Universidad Mayor de San Andrés
Facultad de Ciencias Puras y Naturales
Carrera de Matemática

Controlabilidad de Sistemas Bilineales en el Plano

(Teoría de Control)

Autor: Univ. Victor Hugo Patty Yujra

Tutor: Dr. Efraín Cruz Mullisaca

Tesis de grado presentada para la obtención del título de licenciatura en matemática

La Paz-Bolivia

2010

CONTROLABILIDAD DE SISTEMAS BILINEALES EN EL PLANO

(Teoría de Control)

Por

Victor Hugo Patty Yujra

Remitido en cumplimiento parcial de los requisitos para la obtención del grado de licenciado en matemática en la universidad mayor de san andrés.

La Paz, Bolivia

Junio de 2010

A mis queridos padres,
Dionicio y Andrea.

Índice general

Agradecimientos

Quiero agradecer al profesor Efraín Cruz, mi profesor tutor por todas sus sugerencias y apoyo durante la elaboración de este trabajo, del mismo modo quiero agradecer al Seminario de Teoría de Control de la carrera de Matemática que esta a cargo del Prof. Efraín Cruz, por brindarme la oportunidad de poder desarrollarme matemáticamente. Expresar también mi gratitud al Mgr. Miguel Yucra, Mgr. Willy Condori y Mgr. Hans Nina que siempre confiaron en mi. Y claro, a quienes de una u otra forma contribuyeron en mi formación matemática, a mis profesores.

Por supuesto, este trabajo nunca hubiera sido posible realizar sin el apoyo, paciencia, confianza y amor de mi familia, muy especialmente de mis padres *Dionicio y Andrea*, por quienes mi deuda siempre será infinita.

Quiero agradecer a todos mis amigos, que siempre confiaron en mi Gustavo, Pamela, Iver, Hugo, Ramiro, Freddy, Alizon y tantos otros que sería imposible citarlos aquí pero que jamas olvidaré.

Me gustaría terminar con un agradecimiento muy especial a mis hermanos Lourdes y Wilson por ser más que hermanos, mis mejores amigos. A Claudia por todos los bellos momentos que pasamos juntos, por brindarme su comprensión, apoyo y haberme dado una razón para ser mejor cada día.

Victor H. Patty

Junio, 2010.

Introducción

El problema de controlabilidad de sistemas de control es uno de los problemas abiertos en la teoría geométrica de sistemas de control. Para sistemas lineales, existe un criterio muy simple, que es la condición de rango de Kalman, para sistemas no lineales existen solo en algunas situaciones donde es muy simple verificar condiciones necesarias y suficientes. Al mismo tiempo hacemos mención de los resultados clásicos obtenidos por Jurdjevic y Sussmann, acerca de sistemas de control invariantes a derecha sobre grupos de Lie compactos. Para estos sistemas la condición de rango de álgebras de Lie (LARC) es una condición necesaria y suficiente en el caso en que el sistema de control es analítico. Con el desarrollo de la teoría geométrica de los semigrupos generados por el sistema de control, se obtienen nuevos métodos en el estudio de la controlabilidad de sistemas invariantes sobre grupos de Lie. Uno de esos métodos es utilizado por Hilgert, Hofmann y Lawson, estudiando controlabilidad sobre grupos de Lie solubles y nilpotentes, obteniendo condiciones necesarias y suficientes en estos casos. Controlabilidad sobre grupos nilpotentes es estudiado también por Ayala [1], con diferentes métodos.

El propósito del presente trabajo es realizar el análisis detallado de la controlabilidad de sistemas bilineales de control definidos en el plano $\mathbf{R}^2$, es decir, los sistemas de control de la forma

$$\dot{x} = Ax + uBx \tag{1}$$

con controles admisibles $u \in \mathbf{R}$ (localmente integrables). Aquí $x \in \mathbf{R}^2 - \{0\}$, y A, B son matrices de 2×2. Si consideramos controles constantes por pedazos, entonces tenemos que (1) es controlable si y solo si el semigrupo

$$S_\Sigma = \{e^{t_1 X_1} \cdots e^{t_k X_k} : t_i \geq 0, X_i = A, \pm B, k > 0\}$$

generado por $\Sigma = \{A, \pm B\}$ es transitivo sobre $\mathbf{R}^2 - \{0\}$, en este caso tenemos que dados $x, y \in \mathbf{R}^2 - \{0\}$ existe un $g \in S_\Sigma$ tal que $gx = y$.

Ahora recordemos también que una condición necesaria para la controlabilidad del sistema, es que el grupo

$$G_\Sigma = \{e^{t_1 X_1} \cdots e^{t_k X_k} : t_i \in \mathbf{R}, X_i = A, \pm B, k > 0\}$$

generado por $\Sigma = \{A, \pm B\}$ es transitivo sobre $\mathbf{R}^2 - \{0\}$.

Además, una suposición natural en la teoría de sistemas de control es que G_Σ sea un subgrupo de Lie conexo del grupo $GL(2)$ de todas las matrices inversibles de 2×2 y que $S_\Sigma \subset G_\Sigma$ tenga interior no vacío en la topología de G_Σ, esto se pide para que nuestro sistema sea localmente accesible. Además el álgebra de Lie de G_Σ esté generada por las matrices A y B.

Con esas suposiciones, ahora nos dedicamos a encontrar los subgrupos que son transitivos sobre $\mathbf{R}^2 - \{0\}$. Usando argumentos de la teoría de Lie, probaremos que los únicos grupos conexos transitivos son $GL^+(2)$, el grupo de las matrices con determinante positivo, el grupo $SL(2)$ de las matrices con determinante uno, y $SO(2) \times R^+1$ que es isomorfo al grupo $\mathbf{C}^*$ de números complejos no nulos. El análisis del grupo $GL^+(2)$ se reduce al caso del grupo $SL(2)$, luego el estudio de este grupo es esencial en este trabajo. En este contexto mostraremos que un semigrupo $S \subset SL(2)$, con puntos interiores coincide con $SL(2)$ si y solo si es transitivo sobre el espacio real proyectivo $\mathbf{P}^1$, el conjunto de todos los subespacios de dimensión uno de $\mathbf{R}^2$.

El presente trabajo está dividido en cuatro capítulos y un apéndice.

En el primer capítulo damos los conceptos necesarios sobre la teoría de Lie, más precisamente detallamos los resultados centrales acerca de Grupos de Lie y sus correspondientes Álgebras de Lie, las relaciones que existen entre ambos conceptos y la riqueza de su estudio.

En el segundo capítulo desarrollamos los conceptos fundamentales y resultados de la Teoría de Sistemas de Control. Presentamos los sistemas de control lineales y bilineales en $\mathbf{R}^n$, enunciamos resultados sobre la controlabilidad de los sistemas lineales y mediante un ejemplo mostramos la utilidad de usar los sistemas bilineales en lugar de un sistema lineal con el propósito de obtener controlabilidad. Finalmente consideramos los sistemas de control que tienen como espacio estado a un Grupo de Lie, mostrando cuales son las propiedades que se obtienen al considerar la invariancia de dicho espacio de estados.

