Chapter 2: Non-Linear Systems

1 Multivariable Calculus

Let $f : \mathbb{R}^n \to \mathbb{R}$, also written $f(x_1, x_2, \ldots, x_n)$. The partial derivative $\frac{\partial f}{\partial x_i}$ at $a = (a_1, a_2, \ldots, a_n)$ is

$$\left. \frac{\partial f}{\partial x_i} \right|_a = \lim_{h \to 0} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{h}.$$  

$f$ is differentiable at $a$ if $\exists b \in \mathbb{R}^n$, such that in a nghd $U_a$ of $a$,

$$f(x) = f(a) + b^T (x - a) + \|x - a\| r(x, a)$$

such that $\lim_{x \to a} r(x, a) = 0$. If all $\frac{\partial f}{\partial x_i}$ are continuous is a nghd $U_a$, we say $f$ is $C^1$. Note $C^1$ implies differentiable with $b_i = \left. \frac{\partial f}{\partial x_i} \right|_a$.

**proof:**

$$f(a+h) - f(a) = \sum_i g_i, \quad g_i = f(a_1, \ldots, a_{i-1}, a_i + h_i, \ldots, a_n + h_n) - f(a_1, \ldots, a_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n)$$

$$g_i = \left. \frac{\partial f}{\partial x_i} \right|_{a + \hat{h}_i} h_i$$

$$f(a + h) - f(a) = \sum \left. \frac{\partial f}{\partial x_i} \right|_a h_i + \|h\| \sum_i \left( g_i - \left. \frac{\partial f}{\partial x_i} \right|_a \right) h_i \underbrace{\|h\|}_{r(h)}.$$  

Since $f$ is $C^1$ we have $\lim_{h \to 0} r(h) = 0$. Hence the proof. If all partial derivatives of order $r$ are continuous is a nghd $U_a$, we say $f$ is $C^r$. If partial derivatives of all order are continuous is a nghd $U_a$, we say $f$ is $C^\infty$.

Let $c(t) = f(a + th)$. Then $dc/dt|_{t=0} = \sum \left. \frac{\partial f}{\partial x_i} \right|_a h_i$. My mean value theorem

$$c(1) - c(0) = f(a + h) - f(a) = \sum \left. \frac{\partial f}{\partial x_i} \right|_{a + \hat{h}_i} h_i, \quad \hat{i} \in [0, 1].$$
Now let \( F : \mathbb{R}^n \to \mathbb{R}^m \), where, \( F = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix} \). We say \( F \) is differentiable when \( f_i \) are. \( F \) is \( C^r \) when \( f_i \) are. For \( C^1 \), \( F \), we have,

\[
F(a + h) - F(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \|h\|r(h)
\]

where \( \lim_{h \to 0} r(h) = 0 \). \( DF(a) \) is called Jacobian of mapping \( F \) at \( a \).

**Lemma 1 Chain Rule:** For \( F, G : \mathbb{R}^n \to \mathbb{R}^m, C^1 \), we have

\[
D(F \circ G(a)) = DF(G(a))DG(a).
\]

**proof:** Let’s see for scalar case when \( f : \mathbb{R} \to \mathbb{R}, C^1 \), we have

\[
\frac{f(g(a + h)) - f(g(a))}{h} = f'(y) \frac{g(a + h) - g(a)}{h},
\]

where \( y \in [g(a), g(a + h)] \). Taking the \( \lim_{h \to 0} \), we obtain \((f \circ g)'(a) = f'(g(a)) \cdot g'(a)\) . Now when \( f : \mathbb{R}^n \to \mathbb{R}, C^1 \), we have

\[
\frac{f(G(a + h_k)) - f(G(a))}{h_k} = \sum_i \frac{\partial f}{\partial x_i}(y) \frac{g_i(a + h) - g_i(a)}{h_k},
\]

where \( y \) on line joining \( G(a), G(a + h_k) \). Taking the \( \lim_{h \to 0} \), we obtain \( \frac{\partial (f \circ G)}{\partial x_k}(a) = \sum_i \frac{\partial f}{\partial x_i}(G(a)) \frac{\partial g_i}{\partial x_k}(a) \). Then \( D(F \circ G(a)) = DF(G(a))DG(a) \).

**Corollary 1** For \( F : \mathbb{R}^n \to \mathbb{R}^n \), and \( x(t) \in \mathbb{R}^n \), we have

\[
\frac{d}{dt} F(x(t)) = DF(x(t)) \dot{x}
\]

### 1.1 Inverse Mapping Theorem

**Theorem 1** For \( F, C^1 \), let \( b = F(a) \). If \( DF(a) \) is invertible then \( F \) maps a nghd \( U_a \) one to one and onto nghd \( V_b = F(U_a) \), such that \( F^{-1} \) on \( V_b \) is \( C^1 \).
Proof: Note determinant is a continuous fn. Since say \( \text{det}(DF(a)) > 0 \), we choose a nghd of \( a, U_a(r_o) \) of radius \( r_o \) such that \( \text{det}(F(x)) > \epsilon \) for \( x \in U_a(r_o) \). For \( x \) and \( y \) in \( U_a(r_o) \), we have \( F(x) - F(y) = DF(z)(x - y) \), where \( z \) lies on line joining \( x \) and \( y \). Since \( \text{det}(DF(z)) > \epsilon \), we have \( F \) injective of \( U_a(r_o) \).

Furthermore choose \( U_a(r_o) \) such that \( \|DF^{-1}(x)\| < \frac{\epsilon}{4} \) and \( r(x,y) < \frac{\epsilon}{2} \) for \( (x,y) \in U_a(r_o) \). For \( y \in U_b(r_o) \), let \( x_1 = DF^{-1}(a)(y_1 - b) \). Then \( \|x_1\| < \frac{r_o}{2} \). Let \( y_2 = F(a + x_1) \), the \( \|y_1 - y_2\| < \frac{r_o}{4} \). Now define \( x_2 = DF^{-1}(x_1)(y_1 - y_2) \), then \( \|x_2\| < \frac{r_o}{4} \). We can continue \( y_k = F(a + \sum_{i=1}^{k-1} x_i) \) and \( \|y_1 - y_k\| < \frac{r_o}{2^k} \) and \( x_k = DF^{-1}(x_1)(y_1 - y_k) \) and \( \|x_k\| < \frac{r_o}{2^k} \). Then \( F(a + \sum x_i) \) converges to \( y_1 \) and \( a + \sum x_i \in U_a(r_o) \). Infact we have shown that \( V_b = F(U_a(r_o)) \) is open. We show \( G = F^{-1} \) is continuous on \( V_b \). At \( b \), we can get to within \( r_o \) of \( a = G(b) \) by choosing \( y_1 \), within \( r_1 \) of \( b \). Similarly at other points. By chain rule \( DG(y) = (DF(G(y)))^{-1} \). Since \( F \) is \( C^1 \) and \( G \) continuous, we have \( G \) as \( C^1 \).

