

(43)

Normal Form Reduction

Introduction: Let $\text{vgl}(n)^{\mathbb{C}}$ denote the $\mathbb{C}$ -vector space of complex $n \times n$ (square) matrices. Suppose we want to prove the following statement: for all $A \in \text{vgl}(n)^{\mathbb{C}}$, we have

$$\text{Tr}(A) = \sum_{\substack{\lambda \text{ eigen-} \\ \text{value of } A}} \lambda, \quad \text{where the sum takes into account the algebraic multiplicity of the eigenvalues}$$

Here's a way to do it, in 2 steps:

- 1) Note that, if $A \sim B$ (so $A = CBC^{-1}$ for some invertible $C \in \text{vgl}(n)^{\mathbb{C}}$) then $\text{Tr}(A) = \text{Tr}(CBC^{-1}) = \text{Tr}(BC^{-1}C) = \text{Tr}(B)$, and likewise
- $$\sum_{\substack{\lambda \text{ eigen-} \\ \text{value of } A}} \lambda = \sum_{\substack{\lambda \text{ eigen-} \\ \text{value of } B}} \lambda, \quad \text{so both sides are independent of conjugation.}$$

- 2) If $A = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_k & \\ 0 & & & \ddots & \\ & & & & \lambda_k \end{pmatrix}$ is in Jordan-normal form, then clearly $\text{Tr}(A) = \lambda_1 + \lambda_1 + \dots + \lambda_k$

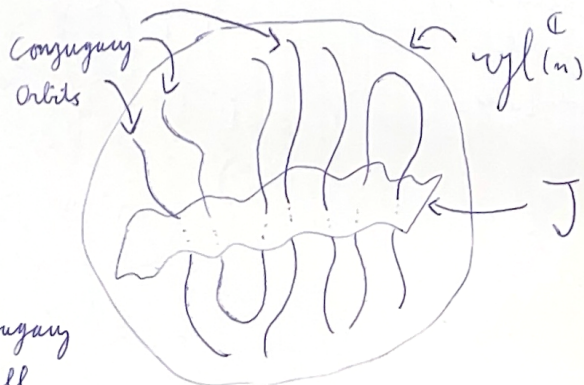
Since every A has a ^{Jordan} normal form in $\text{vgl}(n)^{\mathbb{C}}$, the result follows from 1) + 2).

What did we do here? We picked a subset $J \subseteq \text{vgl}(n)^{\mathbb{C}}$ such that every conjugacy class intersects J at least once

J is the set of matrices in the form ②

Note that, e.g.

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{so conjugacy orbits will often intersect } J \text{ more than once.}$$



(44) But the proof didn't care. in fact: we could have just as easily used J : "upper triangular matrices", so that

$$\text{for } n \geq 2, \#(C(A) \cap J) = \infty, \text{ where } C(A) = \{B \in \text{M}_n(\mathbb{C}) \mid A \sim B\}$$

What matters is that 1) the statement is easy to prove for elements in J

2) the property is independent of conjugacy-class

3) every element in $\text{M}_n(\mathbb{C})$ has an element of J (at least 1) in its conjugacy-class.

Note that, by definition, dynamical properties are independent on dynamical conjugacy, so 2) is sort-of guaranteed for questions about dynamical systems. In fact, how did you learn about the dynamical behavior of linear ODEs?

$\dot{X} = AX$ You first understood it for (real) Jordan normal-forms, and then noted that, if

$$A = CJC^{-1}, \text{ for some invertible } C, \text{ and } J \text{ in real}$$

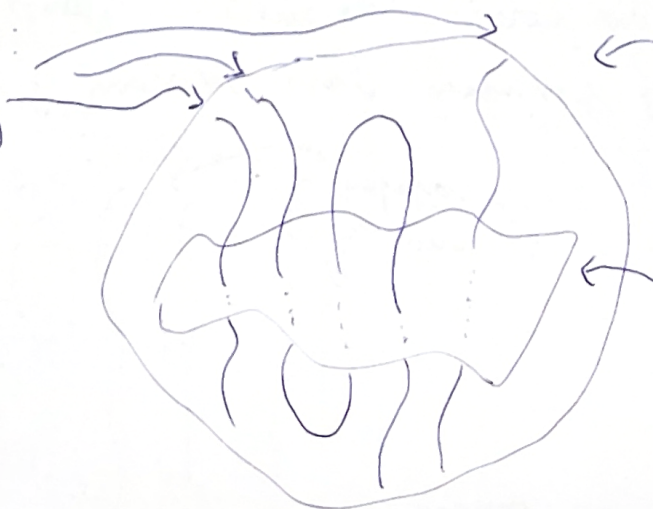
Jordan normal form, we may define new-coordinates

$$y = C^{-1}x, \text{ then}$$

$$\dot{y} = C^{-1}\dot{x} = C^{-1}AX = C^{-1}ACy = Jy, \text{ Done!}$$

The idea of normal-forms (normal form reduction) is to do this for more general (non-linear) ODEs too.

i.e.:
Sets of
(dynamically)
conjugate
vectorfields

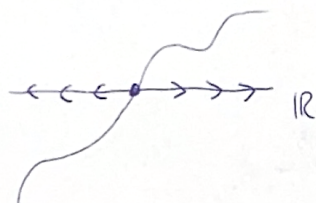


"all ODEs"
on $\mathbb{R}^n$

Subset N of
ODEs that
are easy to
analyse

(15) Often, the class $\mathcal{N}$ is defined by the property that vectorfields on it "look easy". I.e. lots of Taylor-coefficients vanish. Usually, all statements above should be read locally.

Example: Ever wondered why all ODEs on $\mathbb{R}$ of the form
 $\dot{X} = X + a_2 X^2 + a_3 X^3 + \dots$ look alike around $X=0$?



Well, introduce new coordinates: $y = X + bX^2$,
 for some $b \in \mathbb{R}$ (to be determined) by the Implicit/Inverse
 function Theorem, around 0 this is locally a
 diffeomorphism. In fact, if we write $X = c_1 y + c_2 y^2 + \dots$
 then

$$\begin{aligned} y &= X + bX^2 \\ &= c_1 y + c_2 y^2 + \dots + b(c_1 y + c_2 y^2 + \dots)^2 \\ &= c_1 y + c_2 y^2 + \dots + b c_1^2 y^2 + \dots \end{aligned}$$

$$\text{So } c_1 = 1, \quad c_2 + b c_1^2 = 0 \Rightarrow c_2 = -b.$$

$$\text{So } X = y - b y^2 + \dots$$

Then

$$\begin{aligned} \dot{y} &= \dot{X} + 2bX\dot{X} \\ &= X + a_2 X^2 + \dots + 2bX(X + a_2 X^2 + \dots) \\ &= X + (a_2 + 2b)X^2 + \dots \\ &= y - b y^2 + (a_2 + 2b)y^2 + \dots \\ &= y + (a_2 + b)y^2 + \dots \end{aligned}$$

Thus if we choose $b = -a_2$, we'd get the new ODE

$$\dot{y} = y + d_3 y^3 + d_4 y^4 + \dots$$

Thus, no more quadratic terms!

(46)



ODEs on IR of the form
 $\dot{X} = X + a_2 X^2 + a_3 X^3 + \dots$

ODEs on IR of the form
 $\dot{X} = X + d_3 X^3 + d_4 X^4 + \dots$
 (no more quadratic terms)

"Local Conjugacy-classes, around 0"

Now suppose $\dot{y} = y + d_3 y^3 + \dots$

Let $z = y + D y^3$, for some $D \in \mathbb{R}$, as before

$$y = z - D y^3 = z - D(z - D y^3)^3 \\ = z - D z^3 + \text{h.o.t.}$$

$$\begin{aligned} \text{Then } \dot{z} &= \dot{y} + 3D y^2 \dot{y} \\ &= y + d_3 y^3 + 3D y^3 (y + \dots) + \dots \end{aligned}$$

$$\begin{aligned} y = z - D y^3 &\rightarrow \begin{aligned} &= y + d_3 y^3 + 3D y^3 + \dots \\ &= \underbrace{z - D y^3}_{\text{circled}} + d_3 y^3 + 3D y^3 + \dots \end{aligned} \\ &= z + (2D + d_3) y^3 + \text{h.o.t. in } y \\ &= z + (2D + d_3) y^3 + \text{h.o.t. in } z \end{aligned}$$

So if we choose $D = \frac{-d_3}{2}$, we get $\dot{z} = z + l_4 z^4 + l_5 z^5 + \dots$

We can do this as long as we want (! not infinitely long, as we need to worry about convergence of an infinite sequence of locally defined coordinate changes!!)

So the picture is in fact as follows:

Local
conjugacy
classes



ODEs on IR which vanish at 0
and with derivative 1 at 0

ODEs like
 $\dot{X} = X + l_{10^{100}} X^{10^{100}} + \text{h.o.t.}$

(47) Remark. The truncated normal form would be just $X = X$, which all of a sudden has a lot of symmetry: If $\gamma(t)$ is a solution curve, then so is $-\gamma(t)$, and in fact $c\gamma(t)$ for any $c \in \mathbb{R}$. In a sense, we see that we have introduced symmetry in the (truncated) ODE. This will be a theme in the topic of normal form reduction.

A tool we will use a lot in this chapter is the Lie-Bracket between vectorfields $f, g: \mathbb{R}^n \rightarrow \mathbb{R}^n$, denoted by $[f, g]: \mathbb{R}^n \rightarrow \mathbb{R}^n$, which is another vector field on $\mathbb{R}^n$. If we write

$$f(x) = (f_1(x), \dots, f_n(x)) \quad , \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

$$g(x) = (g_1(x), \dots, g_n(x))$$

and $[f, g] = ([f, g]_1, \dots, [f, g]_n)$ then we have explicitly

$$([f, g])_i(x) = \sum_{k=1}^n \left(f_k(x) \frac{\partial g_i(x)}{\partial x_k} - g_k(x) \frac{\partial f_i(x)}{\partial x_k} \right).$$

for $i \in \{1, \dots, n\}$. To make more sense of this, we will now introduce some formalism from differential geometry. Let us denote by $C^\infty(\mathbb{R}^n)$ the $\mathbb{R}$ -vector space of all C^∞ -maps from $\mathbb{R}^n$ to $\mathbb{R}$. In differential geometry, a vectorfield X is an $\mathbb{R}$ -linear map from $C^\infty(\mathbb{R}^n)$ to $C^\infty(\mathbb{R}^n)$ satisfying the Leibniz-rule: $\forall h, p \in C^\infty(\mathbb{R}^n)$:

$$X(hp) = hX(p) + X(h)p$$

Some properties follow readily about X : for $1 \in C^\infty(\mathbb{R}^n)$ the constant function

$$1) \quad X(1) = X(1 \cdot 1) = 1 \cdot X(1) + X(1) \cdot 1 = 2X(1)$$

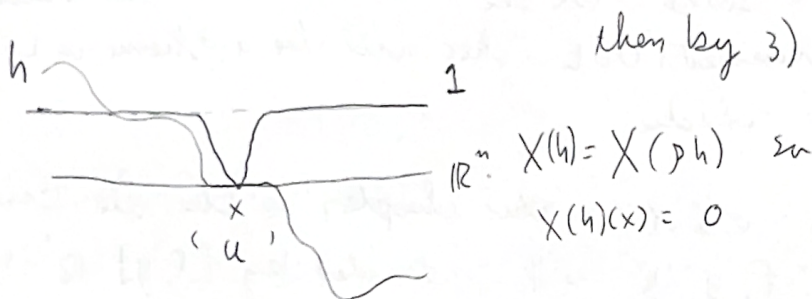
$$\text{thus } X(1) = 2X(1) \Rightarrow X(1) = 0$$

$$2) \quad \text{for any } c \in \mathbb{R}, \quad X(c) = cX(1) = 0, \text{ so } X \text{ vanishes on all constant functions in } C^\infty(\mathbb{R}^n)$$

$$3) \quad \text{for } X \neq 0 \in \mathbb{R}^n, \text{ if } h \in C^\infty(\mathbb{R}^n) \text{ can be written as } h = h_1 \cdot h_2, \text{ for some } h_1, h_2 \in C^\infty(\mathbb{R}^n) \text{ such that } h_1(x) = h_2(x) = 0, \text{ then}$$

$$X(h)(x) = X(h_1 h_2)(x) = h_1(x) X(h_2)(x) + h_2(x) X(h_1)(x) = 0.$$

- 4) for $x \in \mathbb{R}^n$, if we have $x \in U \subseteq \mathbb{R}^n$ open and $h \in C^\infty(\mathbb{R}^n)$ s.t. $h|_U = 0$ then $X(h)(x) = 0$. Because we may find $p \in C^\infty(\mathbb{R}^n)$ such that $ph = h$ and $p(x) = 0$:



- 5) for $x \in \mathbb{R}^n$ if we have $x \in U \subseteq \mathbb{R}^n$ open and $h \in C^\infty(\mathbb{R}^n)$ s.t. $h|_U = c$ (constant) then $X(h)(x) = 0$, as

$$0 \stackrel{\text{by 4)}}{=} X(h-c)(x) = X(h)_{(x)} - X(c)_{(x)} = X(h)(x) = 0 \text{ by 2)}$$

- 6) X is local. I.E. for $x \in \mathbb{R}^n$, if $x \in U \subseteq \mathbb{R}^n$ open and $p, q \in C^\infty(\mathbb{R}^n)$ are such that $p|_U = q|_U$ then $X(p)(x) = X(q)(x)$. Because $X(p-q)_{(x)} = 0$ by 4).
So $X(p)(x) - X(q)(x) = 0$.