En el tercer capítulo realizamos un estudio de los grupos de Lie lineales que son transitivos sobre $\mathbf{R}^2 - \{0\}$, posteriormente buscamos dentro de $SL(2)$ semigrupos con interior no vacío que son transitivos sobre $\mathbf{P}^1$, obteniendo así uno de los resultados centrales de está tesis, que afirma la no existencia de tales semigrupos. Finalmente demostramos que el único semigrupo dentro de $SL(2)$ que actúa transitivamente sobre el plano es $SL(2)$.

En el cuarto capítulo realizamos un análisis para la controlabilidad del sistema bilineal, considerando el caso en que $trA = trB = 0$. Buscamos el criterio mediante el signo de $\det A$ y $\det B$, esto para conocer sus autovalores. Posteriormente mostraremos el resultado central de está tesis, este resultado caracteriza la controlabilidad del sistema (1) con la suposición que A y B son matrices de traza cero, su enunciado asegura que la controlabilidad es equivalente a $\det[A, B] < 0$.

CAPÍTULO I

Grupos y Álgebras de Lie

Los grupos de Lie sin duda son la clase más importante de variedades diferenciables. Los grupos de Lie son variedades que tienen una estructura de grupo compatible con la estructura diferencial. Ejemplos bien conocidos de grupos de Lie (de matrices) son el grupo general lineal, el grupo especial lineal y el grupo ortogonal.

En este capítulo veremos en conjunto los fundamentos necesarios para la comprensión del presente trabajo. De central importancia para la Teória de Lie, es la relación entre un grupo de Lie y su álgebra de Lie, que posteriormente veremos que es el conjunto de campos de vectores invariantes por la izquierda(o por derecha) o isomórficamente el espacio tangente en la identidad. Estudiaremos la correspondencia entre subgrupos y subálgebras; entre homomorfismos de grupos de Lie y homomorfismos de sus álgebras de Lie. También estudiaremos las propiedades de la aplicación exponencial el cual es una generalización a grupos de Lie arbitrarios de la exponenciación de matrices y que provee un vínculo clave entre un grupo de Lie y su álgebra de Lie. Investigaremos la representación adjunta y se póbara el teorema del subgrupo cerrado de Cartan. Finalmente se enunciarán las nociones básicas de álgebras de Lie Solubles y posteriormente el teorema fundamental sobre álgebras de Lie Solubles que será usado de manera fundamental en el capítulo 3. La principal fuente de consulta para la comprensión de este capítulo es [16].

I.1. Definiciones y ejemplos

Definición I.1. *Un grupo de Lie G es una variedad diferenciable dotada de una estructura de grupo tal que las aplicaciones $G \times G \to G$ y $G \to G$ definidas respectivamente por*

$$(x, y) \mapsto xy, \qquad x \mapsto x^{-1}$$

son C^∞, o equivalentemente la aplicación $G \times G \to G$ definida por $(x, y) \mapsto xy^{-1}$ es C^∞.

Ejemplo I.2. *Sea $G = \mathbf{R}^n$, con la estructura natural de variedad diferenciable y la operación de grupo dada por la adición de vectores. En estas condiciones G es un grupo de Lie.*

Ejemplo I.3 (El grupo general lineal). *El conjunto de matrices cuadradas $n \times n$ inversibles con entradas reales denotado por $GL(n, \mathbf{R})$, es un grupo con respecto al producto de matrices. En efecto, recordemos que una matriz $A \in M(n, \mathbf{R})$ es invertible si y sólo si $\det A \neq 0$. Luego si $A, B \in GL(n, \mathbf{R})$ entonces de las propiedades elementales de la aplicación determinante, $\det(AB) = \det A \det B$ y $\det(A^{-1}) = (\det A)^{-1}$, tenemos que $AB \in GL(n, \mathbf{R})$ y $A^{-1} \in GL(n, \mathbf{R})$. Ya que la aplicación determinante es continua por ser n-lineal sobre sus vectores columna, tenemos que el conjunto $GL(n, \mathbf{R}) = \det^{-1}(\mathbf{R} - \{0\})$ es abierto, por tanto hereda la estructura de variedad diferenciable. Ahora el producto de dos matrices es diferenciable pues las entradas de AB es un polinomio con entradas de A y B. La diferenciabilidad de la inversa se sigue similarmente de la regla de Crammer's: $(A^{-1})_{ji} = \frac{\det A^{ij}}{\det A}$ donde A^{ij} es la matriz obtenida de A quitando la fila i y la columna j.*

Ejemplo I.4 (El grupo especial lineal). *Nuevamente consideremos la función real $\det : \mathbf{R}^{n^2} \approx M(n, \mathbf{R}) \to \mathbf{R}$ que es de clase C^∞. Por la expresión general de la derivada de una función n-lineal, se tiene*

$$(\det)'(X)H = \sum_{i=1}^{n} \det(X^1, ...H^i, ..., X^n), \qquad X, H \in M(n, \mathbf{R}).$$

En particular, para $X = I =$ *matriz identidad* $n \times n$,

$$(\det)'(I)H = \sum_{i=1}^{n} \det(e_1, ...H^i, ..., e_n) = \sum_i h_i^i = trH$$

y

$$\frac{\partial \det}{\partial x_s^r}(X) = (\det)'(X)E^{r,s} = (-1)^{r+s} \det X_s^r.$$

Consideremos la restricción $\det : GL(n, \mathbf{R}) \to \mathbf{R}$. *De la expansión del determinante a lo largo de una linea (o columna), se sigue que, dada* $A \in GL(n, \mathbf{R})$, *existe algún determinante menor* $\det(A_s^r) \neq 0$. *Esto muestra que* $\det : GL(n, \mathbf{R}) \to \mathbf{R}$ *es una submersión de clase* C^∞. *En otras palabras, todo número real no nulo* c *es un valor regular de* $\det$.

Se concluye que el conjunto

$$SL(n, \mathbf{R}) = \{X \in GL(n, \mathbf{R}) : \det X = 1\} = (\det)^{-1}(1)$$

es una superficie (por tanto una variedad) de dimensión $n^2 - 1$ *y clase* C^∞ *en* $\mathbf{R}^{n^2}$. $SL(n, \mathbf{R})$ *es llamado el grupo especial lineal o grupo unimodular. Evidentemente, si* $X, Y \in SL(n, \mathbf{R})$ *entonces* $XY \in SL(n, \mathbf{R})$ *y* $X^{-1} \in SL(n, \mathbf{R})$. *O sea,* $SL(n, \mathbf{R})$ *es un subgrupo de* $GL(n, \mathbf{R})$, *que es una variedad* C^∞. *Por lo tanto un grupo de Lie. El espacio tangente a* $SL(n, \mathbf{R})$ *en* I *es el conjunto de todas las matrices de traza cero.*

Ejemplo I.5 (El grupo ortogonal). *El conjunto* $O(n, \mathbf{R})$ *de todas las matrices reales* $n \times n$, X *tales que* $XX^t = I$. *Se tiene que* $O(n, \mathbf{R})$ *es un subgrupo de* $GL(n, \mathbf{R})$. *Geométricamente, un operador lineal en* $\mathbf{R}^n$ *es una isometría (esto es, preserva distancias) si y solo si su matriz con respecto a la base canónica es ortogonal.*

Vamos a demostrar que $O(n, \mathbf{R})$ *es una superficie compacta de dimensión* $\frac{n}{2}(n-1)$ *y clase* C^∞ *en* $\mathbf{R}^n$.