1.2 Implicit Function Theorem

Let \( A \) be a \( m \times n \), \( (m \leq n) \) matrix of rank \( r \), then we ask what are solutions of

\[
Ax = b.
\]

By similarity transformations we can express \( A = P_1BP_2 \), where,

\[
B = \begin{bmatrix}
I_{r \times r} & \times \\
0 & 0
\end{bmatrix}
\]
, then above equation is written as
\[
B \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} c_{r \times 1} \\ 0 \end{bmatrix}.
\]

and \( y = P_2 x \). Then we are free to choose \( y_2 \) and that determines \( y_1 \) uniquely and \( x = P_2^{-1} y \). Then the solution set is parameterized by a \( n - r \) dimensional space \( y_2 \). Every solution \( x \) is in one to one correspondence with \( y_2 \) which are coordinates of \( x \).

There is important nonlinear generalization of this called implicit function theorem. Let \( F : \mathbb{R}^n \to \mathbb{R}^m \), such that \( F \) is \( C^1 \) and \( F(a) = b \). If \( DF \) is of constant rank \( r \) is a nghd \( U_a \), then we can find a nghd \( V_a \subset U_a \) and a map \( G : \mathbb{R}^n \to \mathbb{R}^n \), that maps \( V_a \) one to one onto a nghd \( W \) of origin such that solution set \( S \) of \( F(x) = b \) contained in \( V_a \) is simply
\[
G^{-1}(0, y_2) = S.
\]

for \((0, \ldots, y_{2r}) \in W \). We have a parameterization of \( S \).

**Proof:** Writing
\[
DF(a) = \begin{bmatrix} \frac{\partial F_1}{\partial a_1} & \frac{\partial F_1}{\partial a_2} \\ \vdots & \vdots \\ \frac{\partial F_r}{\partial a_1} & \frac{\partial F_r}{\partial a_2} \end{bmatrix},
\]
where \( a = (a_1, \ldots, a_{n-r}) \), has \( a_1 \) as first \( r \) coordinates and \( a_2 \) as last \( n - r \) coordinates. Similarly the map \( F = (F_1, \ldots, F_r) \), where \( F(a) = (b_1, \ldots, b_{n-r}) \). Since \( DF(a) \) is rank \( r \), W.L.O.G we assume that the top-left, \( r \times r \) block is non-singular. We can find a open nghd around \( a \) on which top-left block is non-singular. Now consider the map \( G(x_1, x_2, \ldots, x_n) = (f_1, \ldots, f_r, x_{r+1}, \ldots, x_n) \). Then
\[
DG(a) = \begin{bmatrix} \frac{\partial f_1}{\partial a_1} & \frac{\partial f_1}{\partial a_2} \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix},
\]
which is full rank and hence by inverse mapping theorem we can find a nghd \( V_a \) such that \( G \) maps one-one onto a nghd \( W \) of \( G(a) \). Now observe \( S \) in \( V_a \) maps to plane \((b_1, \cdot) \) in \( W \), furthermore in \( W \), \( G^{-1}(b_1, \cdot) = S \). In \( W \), given a point on the plane \((b_1, \cdot) \), look at it preimage \( z \) in \( V_a \) and join it by a curve \( C \) to \( a \). On \( V_a \), the last \( n - r \) rows of \( DF(x) \) are dependent on first \( r \) rows. Since the integral of first \( r \) rows along \( C \) is zero so is true for last \( n - r \) rows. Hence \( F(z) = b \) and thus \( z \in S \). Thus we have a parametrization of \( S \), the plane \((b_1, \cdot) \) in \( W \).
2 Manifolds

In implicit function theorem, we saw how the solution set $S$ of $F(x) = b$ has a local parametrization $G^{-1}(b_1, \cdot)$. This is called a manifold $M$. When around every point, we can find a nghd $U$ and a map $\phi$ that maps $U$ one-one, onto a nghd $V$ of orgin such that in $U$, $M = \phi^{-1}(\begin{bmatrix} 0 \\ n-r \\ r \end{bmatrix}, \cdot)$, for $(\begin{bmatrix} 0 \\ n-r \\ r \end{bmatrix}, \cdot) \in V$. We say we have local coordinates for $M$. $r$ is called dimension of $M$.

2.1 Examples

**Sphere:** Let $X = (x, y, z)$ satisfy $F(x, y, z) = x^2 + y^2 + z^2 = 1$, we show $M$ is a manifold. Consider $DF(X) = 2 \begin{bmatrix} x & y & z \end{bmatrix}$. Then DF(X) is rank 1 in a nghd of every $X \in M$ and by implicit function theorem $M$ is a manifold of dimension $3 - 1 = 2$. 
Orthogonal Group $O(n)$: Let $X$ be a $n \times n$ real matrix satisfying $F(X) = X^T X = I$. Then at nonsingular $X$, we can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ matrices. Then $\frac{dF}{dt} = X^T (A^T + A)X$. The null space is $A^T + A = 0$. All skew symmetric matrices of dim $\frac{n(n-1)}{2}$. Then rank is dim $\frac{n(n+1)}{2}$ and by implicit function theorem we have a manifold of dim $\frac{n(n-1)}{2}$. Its called the Orthogonal group.

Special Orthogonal Group $SO(n)$: Let $X \in O(n)$ the $detX = \pm 1$. Then $X$ has two disconnected component $detX = 1$ and $detX = -1$. The component with $detX = 1$ is called $SO(n)$. Its called the special Orthogonal group.

Unitary Group $U(n)$: Let $X$ be a $n \times n$ complex matrix satisfying $F(X) = X'X = I$. Then at nonsingular $X$, we can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ complex matrix. Then

$$\frac{dF}{dt} = X'(A' + A)X.$$ 

The null space is $A' + A = 0$, all skew hermitian matrices of dim $n^2$. Then rank is dim $n^2$ and by implicit function theorem we have a manifold of dim $2n^2 - n^2 = n^2$. Its called the Unitary group.

Special Unitary Group $SU(n)$: Let $X$ be a $n \times n$ complex matrix satisfying $F(X) =$ \[
\begin{bmatrix}
X'X \\
det(X)
\end{bmatrix} = \begin{bmatrix} I \\ 1 \end{bmatrix}. 
\]

Then at nonsingular $X$, we can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ complex matrix. Then

$$\frac{dF}{dt} = \begin{bmatrix} X'(A' + A)X \\ tr(A)det(X) \end{bmatrix}.$$ 

The null space is $A' + A = 0$, all skew hermitian matrices and $tr(A) = 0$ of dim $n^2 - 1$. Then rank is dim $n^2 + 1$ and by implicit function theorem we have a manifold of dim $2n^2 - (n^2 + 1) = n^2 - 1$. Its called the special Unitary group.

Special Linear group $SL(n, \mathbb{R})$: Let $X$ be a $n \times n$ real matrix satisfying $detX = 1$. Then at nonsingular $X$, we can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ complex
matrix. Then
\[ \frac{dF}{dt} = tr(A)X. \]
The null space is $tr(A) = 0$, all traceless matrices of dim $n^2 - 1$. Then rank is dim 1 and by implicit function theorem we have a manifold of dim $n^2 - 1$.