Putting this all together, you can show that a vectorfield X necessarily looks like $X = X_f = \sum_{k=1}^n f_k(x) \frac{\partial}{\partial x_k}$, for some $f = (f_1, \dots, f_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ($f_k \in C^\infty(\mathbb{R}^n)$)

where for $h \in C^\infty(\mathbb{R}^n)$ then

$$X_f(h)(x) = \sum_{k=1}^n f_k(x) \frac{\partial h(x)}{\partial x_k}$$

Thus, there is a 1-to-1 correspondence between vectorfields on $\mathbb{R}^n$, seen as maps $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (C^∞), and vectorfields as in differential geometry on $\mathbb{R}^n$, by sending

$$f = (f_1, \dots, f_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \text{to} \quad X_f(x) = \sum_{k=1}^n f_k(x) \frac{\partial}{\partial x_k}$$

④ Let us denote by $\Gamma(T\mathbb{R}^n)$ the $\mathbb{R}$ -vector space of C^∞ vector fields (say in the diff-geo. definition) on $\mathbb{R}^n$. Given $X, Y \in \Gamma(T\mathbb{R}^n)$, define, ~~for~~

$$[X, Y] : C^\infty(\mathbb{R}^n) \rightarrow C^\infty(\mathbb{R}^n) \text{ by}$$

$$[X, Y](h) = X(Y(h)) - Y(X(h)) \text{ for } h \in C^\infty(\mathbb{R}^n). \text{ clearly } [X, Y] \text{ is a linear map from } C^\infty(\mathbb{R}^n) \text{ to itself. Moreover, for } p, q \in C^\infty(\mathbb{R}^n),$$

$$\begin{aligned} [X, Y](pq) &= X(Y(pq)) - Y(X(pq)) \\ &= X(pY(q) + qY(p)) - Y(pX(q) + qX(p)) \\ &= pXY(q) + \underline{Y(q)X(p)} + qXY(p) + \underline{Y(p)X(q)} \\ &\quad - pYX(q) - \underline{X(q)Y(p)} - qYX(p) - \underline{X(p)Y(q)} \\ &= p(XY(q) - YX(q)) + q(XY(p) - YX(p)) \\ &= p[X, Y](q) + q[X, Y](p). \end{aligned}$$

So $[X, Y]$ satisfies Leibniz rule, so $[X, Y] \in \Gamma(T\mathbb{R}^n)$

• If $X_{(x)} = \sum_{k=1}^n f_k(x) \frac{\partial}{\partial x^k}$ and $Y_{(x)} = \sum_{k=1}^n g_k(x) \frac{\partial}{\partial x^k}$, and $h \in C^\infty(\mathbb{R}^n)$

$$\begin{aligned} \text{then } X(Y(h)) &= X\left(\sum_{k=1}^n g_k(x) \frac{\partial h(x)}{\partial x^k}\right) \\ &= \sum_{k,l=1}^n f_l(x) \frac{\partial}{\partial x^l} \left(g_k(x) \frac{\partial h(x)}{\partial x^k}\right) \\ &= \sum_{k,l=1}^n f_l(x) \left(\frac{\partial g_k(x)}{\partial x^l} \frac{\partial h(x)}{\partial x^k} + g_k(x) \frac{\partial^2 h(x)}{\partial x^k \partial x^l}\right) \end{aligned}$$

Likewise: $Y(X(h)) =$

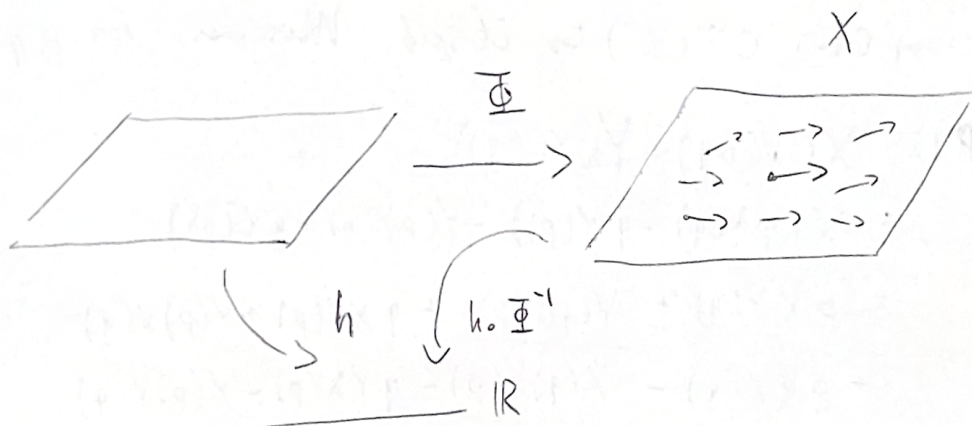
$$\sum_{k,l=1}^n g_l(x) \left(\frac{\partial f_k(x)}{\partial x^l} \frac{\partial h(x)}{\partial x^k} + f_k(x) \frac{\partial^2 h(x)}{\partial x^k \partial x^l}\right)$$

So $X(Y(h)) - Y(X(h)) =$

$$\begin{aligned} &\sum_{k,l=1}^n f_l(x) \frac{\partial g_k(x)}{\partial x^l} \frac{\partial h(x)}{\partial x^k} - g_l(x) \frac{\partial f_k(x)}{\partial x^l} \frac{\partial h(x)}{\partial x^k} \\ &= \sum_{k=1}^n \left(\sum_{l=1}^n f_l(x) \frac{\partial g_k(x)}{\partial x^l} - g_l(x) \frac{\partial f_k(x)}{\partial x^l} \right) \frac{\partial}{\partial x^k} \circ h \end{aligned}$$

50) Thus. $[X_f, X_g] = X_{[f, g]}$ as defined earlier!

Now suppose we have a diffeomorphism $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$
and a vectorfield $X \in \Gamma(T\mathbb{R}^n)$



We may define a new vectorfield on $\mathbb{R}^n$, $\Phi^*X \in \Gamma(T\mathbb{R}^n)$
(called the pullback of X along Φ) defined by what it does
on elements $h \in C^\infty(\mathbb{R}^n)$ by

$$((\Phi^*X)(h))(x) = (X(h \circ \Phi^{-1}))(\Phi(x)), \text{ for } x \in \mathbb{R}^n$$

Note that $\Phi^*X: C^\infty(\mathbb{R}^n) \rightarrow C^\infty(\mathbb{R}^n)$ is $\mathbb{R}$ -linear. Moreover, given
 $p, q \in C^\infty(\mathbb{R}^n)$

$$\begin{aligned} (\Phi^*X)(pq) &= X(pq \circ \Phi^{-1}) \circ \Phi \\ &= X((p \circ \Phi^{-1})(q \circ \Phi^{-1})) \circ \Phi \\ &= ((p \circ \Phi^{-1})X(q \circ \Phi^{-1}) + (q \circ \Phi^{-1})X(p \circ \Phi^{-1})) \circ \Phi \\ &= p \cdot X(q \circ \Phi^{-1}) \circ \Phi + q \cdot X(p \circ \Phi^{-1}) \circ \Phi \\ &= p(\Phi^*X)(q) + q(\Phi^*X)(p) \end{aligned}$$

So $\Phi^*X \in \Gamma(T\mathbb{R}^n)$ (it satisfies Leibniz too!)

We can see what Φ^*X looks like by noting that, if

$$(\Phi^*X)(x) = \sum_{k=1}^n g_k(x) \frac{\partial}{\partial x_k}, \text{ then } ((\Phi^*X)(x_i))(x) = \sum_{k=1}^n g_k(x) \frac{\partial x_i}{\partial x_k} = g_i(x)$$

$X \in C^\infty(\mathbb{R}^n)$

(51) So for $i \in \{1, \dots, n\}$, $((\Phi^* X)(x))_i = X(x_i \circ \Phi^{-1})(\Phi(x))$

Write $X_{\Phi(x)} = \sum_{k=1}^m f_k(x) \frac{\partial}{\partial x^k}$, so $X(x_i \circ \Phi^{-1})(\Phi(x)) = X(\Phi^{-1})_i(\Phi(x))$

$$= \sum_{k=1}^n f_k(\Phi(x)) \frac{\partial \Phi^{-1}_i}{\partial x^k}(\Phi(x)) = \sum_{k=1}^m (D\Phi^{-1}(\Phi(x)))_{ik} f_k(\Phi(x))$$

If we identify $\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\}$ with the standard basis $\{e_1, \dots, e_n\}$ for $\mathbb{R}^n$, then we can write $f_k(\Phi(x)) = (X(\Phi(x)))_k$, so we get

$$((\Phi^* X)(x))_i = \sum_{k=1}^m (D\Phi^{-1}(\Phi(x)))_{ik} (X(\Phi(x)))_k \quad \text{or just:}$$

$$(\Phi^* X)(x) = D\Phi^{-1}(\Phi(x)) X(\Phi(x))$$

Note that $\text{Id} = \Phi^{-1} \circ \Phi$, so $\text{Id} = D\Phi^{-1}(\Phi(x)) \cdot D\Phi(x)$.

so $D\Phi^{-1}(\Phi(x)) = (D\Phi(x))^{-1}$. Thus we may also

write $(\Phi^* X)(x) = (D\Phi(x))^{-1} X(\Phi(x))$. (This may be understood

in terms of tangent spaces too, where $(\Phi^* X)(x) \in \text{span}\{\frac{\partial}{\partial x_1}|_x, \dots, \frac{\partial}{\partial x_n}|_x\} =: T_x \mathbb{R}^d$ and $X(\Phi(x)) \in T_{\Phi(x)} \mathbb{R}^d$ and we have linear maps

$$D\Phi(x): T_x \mathbb{R}^d \rightarrow T_{\Phi(x)} \mathbb{R}^d, (D\Phi(x))^{-1} = D\Phi^{-1}(\Phi(x)): T_{\Phi(x)} \mathbb{R}^d \rightarrow T_x \mathbb{R}^d$$

Why did we introduce this? Well, given $X \in \mathfrak{X}(\mathbb{R}^d)$,

denote by $\varphi_t^X(x)$ its flow. Then for fixed (small) t ,

the map $\varphi_t^X: U \rightarrow \varphi_t(U)$ is a local diffeomorphism (for U open and sufficiently small), given $x \in U$ and t small

enough, we have the family of

$$\text{tangent vectors } (\varphi_t^{X*} Y)(x) \in T_x \mathbb{R}^d$$

⑤2) Lemma $\frac{d}{dt}\big|_{t=0} (\varphi_t^* Y)(x) = [X, Y](x)$

Proof: Fix $h \in C^\infty(\mathbb{R}^d)$, then

$$\begin{aligned} \frac{d}{dt}\big|_{t=0} ((\varphi_t^* Y)(h))(x) &= \frac{d}{dt}\big|_{t=0} Y(h \circ \varphi_{-t}^x)(\varphi_t^x(x)) \\ &= \frac{d}{dt}\big|_{t=0} Y(h \circ \varphi_{-t}^x)(x) + \frac{d}{dt}\big|_{t=0} Y(h)(\varphi_t^x(x)) \\ &= Y(-X(h))(x) + X(Y(h))(x) = X(Y(h))(x) - Y(X(h))(x) \\ &= ([X, Y](h))(x) \quad \square \end{aligned}$$

Convince yourself that, for $F, G: \mathbb{R}^d \hookrightarrow (\text{local}) \text{ diffeomorphisms}$ (with appropriate (co)-domains), we have

$$(G \circ F)^* Y = F^*(G^* Y) \quad \text{for } Y \in \Gamma(T\mathbb{R}^d)$$

```

    graph LR
      Y[Y] -- G --> Gy[G*Y]
      Gy -- F --> Fgy[F*G*Y]
      Y -- "G o F" --> Fgy
  
```

Thus we get as corollary of the previous lemma:

Cor $\frac{d}{dt} \varphi_t^* Y = [X, \varphi_t^* Y]$ (for all t close enough to 0)

Proof $\frac{d}{dt}\big|_{t=s} \varphi_t^* Y = \frac{d}{du}\big|_{u=0} \varphi_{u+s}^* Y = \frac{d}{du}\big|_{u=0} \varphi_u^* (\varphi_s^* Y)$

$= (\text{by lemma}) = [X, \varphi_s^* Y]$, as s is general, we get

$$\frac{d}{dt} \varphi_t^* Y = [X, \varphi_t^* Y]$$

Before we apply this machinery to coordinate transformations, we quickly gather some interesting properties of the Lie-Bracket. These easily follow from the equivalent description, for $X, Y \in \Gamma(T\mathbb{R}^d)$, $h \in C^\infty(\mathbb{R}^d)$, $[X, Y](h) = X(Y(h)) - Y(X(h))$

53) 1) $[X, Y] = -[Y, X]$ for all $X, Y \in \Gamma(\mathbb{R}^d)$
 In particular, $[X, X] = 0$ (antisymmetry)

2) $[sX_1 + rX_2, Y] = s[X_1, Y] + r[X_2, Y]$ &
 $[X, sY_1 + rY_2] = s[X, Y_1] + r[X, Y_2]$ for all $s, r \in \mathbb{R}$,
 for all $X_1, X_2, Y_1, Y_2 \in \Gamma(\mathbb{R}^d)$
 ($\mathbb{R}$ -bilinearity)

3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$
 for all $X, Y, Z \in \Gamma(\mathbb{R}^d)$
 (Jacobi-identity)

4) $[X, fY] = f[X, Y] + X(f)Y$ & for all $X, Y \in \Gamma(\mathbb{R}^d)$
 $[gX, Y] = g[X, Y] - Y(g)X$ and $f, g \in C^\infty(\mathbb{R}^d)$

Properties 1), 2), 3) are reminiscent of the usual commutator between matrices $A, B \in \text{gl}(\mathbb{R}^d)$: $[A, B] := AB - BA$
 These first 3 properties make $\Gamma(\mathbb{R}^d)$ with $[\cdot, \cdot]$ into a Lie-algebra
 (as well as $\text{gl}(\mathbb{R}^d)$ with the commutator)

Infact: Let $X_{A, B} \in \Gamma(\mathbb{R}^d)$ be defined by $X_{A, B}(x) = Ax - Bx$,
 for $A, B \in \text{gl}(\mathbb{R}^d)$. Then $\varphi_t^{X_A}(x) = e^{tA}x$

$$\text{So } \frac{d}{dt}\bigg|_{t=0} (\varphi_t^{X_A*} X_B)(x) = \frac{d}{dt}\bigg|_{t=0} e^{-tA} B e^{tA} = BA - AB$$

So we get the commutator with a minus.

Also property 4) is enough to calculate all Lie brackets, as clearly

$$[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0. \text{ Thus}$$

$$\begin{aligned} & \left[\sum_{i=1}^d f_i(x) \frac{\partial}{\partial x_i}, \sum_{j=1}^d g_j(x) \frac{\partial}{\partial x_j} \right] = \sum_{i,j=1}^d \left[f_i(x) \frac{\partial}{\partial x_i}, g_j(x) \frac{\partial}{\partial x_j} \right] \\ & = \sum_{i,j=1}^d g_j \left[f_i \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right] + f_i \frac{\partial g_j}{\partial x_i} \frac{\partial}{\partial x_j} = \sum_{i,j=1}^d g_j f_i \left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right] \\ & - \sum_{i,j=1}^d g_j \frac{\partial f_i}{\partial x_j} \frac{\partial}{\partial x_i} + \sum_{i,j=1}^d f_i \frac{\partial g_j}{\partial x_i} \frac{\partial}{\partial x_j} = \sum_{i=1}^d \left(\sum_{j=1}^d f_j \frac{\partial g_i}{\partial x_j} - \sum_{j=1}^d g_j \frac{\partial f_i}{\partial x_j} \right) \frac{\partial}{\partial x_i} \end{aligned}$$

(54) Now Suppose we have a vectorfield $Y \in \Gamma(T\mathbb{R}^d)$. Suppose $Y(0) = 0$ (the vectorfield vanishes at the origin) the coordinate transformations we will use will be the time-1 flow of another vectorfield $X \in \Gamma(T\mathbb{R}^d)$. So, let $\varphi_t(x)$ denote the flow of $\dot{x} = X(x)$. We assume $X(0) = 0$ as well, so that $\varphi_t(0) = 0$ (and in particular $\varphi_1(0) = 0$)

Recall from the previous Corollary that

$$\frac{d}{dt} \varphi_t^* Y = [X, \varphi_t^* Y] \quad \star$$

Let us write $\text{ad}_X: \Gamma(T\mathbb{R}^d) \rightarrow \Gamma(T\mathbb{R}^d)$ for the $\mathbb{R}$ -linear map given by $\text{ad}_X Z = [X, Z]$, $Z \in \Gamma(T\mathbb{R}^d)$. Then $\star$ can be written as

$$\frac{d}{dt} (\varphi_t^* Y) = \text{ad}_X (\varphi_t^* Y)$$

$t \mapsto \varphi_t^* Y$ is a curve in $\Gamma(T\mathbb{R}^d)$, and $\text{ad}_X: \Gamma(T\mathbb{R}^d) \rightarrow \Gamma(T\mathbb{R}^d)$ is linear, meaning that (this can be made more rigorous!)

$$\begin{aligned} \varphi_t^* Y &= e^{t \text{ad}_X} (\varphi_0^* Y) = e^{t \text{ad}_X} (Y) \\ &= Y + t[X, Y] + \frac{t^2}{2} [X[X, Y]] + \frac{t^3}{3!} [X[X[X, Y]]] + \dots \end{aligned}$$

Setting $t=1$ then gives $\varphi_1^* Y = Y + [X, Y] + \frac{1}{2} [X[X, Y]] + \dots$

Now let us define, for $k \in \{1, 2, \dots\}$ the space

$\mathcal{P}_k^{\mathbb{R}^d}$ (or simply $\mathcal{P}^k$) as the space of all polynomial maps from $\mathbb{R}^d$ to $\mathbb{R}^d$ with only monomials of degree k .

e.g. $\mathcal{P}_3^2 = \left\{ \begin{aligned} & a_{(2,0,0)}^1 X^2 + a_{(1,1,0)}^1 X Y + a_{(0,1,1)}^1 Y Z + a_{(1,0,1)}^1 X Z + a_{(0,2,0)}^1 Y^2 + a_{(0,0,2)}^1 Z^2 \\ & a_{(2,0,0)}^2 X^2 + a_{(1,1,0)}^2 X Y + a_{(0,1,1)}^2 Y Z + a_{(1,0,1)}^2 X Z + a_{(0,2,0)}^2 Y^2 + a_{(0,0,2)}^2 Z^2 \\ & a_{(2,0,0)}^3 X^2 + a_{(1,1,0)}^3 X Y + \dots \dots \dots a_{(0,0,2)}^3 Z^2 \end{aligned} \right\}$
 $a_{(i,j,k,l)} \in \mathbb{R}$

(55) By the definition of the Lie-Bracket, we have

$$[P^k, P^l] \in P^{k+l-1}$$

Note that, since $x \mapsto AX \in P^1$, for $A: \mathbb{R}^d \rightarrow \mathbb{R}^d$ linear, we have

$$\text{ad} A: P^k \hookrightarrow P^{k-1} \text{ for each } k \geq 1$$

Note that, for $P \in P^k$ (or more general $P \in \Gamma(T\mathbb{R}^d)$) we have

$$\text{ad} A(P)(x) := [A, P](x) = DP(x)AX - AP(x)$$

Now let us write $Y(x) = AX + Y^{(2)}(x) + Y^{(3)}(x) + \dots$

and choose $X(x) = X^{(2)}(x) + X^{(3)}(x) + \dots$

$$\text{for } Y^{(k)}, X^{(k)} \in P^k, k \geq 2$$

Then we have $\varphi_1^* Y = Y + [X, Y] + \frac{1}{2}[X[X, Y]] + \dots$

$$= AX + Y^{(2)}(x) + Y^{(2)}(x) + Y^{(3)}(x) + \dots$$

$$+ [X^{(2)}, A](x) + [X^{(2)}, Y^{(2)}] + [X^{(2)}, Y^{(3)}] + \dots$$

$$+ [X^{(3)}, A](x) + [X^{(3)}, Y^{(2)}] + \dots$$

$$+ [X^{(4)}, A](x) + \dots$$

$$+ \frac{1}{2}[X^{(2)}, [X^{(2)}, A]](x) + \frac{1}{2}[X^{(2)}, [X^{(2)}, Y^{(2)}]](x) + \dots$$

$$\frac{1}{2}[X^{(2)}, [X^{(3)}, A]](x) + \frac{1}{2}[X^{(3)}, [X^{(2)}, A]](x) + \dots$$

$$+ \frac{1}{6}[X^{(2)}, [X^{(2)}, [X^{(2)}, A]]](x) + \dots$$

Thus we have:

linear terms: AX

$$\text{quadratic terms: } Y^{(2)}(x) + [X^{(2)}, A](x) = Y^{(2)}(x) - \text{ad} A(X^{(2)})(x)$$

$$\text{third order: } Y^{(3)}(x) + [X^{(3)}, A](x) + [X^{(2)}, Y^{(2)}](x) + \frac{1}{2}[X^{(2)}, [X^{(2)}, A]](x)$$

$$= Y^{(3)}(x) + [X^{(2)}, Y^{(2)}](x) + \frac{1}{2}[X^{(2)}, [X^{(2)}, A]](x) - \text{ad} A(X^{(3)})(x)$$

56 fourth order:

$$Y^{(4)}(x) + [X^{(2)}, Y^{(3)}] + [X^{(3)}, Y^{(2)}] \\ + \frac{1}{2} \left([X^{(2)}, [X^{(2)}, Y^{(2)}]](x) + [X^{(2)}, [X^{(3)}, A]](x) \right. \\ \left. + [X^{(3)}, [X^{(2)}, A]](x) \right) \\ + \frac{1}{6} [X^{(2)}, [X^{(2)}, [X^{(2)}, A]](x) - \text{ad } A(X^{(4)})$$

In general, the k -th order is given by

$$G^{(k)}(A, Y^{(2)}(x), \dots, Y^{(k)}(x), X^{(2)}(x), \dots, X^{(k-1)}(x)) - \text{ad } A(X^{(k)})$$

for some complicated formula $G^{(k)}$ involving Lie-brackets

This leads to the main result:

Theorem: For each $k \geq 2$, choose a subspace

$$V^k \subseteq \mathcal{P}^k, \text{ such that } \mathcal{P}^k = V^k \oplus \text{Im}(\text{ad } A|_{\mathcal{P}^k})$$

Choose $m \geq 2$. Then there exists a local diffeomorphism $\Psi: \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\Psi(0) = 0$ and

$$(\Psi^* Y)(x) = Ax + Z^{(2)}(x) + Z^{(3)}(x) + \dots \quad (\text{around } 0)$$

$$\text{where } Z^{(k)}(x) \in V^k \text{ for all } 2 \leq k \leq m.$$

Proof. We will choose a vectorfield X such as before:

$$X = X^{(2)} + X^{(3)} + \dots + X^{(m)} \quad \text{and } \Psi = \varphi_1 \text{ for the}$$

corresponding flow.

$$\text{Then } (\Psi^* Y)_{(x)} = (\varphi_1^* Y)_{(x)} = Ax + \text{h.o.t.}$$

The quadratic terms are given by

$$Y^{(2)}(x) - \text{ad } A(X^{(2)})(x), \text{ so write}$$

$$Y^{(2)} = w^{(2)} + Q^{(2)}, \text{ with } w^{(2)} \in V^{(2)}$$

$$Q^{(2)} \in \text{Im}(\text{ad } A|_{\mathcal{P}^2}).$$

by the direct sum decomposition

(57) Now simply choose $X^{(2)}$ such that $\text{ad} A(X^{(2)}) = Q^{(2)}$.

then $Y^{(2)} - \text{ad} A(X^{(2)}) = \omega^{(2)} + Q^{(2)} - Q^{(2)} = \omega^{(2)} \in V^{(2)}$

Now. Suppose we have arranged for $Z^{(2)}, Z^{(3)} \dots Z^{(k-1)}$ to be in $V^{(2)}, V^{(3)} \dots V^{(k-1)}$ by choosing $X^{(2)} \dots X^{(k-1)}$ appropriately. then for k th order, write.

$$G^{(k)}(A, Y^{(2)}(x) \dots Y^{(k)}(x), X^{(2)}(x), \dots, X^{(k-1)}(x)) = \omega^{(k)} + Q^{(k)}$$

$$Q^{(k)} \in \text{Im}(\text{ad} A|P^k), \quad \omega^{(k)} \in V^{(k)}$$

choose $X^{(k)} \in P^k$ such that $\text{ad} A(X^{(k)}) = Q^{(k)}$.

then the k -th term is now given by

$$Z^{(k)} = \omega^{(k)} + Q^{(k)} - Q^{(k)} = \omega^{(k)} \in V^{(k)}. \text{ Thus the}$$

result follows

In fact, we have:

Lemma: Suppose we are given $X^{(k)} \in P^k$ for each $k \geq 2$.