Consideremos la aplicación de clase C^∞

$$f : M(n, \mathbf{R}) \to S(n, \mathbf{R}) \approx \mathbf{R}^{n^2}$$

donde $S(n,\mathbf{R})$ *es el conjunto de matrices simétricas, definida por* $f(X) = XX^t$*. Si mostramos que* $I \in S(n,\mathbf{R})$ *es un valor regular de* f *entonces concluiremos que* $O(n,\mathbf{R}) = f^{-1}(I)$ *es una superficie de dimensión* $n^2 - \frac{n}{2}(n+1) = \frac{n}{2}(n-1)$ *en* $\mathbf{R}^{n^2}$*.*

Sea por tanto $X \in f^{-1}(I) = O(n,\mathbf{R})$*. Queremos probar que la derivada* $f'(X) : M(n,\mathbf{R}) \to S(n,\mathbf{R})$*, dada por* $f'(X)H = XH^t + HX^t$*, es sobreyectiva. Dada* $S \in S(n,\mathbf{R})$*, sea* $V = \frac{SX}{2}$*. Entonces*

$$f'(X)V = X(\frac{SX}{2})^t + (\frac{SX}{2})X^t = (XX^t)\frac{S}{2} + \frac{S}{2}XX^t = S.$$

El espacio tangente a $O(n,\mathbf{R})$ *en la identidad es el conjunto de matrices antisimétricas como se puede ver fácilmente.*

Ejemplo I.6 (El grupo de Heisenberg de dimensión tres)**.** *Se define por*

$$G = \left\{ \begin{pmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{pmatrix} : x_1, x_2, x_3 \in \mathbf{R} \right\}.$$

La aplicación $\varphi : G \to \mathbf{R}^3$ *definida por*

$$\varphi(\begin{pmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{pmatrix}) = (x_1, x_2, x_3)$$

es un difeomorfismo. En particular, G *es subconjunto cerrado de* $GL(3,\mathbf{R})$ *pero no es compacto, además, es conexo y simplemente conexo. La estructura de grupo abstracto de* G *está definida por la multiplicación de matrices. Se induce en* $\mathbf{R}^3$ *una estructura de grupo requiriendo que* φ *sea un homomorfismo de grupos, esta operación está dada por*

$$(x_1, x_2, x_3) * (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3 + x_1 y_2).$$

El primer paso en el estudio de los grupos de Lie consiste en la construcción de las álgebras de Lie asociadas a dichos grupos.

De ahora en adelante siempre que se mencionen espacios vectoriales serán reales.

Definición I.7. *Un álgebra de Lie consiste de un espacio vectorial* $\mathbf{g}$ *dotado de un producto (corchete o conmutador)*

$$[\cdot,\cdot] : \mathbf{g} \times \mathbf{g} \to \mathbf{g}$$

con las siguientes propiedades:

1. *bilinealidad;*
2. *antisimetría, es decir,* $[X,Y] = -[Y,X]$ *para todo* $X, Y \in \mathbf{g}$*, o equivalentemente,* $[X,X] = 0$ *para todo* $X \in \mathbf{g}$*.*
3. *Satisface la identidad de Jacobi, esto es, para todo* $X, Y, Z \in \mathbf{g}$*,*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0.$$

En general, un álgebra es un espacio vectorial $\mathbf{g}$ dotado de un producto, esto es, una aplicación $\mathbf{g} \times \mathbf{g} \to \mathbf{g}$. Cualquier aplicación de este tipo que merece el nombre de producto debe ser bilineal. La antisimetría y la identidad de Jacobi son características de las álgebras de Lie. Otros tipos de álgebras tienen otros tipos de propiedades que la definen. Existen por ejemplo las álgebras asociativas, es decir, las que verifican $x(yz) = (xy)z$. Conviene observar que el corchete no es en general asociativo, pues en cualquier circunstancia $[[X,X],Y] = 0$ en tanto $[X,[X,Y]]$ no siempre se anula.

Definición I.8. *Sea* $\mathbf{g}$ *un álgebra de Lie. Un subálgebra de Lie de* $\mathbf{g}$ *es un subespacio vectorial* $\mathbf{h}$ *de* $\mathbf{g}$ *que es cerrado respecto al corchete, esto es,* $[X,Y] \in \mathbf{h}$ *si* $X, Y \in \mathbf{h}$*.*

Ejemplo I.9. *Álgebras abelianas:*$[\cdot,\cdot] = 0$*. En este caso la estructura de álgebra de Lie no enriquece la estructura de espacio vectorial.*

1. *Todo subespacio vectorial de dimensión 1 de un álgebra de Lie, es una subálgebra abeliana.*
2. *Todo subespacio de un álgebra abeliana es un subálgebra.*

Ejemplo I.10. *Álgebras de dimensión* ≤ 2.

1. $\dim \mathbf{g} = 1$. *Se sigue de la antisimetría del corchete que* $\mathbf{g}$ *es abeliana.*

2. $\dim \mathbf{g} = 2$. *Existen dos posibilidades.*

   a) $\mathbf{g}$ *es abeliana.*

   b) *Existe una base* $\{X, Y\}$ *de* $\mathbf{g}$ *tal que* $[X, Y] = Y$, *luego el corchete de dos elementos cualesquiera de* $\mathbf{g}$ *estará dado por*

$$[aX + bY, cX + dY] = (ad - bc)[X, Y] = (ad - bc)Y.$$

En efecto, supongamos que $\mathbf{g}$ *no sea abeliana y sea* $\{X', Y'\}$ *una base de* $\mathbf{g}$. *Entonces* $[X', Y'] \neq 0$, *pues en caso contrario* $\mathbf{g}$ *sería abeliana. Sea* $Y'' = [X', Y']$ *y se escoge* X'' *de tal forma que* $\{X'', Y''\}$ *es una base de* $\mathbf{g}$. *Entonces* $X'' = aX' + bY'$, $Y'' = cX' + dY'$ *y*

$$[X'', Y''] = \alpha Y''$$

con $\alpha \neq 0$, *en caso contrario el álgebra sería abeliana. Los elementos* $X = \frac{1}{\alpha}X''$, $Y = Y''$ *forman la base requerida. Las álgebras*

$$\mathbf{g} = \{\begin{pmatrix} a & b \\ 0 & -a \end{pmatrix} : a, b \in \mathbf{R}\} \quad y \quad \mathbf{g} = \{\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a, b \in \mathbf{R}\}$$

son ejemplos concretos de álgebras de Lie no abelianas de dimensión dos.

I.2. El álgebra de Lie de un grupo de Lie

Definición I.11. *Sean* G *un grupo de Lie y* $g \in G$. *Las traslaciones izquierda y derecha por* g, *son respectivamente las aplicaciones* $L_g : G \to G$ *y* $R_g : G \to G$ *definidos por*

$$L_g(h) = gh, \quad R_g(h) = hg$$

para todo $h \in G$.