**Symplectic Group $Sp(n, \mathbb{R})$:** Let $X$ be a $2n \times 2n$ real matrix satisfying $X'JX = J$, where $J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$. Then at nonsingular $X$, we can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ matrices. Then $\frac{dF}{dt} = X^T(A^TJ + JA)X$. The null space is $A^TJ + JA = 0$. If $A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$, then $A_1 = -A_4^T$ and $A_2 = A_3^T$ and $A_3 = A_2^T$. The dim of null space is $n^2 + n(n + 1) = 2n^2 + n$. Then rank is dim $2n^2 - n$ and by implicit function theorem we have a manifold of dim $2n^2 + n$.

### 2.2 Tangent Space of $M$

Given a point $p \in M$, we have a nhgd $U_p$ mapped by $\phi$ to $\phi : U \to V_0$ such that $M$ in $U_a$ is mapped to plane $\{(0, y, \ldots, y_r)\}$ in $V_0$, which we denotes as $(0, y)$. Consider the curve $(0, y(t))$, which is mapped to $x(t) = \phi^{-1}(0, y(t))$, passing through $p$. Then $\dot{x} = D\phi^{-1}(0, \dot{y})$, also denoted as $\dot{\phi}^{-1}(0, \dot{y})$. Observe $\dot{y}$ lies in $r$ dimensional vector space so $\dot{x} = D\phi^{-1}(0, \dot{y})$ lies in a $r$ dimensional subspace called the tangent space of $p$ denotes by $T_pM$. Let's compute the tangent space of various manifolds.

**Sphere:** Let $M$ be $(x, y, z)$ satisfying $x^2 + y^2 + z^2 = 1$. Given a $p$ on $M$ a curve through $p$ satisfies $x^2(t) + y^2(t) + z^2(t) = 1$, then we have $x\dot{x} + y\dot{y} + z\dot{z} = 1$. Then $(\dot{x}, \dot{y}, \dot{z})$ is orthogonal to $p$, a two dimensional space.

**Orthogonal Group $O(n)$:** $X \in M$ is a $n \times n$ real matrix satisfying $X^TX = I$. Given a $p$ on $M$ a curve through $p$ satisfies $X^T(t)X(t) = I$, with $X(0) = p$. We can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ matrix. Then $\dot{X}^T(t)X(t) + X^T\dot{X}(t) = 0$, i.e., $X^T(A^T + A)X = 0$ implying $A^T + A = 0$. A is skew symmetric matrix. The tangent space at $p = X(0)$ is of the form $AX(0)$ where $A$ is skew symmetric matrix. Dim of $T_pM$ is $\frac{n(n-1)}{2}$.

**Unitary Group $U(n)$:** $X \in M$ is a $n \times n$ complex matrix satisfying $X'X = I$. Given a $p$ on $M$ a curve through $p$ satisfies $X'(t)X(t) = I$, with $X(0) = p$. We can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ matrix. Then $X'(t)X(t) + X'\dot{X}(t) = 0$, i.e., $X'(A' + A)X = 0$.
implying $A' + A = 0$. $A$ is skew hermitian matrix. The tangent space at $p = X(0)$ is of the form $AX(0)$ where $A$ is skew hermitian matrix. Dim of $T_p\mathcal{M}$ is $n^2$. 

**Special Unitary Group SU(n):** $X \in \mathcal{M}$ is a $n \times n$ complex matrix satisfying $X'X = I$ and $\det X = 1$. Given a $p$ on $\mathcal{M}$ a curve through $p$ satisfies $X'(t)X(t) = I$, $\det X(t) = 1$ with $X(0) = p$. We can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ matrix. Then $\dot{X}'(t)X(t) + X'\dot{X}(t) = 0$, and $tr(A)\det X = 0$ i.e., $X'(A' + A)X = 0$ implying $A' + A = 0$ and $tr(A) = 0$. $A$ is traceless skew hermitian matrix. The tangent space at $p = X(0)$ is of the form $AX(0)$ where $A$ is traceless skew hermitian matrix. Dim of $T_p\mathcal{M}$ is $n^2 - 1$.

**Special Linear group SL(n, R):** $X \in \mathcal{M}$ is a $n \times n$ real matrix satisfying $\det X = 1$. Given a $p$ on $\mathcal{M}$ a curve through $p$ satisfies $\det X(t) = 1$ with $X(0) = p$. We can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ real matrix. Then $\frac{d}{dt}\det X = tr(A)\det X = 0$ i.e., $tr(A) = 0$. The tangent space at $p = X(0)$ is of the form $AX(0)$ where $A$ is traceless real matrix. Dim of $T_p\mathcal{M}$ is $n^2 - 1$.

**Symplectic Group Sp(n, R):** $X \in \mathcal{M}$ is a $2n \times 2n$ real matrix satisfying $X'JX = J$, where $J = \begin{bmatrix} O & -I_n \\ I_n & 0 \end{bmatrix}$. Given a $p$ on $\mathcal{M}$ a curve through $p$ satisfies $X'(t)JX(t) = J$, with $X(0) = p$. We can write a velocity vector $\dot{X}(t) = AX(t)$ for $A \in n \times n$ matrices. Then $X^T(t)(A^TJ + JA)X(t) = 0$, i.e., $A^TJ + JA = 0$. If $A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$, then $A_1 = -A_4^T$ and $A_2 = A_2^T$ and $A_3 = A_3^T$. The dim of these matrices is $n^2 + n(n + 1) = 2n^2 + n$. 

8
3 Lie Groups

The manifolds like $O(n)$, $SO(n)$, $Sl(n, R)$, $U(n)$, $SU(n)$, $Sp(n, R)$ besides being manifolds are also groups under matrix multiplication and are closed under multiplication and inverse. In particular $M$ has identity element $I$. They are called Lie Groups and denoted by $G$. As an example, let $G = Sp(n, R)$. If $X, Y \in G$, then $(XY)^T JXY = J$ and $(X^{-1})^T JX^{-1} = J$.

**Lie Algebra:** If we look at tangent space at $X$, it takes the form $AX$, where $A$ belongs to a vector space, lets call $g$. Then tangent space at $I$, is simply $A$. It is no coincidence that for the group $G$, the tangent space at Identity and at any arbitrary element $X$ are related by left multiplication by $X$. This is because if $p(t)$ is curve passing through $I$ ($p(0) = I$) then $Y(t) = p(t)X$ is curve passing through $X$, then if $\dot{p}(0) = A$ lies in a vector space, then $\dot{Y}(0) = AX$. Then $g$ is called the Lie algebra of $G$.

**Vector Field:** Now let $A \in g$ and consider the differential eq. $\dot{X} = AX$, with $X(0) = I$. Then $X(t) \in G$. Consider a nghd $U$ of $I$ in $M$ and $U_1$ is its intersection with $M$. Let $V = \phi(U)$ and $V_1 = \phi(U_1)$ is the horizontal plane, the coordinates of $M$. $U_1$ is called nghd of $I \in M$ and $V_1$ is called its coordinate chart. Observe, $X \in U_1$, the vector $AX$ assigns a vector at each $X$ and is called a vector field. Then $\phi_*(AX) = f(y)$ is a vector field on $V_1$, which is horizontal. Now consider the evolution $\dot{y} = f(y)$, as $f(y)$ is horizontal, $y(t) \in V_1$ for $t \in [-\delta, \delta]$ and it preimage satisfies $x(t) \in U_1$ and $\dot{x} = AX$. Thus we can say that for
\( t \in [-\delta, \delta], \exp(At) \in G \) and hence \( \exp(At) \in G \) for all \( t \).