Then there exists a (C^∞) vectorfield $X \in \Gamma(T(\mathbb{R}^d))$

such that it's (barnal) Taylor expansion around

$x=0$ is given by $X(x) = X^{(2)}(x) + X^{(3)}(x) + \dots$

and the time- t flow of X is defined on all of $\mathbb{R}^d$ for all time.

Proof The last property can always be arranged for, without destroying the first, by multiplying X by a bump-function. Then $X(x) = 0$ for $\|x\| \geq C$ for some C , so the flow is stationary there.

Now let $\chi: \mathbb{R}^d \rightarrow \mathbb{R}$ be a bump function, (C^∞) ,

assume $0 \leq \chi(x) \leq 1$ for all $x \in \mathbb{R}^d$, $\chi(x) = 1$ if $\|x\| \leq 1$
 $\chi(x) = 0$ if $\|x\| \geq 2$

(58) define $\phi_k(x) = \mathcal{X}(x) X^{(k)}(x)$, then,

let $M_k := \sup_{\substack{x \in \mathbb{R}^d \\ 0 \leq m \leq k}} \|D^m \phi_k(x)\| < \infty$, and $p_k = \min(1, 2^{-k-1} M_k^{-1})$

Then: set $\psi_k(x) = \mathcal{X}(p_k^{-1}x) X^{(k)}(x)$
 $= \mathcal{X}(p_k^{-1}x) X^{(k)}(p_k p_k^{-1}x)$
 $= p_k^K \mathcal{X}(p_k^{-1}x) X^{(k)}(p_k^{-1}x)$
 $= p_k^K \phi_k(p_k^{-1}x)$

because $X^{(k)}$ has as its components polynomials of all terms of degree k

Then for fixed $m \geq 0$, for all $k > m$, if $k \geq 2$:

$$\|D^m \psi_k(x)\| \leq p_k^{k-m} \sup_{y \in \mathbb{R}^d} \|D^m \phi_k(y)\| \leq p_k M_k = 2^{-k}$$

(as $p_k \leq 1$ and

$k-m > 0$, we have

$$p_k^{k-m} \leq p_k)$$

Thus for each $m \geq 0$, the series

$$\psi^{(m)}(x) := \sum_{k=2}^{\infty} D^m \psi_k(x) \text{ converges uniformly in } x$$

then $\psi^{(0)}(x)$ will satisfy the conditions, as (by standard results) $\psi^{(m)}(x)$ is a differentiable map with derivative $\psi^{(m+1)}(x)$, so $\psi_{\infty}^{(0)}$ is C^{∞} , with derivatives $\psi^{(1)}, \psi^{(2)}$, etc.

Also $\psi^{(0)}(0) = \sum_{k=2}^{\infty} \psi_k(0) = 0$,

$$D\psi^{(0)}(0) = \psi^{(1)}(0) = \sum_{k=2}^{\infty} D\psi_k(0) = 0$$

$\vdots$

$$D^m \psi^{(0)}(0) = \psi^{(m)}(0) = \sum_{k=2}^{\infty} D^m \psi_k(0) = D^m \psi_{\infty}(0) = D^m X^{(m)}(0)$$

So the derivatives agree, for all orders, at $x=0$ $\square$

(59) Thus we obtain:

Theorem Given $Y(x) = Ax + Y^{(2)}(x) + Y^{(3)}(x) \in \Gamma(T\mathbb{R}^d)$ of class C^∞

For each $k \geq 2$, choose a subspace $V^k \subseteq \mathcal{P}^k$ such that $\mathcal{P}^k = V^k \oplus \text{Im}(\text{ad}_A|_{\mathcal{P}^k})$. Then there exists a C^∞ global diffeomorphism $\Psi: \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\Psi(0) = 0$ and

$$(\Psi^* Y)(x) = Ax + Z^{(2)}(x) + Z^{(3)}(x) + \dots \quad (\text{around } x=0)$$

where $Z^{(k)}(x) \in V^k$ for all $k \geq 2$.

(In the sense of Taylor expansion i.e. formally)

Proof: Choose $X^{(2)} \in \mathcal{P}^2$, $X^{(3)} \in \mathcal{P}^3$ etc. satisfying

$G^{(k)}(A, Y^{(2)}, \dots, Y^{(k)}, X^{(2)}, \dots, X^{(k-1)}) - \text{ad}_A(X^{(k)}) \in V^k$, iteratively, as before. Then, as by the previous lemma, choose $X \in \Gamma(T\mathbb{R}^d)$ smooth with $X^{(2)}, X^{(3)}, \dots$ as its formal power-series at $x=0$. The corresponding time 1 flow $\Psi := \varphi_{t=1}: \mathbb{R}^d \rightarrow \mathbb{R}^d$ then does the trick. $\square$

We may also incorporate parameters:

Theorem Let $A \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ be a given linear map. As before, choose $V^k \subseteq \mathcal{P}^k$ such that $\mathcal{P}^k = V^k \oplus \text{Im}(\text{ad}_A|_{\mathcal{P}^k})$, $k \geq 1$. Let

$Y: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$ be a C^∞ , d -parameter family of vector fields on $\mathbb{R}^n$ such that $Y(0; \lambda) = 0$ for all $\lambda \in \mathbb{R}^d$ and

$D_x Y(0, 0) = A$. Choose $m \geq 2$. Then there exists an

open neighborhood $U \subseteq \mathbb{R}^d$ (parameter space), $0 \in U$, and a C^∞ map $\Psi: \mathbb{R}^n \times \widehat{U} \rightarrow \mathbb{R}^n$ such that $\Psi(0; \lambda) = 0$

and for any fixed $\lambda \in U$, $\Psi(\cdot, \lambda): \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a (global) diffeomorphism, and for all $\lambda \in U$, $\Psi_\lambda^* Y$, where $\Psi_\lambda(x) = \Psi(x, \lambda)$, has the Taylor expansion

$$(\Psi_\lambda^* Y)(x) = B_\lambda x + Z_\lambda^{(2)}(x) + \dots + Z_\lambda^{(m)}(x) + \dots$$

where for fixed values of $\lambda \in U$, we have

(60)

$$1) B_\lambda - A \in V^1, B_0 = A,$$

$$2) Z_\lambda^k \in V^k \text{ for all } 2 \leq k \leq m$$

So in essence, we have

$$\begin{aligned}
 (\varphi^* Y)(x) = & A x + Z_0^{(1)}(x) + Z_0^{(2)}(x) + \dots \\
 & + B_1^{(1)}(\lambda) x + Z_1^{(2)}(\lambda, x) + Z_1^{(3)}(\lambda, x) + \dots \\
 & + B_2^{(1)}(\lambda) x + Z_2^{(2)}(\lambda, x) + Z_2^{(3)}(\lambda, x) + \dots \\
 & \vdots
 \end{aligned}$$

with $\lambda \mapsto B_k(\lambda)x$ is homogeneous k th order into $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$
 $(\lambda, x) \mapsto Z_l^{(k)}(\lambda, x)$ is l th order in λ & k th order in x
 $\wedge V^k$ (with slight abuse of notation)

Proof

$$\text{While } X_\lambda = X_\lambda^{(1)}(x) + X_\lambda^{(2)}(x) + \dots$$

where $X_\lambda^{(k)} \in \mathcal{P}^k$, for fixed λ .

as before, if $\varphi_\lambda^*(x)$ is the time 1-flow for X_λ , for fixed λ , we have

$$\begin{aligned}
 (\varphi_\lambda^{1*} Y) &= Y + [X, Y] + \frac{1}{2} [X, [X, Y]] + \dots \\
 &= (A + C_\lambda)x + Y_\lambda^{(1)}(x) + Y_\lambda^{(2)}(x) + \dots \\
 &\quad + [X_\lambda^{(1)}, (A + C_\lambda)x] + [X_\lambda^{(1)}, Y_\lambda^{(1)}] + \dots \\
 &\quad + [X_\lambda^{(1)}, (A + C_\lambda)x] + [X_\lambda^{(2)}, Y_\lambda^{(1)}] + \dots \\
 &\quad + \frac{1}{2} [X_\lambda^{(1)}, [X_\lambda^{(1)}, (A + C_\lambda)x]] + \dots
 \end{aligned}$$

where $C_\lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ for fixed value of λ , and $C_0 = 0$.

We first consider the linear terms, which are of course just given by

$$⑥ A + C\lambda + [X_\lambda^{(1)}, A + C\lambda] + \frac{1}{2} [X_\lambda^{(2)}, [X_\lambda^{(1)}, A + C\lambda]] + \dots$$

Thus, we want

$$C\lambda + [X_\lambda^{(1)}, A + C\lambda] + \frac{1}{2} [X_\lambda^{(2)}, [X_\lambda^{(1)}, A + C\lambda]] + \dots$$

to lie in V' . In other words, let $\pi': P' \rightarrow P'$ be the projection onto V' along $\text{Im}(\text{ad} A|_{P'})$, then we want

$$(1 - \pi') (C\lambda + [X_\lambda^{(1)}, A + C\lambda] + \frac{1}{2} [X_\lambda^{(2)}, [X_\lambda^{(1)}, A + C\lambda]] + \dots) = 0.$$

Define $\Omega: P' \times \mathbb{R}^d \rightarrow \text{Im}(\text{ad} A|_{P'})$ by

$$\Omega(D, \lambda) = (1 - \pi') (C\lambda + [D, A + C\lambda] + \frac{1}{2} [D, [D, A + C\lambda]] + \dots)$$

$$\begin{aligned} \Omega(D, \lambda) &= (1 - \pi') (C\lambda + [D, A + C\lambda] + \frac{1}{2} [D, [D, A + C\lambda]] + \dots) \\ &= (1 - \pi') (C\lambda + [D, A + C\lambda] + \frac{1}{2} [D, [D, A + C\lambda]] + \dots). \end{aligned}$$

(using that $[D, Z] = 0$ for all $Z \in P'$)

$$\text{Then } \Omega(0, 0) = (1 - \pi')(0) = (1 - \pi')(0) = 0.$$

$$\text{and } D\Omega_{(0,0)} V = (1 - \pi') * [V, A] = [V, A].$$

Thus $D\Omega_{(0,0)}: P' \rightarrow \text{Im}(\text{ad} A|_{P'})$ is surjective, so by the implicit function theorem, we can find

a family $D(\lambda)$ such that $D(0) = 0$ and
 local $X_\lambda^{(1)} = \text{Id} + D(\lambda)$ makes sure that the linear parts at $(\varphi_i^{\lambda*} \gamma)$, for fixed λ , are given by

$$A + C\lambda + [X_\lambda^{(1)}, A + C\lambda] + \dots$$

$$= A + \tilde{B}(\lambda), \text{ with } \tilde{B}(\lambda) \in V' \text{ and } \tilde{B}(0) = 0$$

Note that D is only defined locally in $\mathbb{R}^d$,
 i.e. for some neighbourhood of 0.

The other terms go similarly.

(62) We will have to solve.

$$G^{(k)}(A + C_\lambda, Y_\lambda^{(2)}, \dots, Y_\lambda^{(k)}, X_\lambda^{(1)}, \dots, X_\lambda^{(k-1)}) - \text{ad}(A + C_\lambda)(X_\lambda^{(k)}) \in V^k,$$

So if $\pi^k : \mathcal{P}^k \rightarrow \mathcal{P}^k$ is the projection with $\text{Im}(\pi^k) = V^k$,
 $\ker(\pi^k) = \text{Im}(\text{ad}_A|_{\mathcal{P}^k})$,

we want to solve $(1 - \pi^k)(G^{(k)}(\text{---}) - \text{ad}(A + C_\lambda)(X_\lambda^{(k)})) = 0$.
as $C_0 = 0$, there is a solution for $\lambda = 0$. as

$\mathcal{P} \mapsto \text{ad}_A(\mathcal{P})$, $\mathcal{P}^k \rightarrow \text{Im}(\text{ad}_A|_{\mathcal{P}^k})$ is surjective,
we get a smooth solution $X_\lambda^{(k)}$, defined only locally
in λ .