Estas aplicaciones son diferenciables pues resultan de la composición de aplicaciones diferenciables. En realidad, ambas traslaciones son difeomorfismos, ya que

$$L_g \circ L_{g^{-1}} = R_g \circ R_{g^{-1}} = I_G.$$

Definición I.12. *Un campo de vectores* X *en* G *se dice invariante a izquierda, si* X *es* L_g*-relacionado consigo mismo para todo* $g \in G$*, es decir*

$$D(L_g) \circ X = X \circ L_g$$

para todo $g \in G$*. Similarmente definimos los campos de vectores invariantes a derecha, el campo vectorial* X *es invariante a derecha si es* R_g*-relacionado con sigo mismo para todo* $g \in G$*.*

Los campos de vectores invariantes a izquierda y derecha están completamente determinados por su valor en la identidad; pues, para todo $g \in G$ la condición de invariancia a izquierda, por ejemplo, implica que

$$X(g) = D(L_g)_1(X(1)).$$

Luego cada elemento de T_1G determina un único campo de vectores invariante a izquierda y un único campo de vectores invariante a derecha. Dado $A \in T_1G$ denotaremos por A^l y A^r los campos vectoriales a izquierda y derecha definidos por:

$$A^l(g) = D(L_g)_1(A), \quad A^r(g) = D(R_g)_1(A).$$

Denotaremos por Inv^l el conjunto de los campos de vectores invariantes a izquierda y por Inv^r el de los campos invariantes a derecha. Estos conjuntos son subespacios vectoriales del espacio de todos los campos de vectores en G (y ambos son generalmente distintos), ya que $D(L_g)$ y $D(R_g)$ son aplicaciones lineales sobre los campos de vectores. Las aplicaciones $A \in T_1G \mapsto A^l \in Inv^l$ y $A \in T_1G \mapsto A^r \in Inv^r$ son ambos isomorfismos entre espacios vectoriales.

Es fácil demostrar que si X e Y son campos de vectores invariantes a izquierda en un grupo de

Lie G, $[X,Y]$ es invariante a izquierda; en efecto, al ser $L_g : G \to G$ un difeomorfismo se cumple que $D(L_g)[X,Y] = [D(L_g)(X), D(L_g)(Y)] = [X,Y]$ de donde concluimos que $[X,Y]$ es invariante a izquierda. Análogamente, si dos campos X e Y son invariantes a derecha entonces $[X,Y]$ es invariante a derecha.

Los espacios Inv^l y Inv^r son, luego, subálgebras de Lie del álgebra de Lie de los campos de vectores en G. Así, ambas son álgebras de Lie. Además, si $A \in T_1G$ se cumple que $D(i)(A^r) = (-A)^l$, donde $i : g \to g^{-1}$. Luego $D(i)$ es un isomorfismo entre Inv^l y Inv^r. Por otra parte, como vimos T_1G es isomorfo tanto a Inv^l como a Inv^r. A través de estos isomorfismos se inducen naturalmente los corchetes de Lie sobre T_1G. Esos corchetes son dados por $[A,B]_l = [A^l, B^l](1)$ y $[A,B]_r = [A^r, B^r](1)$. Evidentemente $[A,B]_r = -[A,B]_l$, luego ambas estructuras de álgebras de Lie sobre T_1G inducidas por los corchetes mencionados son isomorfas por la aplicación $-I : T_1G \to T_1G$.

Definición I.13. *El álgebra de Lie de G, denotada por $\mathbf{g}$ o $L(G)$, es cualquiera de las álgebras isomorfas Inv^l, Inv^r, $(T_1G, [\cdot,\cdot]_l)$, $(T_1G, [\cdot,\cdot]_r)$.*

Ejemplo I.14. *Sea $G = GL_n(\mathbf{R})$ el grupo de las transformaciones lineales inversibles de $\mathbf{R}^n$, o lo que es lo mismo, el grupo de las matrices $n \times n$ inversibles. Fijando $g \in G$, las traslaciones a izquierda a derecha L_g y R_g pueden ser extendidas a todo el espacio de las matrices, dando origen a transformaciones lineales de $M_n(\mathbf{R})$ y luego el fibrado tangente a G se identifica con $G \times M_n(\mathbf{R})$. De ahí que un campo de vectores X en G no es nada mas que una aplicación $X : G \to M_n(\mathbf{R})$. Además, por esta identificación, las trasformaciones lineales L_g y R_g satisfacen $D(L_g)_h = L_g$ y $D(R_g)_h = R_g$ para cualesquiera $h, g \in G$.*

A partir de esas observaciones es posible describir los campos invariantes en G. Suponga que $X : G \to M_n(\mathbf{R})$ es invariante a izquierda. Entonces, para todo $g \in G$,

$$X(g) = D(L_G)_1(X(1)) = L_g(X(1)) = gX(1).$$

Luego, los campos invariantes por izquierda son de la forma $X_A(g) = gA$ con A una matriz en T_1G. En coordenadas locales el corchete de Lie de dos campos vectoriales es dado por

$$[X,Y] = dY(X) - dX(Y).$$

Para una matriz A, X_A se extiende a una aplicación lineal en el espacio de las matrices. Por tanto, $DX_A = X_A$. Así, aplicando esa fórmula del corchete a X_A y X_B, se obtiene

$$[X_A, X_B](g) = (gA)B - (gB)A$$

esto es, $[X_A, X_B] = X_{AB-BA}$.

Ejemplo I.15. *Todo espacio vectorial de dimensión finita V admite una estructura natural de variedad diferenciable. Sean $\{e_i\}$ una base de V y $\{r_i\}$ su base dual. Entonces existe una identificación natural del espacio tangente T_pV a V en p con el mismo V, dada por*

$$\Sigma a_i \frac{\partial}{\partial r_i}|_p \longleftrightarrow \Sigma a_i e_i.$$

Se sigue que este espacio vectorial con su operación suma es un grupo de Lie.

Ejemplo I.16. *Álgebra de Lie de $\mathbf{R}^n$. Sea $G = (\mathbf{R}^n, +)$. Fijando $v \in \mathbf{R}^n$, las traslaciones a izquierda y derecha son*

$$L_v(x) = R_v(x) = x + v.$$

Por tanto, $D(L_g)y = D(R_g)y = I$ para todo $y \in \mathbf{R}^n$. Luego tenemos que los campos X invariantes a izquierda son constantes, pues $X(x) = D(L_x)_0 X(0) = X(0) = v$. Como el corchete de campos constantes es cero, la álgebra de Lie del grupo abeliano $(\mathbf{R}^n, +)$ es abeliana, esto es, $[X, Y] = 0$ para todo X, Y campos invariantes en $\mathbf{R}^n$.

Ejemplo I.17. *Sean G y H grupos de Lie. Entonces, el producto cartesiano $G \times H$ admite la estructura de variedad producto y la estructura de grupo producto, haciendo de este un grupo de Lie. Si $\mathbf{g}$ y $\mathbf{h}$ son sus álgebras de Lie, entonces la álgebra de Lie de $G \times H$ es $\mathbf{g} \times \mathbf{h}$, donde el corchete está dado por*

$$[(X_1, Y_1), (X_2, Y_2)] = ([X_1, X_2], [Y_1, Y_2]).$$

I.3. La aplicación exponencial

La aplicación exponencial $\exp : \mathbf{g} \to G$ es el objeto central para transportar al grupo de Lie G propiedades de su álgebra de Lie $\mathbf{g}$. La idea básica de su construcción es que, por definicón los elementos de $\mathbf{g}$ son ecuaciones diferenciales ordinarias en G (campos invariantes), que poseen flujos, los cuales son formados por difeomorfismos locales de G. Estos elementos se identifican naturalmente con elementos de G, permitiendo construir, a partir de $X \in \mathbf{g}$, un subgrupo de G parametrizado por $t \in \mathbf{R}$.