**Exponential Coordinates:** At \( I \in G \), we have a nghd \( U_1 \) mapped to \( V_1 \), the cordinates of \( U_1 \). There is a natural choice of cordinates in \( G \) called exponential cordinates. Let \( A_i \) be a basis of \( r \) dimesional \( g \). Consider the map

\[
\psi(t_1, \ldots, t_r) = \exp(\sum_i t_i A_i).
\]

Then \( \phi \circ \psi(t_1, \ldots, t_r) \to V_1 \) such that \( \phi \circ \psi(0) = 0 \). Note \( \frac{\partial \psi}{\partial t_i}|_{t=0} = A_i \).

Then observe

\[
(\phi \circ \psi)_*(0) = \phi_* \psi_*(0) = \phi_* \begin{bmatrix} A_1 & \ldots & A_r \end{bmatrix},
\]

is full rank. Thus \( \phi \circ \psi \) maps onto a nghd or origin in \( V_1 \) and hence \( \psi(t_1, \ldots, t_r) \) maps onto a nghd \( U_e \) of \( I \in G \). We call \( U_e \) exponential nghd. For \( g \in G \), \( U_e \ g \) is a nghd around \( g \).

### 4 Non-Linear Controllability

We now start talking about nonlinear control systems. The most common ones are of the form

\[
\dot{x} = \sum_{i=1}^{m} u_i(t) g_i(x), \quad x \in \mathbb{R}^n, \ g_i: \mathbb{R}^n \to \mathbb{R}^n,
\]

where we assume \( g_i \) are smooth functions and are called vector fields. We ask can we steer this system between points of interest by choice of \( u_i(t) \). We can write the above system as

\[
\dot{x} = \sum_i u_i(t) g_i(x) = \begin{bmatrix} g_1(x) & g_2(x) & \ldots & g_m(x) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}
\]

\( G(x) \) is a collection of vector fields. If they span a \( r \) dimensional space at each point we call it a rank \( r \) distribution. If \( r = n \), we have a controllable system. We can just follow the velocity of a path. Interesting case is when \( r < n \). As an example take the following system called *nonholonomic integrator* which models kinematics of a mobile robot.
\[
\begin{bmatrix}
\dot{x} \\
\dot{y} \\
\dot{z}
\end{bmatrix} = \begin{bmatrix}
u \\
v \\
xv - yu
\end{bmatrix} = \begin{bmatrix}1 \\
0 \\
-y
\end{bmatrix} u + \begin{bmatrix}0 \\
1 \\
x
\end{bmatrix} v.
\]

Now we are in \( \mathbb{R}^3 \) and have two dimensional distribution spanned by \( f, g \). At each point we can move in two directions \( f \) and \( g \). Is the system controllable? If we switch on \( u = +1 \) we move along \( f \) and if we switch on \( u = -1 \) we move along \(-f \). Similarly if we switch on \( u = +1 \) we move along \( g \) and if we switch on \( v = -1 \) we move along \(-g \). Let's go along \( f \), then \( g \) and \(-f\) and \(-g \). Let's evaluate what happens then. We use the notation \( \phi^\Delta_t(x_0) \) to denote the final point as the initial point \( x_0 \) moves along \( f \) for time \( \Delta t \). Then we want to find out what is \( \phi^\Delta_{-g} \circ \phi^\Delta_f \circ \phi^\Delta_g \circ \phi^\Delta_f(x_0) \)

Observe

\[
x_1 = x(\Delta t) = x_0 + f(x_0)\Delta t + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2. \tag{2}
\]
\[
x_2 = x(2\Delta t) = x_1 + g(x_1)\Delta t + \frac{1}{2} \frac{\partial g}{\partial x} g(x_1) \Delta t^2. \tag{3}
\]
\[
x_3 = x(3\Delta t) = x_2 - f(x_2)\Delta t + \frac{1}{2} \frac{\partial f}{\partial x} f(x_2) \Delta t^2. \tag{4}
\]
\[
x_4 = x(4\Delta t) = x_3 - g(x_3)\Delta t + \frac{1}{2} \frac{\partial g}{\partial x} g(x_3) \Delta t^2. \tag{5}
\]
\[
x_2 = x_0 + f(x_0)\Delta t + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2 + g(x_0)\Delta t + \frac{\partial g}{\partial x} f(x_0) \Delta t^2 + \frac{1}{2} \frac{\partial g}{\partial x} g(x_0) \Delta t^2 - f(x_0) \Delta t - \frac{\partial f}{\partial x} (f(x_0) + g(x_0)) \Delta t^2 + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2 + o(\Delta t^3). \tag{6}
\]
\[
x_3 = x_0 + f(x_0)\Delta t + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2 + g(x_0)\Delta t + \frac{\partial g}{\partial x} f(x_0) \Delta t^2 + \frac{1}{2} \frac{\partial g}{\partial x} g(x_0) \Delta t^2 - \frac{\partial f}{\partial x} (f(x_0) + g(x_0)) \Delta t^2 + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2 + o(\Delta t^3). \tag{7}
\]
\[
x_4 = x_0 + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2 + \frac{\partial g}{\partial x} f(x_0) \Delta t^2 + \frac{1}{2} \frac{\partial g}{\partial x} g(x_0) \Delta t^2 - \frac{\partial f}{\partial x} (f(x_0) + g(x_0)) \Delta t^2 + \frac{1}{2} \frac{\partial f}{\partial x} f(x_0) \Delta t^2 + \frac{1}{2} \frac{\partial g}{\partial x} g(x_0) \Delta t^2 + o(\Delta t^3). \tag{8}
\]
\[
x_4 = x_0 + (\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x} g(x_0)) \Delta t^2 + o(\Delta t^3). \tag{9}
\]
\[ [f,g](x) = \frac{\partial g}{\partial x}f - \frac{\partial f}{\partial x}g. \]

When we make the maneuver we proposed. We donot return back to \( x_0 \) instead we make a leading order motion in direction given by Lie bracket of \( f \) and \( g \) denoted as \([f,g]\).

For the vector fields given
\[
\frac{\partial g}{\partial x}f - \frac{\partial f}{\partial x}g = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -y \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = h.
\]

We can not only go in direction \( f \) and \( g \), we can also go in direction \( h \) by making the maneuver. Then we have three independent directions of motion, which suggest we have controllability because I can move everyway.

### 4.1 Frobenius Theorem

We introduce the notation \( \exp(tf)x_0 \), evolves \( x_0 \) under \( f \) for time \( t \) or evolves \( x_0 \) under \( tf \) for time 1. In this notaion \( \exp(tf)x_0 = \phi_f^t(x_0) \).