Taking $U \subseteq \mathbb{R}^d$ the intersection of all the open neighbourhoods,
we get the result. all diffeomorphisms can be made global,
by giving X compact support. But, note that it is
local in λ . This neighborhood for λ may shrink to zero, as
 $m \rightarrow \infty$, meaning we cannot do better (except apply the
previous theorem to the system for $\lambda = 0$)

Okay. So what are these spaces V^k that we use?
there are, in general, 3 methods people use:

- 1) Based on $\ker(\text{ad}_A^T)$
- 2) Based on the Chevalley decomposition $\mathfrak{a} = \mathfrak{N} + \mathfrak{S}$
- 3) Based on $\mathfrak{sl}(2; \mathbb{C})$ triples

We start with method 1):

63) Given $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, recall that $\mathbb{R}^n = \text{Im } B \oplus \ker B^T$.

(quick prove, note that $\dim(\text{Im } B) = \text{rank } B = \text{rank}(B^T)$
 $= \dim(\text{Im } B^T)$, by

rank nullity $\dim(\text{Im } B) + \dim(\ker B) = n$, so

$$\dim(\ker B^T) = n - \dim(\text{Im } B^T) = n - \dim(\text{Im } B)$$

But let $\langle \cdot, \cdot \rangle$ be the standard inner product on $\mathbb{R}^n$. then

$$\mathbb{R}^n = \text{Im } B \oplus (\text{Im } B)^\perp \text{ \& } \dim((\text{Im } B)^\perp) = n - \dim(\text{Im } B) = \dim(\ker B^T)$$

But If $v \in \ker B^T$, let $w \in \text{Im}(B)$, say $Bu = w$.

$$\text{then } \langle v, w \rangle = \langle v, Bu \rangle = \langle B^T v, u \rangle \stackrel{v \in \ker B^T}{=} \langle 0, u \rangle = 0$$

So $\ker B^T \subseteq (\text{Im } B)^\perp$. By dimensions $(\text{Im } B)^\perp = \ker(B^T)$.

So indeed $\mathbb{R}^n = \text{Im } B \oplus \ker B^T$

Now define, on $\mathcal{P}^k$, $k \geq 1$, the inner-product

$$\left\langle v, X_1^{k_1} X_2^{k_2} \dots X_n^{k_n}, w, X_1^{l_1} X_2^{l_2} \dots X_n^{l_n} \right\rangle_k =$$

$$\begin{cases} 0 & \text{if not } k_1=l_1, k_2=l_2, \dots, k_n=l_n \\ k_1! k_2! \dots k_n! \cdot \langle v, w \rangle & \text{if } k_1=l_1, \dots, k_n=l_n \end{cases}$$

where $v, w \in \mathbb{R}^n$ and with $\langle \cdot, \cdot \rangle$ the usual inner-product on $\mathbb{R}^n$. Note that this is indeed an inner-product; an orthogonal basis set is given by

$$\left\{ e_i X_1^{k_1} X_2^{k_2} \dots X_n^{k_n}; e_i \text{ the } i\text{th basis for } \mathbb{R}^n, 1 \leq i \leq n \right. \\ \left. k_1 + \dots + k_n = k \right\}$$

⑥④ Lemma: $\langle (\text{ad } A^T) \cdot p, q \rangle_K = \langle p, (\text{ad } A) \cdot q \rangle_K$

for all $p, q \in P^K$ so $(\text{ad } A|_{P^K})^T = (\text{ad } A^T|_{P^K})$

Proof. Suppose we prove that, for any $C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$,

$\langle p, q \cdot C \rangle_K = \langle p \cdot C^T, q \rangle_K$, for all $p, q \in P^K$ ⑥

Then $\langle p, C^{-1} \cdot q \cdot C \rangle_K = \langle C^{-T} \cdot p, q \cdot C \rangle_K$

$\langle C^{-T} \cdot p \cdot C^T, q \rangle_K$, so setting

$C = e^{At}$ we get $\langle p, e^{-At} q(e^{At}x) \rangle_K$
 $= \langle e^{-A^T t} p \cdot e^{tA^T}, q \rangle_K$

Now note that $\frac{d}{dt} \big|_{t=0} e^{-At} q(e^{At}x) = Dq(x)Ax - Aq(x)$
 $= [q, A](x)$
 $= (\text{ad } A q)(x)$

So differentiating at $t=0$ gives

$\langle p, \text{ad } A(q) \rangle_K = \langle \text{ad } A^T(p), q \rangle_K$

★ Note that, by definition of $\langle \cdot, \cdot \rangle_K$, $\langle p, Cq \rangle = \langle C^T p, q \rangle$
 (check on basis elements).

Remains to prove ⑥

for two polynomial maps $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$

if $f(x) = \sum_{\sigma} a_{\sigma} x_1^{\sigma_1} \dots x_n^{\sigma_n}$, where $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathbb{Z}_{\geq 0}^n$,

we define $(f(\partial) \cdot g)(x) = \sum_{\sigma} a_{\sigma} \frac{\partial^{\sigma_1}}{\partial x_1^{\sigma_1}} \frac{\partial^{\sigma_2}}{\partial x_2^{\sigma_2}} \dots \frac{\partial^{\sigma_n}}{\partial x_n^{\sigma_n}} g$

and set $\langle f, g \rangle = (f(\partial) \cdot g)(0)$

(65) Note that

$$(X_1^{o_1} \dots X_n^{o_n}, X_1^{\tau_1} \dots X_n^{\tau_n}) = \sigma_1! \dots \sigma_n! \delta_{\sigma, \tau}$$

$$\text{where } \delta_{\sigma, \tau} = \delta_{\sigma_1, \tau_1} \delta_{\sigma_2, \tau_2} \dots \delta_{\sigma_n, \tau_n} = \begin{cases} 1 & \text{if } (o_1, \dots, o_n) = (\tau_1, \dots, \tau_n) \\ 0 & \text{else} \end{cases}$$

So if $p \in P^k$ has n components $\{e_i, p_i\}_{i=1}^n$, then

$$\langle p, q \rangle_k = \sum_{i=1}^n (p_i, q_i) = \sum_{i=1}^n (p_i(\partial), q_i(\partial))$$

Now note that

$$\frac{\partial}{\partial x_j} (q \circ A)(x) = \sum_{l=1}^n \frac{\partial q}{\partial x_l} (Ax) A_{lj} = \sum_{l=1}^n A_{jl}^T \left(\frac{\partial q}{\partial x_l} \circ A \right)$$

for any q .

$$\text{Thus } x_j(\partial)(q \circ A) = ((A^T x)_j(\partial) \cdot q) \circ A.$$

$$\text{By induction: } p_i(\partial)(q \circ A) = ((p_i \circ A^T)(\partial) \cdot q) \circ A.$$

$$\begin{aligned} \text{Thus } (p_i, q_i)_{\circ A} &= (p_i(\partial)(q \circ A))(\partial) = (p_i \circ A^T(\partial) \cdot q)(A\partial) \\ &= (p_i \circ A^T(\partial) \cdot q)(\partial) \\ &= (p_i \circ A^T, q_i) \end{aligned}$$

$$\text{Thus indeed } \langle p, q \circ A \rangle_n =$$

$$\sum_{i=1}^n (p_i, q_i \circ A) = \sum_{i=1}^n (p_i \circ A^T, q_i) = \langle p \circ A^T, q \rangle_k$$

This completes the proof. $\square$

So we can take $V^k = \ker(\text{ad } A^T|_{P^k})$. The following result gives a nice interpretation of this.

Lemma. Given $p \in P^k$, $k \geq 1$, we have $\text{ad } A^T p = 0$,

$$\text{or } p \in \ker(\text{ad } A^T|_{P^k}), \text{ iff } p \circ_e A^T = e^T \circ p.$$

(66) Proof: Recall that $\text{ad}_{A^T} p = [A^T, p]_{(x)} = \frac{d}{dt} \Big|_{t=0} e^{tA^T} p(e^{-tA^T} x)$

$$= A^T p(x) - Dp(x) A^T x$$

So if $p \circ e^{tA^T} = e^{tA^T} \circ p$, then also $e^{tA^T} p(e^{-tA^T} x) = p(x)$

So $\frac{d}{dt} \Big|_{t=0} e^{tA^T} p(e^{-tA^T} x) = 0 = [A^T, p]$

Conversely, if $[A^T, p]_{(x)} = 0$ for all x , then

$$\frac{d}{dt} \Big|_{t=0} e^{tA^T} p(e^{-tA^T} x) = 0$$

$$= \frac{d}{dt} \Big|_{t=0} e^{(s+t)A^T} p(e^{-(s+t)A^T} x) =$$

$$\frac{d}{du} \Big|_{u=0} e^{uA^T} p(e^{-uA^T} x)$$

So $e^{uA^T} p(e^{-uA^T} x)$ is constant, so equal to $p(x)$.

Thus

$$p \circ e^{-tA^T} = e^{-tA^T} \circ p, \forall t \Rightarrow p \circ e^{tA^T} = e^{tA^T} \circ p. \quad \square$$

Example: Consider the ODE on $\mathbb{R}$, $\dot{x} = ax + \mathcal{O}(|x|^2)$, for $a \in \mathbb{R}$, $a \neq 0$. Then $p^k = \langle x^k \rangle_{\mathbb{R}}$ and

$$\text{ad}_a x^k = ax \cdot kx^{k-1} - x^k \cdot a = a(k-1)x^k \neq 0 \text{ if } k \geq 2$$

So $\text{Im}(\text{ad}_a | p^k) = p^k$ for $k \geq 2$, so necessarily

we have $V^k = \{0\}$ for $k \geq 2$, so we may transform

$$\dot{x} = ax + \mathcal{O}(|x|^2) \text{ to}$$

$$\dot{y} = ay + \mathcal{O}(|y|^N) \text{ with } N \text{ arbitrarily high}$$

and even have the form $\dot{y} = ay + h(y)$

with $h(0)=0$, $\frac{dh}{dy}(0)=0$, $\frac{d^2h}{dy^2}(0)=0$, ...

(67) Note that $e^{ta^T} = e^{ta}$, and $e^{ta}|_{t=\frac{\ln(2)}{a}} = 2$.

Beut $h(2x) \neq 2h(x)$, id $h(x) = x^k$ für $k \geq 2$.

$$2^k h(x)$$

So from that we also see that $V^k = \{0\}$ for $k \geq 2$.

Example Let $A = A^T = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, then $e^{tA^T} = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{4t} \end{pmatrix}$

Write $H(x, y) = \begin{pmatrix} x^l y^{k-l} \\ 0 \end{pmatrix}$, then $H(e^{tA^T}(x, y)) = \begin{pmatrix} e^{t(2l+4(k-l))} x^l y^{k-l} \\ 0 \end{pmatrix}$

$\leftarrow (0 \leq l \leq k)$

$= e^{2t(l+2(k-l))} \begin{pmatrix} x^l y^{k-l} \\ 0 \end{pmatrix} \neq \begin{pmatrix} e^{2t} 0 \\ 0 e^{4t} \end{pmatrix} \begin{pmatrix} x^l y^{k-l} \\ 0 \end{pmatrix}$ as soon as $k \geq 2$

Now write $G(x, y) = \begin{pmatrix} 0 \\ x^l y^{k-l} \end{pmatrix}$ then

$G(e^{tA^T}(x, y)) = e^{t(2l+4(k-l))} \begin{pmatrix} 0 \\ x^0 y^{k-l} \end{pmatrix}$. This can be equal to $e^{4t} \begin{pmatrix} 0 \\ x^0 y^{k-l} \end{pmatrix}$ only if $2l+4(k-l) = 4$

So $l + 2(k-l) = 2$. This can be checked easily, as

we need $l=0, \quad k-l=2$ (not possible, as $k \geq 2$)
 $2(l-l)=?$ (not possible)

or $l=2$ $2(l-1)=2$ (not possible)

$l=1$ $2(k-1) = 1 \Rightarrow k = 1.5$ (not integer)
 $l=2$ $k-1=0 \Rightarrow k=1$ $\text{So } \begin{pmatrix} 0 \\ x^2 \end{pmatrix}$

Thus $V^2 = \left\langle \begin{pmatrix} 0 \\ x^2 \end{pmatrix} \right\rangle_{\mathbb{R}}, V^3 = \{0\}, V^4 = \{0\}, \dots$

Thus we may transform

$$x = 2x + h.o.e$$

$$\dot{y} = 4y + h.o.t.$$

t_0 (in formal power series)

$\dot{X} = 2X$

$$\dot{y} = 44 + ax^2$$

for some $a \in \mathbb{R}$.