Para colocar esos comentarios de manera precisa, sea X un campo invariante (a izquierda o a derecha en G). La invariancia de X tiene por consecuencia la siguiente simetría del flujo X_t: supongamos que X es un campo invariante a izquierda, tomemos $g, h \in G$ y consideremos la curva

$$\alpha(t) = L_g(X_t(h)) = gX_t(h).$$

Su dominio es un intervalo abierto de $\mathbf{R}$ conteniendo a 0. Como $X_0(h) = h$, $\alpha(o) = gh$. Además,

$$\alpha'(t) = D(L_g)_{X_t(h)}(X(X_t(h))) = X(gX_t(h)) = X(\alpha(t)),$$

por la invariancia de X. Así, α es solución de

$$\frac{dg}{dt} = X(g)$$

con condición inicial $\alpha(0) = gh$, esto es, $\alpha(t) = X_t(gh)$. Eso significa que

$$X_t(gh) = gX_t(h), \quad X \in Inv^l.$$

Tomando $h = 1$, $X_t(g) = gX_t(1)$. Esto es, la solución que pasa por g es obtenida por traslación de la solución que pasa por el elemento neutro. Así, todas las trayectorias se prolongan por el mismo intervalo de $\mathbf{R}$. Eso, permite demostrar que los campos invariantes son completos.

Proposición I.18. *Un campo invariante (a izquierda o a derecha) es completo. Además, si* $A \in$

T_1G, entonces $(A^r)_t(1) = (A^l)_t(1)$ para todo $t \in \mathbf{R}$ y la aplicación

$$t \in \mathbf{R} \longmapsto (A^r)_t(1) \in G$$

es un homomorfismo del grupo aditivo $\mathbf{R}$ a valores en G.

Demostración. Sea X un campo invariante a izquierda. Para mostrar que X es completo basta mostrar que la trayectoria $X_t(1)$ que pasa por el elemento neutro en $t = 0$ se prolonga para todo $\mathbf{R}$. Sea (α, β) el dominio de definición de la solución maximal $t \mapsto X_t(1)$. Supongamos que $\beta < \infty$. Entonces, $X_{\frac{\beta}{2}}(1)$ está bien definida y la trayectoria $s \mapsto X_s(X_{\frac{\beta}{2}}(1))$ se prolonga sobre (α, β), pues las trayectorias de un campo invariante se prolongan sobre el mismo intervalo. Ese prolongamiento permite definir la curva $\sigma : (\alpha, 3\beta/2) \to G$ por

$$\sigma(t) = \begin{cases} X_t(1) \\ X_{t-\beta/2}(X_{\beta/2}) \end{cases}.$$

Es claro que $\sigma(0) = 1$, y por definición se ve que σ es una solución de la ecuación $\frac{dg}{dt} = X(g)$ que se prolonga a $(\alpha, 3\beta/2)$ contradiciendo el hecho que (α, β) es el intervalo de definición máximo que pasa por 1. Con eso queda demostrado que el campo es completo. $\square$

Con esta proposición tenemos la definición siguiente.

Definición I.19. *Sea $A \in T_1G$. Entonces,* $\exp A = (A^r)_1(1) = (A^l)_1(1)$.

La aplicación exponencial es de la forma $\exp : \mathbf{g} \to G$. Por las identificaciones anteriores, si X es un campo invariante $\exp X$ tiene sentido y es el valor en $1 \in \mathbf{R}$ de la solución de X que pasa por el elemento neutro cuando $t = 0$. Por la proposición anterior, la aplicación $t \mapsto \exp(tX)$, $X \in \mathbf{g}$, es un homomorfismo. Por tanto,

$$\{\exp(tX) : t \in \mathbf{R}\}$$

es un subgrupo de G, denominado *subgrupo a 1-parámetro generado por X.*

A continuación se enuncian algunas propiedades de la aplicación exponencial.

Proposición I.20. *Valen las siguientes afirmaciones.*

1. *Si X es invariante a derecha entonces $X_t = L_{\exp(tX)}$, esto es, $X_t(g) = \exp(tX)g$.*

2. *Si X es invariante a izquierda entonces $X_t = R_{\exp(tX)}$, esto es, $X_t(g) = g\exp(tX)$.*

3. $\exp 0 = 1$.

4. *Para todo $X \in \mathbf{g}$ y $t, s \in \mathbf{R}$,*

$$\exp(t+s)X = \exp(tX)\exp(sX) = \exp(sX)\exp(tX),$$

   esto es, los elementos del subgrupo $\{\exp(tX) : t \in \mathbf{R}\}$ conmutan entre si.

5. *Sean $X, Y \in \mathbf{g}$. Entonces $[X, Y] = 0$ si y sólo si, $\exp(tX)\exp(sY) = \exp(sY)\exp(tX)$.*

La aplicación exponencial fue definida a través de la solución de la ecuación diferencial $dg/dt = X(g)$, con X invariante. El conjunto de las ecuaciones definidas (por ejemplo) por los campos invariantes a derecha puede ser colocada en una única ecuación independiente del parámetro $A \in T_1G$, escribiendo

$$\frac{dg}{dt} = f(A, g)$$

donde $f : T_1G \times G \to G$ es dada por $f(A, g) = D(R_g)_1(A)$ y es de clase C^∞. Por tanto las soluciones de la ecuación dependen diferenciablemente del parámetro A. Así, $\exp : \mathbf{g} \to G$ es una aplicación diferenciable. Además es fácil ver que $D(\exp)_0 = Id$, luego por el teorema de la función inversa, existe un entorno U de $0 \in \mathbf{g}$ y un entorno V de 1 en G tal que $\exp|_U : U \to V$ es un difeomorfismo. Y si G es un grupo de Lie conexo y $g \in G$, existen $X_1, ..., X_r \in \mathbf{g}$ tal que

$$g = \exp(X_1)\cdots\exp(X_r).$$

Pues en un grupo de Lie conexo tenemos que si U es un entorno de la identidad entonces

$$G = \bigcup_{n=1}^{\infty} U^n,$$

como se demuestra en la proposición de abajo.

Teorema I.21. *Sea G un grupo de Lie conexo y sea U una vecindad de e. Entonces,*

$$G = \bigcup_{n=1}^{\infty} U^n$$

donde U^n consiste de todos los productos de n elementos de U. (Se dice que U genera G.)