If \( f \) and \( g \) commute, then
\[
\exp(tf) \exp(sg)x_0 = \exp(sg) \exp(tf)x_0 = \exp(tf + sg)x_0. \tag{10}
\]

To see this if \( f \) and \( g \) commute, then flow of \( f \) preserves \( g \), i.e.
\[
g(\phi_f^t(x_0)) = (\phi_f^t)_*g(x_0). \tag{11}
\]

Note
\[
\frac{dg(x(t))}{dt} = \frac{\partial g}{\partial x}f. \tag{12}
\]
as \( f \) and \( g \) commute, we have
\[
\frac{dg(x(t))}{dt} = \frac{\partial f}{\partial x}g(x(t)), \tag{13}
\]
but this is the equation of \( (\phi_f^t)_*g(x_0) \). Now consider the curve
\[
x(t) = \phi_g^{-s} \phi_f^t \phi_g^s x_0 \tag{14}
\]
Then,
\[
\dot{x}(t) = (\phi^{-s}_g)_* f(\phi^s_g x_0) = f(x(t)) \quad (15)
\]

\[
x(t) = \phi^t_f = \phi^{-s}_g \phi^t_f \phi^s_g x_0. \quad (16)
\]

hence the proof.

Similarly, consider the curves

\[
x_1(t) = \exp(tf) \exp(tg) x_0, \quad x_2(t) = \exp(t(f + g)) x_0.
\]

and

\[
\dot{x}_2(t) = (f + g)(x_2(t)).
\]

From above discussion on preserving the vector fields, we have

\[
\exp(tf)_* g(\exp(tg) x_0) = g(\exp(tf) \exp(tg) x_0) = g(x_1(t)).
\]

Therefore \(\dot{x}_1(t) = (f + g)(x_1(t))\). By uniqueness, we have,

\[
\exp(tf) \exp(tg) x_0 = \exp(t(f + g)) x_0.
\]

Now consider a \(r\) dimensional distribution \(\Delta = \{f_1, \ldots, f_r\}\) such that \(r < n\) and \(f_i\) commute. Then we claim we cannot go everywhere. Our motion is restricted to a \(r\) dimensional manifold.

Let \(\{f_1, \ldots, f_r, e_{r+1}, \ldots, e_n\}\) span \(\mathbb{R}^n\) at \(x_0\). Consider the map

\[
\Phi(t_1, \ldots, t_n) = \prod_{i=1}^r \exp(t_i f_i) \prod_{i=r+1}^n \exp(t_i e_i) x_0.
\]

where the map \(\exp(tf) x_0\), evolves \(x_0\) under \(f\) for time \(t\). Then \(\frac{d}{dt} \exp(tf) x_0|_0 = f(x_0)\). Then

\[
\frac{\partial \Phi}{\partial t_i}(0) = f_i(x_0), \quad i = 1, \ldots, r
\]

\[
\frac{\partial \Phi}{\partial e_i}(0) = e_i(x_0), \quad i = r + 1, \ldots, n
\]

By inverse mapping theorem, we have a one-one onto map that maps nghd \(V\) to nghd \(U\) of \(x_0\) such that the plane \((t, \ldots, t_r)\) to \(\mathcal{M}\).
Infact,

\[
\frac{\partial \Phi}{\partial t_i}(t_1, \ldots, t_n) = f_i(x), \ i = 1, \ldots, r
\]  \hspace{1cm} (17)

Thus in \( U \), \( \Phi^{-1}_x(f_i) \) are horizontal vector fields in \( V \). Hence if we consider the equation 
\( \dot{x} = \sum_i u_i f_i \), we are always evolving horizontally in \( V \). Then starting from 0 in \( V \), we stay 
on the horizontal plane and hence \( x(t) \) evolves on \( \mathcal{M} \), which is called the integral manifold of 
vector fields \( \{f_i\} \). Hence our motion is restricted to a \( r \) dimensional manifold \( \mathcal{M} \).

![Figure 4](image)

Now lets relax the constraint that \( \{f_i\} \) commute but instead \( [f_i, f_j] \in \Delta \), i.e., at every \( x \),
\( [f_i, f_j](x) = \sum_k \alpha_k(x) f_k \). Then we say our distribution \( \Delta \) is involutive. Given an involutive 
distribution \( \Delta \), there is a \( r \) dimensional manifold \( \mathcal{M} \) such that \( \Delta \) is tangent to \( \mathcal{M} \) at each 
\( x \in \mathcal{M} \). \( \mathcal{M} \) is called the integral manifold of \( \Delta \).

Written as a \( n \times r \) matrix

\[
\begin{bmatrix} f_1 & \cdots & f_r \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix}
\]

These \( r \) columns are independent. At \( x_0 \), we assume, w.l.o.g the top \( r \) rows are independent. 
Then the top \( r \times r \) matrix is invertible at \( x_0 \) and so in a nghd \( U \) of \( x_0 \). Then define

\[
G = \begin{bmatrix} g_1 & \cdots & g_r \end{bmatrix} = F \ast A^{-1} = \begin{bmatrix} I \\ C \end{bmatrix}
\]  \hspace{1cm} (19)

Now observe given a function \( h(x) \), we have

\[
[h f, g] = h \left( \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g \right) - \left( \frac{\partial h}{\partial x} g \right) f = h[g, f] - h_1 f
\]
Therefore if $\Delta = \{ f_i \}$ is involutive, we have for arbitrary functions $h_i$, $[h_i f_i, h_j f_j] \in \Delta$. Therefore, Eq. (19), $\{g_i\}$ is involutive and is same as $\Delta$. Now take bracket of $[g_i, g_j]$, it is of the form $[g_i, g_j] = \begin{bmatrix} 0 & c(x) \\ c(x) & 0 \end{bmatrix}$. But being involutive, it should in span of $g_i$, implying $c(x) = 0$ and $g_i$ commute.

Therefore, we are back to the situation of commuting vector fields.

$$\mathcal{M} = \prod_{i=1}^{r} \exp(t_i g_i) x_0,$$

is a integral manifold passing through $x_0$ such that our motion under the flow $\dot{x} = \sum_i u_i f_i$ is restricted to $\mathcal{M}$. This is called Frobenius Theorem.

This means in our control system

$$\dot{x} = \sum_{i=1}^{r} u_i f_i(x),$$

when vector fields $f_i$ donot span $\mathbb{R}^n$ and on bracketting donot generate a new vector field, then our motion is restricted to a $r$ dimensional integral manifold $\mathcal{M}$.

### 4.2 Chows Theorem

Recall from Eq. (9), we calculated the map

$$\Phi_{[f,g]}^{\Delta t} x_0 = \phi_g^{\Delta t} \circ \phi_f^{\Delta t} \circ \phi_g^{\Delta t} \circ \phi_f^{\Delta t}(x_0).$$

We found to leading order we get

$$\Phi_{[f,g]}^{\Delta t} x_0 = x_0 + [f, g](x_0) \Delta t^2 + o(\Delta t^3).$$

To leading order we proceed in direction $[f, g]$. We say, we generate the first brackett.