(68) Note that $[A, G] = \begin{pmatrix} 0 & 0 \\ 2x \frac{\partial}{\partial x} X^2 - X^2 \frac{\partial}{\partial y} & (4y) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 4x^2 - 4y^2 & \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$

If $G(x) = \begin{pmatrix} 0 \\ x^2 \end{pmatrix}$

Example

$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, We get the normal forms

$$\begin{aligned} \dot{x} &= x + a_1 x^2 y + a_2 x(xy)^2 + a_3 x(xy)^3 + \dots \\ \dot{y} &= -y + b_1 xy^2 + b_2 y(xy)^2 + b_3 y(xy)^3 + \dots \end{aligned}$$

As you are looking for polynomials that commute with all $\begin{pmatrix} \theta & 0 \\ 0 & \theta^{-1} \end{pmatrix} \quad \theta \in \mathbb{R} \setminus \{0\}$

Example

$A = \begin{pmatrix} -3 & 0 \\ 0 & 3 \end{pmatrix}$

we get

$$\begin{aligned} \dot{x} &= -3x + a_1 x(x^2 y^3) + \dots \\ \dot{y} &= 3y + b_1 y(x^2 y^3) + \dots \end{aligned}$$

I think ...

We now move to Method 2, to find the V^k 's.

Recall that, given $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, there are unique N, S

$\in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ s.t. 1) $A = N + S$

2) $NS = SN$

3) S is semisimple

4) N is nilpotent

N & S are referred to as (the semisimple, nilpotent part of) the Chevalley decomposition of A - we have:

Theorem : Given the Chevalley-decomposition $A = N + S$, we have the Chevalley-decomposition

$$\text{ad } A|_{p^k} = \text{ad } N|_{p^k} + \text{ad } S|_{p^k}$$

$k \geq 2$.

69) As Proof As expected, the proof consists of 4 parts.

$$1) \text{ Clearly } \text{ad}_A(p) = [A, p] = [N+S, p] = [N, p] + [S, p] \\ = \text{ad}_N(p) + \text{ad}_S(p), \text{ for all } p \in V^k$$

$$\text{So } \text{ad}_A/p^k = \text{ad}_N/p^k + \text{ad}_S/p^k$$

2) Recall the Jacobi-identity:

$$[N, [S, p]] + [S, [p, N]] + [p, [N, S]] = 0. \quad \forall p \in p^k$$

But $[N, S] = 0$, so we get

$$[N, [S, p]] + [S, [p, N]] = 0 \quad \text{So}$$

$$[N, [S, p]] = -[S, [N, p]] = 0$$

$$[N, [S, p]] = [S, [N, p]]$$

$$\text{ad}_N(\text{ad}_S p) = \text{ad}_S(\text{ad}_N p) \quad \text{So } \text{ad}_N \circ \text{ad}_S = \text{ad}_S \circ \text{ad}_N$$

3) In general, choose any basis for $\mathbb{R}^n$, $\{e_1, \dots, e_n\}$, and define $X_i \in (\mathbb{R}^n)^*$ by $X_i(a_1 e_1 + a_2 e_2 + \dots + a_n e_n) = a_i$, for $a_1, \dots, a_n \in \mathbb{R}$

Then a basis for p^k is given by

$$\left\{ X_1^{\sigma_1} X_2^{\sigma_2} \dots X_n^{\sigma_n} \cdot e_i \mid 0 \leq \sigma_1, \dots, \sigma_n \leq k, \sigma_1 + \dots + \sigma_n = k, 1 \leq i \leq n \right\}$$

This is also a basis over $\mathbb{C}$ for $\mathbb{C} \otimes p^k =: p_{\mathbb{C}}^k$

Now we let $H_{\sigma_i}(x) = X_1^{\sigma_1} \dots X_n^{\sigma_n} \cdot e_i$

$$\text{Then } (\text{ad}_A H_{\sigma_i})(x) = \sum_{j=1}^n (Ax)_j \frac{\partial (H_{\sigma_i})_l(x)}{\partial x_j} - (H_{\sigma_i}(x))_j \frac{\partial (Ax)_l}{\partial x_j}$$

$$(70) = \sum_{j=1}^n \sum_{p=1}^n A_{jp} X_p \sigma_j X_1^{v_1} \dots X_j^{v_j-1} \dots X_n^{v_n} (e_i)_l$$

$$= \sum_{j=1}^n X_1^{v_1} \dots X_n^{v_n} (e_i)_j A_{lj}$$

$$= \sum_{j,p=1}^n A_{jp} X_p \sigma_j X_1^{v_1} \dots X_j^{v_j-1} \dots X_n^{v_n} (\delta_{il}) = \sum_{j=1}^n X_1^{v_1} \dots X_n^{v_n} A_{lj} (\delta_{ij})$$

$$= \sum_{j=1}^n \cancel{A_{jp}} \sigma_j (Ax)_j X_1^{v_1} \dots X_j^{v_j-1} \dots X_n^{v_n} \delta_{il} = A_{li} X_1^{v_1} \dots X_n^{v_n}$$

$$\text{Thus } (\text{ad} A H_{\alpha_i})(x) = \left(\sum_{j=1}^n \sigma_j (Ax)_j X_1^{v_1} \dots X_j^{v_j-1} \dots X_n^{v_n} \right) \cdot e_i$$

$$= (X_1^{v_1} \dots X_n^{v_n}) \cdot (A e_i)$$

• Note as well the structure of $\text{ad} A$: we may write $P^k = \text{Sym}(\mathbb{R}^{n*} \otimes \dots \otimes \mathbb{R}^{n*}) \otimes \mathbb{R}^n$, where

$\text{Sym}(\mathbb{R}^{n*} \otimes \dots \otimes \mathbb{R}^{n*})$ ^{k times} denotes polynomials from $\mathbb{R}^n$ to $\mathbb{R}$ and " $\otimes \mathbb{R}^n$ " is there because we have n of these

Then we may write $\text{ad} A = \mathbb{Q} A \otimes \text{Id} - \text{Id} \otimes A$

where $\mathbb{Q} A: \text{Sym}(\mathbb{R}^{n*} \otimes \dots \otimes \mathbb{R}^{n*}) \rightarrow$ is defined on basis elements by

$$\mathbb{Q} A (X_1^{v_1} \dots X_n^{v_n}) = \sum_{j=1}^n \sigma_j (Ax)_j X_1^{v_1} \dots X_j^{v_j-1} \dots X_n^{v_n}$$

Now plug in first $A = S$ in the above. assume

$\{e_1, \dots, e_n\} \in \mathbb{C}^n$ is a basis of eigenvectors for S

say $S e_j = \lambda_j e_j$ for all $1 \leq j \leq n$

⑦) Then by the above

$$(\text{ad}_S H_{0,i})(x) = \left(\sum_{j=1}^n \sigma_j (Sx)_j X_1^{\sigma_1} \dots X_j^{\sigma_j-1} \dots X_n^{\sigma_n} \right) \cdot e_i$$

$$= (X_1^{\sigma_1} \dots X_n^{\sigma_n}) \cdot (\sum_{j=1}^n \sigma_j e_j)$$

$$= \left(\sum_{j=1}^n \sigma_j \lambda_j X_j X_1^{\sigma_1} \dots X_j^{\sigma_j-1} \dots X_n^{\sigma_n} \right) \cdot e_i$$

$$= (X_1^{\sigma_1} \dots X_n^{\sigma_n}) (\lambda_i e_i)$$

$$= \left(\sum_{j=1}^n \sigma_j \lambda_j (X_1^{\sigma_1} \dots X_n^{\sigma_n}) \right) \cdot e_i = \left(\left(\sum_{j=1}^n \sigma_j \lambda_j \right) - \lambda_i \right) H_{0,i}(x)$$

$$= \lambda_i (X_1^{\sigma_1} \dots X_n^{\sigma_n}) \cdot e_i$$

So $\text{ad}_S|_{p^k}$ is also semisimple: with basis of eigenvectors the $H_{0,i}(x)$ and eigenvalues these

"resonances" $\left(\sum_{j=1}^n \sigma_j \lambda_j - \lambda_i \right)$

4) — Now instead use $A=N$. Suppose we prove that Q_N is nilpotent. Say $Q_N^M = 0$, if $N^K = 0$, then ad_N is nilpotent, as

$$\text{ad}_N = Q_N \otimes \text{Id} - \text{Id} \otimes N, \text{ so}$$

$$\text{ad}_N^{M+K} = \text{" terms } Q_N^{P_1} \otimes N^{P_2} \text{ with either } P_2 \geq K \text{ or } P_1 \geq M \text{"}$$

$$= \text{" terms } 0 \otimes N^P + Q^P \otimes 0 \text{"} = 0.$$

To see that Q_N is nilpotent, choose a basis

92) $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$ such that $Ne_i = K_i e_{i+1}$ for $1 \leq i < n$, and with $K_i \in \{0, 1\}$, and $Ne_n = 0$.
So we are writing

$$N = \begin{pmatrix} 0 & & & & 0 \\ 1 & 0 & & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & 0 \\ & & & & \ddots & \ddots & \ddots & \ddots & 0 \end{pmatrix}$$

$\nwarrow$ the K_i

Now order the $\sigma = (\sigma_1, \dots, \sigma_n)$

such that $\sigma \succ \tau = (\tau_1, \dots, \tau_n)$ if $\sigma_1 > \tau_1$ or
 $\sigma_1 = \tau_1$ and $\sigma_2 > \tau_2$ or
 $\sigma_1 = \tau_1, \sigma_2 = \tau_2, \sigma_3 > \tau_3$
 or $\dots$

(So Lexicographically)

$$\sigma_1 = \tau_1, \sigma_2 = \tau_2, \dots, \sigma_{n-1} = \tau_{n-1}, \sigma_n > \tau_n$$

write $E_\sigma = X_1^{\sigma_1} \dots X_n^{\sigma_n} \in \text{Sym}(\mathbb{R}^{n*} \otimes \dots \otimes \mathbb{R}^{n*})$

$$\begin{aligned} \text{Then } Q_N(E_\sigma) &= \sum_{j=1}^n \alpha_j(Nx) X_1^{\sigma_1} \dots X_j^{\sigma_j-1} X_{j+1}^{\sigma_{j+1}} \dots X_n^{\sigma_n} \\ &= \sum_{j=1}^n \alpha_j \hat{K}_j X_{j-1}^{\sigma_1} \dots X_j^{\sigma_j-1} \dots X_n^{\sigma_n} \end{aligned}$$

$$\text{for some } \hat{K}_j \in \{0, 1\} \rightarrow \sum_{j=1}^n \alpha_j \hat{K}_j X_1^{\sigma_1} X_2^{\sigma_2} \dots X_{j-1}^{\sigma_{j-1}+1} X_j^{\sigma_j-1} \dots X_n^{\sigma_n}$$

= "expression of E_τ , with $\tau \succ \sigma$ "

Thus Q_N is in strict lower triangular form with respect to this ordered basis $\{E_\sigma\}$. So Q_N is indeed nilpotent. Thus so is $\text{ad}_N|_{p^k}$

Combining 1), 2), 3) & 4) with the fact that Chevalley decomposition is unique, we see that this decomposition is indeed given by $\text{ad}_A|_{p^k} = \text{ad}_N|_{p^k} + \text{ad}_S|_{p^k} \square$

(73) Here are two useful conclusions we may draw from this:

Corollary 1: If A is semisimple (diagonalisable over $\mathbb{C}$), then the Jordan-Chevalley decomposition of A is just

$$A = \underset{\substack{\uparrow \\ \text{semisimple}}}{A} + \underset{\substack{\nwarrow \\ \text{nilpotent}}}{O}$$

Thus the Jordan-Chevalley decomposition of $\text{ad} A|_{\mathcal{P}^k}$ is given by

$$\text{ad} A|_{\mathcal{P}^k} = \underset{\substack{\uparrow \\ \text{semisimple}}}{\text{ad} A|_{\mathcal{P}^k}} + \underset{\substack{\uparrow \\ \text{nilpotent}}}{\text{ad} O|_{\mathcal{P}^k}}$$

$\text{ad} O|_{\mathcal{P}^k} \equiv 0$.

Thus $\text{ad} A|_{\mathcal{P}^k}$ is itself semisimple, for any $k \geq 1$.

for semisimple maps $S: \mathbb{R}^N \rightarrow \mathbb{R}^N$, we have $\mathbb{R}^N = \text{Im}(S) \oplus \text{ker}(S)$.

Thus $\mathcal{P}^k = \text{Im}(\text{ad} A|_{\mathcal{P}^k}) \oplus \text{ker}(\text{ad} A|_{\mathcal{P}^k})$ in this case.

So we may choose $V^{(k)} = \text{ker}(\text{ad} A|_{\mathcal{P}^k})$. As before, for $p \in \mathcal{P}^k$

$p \in \text{ker}(\text{ad} A|_{\mathcal{P}^k}) \Leftrightarrow p \cdot e^{tA} = e^{tA} \cdot p \quad \forall t \in \mathbb{R}$. Note that also $A e^{tA} = e^{tA} A$, so the whole truncated normal form may be arranged to commute with e^{tA} .