Demostración. Sea V un subconjunto abierto de U que contenga a e tal que $V = V^{-1}$ (donde $V^{-1} = \{x^{-1} : x \in V\}$). Por ejemplo, $V = U \cap U^{-1}$ verifica tal propiedad. Sea

$$H = \bigcup_{n=1}^{\infty} V^n \subset \bigcup_{n=1}^{\infty} U^n.$$

Entonces H es un subconjunto abstracto de G y es un subconjunto abierto de G pues $x \in H$ implica que $xV \subset H$ Así cada clase mod H es abierto en G. Ahora, H es el complemento en G de la unión de todas las clases mod H diferentes del mismo H. Luego H es un subconjunto cerrado de G. Como G es conexo y H es no vacío entonces $G = H$. De esto último con la inclusión de encima, tenemos el resultado. □

Ejemplo I.22. *Como fue visto los campos de vectores invariantes por izquierda en $GL_n(\mathbf{R})$ son de la forma $X(g) = gA$, con A una matriz de $n \times n$. La ecuación diferencial asociada a X es el sistema lineal*

$$\frac{dg}{dt} = gA$$

en el espacio de las matrices. Su solución fundamental está dada por la exponencial de matrices, que coincide, por tanto, con la aplicación exponencial en $GL_n(\mathbf{R})$.

Ejemplo I.23. *Vimos que en $(\mathbf{R}^n, +)$ los campos invariantes son constantes: $X(x) = v$. El flujo de estos campos está dado por las traslaciones $X_t(x) = x + tv$. Tomando $x = 0$ se tiene que*

$$\exp(tv) = tv.$$

En particular, $\exp(v) = v$ y $\exp = Id$.

Ejemplo I.24. *Si G y H son grupos de Lie con álgebras de Lie $\mathbf{g}$ y $\mathbf{h}$ respectivamente, entonces $\exp(X, Y) = (\exp X, \exp Y)$ si $(X, Y) \in \mathbf{g} \times \mathbf{h}$, el álgebra de Lie de $G \times H$.*

I.4. Homomorfismos

Sean G y H dos grupos de Lie. Un *homomorfismo de grupos de Lie* $\phi : G \to H$, es un homomorfismo entre grupos que es diferenciable. La misma terminología se aplica a isomorfismos y automorfismos entre grupos de Lie. Teniendo en cuenta el hecho de que los grupos de Lie deben ser estudiados a través de sus álgebras de Lie, los homomorfismos entre grupos de Lie serán descritos a través de homomorfismos entre sus álgebras de Lie.

Un homomorfismo entre las álgebras de Lie $\mathbf{g}$ y $\mathbf{h}$ es una transformación lineal $\theta : \mathbf{g} \to \mathbf{h}$ que satisface $\theta[X, Y] = [\theta X, \theta Y]$ para todo X, Y en $\mathbf{g}$. La relación entre los homomorfismos de grupos y álgebras de Lie es dada por la diferencial en el elemento neutro.

Proposición I.25. *Sean G y H grupos de Lie con álgebras de Lie $\mathbf{g}$ y $\mathbf{h}$ respectivamente. Sea $\theta : G \to H$ un homomorfismo diferenciable y tomemos $X \in \mathbf{g}$. Entonces,*

$$\theta(\exp X) = \exp(d\theta_1(X)).$$

Proposición I.26. *Sean G y H grupos de Lie con álgebras de Lie $\mathbf{g}$ y $\mathbf{h}$ respectivamente. Sea $\theta : G \to H$ un homomorfismo diferenciable y tomemos $X \in \mathbf{g}$. Entonces, $d\theta_1 : \mathbf{g} \to \mathbf{h}$ es un homomorfismo.*

La proposición anterior significa que homomorfismos de grupos de Lie inducen homomorfismos de Lie entre álgebras de Lie correspondientes. El procedimiento inverso, la construcción de homomorfismos de grupos que extienden homomorfismos de álgebras de Lie no siempre es posible, esto es, si G y H son grupos de Lie con álgebras de Lie $\mathbf{g}$ y $\mathbf{h}$ respectivamente, y $\theta : G \to H$ un homomorfismo, no siempre existe un homomorfismo $\phi : G \to H$ tal que $\theta = d\phi_1$. Como veremos en el caso en que G es simplemente conexo, todo homomorfismo cuyo dominio es el álgebra de Lie $\mathbf{g}$ de G se extiende a un homomorfismo diferenciable con dominio G. Sin embargo, independiente de la existencia de homomorfismos globales, siempre es posible realizar una construcción local.

Remarca I.27. *Tomemos $G = S^1$ y $H = \mathbf{R}$. Entonces, $\dim \mathbf{g} = \dim \mathbf{h} = 1$; por tanto existe un isomorfismo entre $\mathbf{g}$ y $\mathbf{h}$. Mas, ningún isomorfismo es de la forma $d\phi_1$, pues el único homomorfismo*

$G \to H$ es constante, ya que S^1 es compacto y $\{0\}$ es el único subgrupo de $\mathbf{R}$ contenido en un compacto. Y como se ve claramente lo que está en juego aquí son las propiedades topológicas globales del grupo G.

I.5. Representaciones

Definición I.28. *Sea V un espacio vectorial y $\mathbf{gl}(V)$ el álgebra del grupo de Lie de las transformaciones lineales de V. Si $\mathbf{G}$ un álgebra de Lie, una representación de $\mathbf{G}$ en V es un homomorfismo $\rho : \mathbf{G} \to \mathbf{gl}(V)$. El álgebra V se denomina espacio de representación en cuanto que su dimensión es la dimensión de la representación. Una representación ρ es fiel si $\ker \rho = \{0\}$.*

La noción de representación está asociada a la idea de describir las álgebras de Lie como subálgebras de las transformaciones lineales. En el caso de las representaciones fieles, $\mathbf{g}$ es isomorfa a im $\mathbf{g}$. En el caso de dimensión finita toda álgebra de Lie se puede considerar de transformaciones lineales, es decir, de matrices. Ese resultado es el teorema

Teorema I.29 (Ado)**.** *Toda álgebra de Lie de dimensión finita admite una representación fiel de dimensión finita.*

La demostración de este teorema puede ser encontrada en [16].

Ejemplo I.30. *Sea $\mathbf{g}$ un álgebra de Lie no abeliana de dimensión dos y sea $\{X, Y\}$ una base tal que $[X, Y] = Y$. La única transformación lineal $\rho : \mathbf{g} \to \mathbf{gl}(2)$ que satisface*

$$\rho(X) = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \qquad \rho(Y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

define una representación fiel de $\mathbf{g}$ de dimensión dos. Su imagen es

$$im\rho = \{ \begin{pmatrix} a & b \\ 0 & -a \end{pmatrix} : a, b \in \mathbf{R}\}$$

Ejemplo I.31. *Sean* $\mathbf{g}$ *un álgebra de Lie de dimensión 3 con base* $\{X', Y', Z'\}$ *sujeta a las relaciones* $[Z', X'] = 2X'$, $[Z', Y'] = -2Y'$ *y* $[X', Y'] = Z'$. *Sea* $\rho : \mathbf{g} \to \mathbf{gl}(2)$ *la única transformación lineal tal que*

$$\rho(X') = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ \rho(Y') = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ \rho(Z') = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Entonces, ρ *define una representación fiel, obsérvece que im*$\mathbf{g} = \mathbf{sl}(2)$.

