting \( u(x, y) = 0_m \) we procure the relation we referenced a couple of slides ago.

Assume we have a point \((a_1, \ldots, a_n, b_1, \ldots, b_m)^T \in \mathbb{R}^{n+m}\) that satisfy our relation; i.e., \( u(a, b) = 0_m \).

What we want to accomplish here is to find, in the neighborhood of the point \( x = a \), the function of several slides ago that is induced by the relation \( u(x, y) = 0_m \). We will call this function \( v(x) \), and note that \( v : N_a \to \mathbb{R}^m \), where \( N_a \) is a neighborhood of the point (open set of) \( x = a \).

Note that if this is indeed the function induced by the relation, then we must have that \( u(x, v(x)) = 0_m \); that is, \( y = v(x) \).
The Implicit Function Theorem: Bringing in the Jacobian

- We will need the Jacobian of the relation we were provided. Specifically, consider the partial derivative of $\mathbf{u}$:

$$\frac{\partial \mathbf{u}}{\partial z} = \begin{bmatrix}
\frac{\partial u_1}{\partial x_1}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial u_1}{\partial x_n}(\mathbf{a}, \mathbf{b}) & \frac{\partial u_1}{\partial y_1}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial u_1}{\partial y_m}(\mathbf{a}, \mathbf{b}) \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial u_m}{\partial x_1}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial u_m}{\partial x_n}(\mathbf{a}, \mathbf{b}) & \frac{\partial u_m}{\partial y_1}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial u_m}{\partial y_m}(\mathbf{a}, \mathbf{b})
\end{bmatrix} = [ \mathbf{u}_x(\mathbf{a}, \mathbf{b}) \mathbf{u}_y(\mathbf{a}, \mathbf{b}) ]$$

- Note that this partial derivative was evaluated at the 'good' point $(\mathbf{a}, \mathbf{b})$; i.e., $\mathbf{a}$, and $\mathbf{b}$, satisfy the relation $\mathbf{u}(\mathbf{a}, \mathbf{b}) = 0_m$.

- We concentrate on the $m \times m$ matrix $\mathbf{u}_y(\mathbf{a}, \mathbf{b})$. Specifically, the determinant of this submatrix will be assumed to nonzero; i.e., $\mathbf{u}_y(\mathbf{a}, \mathbf{b})$ is nonsingular.
The Implicit Function Theorem: The Actual Formal Thing

**Theorem:** Assume that the function $u(x, y)$ introduced two slides ago is continuously differentiable, and assume that at the ’good’ point $(a, b)$, we have that $u(a, b) = 0_m$ and $\text{det}[u_y(a, b)] \neq 0$. Then there exists

- an open set $N_a$ (neighborhood of $a$)
- an open set $N_b$ (neighborhood of $b$)
- a *unique* continuously differentiable function $v : N_a \to N_b$

such that,

$$u(x, v(x)) = 0_m \quad \text{for any} \quad x \in N_a$$

**Important observation regarding differentiability of $v(x)$:**

It can be proved that whenever we have the additional hypothesis that $u$ is continuously differentiable up to $k$ times inside $N_a \times N_b$, then the same holds true for the induced function $v$ inside $N_a$. For instance, if $k = 1$, we have that for any $x \in N_a$

$$\frac{\partial v}{\partial x_j}(x) = -[u_y(x, v(x))]^{-1} \cdot \frac{\partial u}{\partial x_j}(x)$$
Recall our original example, where the relation $u(x, y) = 0$ was provided using $u(x, y) = x^2 + y^2 - 8$. Note that in this case $n = m = 1$.

The Jacobian of interest is simple in this case: $J = [2x \quad 2y]$

The subjacobian of interest is $u_y(x, y) = 2y$ (in fact for this simple case, $u_y$ doesn’t even depend on $x$).

Note that for any value $y \neq 0$, one has that $det(u_y) \neq 0$, which is what we need to get a unique function $v(x)$ defined in the neighborhood of any point $a$, where $(a, b)$ is chosen such that is satifies $u(a, b) = 0$. For instance, $(1, \sqrt{7})$ qualifies as such a point.