Ex: take $A = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$, $\omega \in \mathbb{R}$, so $e^{tA} = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}$.

identify $\mathbb{R}^2 \cong \mathbb{C}$, then applying e^{tA} can be identified by multiplying by the complex number $\cos(\omega t) + i \sin(\omega t)$ to we need elements of $\mathcal{P}^k$ to commute, in this way, with any element in $\mathbb{C}$ with norm 1. (call this set $\mathbb{C}^*$) thus, given $\lambda \in \mathbb{C}^*$, we need $p(\lambda z) = \lambda p(z)$.

Note that for $\lambda \in \mathbb{C}^*$, $\bar{\lambda} = \frac{\lambda}{\lambda \bar{\lambda}} = \frac{1}{\lambda}$. So if $p(z) = z^{k-l} \bar{z}^l$ we have

$$\begin{aligned} p(\lambda z) &= (\lambda z)^{k-l} (\overline{\lambda z})^l = \lambda^{k-l} \bar{\lambda}^l z^{k-l} \bar{z}^l \\ &= \lambda^{k-l} \frac{1}{\lambda^l} z^{k-l} \bar{z}^l = \lambda^{k-l-l} z^{k-l} \bar{z}^l = \lambda^{k-2l} z^{k-l} \bar{z}^l \end{aligned}$$

74) Thus we need $\lambda(k-2l) = 1$ for all $\lambda \in \mathbb{C}^*$, so
 $k-2l = 1$, $l = \frac{k-1}{2}$.

This is not possible if k is even, so $V^k = \{0\}$ for k even,

and we also see that $V^k = \left\langle z^{\frac{k+1}{2}} \bar{z}^{\frac{k-1}{2}} \right\rangle_{\mathbb{R}}$
 $= \left\langle z(z\bar{z})^{\frac{k-1}{2}} \right\rangle_{\mathbb{R}}$
 $= \left\langle z |z|^{k-1} \right\rangle_{\mathbb{R}}$

So a normal form looks like

$$g(z) = \underbrace{-i\omega z}_{\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}} + \beta_1 z|z|^2 + \beta_2 z|z|^4 + \beta_3 z|z|^6 + \dots$$

with $\beta_1, \beta_2, \dots \in \mathbb{C}$

Corollary 2. In general, If A is not semisimple, we may still decompose $A = S + N$, (S semisimple, N nilpotent, $SN = NS$)

As S is semisimple, so is $\text{ad}_S|_{\mathfrak{p}^k}$, thus $\forall k \geq 1$,

$$\mathfrak{p}^k = \text{Im}(\text{ad}_S|_{\mathfrak{p}^k}) \oplus \ker(\text{ad}_S|_{\mathfrak{p}^k}). \text{ Now, If } B \text{ is}$$

any matrix with Jordan-Chevalley decomposition $B = S_B + N_B$, it is not hard to see that $\text{Im}(\text{ad}_S) \subseteq \text{Im}(B)$.

for example, if we put B & S_B in Jordan-normal form simultaneously.

$$B \sim \left(\begin{array}{c|c} \begin{matrix} \diagup & & \\ 0 & 1 & 0 \\ \diagdown & & \end{matrix} & 0 \\ \hline 0 & \begin{matrix} \lambda \neq 0 & 1 & 0 \\ & \ddots & \ddots \end{matrix} \end{array} \right), \quad S_B = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \begin{matrix} \lambda \neq 0 & 0 \\ & \diagdown \end{matrix} \end{array} \right)$$

(75) We see that $\text{Im}(S_\beta) \subseteq \text{Im}(B)$, (sometimes strict!)
applying this to $B = \text{ad} A|_{p^k}$, we get

$$\text{Im}(\text{ad} S|_{p^k}) \subseteq \text{Im}(\text{ad} A|_{p^k}). \text{ Thus}$$

$$p^k = \text{Im}(\text{ad} A|_{p^k}) + \ker(\text{ad} S|_{p^k}).$$

Thus we may choose $V^k \subseteq \ker(\text{ad} S|_{p^k})$ (maybe strict)
such that

$$p^k = \text{Im}(\text{ad} A|_{p^k}) \oplus V^k.$$

$$\text{Now we have } q \in V^k \Rightarrow q \cdot e^t S = e^t S \cdot q \quad \forall t \in \mathbb{R}.$$

Note that as well, $AS = (N+S)S = NS + S^2 = SN + S^2 = S(N+S) = SA$.
So also $Ae^{tS} = e^{tS}A \quad \forall t \in \mathbb{R}$. Thus we get the

important Observation:

Normal Form reduction can always be performed in such a way that the (truncated) normal form commutes with e^{tS} , with S the semisimple part of $A = Df(0)$

This is very useful. For one, note that the set of semisimple matrices is dense in the set of all $(n \times n)$ matrices.

However, for nilpotent matrices $A = N$ it is not so useful.

For, if A is nilpotent, we have in the Jordan-Chevalley decomp.

$$A = N, S = 0. \text{ Thus } \text{ad} S|_{p^k} \equiv 0, \text{ so } \ker(\text{ad} S|_{p^k}) = p^k.$$

Put differently, $e^{tS} = e^0 = \text{Id}$. So we learn nothing;

$\ker(\text{ad} S|_{p^k})$ is way too big!

This leads us to method #3, Based on $\mathfrak{sl}(2; \mathbb{R})$ -triples

76 Method 3 (briefly). Suppose now $A=N$ is nilpotent

Definition An $\mathfrak{sl}_2$ -triple is a triple of ~~matrices~~ matrices $(n \times n)$

$\{e, h, f\}$, satisfying

$$[e, f] = h$$

$$[h, e] = 2e$$

$$[h, f] = -2f$$

Example take $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

Note that $\langle h, e, f \rangle_{\mathbb{R}} = \left\{ 2 \times 2 \text{ matrices } A \text{ s.t. } \text{Tr}(A) = 0 \right\}$
 $=: \mathfrak{sl}(2, \mathbb{R}) (= \mathfrak{sl}_2)$

Thm. (Jacobson-Morozov). Let N be a nilpotent $n \times n$ matrix, there exist $n \times n$ matrices M & H such that $h=H$, $e=N$ & $f=M$ form an $\mathfrak{sl}_2$ -triple

Thm. if $\{N, H, M\}$ is an $\mathfrak{sl}_2$ -triple, then so is

$$\{ad_N|_{p^k}, ad_H|_{p^k}, ad_M|_{p^k}\}.$$

$$= [ad_N, ad_M](p)$$

Proof: $ad_N \circ ad_M(p) - ad_M \circ ad_N(p) \leftarrow$

$$= [N, [M, p]] - [M, [N, p]]$$

$$= [N, [M, p]] + [M, [p, N]]$$

$$= -[p, [N, M]] = [[N, M], p] = [H, p] = ad_H(p).$$

Jacobi-id.

The other cases are similar.

□.

Now with some more Lie-algebra manipulations, it can be shown

that $p^k = \text{Im}(ad_N|_{p^k}) \oplus \text{ker}(ad_M|_{p^k})$. using this fact for general $\mathfrak{sl}_2$ -triples (and the fact that $\{ad_N|_{p^k}, \dots, ad_M|_{p^k}\}$ is such an $\mathfrak{sl}_2$ -triple)

The Symmetry Lectures

Def: (Informal) A symmetry of a dynamical system is a transformation of the phase space (a map of the phase space to itself) that sends solutions to solutions.

Def: (In our setting): Let $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a (smooth) vectorfield, and $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ a diffeomorphism. Then Φ is a symmetry of the ODE $\dot{x} = f(x)$ if, whenever $t \mapsto \gamma(t)$ is a solution curve, so is $t \mapsto \Phi \gamma(t)$. Thus $\Phi \gamma_t(x) = \gamma_t(\Phi x)$.

Lemma: Let f & Φ as above, then Φ is a symmetry of $\dot{x} = f(x)$, if and only if, for all $x \in \mathbb{R}^n$, $D\Phi(x)f(x) = f(\Phi(x))$.

Proof. Suppose Φ is a symmetry, ^{fix $x \in \mathbb{R}^n$} and let $\gamma(t)$ be a (local) solution, so $\frac{d}{dt} \gamma(t) = f(\gamma(t))$, satisfying $\gamma(0) = x$.

Then $\Phi \gamma(t)$ is a solution too, so

$$\left. \frac{d}{dt} \right|_{t=0} \Phi \gamma(t) = D\Phi(\gamma(0)) \cdot \dot{\gamma}(0) = D\Phi(x) f(\gamma(0)) = D\Phi(x) f(x)$$

$\parallel$ solution

$$f(\Phi \gamma(t)) \Big|_{t=0} = f(\Phi \gamma(0)) = f(\Phi(x)).$$

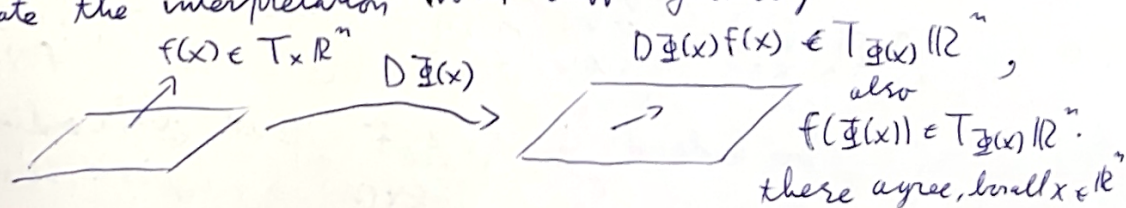
Now suppose $D\Phi(x)f(x) = f(\Phi(x)) \quad \forall x \in \mathbb{R}^n$. Let γ be a solution.

$$\text{then } \frac{d}{dt} \Phi \gamma(t) = D\Phi(\gamma(t)) \dot{\gamma}(t) = \underbrace{D\Phi(\gamma(t))}_\gamma \underbrace{f(\gamma(t))}_\gamma \quad \text{if } \gamma = \gamma(t).$$

$$= D\Phi(\gamma) f(\gamma) = f(\Phi(\gamma)) = f(\Phi \gamma(t)) \quad \square.$$

So $\Phi \gamma$ is a solution too.

Remark: Note the interpretation from diff. geometry:



(78) Remark: If Φ is linear, so $\Phi(x) = Ax$ for some $A \in \text{Gl}(n, \mathbb{R})$.
 then Φ is a symmetry of $\dot{x} = f(x)$, if and only if

$$Af(x) = f(Ax) \quad \forall x \in \mathbb{R}^n, \quad \text{so if and only if}$$

$$A \circ f = f \circ A :$$

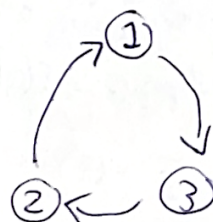
$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^n \\ \downarrow A & & \downarrow A \\ \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^n \end{array}$$

Why do we care about the (very specific?) situation where Φ is linear?

Reason #1: "Structurally induced" symmetry:

Consider 3 identical dynamical systems coupled together in a "triangle configuration"

$$\begin{aligned} \dot{x}_1 &= f(x_1) + h(x_2 - x_1) \\ \dot{x}_2 &= f(x_2) + h(x_3 - x_2) \\ \dot{x}_3 &= f(x_3) + h(x_1 - x_3) \end{aligned}$$



So $f: \mathbb{R} \rightarrow \mathbb{R}$ is (identical)

internal dynamics, and $h: \mathbb{R} \rightarrow \mathbb{R}$ is the coupling function.

Denote the vectorfield by $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, so

$$F \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} f(x_1) + h(x_2 - x_1) \\ f(x_2) + h(x_3 - x_2) \\ f(x_3) + h(x_1 - x_3) \end{pmatrix} \quad \text{Moreover, define } A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

so that $A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_1 \end{pmatrix}$ Now note that

$$F \circ A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = F \begin{pmatrix} x_2 \\ x_3 \\ x_1 \end{pmatrix} = \begin{pmatrix} f(x_2) + h(x_3 - x_2) \\ f(x_3) + h(x_1 - x_3) \\ f(x_1) + h(x_2 - x_1) \end{pmatrix}, \quad \text{while}$$

$$A F \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} f(x_1) + h(x_2 - x_1) \\ f(x_2) + h(x_3 - x_2) \\ f(x_3) + h(x_1 - x_3) \end{pmatrix} = \begin{pmatrix} f(x_2) + h(x_3 - x_2) \\ f(x_3) + h(x_1 - x_3) \\ f(x_1) + h(x_2 - x_1) \end{pmatrix}$$

So indeed $F \circ A = A \circ F$, put differently, if $(x_1(t), x_2(t), x_3(t))$ is a solution, then so is $(x_2(t), x_3(t), x_1(t))$

79) Note that by applying the result again, we see that $A \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = (x_3(t), x_1(t), x_2(t))$ is also a solution. This is part of a more general observation:

Lem. Given an ODE $\dot{x} = f(x)$ on $\mathbb{R}^n$, the set of symmetries as defined above forms a group under composition.