I.6. Representación Adjunta

Definición I.32. *Sea* V *un espacio vectorial y* $\mathbf{gl}(V)$ *el álgebra del grupo de Lie de las transformaciones lineales de* V. *Si* G *es un grupo de Lie, una representación de* G *en* V *es un homomorfismo* $\rho : G \to \mathbf{gl}(V)$. *El álgebra* $\mathbf{gl}(V)$ *se denomina espacio de representación en cuanto que su dimensión es la dimensión de la representación.*

Existe una representación natural del grupo de Lie G en su álgebra de Lie. En efecto, un elemento $g \in G$ define el automorfismo interno $C_g : G \to G$ definido por

$$C_g(x) = gxg^{-1}.$$

Como $C_g(1) = 1$ y la aplicación es un difeomorfismo, $d(C_g)_1$ es una aplicación lineal de $\mathbf{g} \to \mathbf{g}$ y es un isomorfismo. De ahí que $g \mapsto d(C_g)_1$ es una representación de G en $\mathbf{g}$.

Definición I.33. *La representación adjunta* $Ad : G \to GL(\mathbf{g})$, *de* G *en su álgebra de Lie* $\mathbf{g}$ *es definida por*

$$Ad(g) = d(C_g)_1 = d(L_g \circ R_{g^{-1}})_1 = d(R_g \circ L_{g^{-1}})_1$$

$Ad(g) = d(C_g)_1$ es un homomorfismo de $\mathbf{g}$, además vale la siguiente fórmula, bastante usada en relaciones que involucran a la adjunta:

$$g \exp(X) g^{-1} = \exp(Ad(g)X).$$

La representación infinitesimal de Ad es una representación del álgebra de Lie $\mathbf{g}$ en si misma. La representación infinitesimal es la representación adjunta de $\mathbf{g}$.

Definición I.34. *Sea* $\mathbf{g}$ *un álgebra de Lie. Su representación adjunta es la aplicación* $ad : \mathbf{g} \to \mathbf{gl}(\mathbf{g})$ *definida por*

$$ad(X)Y = [X, Y].$$

La aplicación ad es de hecho un homomorfismo de álgebras de Lie, donde el corchete en $\mathbf{gl}(\mathbf{g})$ es dado por el conmutador. De hecho $ad(X)$, $X \in \mathbf{g}$ son derivaciones de $\mathbf{g}$. Ellas son denominadas derivaciones internas de $\mathbf{g}$.

Proposición I.35. *Sea G un grupo de Lie, con álgebra de Lie* $\mathbf{g}$*, con el corchete dado por los campos invariantes a izquierda. Entonces, la representación infinitesimal asociada a su representación adjunta Ad es la representación adjunta ad de* $\mathbf{g}$*. En otras palabras:*

1. $d(AD)_1(X) = ad(X), X \in \mathbf{g}$.

2. $Ad(\exp X) = \exp(ad(X))$.

Una importante consecuencia de esta proposición es que en el caso en que G es conexo las fórmulas anteriores nos ayudan a caracterizar $Ker(Ad)$.

Proposición I.36. *Sea*

$$Z(G) = \{g \in G : gh = hg, \forall h \in G\}$$

el centro de G. Entonces, $Z(G) \subset Ker(Ad)$. Suponga que el grupo de Lie G es conexo, entonces $Z(G) = Ker(Ad)$.

Esta proposición está en concordancia con el hecho que el núcleo $Ker(Ad)$ de la representación adjunta de $\mathbf{g}$ es su centro

$$\mathbf{z}(\mathbf{g}) = \{X \in \mathbf{g} : [X, Y] = 0, \forall Y \in \mathbf{g}\}.$$

Será mostrado posteriormente que $Z(G)$ es un subgrupo de Lie de G, cuya subálgebra de Lie es el centro de $\mathbf{g}$.

Ejemplo I.37. *En $GL_n(\mathbf{R})$, Ad($\mathbf{g}$) coincide con la conjugación C_g, pues C_g se extiende a una transformación lineal en el espacio de las matrices, por tanto coincide con Ad($\mathbf{g}$) que es su diferencial en su identidad. En otras palabras, si $X \in \mathbf{gl}(\mathbf{R})$ y $g \in GL_n(\mathbf{R})$ entonces*

$$Ad(\mathbf{g})X = gXg^{-1}.$$

En el centro de $GL_n(R)$ es el grupo de las matrices escalares aI, $a \neq 0$. Apesar que $GL_n(\mathbf{R})$ tenga dos componentes conexas, la expresión para Ad($\mathbf{g}$) confirma de inmediato que

$$Z(GL_n(\mathbf{R})) = ker(Ad).$$

I.7. Subgrupos de Lie

Un subgrupo de Lie de un grupo de Lie G es un subgrupo H que, al mismo tiempo, es una subvariedad quasi-regular. La condición de quasi-regular garantiza que el producto en H es diferenciable en relación a su estructura intrínsica, pues H es localmente conexo. Un subgrupo de Lie, luego, es a su vez un grupo de lie tal que la inclusión $H \to G$ es un homomorfismo diferenciable. Las componentes conexas de la identidad de un subgrupo de Lie son a su vez subgrupos de Lie.

I.8. Subálgebras de Lie

El principio que dirige la construcción de la teoría de grupos de Lie es el de obtener información sobre la estructura de los grupos de Lie a partir de las álgebras de Lie. Siguiendo ese principio, los subgrupos de Lie de G con álgebra de Lie $\mathbf{g}$ son estudiados relacionándolos con subálgebras de $\mathbf{g}$, estos últimos son subespacios vectoriales cerrados por el corchete.

La inclusión $i : H \to G$ de un subgrupo de Lie H del grupo de Lie G es un homomorfismo diferenciable. Como fue visto, la diferencial di_1 es un homomorfismo de álgebras de Lie. Entre tanto, di_1 es la inclusión del espacio tangente T_1H en el espacio tangente T_1G. Por tanto, $\mathbf{h}$ es una subálgebra de Lie de $\mathbf{g}$, lo que significa que la álgebra de Lie de un subgrupo de Lie se identifica (es isomorfa) a una subálgebra de Lie del álgebra de Lie del grupo G.

Proposición I.38. *Sea H un subgrupo de Lie de G y denotemos por $\mathbf{h}$ su álgebra de Lie. Suponga que $g \in G$ normaliza H, esto es, $gHg^{-1} \subset H$. Entonces, $Ad(\mathbf{g})\mathbf{h} \subset \mathbf{h}$ y por tanto, $Ad(\mathbf{g})\mathbf{h} = \mathbf{h}$*

Un ideal de un álgebra de Lie $\mathbf{g}$ es un subespacio $\mathbf{h} \subset \mathbf{g}$ tal que $[X, Y] \in \mathbf{h}$ para todo $X \in \mathbf{g}$ e $Y \in \mathbf{h}$. Ese concepto es semejante al de subgrupos normales y de hecho los ideales del álgebra de Lie de un grupo de Lie corresponden a los subgrupos normales.

Proposición I.39. *Si H es un subgrupo de Lie normal de G entonces su álgebra de Lie $\mathbf{h}$ es un ideal del álgebra de Lie $\mathbf{g}$ de G.*

En general puede ocurrir que las subálgebras de Lie de subgrupos de Lie distintos coincidan. En efecto, si H es un subgrupo de Lie no conexo entonces la componente conexa de la identidad H_0 también es un subgrupo de Lie y tanto H como H_0 tienen la misma álgebra de Lie.