It’s interesting to take a close look to see what happens when actually $det(u_y) = 0$. One such a point, but not the only one, in our case would be $(2\sqrt{2}, 0)$. In this case one can easily notice that there is a loss of uniqueness since both $v(x) = \sqrt{8 - x^2}$ and $v(x) = -\sqrt{8 - x^2}$ are equally good candidates.
The Implicit Function Theorem: Concluding Remarks

- Implicit Function Theorem is one of the deep theorems of Applied Math
- Make friends with Implicit Function Theorem
Degrees of Freedom Count, Orientation

- The rotation matrix $A$ has nine direction cosines.

- Recall the story of the birth of the $A$ matrix: we started with a L-RF attached to a rigid body. Having a L-RF means that you have a triplet $\vec{f}$, $\vec{g}$, and $\vec{h}$. Having the triplet allowed us to generate the matrix $A$ (since we had its columns).

- Note the following: for each orientation (attitude) of the right body, you have a matrix $A$. Likewise, for each matrix $A$, you have an attitude of the rigid body.

- Conclusion: As soon as you decide on the value of the direction cosines you basically determined the orientation of the rigid body that that L-RF is rigidly attached to.
Degrees of Freedom Count, Orientation

- QUESTION: How many of the nine directions cosines can we specify?
  - In order words, how many free parameters do I have in conjunction with the rotation matrix $A$?

- ANSWER: 3
  - I’m going to show that I can always choose three of the direction cosines and express the other six as a function of the three chosen ones

- Why can I do “six function of three” trick?
  - I can do it because of these six conditions that the direction cosines satisfy:
    \[
    f^T f = g^T g = h^T h = 1 \quad \& \quad f^T g = g^T h = h^T f = 0
    \]

- How am I going to prove the “six function of three” trick?
  - I’m going to use the Implicit Function Theorem
Consider the relation \( u(a_{11}, a_{12}, \ldots, a_{33}) = 0_6 \), where \( u \) is defined as:

\[
\begin{bmatrix}
\frac{1}{2} f^T f - \frac{1}{2} \\
\frac{1}{2} g^T g - \frac{1}{2} \\
\frac{1}{2} h^T h - \frac{1}{2} \\
f^T g \\
g^T h \\
h^T f
\end{bmatrix}
\]

I have six scalar relations that capture the interplay between nine direction cosines.

I’m going to apply the Implicit Function Theorem to show that I can always express six of them as a function \( v \) that depends on the other three of them.
First, introduce the following notation to simplifying the math:

\[
q = \begin{bmatrix} f \\ g \\ h \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{33} \end{bmatrix}
\]

Second, I’ll compute the Jacobian \( u_q \), since I’m about to use the Implicit Function Theorem:

\[
u_q = \begin{bmatrix} f^T & 0 & 0 \\ 0 & g^T & 0 \\ 0 & 0 & h^T \\ g^T & f^T & 0 \\ 0 & h^T & g^T \\ h^T & 0 & f^T \end{bmatrix}_{6 \times 9}
\]

Third, the Jacobian \( u_q \) has full row rank. You can prove this by showing that the only linear combination of the rows that can produce a zero vector requires that the coefficients be all zero. Specifically, I will prove that \( u_q^T x = 0 \Rightarrow x_1 = x_2 = \ldots = x_6 = 0 \).
Fourth, note that we just proved that the rank of \( \text{rank}(u_q) = 6 \). This means that its column rank is also six, which means that out of the nine columns of \( u_q \) I can choose six that are linearly independent. I’m going to group these six columns together to make up a matrix that I’m going to call \( V \). For this \( 6 \times 6 \) matrix, I have that \( \det(V) \neq 0 \).

Fifth, get the remaining three (out of nine) columns, and organize them as matrix \( U \). At this point, we have the following: a relation \( u(q) = 0_6 \), and the Jacobian \( u_q = [U \ V] \), with \( \det(V) \neq 0 \). According to the Implicit Function Theorem, the six direction cosines that are associated with the columns that were used to make up \( V \) can be expressed as a function that depends on the other three direction cosines.

**IMPORTANT CONCLUSION:** You need three variable to express all the other six entires of the orientation matrix \( A \). In other words, once you specify the value of the three direction cosines, you actually define the orientation of the rigid body that the L-RF associated with \( A \) is attached to.