Proof. Clearly $\Phi(x) = x$ (so $\Phi = \text{Id}$) is a symmetry

• If Φ_1 & Φ_2 send solutions to solutions, then so does $\Phi_1 \circ \Phi_2$

• If Φ is a symmetry, then $D\Phi(x)f(x) = f(\Phi(x))$ for all $x \in \mathbb{R}^n$. In particular, for all $y \in \mathbb{R}^n$, we have (setting $x = \Phi^{-1}(y)$)

$$D\Phi(\Phi^{-1}(y))f(\Phi^{-1}(y)) = f(\Phi(\Phi^{-1}(y))) = f(y).$$

Now apply $D\Phi^{-1}(y)$ to both sides; to get

$$\begin{aligned} D\Phi^{-1}(y)f(y) &= \underbrace{D\Phi^{-1}(y)D\Phi(\Phi^{-1}(y))}_{\text{Id}} f(\Phi^{-1}(y)) \\ &= f(\Phi^{-1}(y)) \end{aligned}$$

So Φ^{-1} is also a symmetry. So indeed, these Φ 's form a group. $\square$

In particular: the set of linear symmetries form a group. Many important groups of matrices form so-called Lie-groups:

Def: A Lie group G is a (smooth) manifold G , together with a group structure " $\cdot$ ", such that the operations

- 1) " $\cdot$ " $G \times G \rightarrow G$, $(g, h) \mapsto gh$
- 2) " $^{-1}$ " $G \rightarrow G$, $g \mapsto g^{-1}$ are both smooth

(90) Property 2) is actually redundant, for the following reason:
 denote, for a given $g \in G$, by $L_g: G \rightarrow G$ the operation that is
 left multiplication by g , so $L_g(h) = gh$. Note that
 $L_{g^{-1}}(g) = e$, for any $g \in G$. then, Consider the map
 $F: G \times G \rightarrow G$, $F(g, h) = gh$ and take the derivative in the
 h -direction at (g^{-1}, g) $= L_g(h)$

$$D_h F(g^{-1}, g) = D L_{g^{-1}}(g) \quad \text{Now note that } L_{g^{-1}} \circ L_g = \text{Id},$$

so $D L_{g^{-1}}(g) L_g(e) = \text{Id}$, so $D L_{g^{-1}}(g)$ is invertible. Thus by
 the implicit function theorem, so is a smooth map

$\mu: G \rightarrow G$ (but locally defined), such that ~~that~~

μ locally solves $\{(g, h) \in G \times G, F(g, h) = e\}$

$= \{(g, \mu(g))\}$ and such that $\mu(g^{-1}) = g$.

Of course $\mu(g) = g^{-1}$. so the inverse is smooth around (g^{-1}, g) .

But g was chosen arbitrarily, so the inverse is indeed
 smooth. □

Some examples of Lie groups: (say $n \geq 1$)

1) $GL(n, \mathbb{R})$ "the invertible matrices"

2) $SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}); \det(A) = 1\}$

3) $O(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}); A^{-1} = A^T\}$

4) $SO(n, \mathbb{R}) = O(n, \mathbb{R}) \cap SL(n, \mathbb{R})$

5) $Sp(2n, \mathbb{R}) = \{A \in GL(2n, \mathbb{R}), A^T J A = J\}$

where $J = \begin{pmatrix} 0 & \text{Id}_n \\ -\text{Id}_n & 0 \end{pmatrix}$

6) $GL(n, \mathbb{C})$.

7) $\{a+bi+cj+dk \mid (a+bi+cj+dk)(a-bi-dj-dk) = 1\}$ (unit quaternions)

⑧ It's often useful when a Lie group is compact (as a topological set) this is the case for $3), 4), 7)$ (& 2) only when $n=1$).

Motivated by the above, we now take another approach.

We now fix a group G . to each element $g \in G$ we associate a diffeomorphism Φ_g in such a way that for $g, h \in G$, we have

$$\Phi_g \circ \Phi_h = \Phi_{gh} \quad \& \quad \Phi_e = \text{Id}_{\mathbb{R}^n}, \quad \text{where } e \text{ is the unit of } G.$$

Note that $\Phi_g \circ \Phi_g^{-1} = \Phi_{gg^{-1}} = \Phi_e = \text{Id}$, so $\Phi_g^{-1} = \Phi_{g^{-1}}$

Suppose now that $p \in \mathbb{R}^n$ is a fixed point for all the Φ_g :

$\Phi_g(p) = p$ for all $g \in G$. Then the linear maps $D\Phi_g(p): \mathbb{R}^n \rightarrow \mathbb{R}^n$ behave similarly to the (potentially non-linear) Φ_g . That is, we have $D\Phi_{gh}(p) = D(\Phi_g \circ \Phi_h)(p) = D\Phi_g(p) \cdot D\Phi_h(p) \quad \forall g, h \in G.$

and also $D\Phi_e(p) = \text{Id}$.

We now have: ~~Suppose~~

Thm: Bochner Linearisation Theorem: Suppose G is a compact Lie-group (or for instance, G is finite). Suppose the map $\Phi: G \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(g, x) \mapsto \Phi_g(x)$ is differentiable ~~smooth~~, then locally around p there is a near-identity (+translation) map for coordinate transformation to a neighbourhood of 0, under which Φ transforms to $\hat{\Phi}$, $(g, x) \mapsto D\Phi_g(p)$. So all the maps become linear.

Motivated by the above, we will from here on out only look at the case where the symmetries are by linear (invertible) maps:

(82) Def. Let G be a group and V a finite dimensional (real or complex) vector space. A representation of G on V is a map

$$A: G \rightarrow GL(V) \hookrightarrow \{ \text{linear invertible maps from } V \text{ to } V \}$$

Satisfying 1) $A(g)A(h) = A(gh) \quad \forall g, h \in G'$

2) $A(e) = \text{Id}_V$ für e unitäres Element ab G

as before, we also get $A(g') = (A(g))^{-1} \quad \forall g \in G$. (we often write $A_g := A(g)$)

Def: A linear map $B: V \rightarrow V$ is called

equivariant if $B \circ A_g = A_g \circ B$ for all $g \in G$.

Example: Let $\dot{x} = f(x)$ be an ODE on V . Assume $f(0) = 0$ and suppose $\{A_g\}_{g \in G}$ is a representation of a group G on V , ~~such that~~ A_g is symmetric A_g at $\dot{x} = f(x)$. Then $DF(0)$ is an equivariant map on V .

Proof. by our previous lemma: $f(A_g x) = A_g f(x) \quad \forall x \in V, \forall g \in G$.
differentiate w.r.t. x : $Df(A_g x) A_g = A_g Df(x)$.
if $x \in \text{Fix}(A_g)$ then $A_g x = x$ (so $x \in \text{Fix}(A_g)$) $\forall g \in G$.

differentiate w.r.t. x . $Df(A_g x) \cdot A_g$
Now if $x \in V$ is such that $A_g x = x$ (so $x \in \text{Fix}(A_g)$) $\forall g \in G$.
(for instance $x=0$), we get $Df(x) \cdot A_g = A_g \cdot Df(x)$.

So indeed $Df(0) \cdot A_g = A_g \cdot Df(0)$

Def. A subrepresentation of the representation $(V, \{A_g\}_{g \in G})$ is a linear subspace $u \subseteq V$ s.t. $A_g(u) \subseteq u \quad \forall g \in G$.

Note that a subrepresentation $U \subseteq V$ gives rise to its own representation $(U, \{A_g|_U\}_{g \in G})$.

also, if $U_1, U_2, \dots, U_k \in V$ are subrepresentations of V
then so is $U_1 + U_2 + \dots + U_k$.

A special case: if $V = U_1 \oplus U_2 \oplus U_3 \oplus \dots \oplus U_k$
and the $U_1, \dots, U_k$ are subrepresentations of V
then we say $V = U_1 \oplus \dots \oplus U_k$ as representations of G

⑧ Example: Let B be an equivariant, ^{linear} map from V to V and λ an ^{real} eigenvalue of B then

$$E_\lambda = \text{span} \{ v \in V \mid Bv = \lambda v \} \subseteq V \quad \&$$

$$\hat{E}_\lambda = \text{span} \{ \text{generalised eigenvectors of } B \text{ for } \lambda \} \subseteq V$$

are subrepresentations of V . The same holds for the obvious generalisations to complex eigenvalues

Proof (Sketch). If $\lambda \in \mathbb{R}$ and $Bv = \lambda v$, then for any $g \in G$

$$B(Ag v) = Ag Bv = Ag(\lambda v) = \lambda(Ag v), \text{ so}$$

$Ag v$ is an eigenvector for λ as well.

if $\lambda \in \mathbb{C}$, suppose $v \in \ker((B - \lambda \text{Id})(B - \bar{\lambda} \text{Id}))$

$$\text{then } (B - \lambda \text{Id})(B - \bar{\lambda} \text{Id})v = (B^2 - 2\text{Re}(\lambda)B + |\lambda|^2 \text{Id})v = 0$$

$$\begin{aligned} \text{But then } (B^2 - 2\text{Re}(\lambda)B + |\lambda|^2 \text{Id})(Ag v) &\stackrel{\text{Note}}{=} B^2 Ag v \\ &= Ag(B^2 - 2\text{Re}(\lambda)B + |\lambda|^2 \text{Id})v = 0. \end{aligned} \quad \begin{aligned} &= B^2 Ag v \\ &= Ag B^2 v \end{aligned}$$

$\forall g \in G$. For $\hat{E}_\lambda$, use the same argument, but for $((B - \lambda \text{Id})(B - \bar{\lambda} \text{Id}))^{\dim(V)} : V \rightarrow V$.

Thus we get, as representations of G :

$$V = \bigoplus_{\lambda \text{ eigenvalue of } B/v} \hat{E}_\lambda \quad (\text{where we identify } \lambda \text{ with } \bar{\lambda} \text{ under } v)$$

This is one way in which symmetry forces complicated bifurcation scenarios: the (generalised) kernel and center subspaces are sometimes forced to become high-dimensional

⑧4 Example. Let $G = \mathbb{Z}/4\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$.

$$A_{\bar{1}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad A_{\bar{2}} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_{\bar{3}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad A_{\bar{0}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ on } \mathbb{R}^2$$

$$\dot{\underline{X}} = F(\underline{X}), \quad \underline{X} = (x, y)^T, \quad F(x, y) = \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}.$$

$$A_{\bar{1}} F(\underline{X}) = F(A_{\bar{1}} \underline{X}) \quad \text{gives} \quad \begin{pmatrix} -g(x, y) \\ f(x, y) \end{pmatrix} = \begin{pmatrix} f(-y, x) \\ g(-y, x) \end{pmatrix}.$$

$$\text{So } g(x, y) = -f(-y, x), \text{ while } f(x, y) = -f(-x, -y).$$

$$\text{Thus } F(x, y) = \begin{pmatrix} f(x, y) \\ f(y, -x) \end{pmatrix} \text{ with } f(-x, -y) = -f(x, y).$$

$$(\text{Check: } F(-y, x) = \begin{pmatrix} f(-y, x) \\ f(x, y) \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} F(x, y) = \begin{pmatrix} -f(y, -x) \\ f(x, y) \end{pmatrix} = \begin{pmatrix} f(-y, x) \\ f(x, y) \end{pmatrix})$$

$$\text{So if } \frac{\partial f}{\partial x}(0,0) = a, \quad \frac{\partial f}{\partial y}(0,0) = b, \text{ we have } DF(0,0) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

$$\text{Now, } \det DF(0,0) = a^2 + b^2, \text{ so } \det DF(0,0) = 0 \iff DF(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\text{Thus } \ker DF(0,0) = \{0\} \text{ or } \mathbb{R}^2$$

In fact, we never expect a non-trivial kernel in 1-parameter bifurcation, as we need both $a=0$ & $b=0$

Another way to see that bifurcations tend to become more intricate: if $\gamma(t)$ is a bounded solution

(e.g. periodic or steady state) so is $A_g \gamma(t) \forall g \in G$.

e.g. if in $\dot{x} = f(x)$ we have $f(x_0) = 0$, then

$$f(A_g x_0) = A_g f(x_0) = A_g 0 = 0. \quad \forall g \in G$$

So $\{A_g x_0\}_{g \in G}$ is a set of steady states (not necessarily all different)