Let see, how to generate second brackett, say $[h[f, g]]$. For this consider the map

$$\phi_{h}^{\Delta s} \phi_{[f,g]}^{\Delta t} \phi_{h}^{\Delta s} x_0,$$

Let $\phi_{h}^{\Delta s} x_0 = x_1$. Then we have

$$x_2 = \Phi_{[f,g]}^{\Delta t} x_1 = x_1 + [f, g](x_1) \Delta t^2 + o(\Delta t^3).$$

$$\phi_{-h}^{\Delta s}(x_1 + \epsilon) = \phi_{-h}^{\Delta s}(x_1) + \epsilon - \frac{\partial h}{\partial x} \epsilon \Delta s + o(\Delta t^2 \Delta s^2).$$

using
\[ x_1 = x_0 + h(x_0) \Delta s + o(\Delta s^2) \]

\[ x_2 = x_0 + [f, g](x_0) \Delta t^2 + \left( \frac{\partial [f, g]}{\partial x} h - \frac{\partial h}{\partial x} [f, g] \right)(x_0) \Delta t^2 \Delta s + o(\Delta t^3) + o(\Delta t^2 \Delta s^2) \]

Now evaluate

\[ \Phi_{[g,f]}^{\Delta t} x_2 = x_2 + [g, f](x_2) \Delta t^2 + o(\Delta t^2) \]

\[ = x_0 + [h[f, g]](x_0) \Delta t^2 \Delta s + o(\Delta t^3) + o(\Delta t^2 \Delta s^2) \]

Choosing \( \Delta s = \sqrt{\Delta t} \).

\[ \Psi_{[h, [g, f]]}^{\Delta t} = \Phi_{[g,f]}^{\Delta t} \phi_{h}^{\Delta s} \Phi_{[f, g]}^{\Delta t} \phi_{h}^{\Delta s} x_0 = x_0 + [h[f, g]](x_0) \Delta t^\frac{3}{2} + o(\Delta t^3) \]

The leading order term is \( [h[f, g]](x_0) \) a second order bracket. If we want one more bracket \( [e[h[f, g]]] \), then we do

\[ \phi_{-e}^{\Delta s} \phi_{[h, [g, f]]}^{\Delta t} \phi_{h}^{\Delta s} = x_0 + [h[f, g]](x_0) \Delta t^{5/2} + [e[h[f, g]]](x_0) \Delta t^{5/2} \Delta s + o(\Delta t^3) + o(\Delta t^{5/2} \Delta s^2) \]

Now choose \( \Delta s = \Delta t^{\frac{1}{4}} \), then

\[ \Sigma_{[e[h, [g, f]]]}^{\Delta t} = \Psi_{[-h, [g, f]]}^{\Delta t} \phi_{-e}^{\Delta s} \Psi_{[h, [g, f]]}^{\Delta t} \phi_{h}^{\Delta s} = x_0 + [e[h[f, g]]](x_0) \Delta t^{11/4} + o(\Delta t^3) \]

The leading order term is \( [e[h[f, g]]](x_0) \) a second order bracket. In fact if we choose \( t = (\Delta t)^{\frac{4}{11}} \), we have

\[ \Sigma_{[e[h, [g, f]]]}^{t} x_0 = x_0 + [e[h[f, g]]](x_0) t + o(t^{1+\alpha}), \quad \alpha > 0. \]

Now let’s say we start with \( r \) independent vector fields \( \{f_i\} \). By taking brackets we can generate new vector fields \( X_k \) such that \( \{f_i, X_k\} \), span all of \( \mathbb{R}^n \). As above, we have shown how to construct a map

\[ \Phi_{X}^{t} x_0 = x_0 + X(x_0) t + o(t^{1+\alpha}), \quad \alpha > 0. \]

Now consider map

\[ F(t_1, \ldots, t_r, \ldots, t_n) = \prod_{i=r+1}^{n} \Phi_{X_i}^{t_i} \prod_{i=1}^{r} \exp(f_i t_i) x_0 \]
Then by construction $DF(0) = \{f_i, X_k\}$, a full rank matrix. Hence by inverse function theorem $F$ maps a nghd $U = (t_1, \ldots, t_n)$ of origin to a nghd $V$ of $x_0$. We can go anywhere in $R^n$, we have controllability. Suppose we want to go from $x$ to $y$, then choose a path as shown in the figure 6 below and go from $x$ to $y$ by overlapping ngds. Now we can go withing ngds and go from $x$ to $y$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Figure shows how a path from $x$ to $y$ can be tracked through a sequence of overlapping nghds}
\end{figure}

**Example 1** Let us go back to the non-holonomic integrator

\[
\begin{bmatrix}
\dot{x} \\
\dot{y} \\
\dot{z}
\end{bmatrix} = \begin{bmatrix}
1 \\
0 \\
-y
\end{bmatrix} u + \begin{bmatrix}
0 \\
x \\
g
\end{bmatrix} v.
\]

We also write

\[
f = \frac{\partial}{\partial x} - y \frac{\partial}{\partial z}; \quad g = \frac{\partial}{\partial y} + x \frac{\partial}{\partial z}
\]

Then $h = [f, g] = \begin{bmatrix}
0 \\
0 \\
2
\end{bmatrix}$. Then $f, g, h$ space $R^3$ and we have controllability.

**Example 2** Consider the system

\[
\begin{bmatrix}
\dot{x} \\
\dot{y} \\
\dot{z} \\
\dot{m} \\
\dot{n}
\end{bmatrix} = \begin{bmatrix}
1 \\
0 \\
-y \\
0 \\
y^2
\end{bmatrix} u + \begin{bmatrix}
0 \\
x \\
x^2 \\
x^2 \\
0
\end{bmatrix} v.
\]
Then

\[
[f,g] = \begin{bmatrix}
0 \\
0 \\
2 \\
x \\
y
\end{bmatrix}; \quad [f,[f,g]] = \begin{bmatrix}
0 \\
0 \\
0 \\
1 \\
0
\end{bmatrix}; \quad [[f,g]g] = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
1
\end{bmatrix}.
\]

Then \(f, g, [f, g], [f, [f, g]], [[f, g], g]\) span the space \(\mathbb{R}^5\) and we have controllability.

### 4.3 Non-linear controllability on Lie groups

We now consider control systems of the form

\[
\dot{x} = \left( \sum_i u_i \Omega_i \right) x, \quad x(0) = I
\]

Where \(\Omega_i \in g\) the Lie algebra of a Lie Group \(G\). Then as shown before, \(x(t) \in G\). Suppose span \(\{\Omega_i\} = g\). Then starting from say \(x(0) = I\), we can go to any point in its exponential nghd \(U_e\) by just using the evolution \(\exp(\sum_i u_i \Omega_i)\) for constant \(u_i\). If \(G\) is connected any two points \(X\) and \(Y\) can be joined by a path in \(G\). Around each point draw a exponential nghd and choose a finite cover so that we go from \(X\) to \(Y\) in overlapping nghds. Then we can go from \(X\) to \(Y\).

\[\text{Figure 6: Figure shows how a path from } x \text{ to } y \text{ can be tracked through a sequence of overlapping nghds}\]

Suppose the given \(\Omega_i\) donot span \(g\). Then we cannot go to all points in \(U_e\). Again we propose the maneuver we use before. We go in direction of \(\Omega_1\), then \(\Omega_2\) and then \(-\Omega_1\) and \(-\Omega_2\), for small time \(\Delta t\).