**NOTE:** I do not know which three out of the nine direction cosines can be used to express the other six as a function of. In fact, chances are that as \( A \) changes in time, every once in a while you’d have to change the three direction cosines, since you always have to have a \( V \) that is nonsingular.
Remarks,
Orientation Degrees of Freedom

- **QUESTION:**
  - Why do we obsess about how many degrees of freedom do we have?

- **ANSWER:**
  - We need to know how many generalized coordinates we’ll have to include in our set of unknowns when solving for the time evolution of a dynamic system.

  In this context, we arrived to the conclusion that three direction cosines are independent and need to be accounted for. The other six can be immediately computed once the value of the three independent direction cosines becomes available.
  - You can say that you solve for three, and recover the other six.
**Remarks, Orientation Degrees of Freedom [Cntd.]**

- **QUESTION:**
  - Do I really have to choose three direction cosines and include them in the set of generalized coordinates that I’ll use to understand the time evolution of the mechanical system?

- **ANSWER:**
  - No, in fact I haven’t heard of anyone who does this.
  - What is important is the number of generalized coordinates that are needed.
  - Specifically:
    - I can choose three other quantities, call them $\theta$, $\phi$, $\gamma$, that I decide to adopt as my three rotation generalized coordinates as long as there is a ONE-TO-ONE mapping between these three generalized coordinates and *any set of three out of the nine* direction cosines of the rotation matrix $A$. 


Remarks,
Orientation Degrees of Freedom [Cntd.]

Here are a couple of possible scenarios:

- I indeed choose three generalized coordinates: this is what Euler did, when he chose the Euler Angles to define the entries of $A$ and thus capture the orientation of a L-RF with respect to the G-RF.

- I can choose a set of quaternion, or Euler parameters. There is four of them: $e_0$, $e_1$, $e_2$, and $e_3$, but they are related through a normalization condition:

  $$e_0^2 + e_1^2 + e_2^2 + e_3^2 = 1$$

- Quaternions: born on Monday, October 16, 1843 in one of Sir William Rowan Hamilton’s moments of inspiration.

- If you want to be extreme, you do what people in Spain did: they chose all of the nine direction cosines as generalized coordinates but also added to the equations of motion the following set of six algebraic constraints:

  $$f^T f = g^T g = h^T h = 1 \quad \& \quad f^T g = g^T h = h^T f = 0$$
Hopping from RF to RF

• The discussion framework:
  • Recall that when going from one A-RF$_2$ to a different A-RF$_1$, there is a transformation matrix that multiplies the representation of a geometric vector in A-RF$_2$ to get the representation of the geometric vector in A-RF$_1$:

  \[ \bar{s}_1 = A_{12} \cdot \bar{s}_2 \]

• Question: What happens if you want to go from A-RF$_3$ to A-RF$_2$ and then eventually to the representation in A-RF$_1$?

• Why are we curious?
  • Comes into play when dealing with Euler Angles
Hopping from RF to RF

- Going from A-RF$_3$ to A-RF$_2$ to A-RF$_1$:
  \[ \bar{s}_2 = A_{23} \cdot \bar{s}_3 \oplus \bar{s}_1 = A_{12} \cdot \bar{s}_2 \Rightarrow \bar{s}_1 = A_{12}A_{23} \cdot \bar{s}_3 \]

- The basic idea is clear, you keep multiplying rotation matrices like that to hop from RF to RF until you arrive to your final destination.

- However, how would you actually go about computing $A_{ij}$ if you have $A_i$ and $A_j$ (that is the two rotation matrices from RF$_i$ to RF$_j$, respectively, into the G-RF)?
  - This means that you hop from A-RF$_j$ to A-RF$_i$.
  - Keep in mind the invariant here, that is, the geometric vector $\bar{s}$ whose representation you are playing with:

\[
\begin{align*}
\{ s = A_i \bar{s}_i \\ s = A_j \bar{s}_j \} & \Rightarrow A_i \bar{s}_i = A_j \bar{s}_j \\
& \Rightarrow \bar{s}_i = A_i^T A_j \bar{s}_j \\
& \Rightarrow A_{ij} = A_i^T A_j
\end{align*}
\]