This generates the evolution

18
\[ U_{[\Omega_1, \Omega_2]}(\Delta t) = \exp(-\Omega_2 \Delta t) \exp(-\Omega_1 \Delta t) \exp(\Omega_2 \Delta t) \exp(\Omega_1 \Delta t) \]

\[ = \exp(-\Omega_2 \Delta t) \exp(-\Omega_1 \Delta t) (I + \Omega_2 \Delta t + \Omega_2^2 \frac{\Delta t^2}{2} + o(\Delta t^3)) \exp(\Omega_1 \Delta t) \]

\[ = \exp(-\Omega_2 \Delta t) (I + \Omega_2 \Delta t - [\Omega_1, \Omega_2] \Delta t^2 + \Omega_2^2 \frac{\Delta t^2}{2} + o(\Delta t^3)) \]

\[ = (I - \Omega_2 \Delta t + \Omega_2^2 \frac{\Delta t^2}{2} + o(\Delta t^3))(I + \Omega_2 \Delta t - [\Omega_1, \Omega_2] \Delta t^2 + \Omega_2^2 \frac{\Delta t^2}{2} + o(\Delta t^3)) \]

\[ = I - [\Omega_1, \Omega_2] \Delta t^2 + o(\Delta t^3) \]

To leading order we generate a motion in the matrix commutator \([\Omega_2, \Omega_1]\). First note that \([\Omega_2, \Omega_1] \in \mathfrak{g}\) because for \(\Delta t = \sqrt{t}\), we have \(U(\sqrt{t}) = I - [\Omega_1, \Omega_2]t + o(t^{3/2})\), is a path in \(G\) and its derivative at \(t = 0\) is an element of \(\mathfrak{g}\) which is \([\Omega_2, \Omega_1] \in \mathfrak{g}\).

As before by making a maneuver, we have been able to generate a bracket. We can now generate more brackets

\[ \exp(\Delta s \Omega_3) U_{[\Omega_1, \Omega_2]}(\Delta t) \exp(-\Delta s \Omega_3) = I - [\Omega_1, \Omega_2] \Delta t^2 + o(\Delta t^3) - [\Omega_3, [\Omega_1, \Omega_2]] \Delta s \Delta t^2 + o(\Delta s^2 \Delta t^2) \]

Choose as before \(\Delta s = \sqrt{\Delta t}\). Then,

\[ U_{[\Omega_2, \Omega_1]} \exp(\Delta s \Omega_3) U_{[\Omega_1, \Omega_2]}(\Delta t) \exp(-\Delta s \Omega_3) = I - [\Omega_3, [\Omega_1, \Omega_2]] \Delta t^\frac{5}{2} + o(\Delta t^3). \]

Thus we have map

\[ \Phi_{[\Omega_2, \Omega_1]}(t) = 1 - [\Omega_3, [\Omega_1, \Omega_2]]t + O(t^{1+\alpha}) \]

Thus we start with set of generators \(\{\Omega_1, \ldots, \Omega_r\}\). They need not span \(\mathfrak{g}\). Then we have shown how to generate brackets of \(\Omega_i\). Suppose we generate new generators call \(X_k\) such that \(\Omega_i, X_k\) span \(\mathfrak{g}\). Then consider the map

\[ F(t_1, \ldots, t_r, \ldots, t_n) = \prod_{i=r+1}^{n} \Phi_{X_i}^{t_i} \prod_{i=1}^{r} \exp(\Omega_i t_i). \]

Then by construction \(DF(0) = \{\Omega_i, X_k\} = \mathfrak{g}\). Hence by inverse function theorem, \(F\) maps onto an exponential nghd \(U_e\) of \(I\). We can go anywhere in \(U_e\) using the map \(F\) and then we have controllability as described before.

In nutshell, if commutators of the generators span \(\mathfrak{g}\), we have controllability.
Example 3 Let
\[ \dot{\Theta} = (u\Omega_x + v\Omega_y)\Theta, \quad \Theta(0) = I, \]
where \( \Omega_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix} \) and \( \Omega_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \). Then our system evolves on \( SO(3) \), which has tangent space \( g = so(3) \), \( 3 \times 3 \) skew symmetric matrices. By Lie bracket, we get \( [\Omega_x, \Omega_y] = \Omega_z \). Now we have all three generators \( (\Omega_x, \Omega_y, \Omega_z) \) of \( so(3) \), we span \( g \). We can steer the system anywhere on \( SO(3) \).

Example 4 Let
\[ \dot{\Theta} = \begin{bmatrix} 0 \\ -u^T \\ \mathbf{O} \end{bmatrix} \Theta, \quad \Theta(0) = I, \]
where \( u \in \mathbb{R}^{n-1} \) is our control vector and \( \mathbf{O} \) is \( n-1 \times n-1 \) matrix. First note our system evolves on \( SO(n) \). Let \( \Omega_{ij} \) be skew symmetric with 1 in the \( i,j \) spot, with \( i < j \). Then we have \( \frac{n(n-1)}{2} \) such generators that span \( g = so(n) \), space of skew symmetric matrices. We only have as control, generators of the form \( \Omega_{1k}, k > 1 \). There are \( n-1 \) of them. But see \( [\Omega_{1j}, \Omega_{1k}] = \Omega_{jk} \). Thus we get all the generators by commutators and we have controllability.

Example 5 Let
\[ \dot{\Theta} = \begin{bmatrix} 0 & u_1 & 0 & 0 & 0 \\ -u_1 & 0 & u_2 & 0 & 0 \\ 0 & -u_2 & 0 & u_3 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & -u_{n-1} \end{bmatrix} \Theta, \quad \Theta(0) = I, \]
Again, our system evolves on \( SO(n) \). Now we have \( n-1 \) control generators \( \{\Omega_{12}, \Omega_{23}, \ldots, \Omega_{n-1,n}\} \). Observe \( [\Omega_{12}, \Omega_{23}] = \Omega_{13}, [\Omega_{13}, \Omega_{34}] = \Omega_{14}, \) etc. We can this way generate \( \Omega_{1k} \), and from previous example all of \( so(n) \). Hence controllability.

We now return to control system
\[ \dot{x} = (\sum_i u_i \Omega_i)x, \quad x(0) = I \] (20)
Where \( \Omega_i \in g \) the Lie algebra of a Lie Group \( G \). \( x(t) \in G \). Suppose span \( \{\Omega_i\} \neq g \). Furthermore the Lie algebra generated of \( \Omega_i \) denotes as \( \mathfrak{h} = \{\Omega_i\}_{LA} \) is a proper subalgebra of \( g \). What can we say about controllability now. Lets say \( A_i \) span \( \mathfrak{h} \) and let remaining \( B_i \) span all of matrices. Consider a nghd \( U \),
\[ \Phi(t_1, \ldots, t_r, \ldots, t_n) = \prod_{i=1}^{l} \exp(A_i t_i) \prod_{i=l+1}^{n} \exp(B_i t_i). \]

Then once again \( \Phi^{-1}_z((\sum_i u_i \Omega_i)x)) \) on \( U \) is a horizontal vector field and hence the flow in Eq. (20), stays on \( \prod_{i=1}^{l} \exp(A'_i t_i) \). This a subset of \( \prod_{i=1}^{r} \exp(A'_i t_i) \). where \( A'_i \) span \( g \). Hence our flow is restricted.

### 4.4 Control Systems with Drift

We now consider control systems of the form

\[ \dot{x} = (A + \sum_i u_i B_i)x, \quad x(0) = I \]

Where \( A, B_i \in g \) the Lie algebra of a Lie Group \( G \). But now we have no control on \( A \), its called drift as its drift of its own. We now controllability of such systems. Note before, \( x(t) \in G \).

First result appears when \( A \) is periodic, i.e. this is a \( T \) such that \( \exp(A T) = I \), the flow of \( A \) returns you back. Then we have a similar results on controllability. If \( \{A, B_i\}_{LA} = g \), we have controllability on connected Lie group \( G \). We have to show that we can generate Lie brackets.

The main observation is that \( \exp(A(T - \tau)) = \exp(-A \tau) \), we can go backwards in direction of \( A \). Furthermore

\[ \exp(A(T - \Delta)) \exp((A + B)\Delta) = \exp(-A \Delta) \exp((A + B)\Delta) = I + B \Delta + O(\Delta^2) \sim \exp(B \Delta). \]

Then as before we can generate brackets of \( \{A, B_i\} \) and we have controllability when \( \{A, B_i\}_{LA} = g \).

Suppose \( A \) is not periodic. Then we focus on Lie groups \( G \), which are compact (bounded), like \( SO(n), SU(n) \) etc. Then given \( \epsilon > 0 \), there exists a time \( T \) such that \( \mid \exp(\epsilon T) - I \mid < \epsilon \), i.e., if we wait long enough, we almost come back to origin. Then \( \exp(A(T - \tau)) \sim \exp(-A \tau) \) and we can go backwards in direction of \( A \). This result is based on Kronecker’s theorem, which states that

**Theorem 2 Kronecker:** Given real number \( \alpha_i \), independent, i.e, \( \sum_i n_i \alpha_i \not= Z \) for integers \( n_i \), not all zero. For any \( \epsilon > 0 \), there exist integers \( m_i \) and \( N \) such that \( |\alpha_i N - m_i| < \epsilon \).

**Example 6** Let \( \dot{\Theta} = (\Omega_z + u \Omega_x)\Theta, \quad \Theta(0) = I \),
Then our system evolves on $SO(3)$, which has tangent space $\mathfrak{g} = \mathfrak{so}(3)$, $3 \times 3$ skew symmetric matrices. $\Omega_z$ is drift which is periodic. By Lie bracket, we get $[\Omega_z, \Omega_x] = \Omega_y$. Now we have all three generators $(\Omega_x, \Omega_y, \Omega_z)$ of $\mathfrak{so}(3)$, we span $\mathfrak{g}$. We can steer the system anywhere on $SO(3)$.

**Example 7**

Let

$$
\dot{\Theta} = \begin{pmatrix}
0 & -1 & 0 \\
1 & 0 & \sqrt{2} \\
0 & -\sqrt{2} & 0
\end{pmatrix} A
+ u
\begin{pmatrix}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix} B
\Theta, \quad \Theta(0) = I,
$$

Then our system evolves on $SO(3)$, which has tangent space $\mathfrak{g} = \mathfrak{so}(3)$, $3 \times 3$ skew symmetric matrices. Note drift is not periodic, but we are on a compact manifold. By Lie bracket, we get $[A, B] = \Omega_y$. Now we have all three independent generators $(A, B, \Omega_z)$ of $\mathfrak{so}(3)$, we span $\mathfrak{g}$. We can steer the system anywhere on $SO(3)$.

5 **Excercises**

1. **Stiefel Manifolds:** Consider for $m \leq n$, $n \times m$ real matrices $\Theta$, such that $\Theta^T \Theta = I_m$.
   Show $\Theta$ is a manifold. Find it dimension and tangent space.

2. Let $V$ be a vector space of real lower triangular matrices. Show $V$ is an Lie algebra. Show same when $V$ is strictly lower triangular.

3. Let us look at variant of non-holonomic integrator

   $$
   \begin{bmatrix}
   \dot{x} \\
   \dot{y} \\
   \dot{z}
   \end{bmatrix} =
   \begin{bmatrix}
   1 & 0 \\
   0 & x \\
   f & y
   \end{bmatrix}
   u
   +
   \begin{bmatrix}
   0 \\
   1 \\
   g
   \end{bmatrix}
   v.
   $$

   Is the system controllable. Find an integral manifold passing through origin on which system lives.

4. For $i = 1, \ldots, n$, let

   $$
   \begin{align*}
   \dot{x}_i &= u_i \\
   \dot{z}_{ij} &= x_i u_j - u_i x_j, \ i < j
   \end{align*}
   $$

   We have $n$ variables $x_i$ and $\frac{n(n-1)}{2}$ variables $z_{ij}$. We have $n$ controls and $\frac{n(n+1)}{2}$ state variables. Is the system controllable?
5. Consider the following system with two controls

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ x_2u_1 \\ \vdots \\ x_{n-1}u_1 \end{bmatrix}.$$ 

Is it controllable.

6. Consider the following system with two controls

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ x_1u_2 \\ x_3u_1 \\ x_3u_2 \end{bmatrix}.$$ 

Is it controllable.

7. Consider the matrix equation

$$\dot{U} = -i(u\sigma_x + v\sigma_y)U, \ U(0) = I,$$

where $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\sigma_x = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$. Show that system evolves on $SU(2)$. Is it controllable.

8. Let $A = -i \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & \ldots & \iddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}$. Consider the control system

$$\dot{U} = (A + \sum_{k=2}^{n} u_k\Omega_{1k})U, \ U(0) = I$$

with skew symmetric $\Omega_{1k}$ as defined in the main text. Show system always evolves on $SU(n)$. Show it is controllable if $A \neq 0$. 

23
9. Let \( A = -i \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & \ldots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix} \) and \( B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & \ddots \\ 0 & 0 & \ldots & -1 & 0 \end{bmatrix} \). Consider the control system
\[
\dot{U} = (A + uB)U, \quad U(0) = I
\]
with \( \lambda_{k+1} - \lambda_k \neq \lambda_{j+1} - \lambda_j \). Show system always evolves on \( SU(n) \). Show it is controllable.

10. Let \( H \) and \( U \) be real symmetric matrices and \( Y \) skew symmetric. Consider the system
\[
\begin{align*}
\dot{H} &= U \\
\dot{Y} &= [H, U]
\end{align*}
\]
Is the system controllable over \( H, Y \) for choice of control \( U \).