La idea de la construcción de subgrupos a partir de subálgebras viene de las siguientes observaciones. Sea $H \subset G$ un subgrupo de Lie conexo cuya álgebra de Lie es $\mathbf{h}$. Para todo $g \in H$ la traslación a la derecha R_g deja H invariante, esto es, $R_g H \subset H$ y la restricción de R_g a H es un difeomorfismo de H. Luego, para todo $h \in H$ la imagen por $d(R_g)_h$ del espacio tangente $T_h(H) \subset T_h G$ es el subespacio $T_{hg}H$. En particular, el espacio tangente a H en g es $d(R_g)_1\mathbf{h}$. De la misma forma, R_g es un difeomorfismo entre H y las clases laterales Hg y de ahí que el espacio tangente a Hg en G es $d(R_g)_1\mathbf{h}$.

Las expresiones para los espacios tangentes muestran que el subgrupo H, así como las clases laterales Hg, son subvariedades integrales conexas de la distribución $\Delta_{\mathbf{h}}$ en G definida por

$$\Delta_{\mathbf{h}}(g) = d(R_g)_1\mathbf{h}.$$

Esta distribución depende apenas de $\mathbf{h}$. La idea para construir subgrupos de Lie a partir de subálgebras consiste en revertir esos argumentos. Dada una subálgebra de Lie $\mathbf{h}$ se puede definir la distribución $\Delta_{\mathbf{h}}$ y a partir de ahí obtener el subgrupo de Lie conexo asociado a $\mathbf{h}$ como variedad integral maximal de $\Delta_{\mathbf{h}}$.

Proposición I.40. *Sea G un grupo de Lie con álgebra de Lie $\mathbf{g}$. Dada una subálgebra de Lie $\mathbf{h} \subset \mathbf{g}$, definamos la distribución*

$$\Delta_{\mathbf{h}}(g) = d(R_g)_1 \mathbf{h}.$$

Entonces $\Delta_{\mathbf{h}}$ es integrable.

Ahora solo resta probar que la variedad integral maximal que pasa por 1 es un subgrupo de Lie cuya álgebra de Lie es $\mathbf{h}$. Denotemos por H esa variedad. Al ser H una subvariedad conexa quasi-regular, todo se reduce a ver que H es subgrupo de G. Lo cual es hecho utilizando exponenciales de elementos de $\mathbf{h}$. Así por la conexidad de H, para todo $h \in H$ existen $X_1, X_2, ..., X_s \in \mathbf{h}$ tal que

$$h = e^{X_1} \cdots e^{X_s}.$$

Teorema I.41. *Dado un grupo de Lie G con álgebra de Lie $\mathbf{g}$, sea $\mathbf{h} \subset \mathbf{g}$ una subálgebra. Definamos la distribución $\Delta_{\mathbf{h}}$ en G por $\Delta_{\mathbf{h}}(g) = d(R_g)_1 \mathbf{h}$. Entonces, $\Delta_{\mathbf{h}}$ es integrable, la subvariedad integral maximal que pasa por el elemento neutro, es un subgrupo de Lie con álgebra de Lie $\mathbf{h}$ y las demás subvariedades integrales son las clases laterales Hg, $g \in G$. Además, $h \in G$ esta en H si y sólo si existen $X_1, X_2, ..., X_s \in \mathbf{h}$ tal que*

$$h = e^{X_1} \cdots e^{X_s}.$$

Demostración. Dos elementos $h, g \in H$ son producto de exponenciales de elementos de $\mathbf{h}$. Entonces es claro que gh y g^{-1} también son producto de exponenciales de elementos de $\mathbf{h}$ y, por tanto, están en H. El álgebra de Lie de H es el espacio tangente a H en el elemento neutro, que por construcción es $\Delta_{\mathbf{h}}(1) = d(R_1)(\mathbf{h}) = \mathbf{h}$. También por construcción, cada clase lateral Hg es una subvariedad integral conexa de $\Delta_{\mathbf{h}}$. Ellas son maximales, pues si una variedad integral conexa I contiene a Hg entonces $R_{g^{-1}}(I)$ es una variedad integral que contiene a H. Como H es maximal, se sigue que $H = R_{g^{-1}}(I)$, lo que implica que $Hg = I$. □

Corolario I.42. *Sea G un grupo de Lie con álgebra de Lie $\mathbf{g}$. Entonces, para cualquier subálgebra de Lie $\mathbf{h} \subset \mathbf{g}$, existe un único subgrupo de Lie conexo $H \subset G$, cuya álgebra de Lie es $\mathbf{h}$. Además si G es conexo y $\mathbf{h}$ un ideal, entonces H es un subgrupo normal.*

I.9. Subgrupos Cerrados

El teorema del subgrupo cerrado de Cartan asegura que cualquier subgrupo cerrado H de un grupo de Lie G es un subgrupo de Lie. Ese es uno de los resultados de la teoría de grupos de Lie y es ampliamente utilizado en las más diversas situaciones.

La estrategia para demostrar el teorema del subgrupo cerrado consiste en definir una subálgebra de Lie $\mathbf{h} \subset \mathbf{g}$ de tal forma que la componente de la identidad del subgrupo cerrado H sea generado por $\exp \mathbf{h}$. Esa subálgebra es definida como siendo el conjunto de los $X \in \mathbf{g}$ tales que $\exp X \in H$. La parte mas delicada de la demostración esta en verificar que este conjunto es de hecho subálgebra de $\mathbf{g}$. Y eso es demostrado usando algunas fórmulas clásicas, que envuelven productos de exponenciales.

Proposición I.43. *Dado un grupo de Lie G con álgebra de Lie $\mathbf{g}$, sea $H \subset G$ un subgrupo cerrado y defina*

$$\mathbf{h}_H = \{X \in \mathbf{g} : \forall t \in \mathbf{R}, \exp(tX) \in H\}.$$

Entonces, $\mathbf{h}_H$ es subálgebra de Lie de $\mathbf{g}$.

Demostración. El conjunto $\mathbf{h}_H$ es no vacío pues $0 \in \mathbf{h}_H$. Tomemos ahora $X, Y \in \mathbf{h}_H$. Los subgrupos a 1-parámetro generados pos X y aX coinciden si $a \neq 0$, lo que es inmediato de la definición de sus flujos. Por tanto, $aX \in \mathbf{h}_H$, para $a \in \mathbf{R}$. Por la fórmula

$$\exp(X+Y) = \lim_{n\to\infty} (\exp(\frac{X}{n}) \exp(\frac{Y}{n}))^n$$

y aplicando el hecho que exp es continua y H cerrado tenemos que el límite $\exp(X+Y)$ esta en $\mathbf{h}_H$. De la misma manera, la fórmula

$$\exp(-[X,Y]) = \lim_{n\to\infty} (e^{\frac{X}{n}} e^{\frac{Y}{n}} e^{-\frac{X}{n}} e^{-\frac{Y}{n}})^{n^2}$$

garantiza que $[X, Y] \in \mathbf{h}_H$, mostrando así que $\mathbf{h}_H$ es una subálgebra de Lie. $\square$

Una consecuencia inmediata de este hecho, es el siguiente teorema que muestra que los subgrupos cerrados son de Lie.

Teorema I.44. *Todo subgrupo cerrado H de un grupo de Lie G es subgrupo de Lie. Más precisamente; un subgrupo cerrado H admite una estructura de variedad que lo torna un subgrupo de Lie.*

Proposición I.45. *Si $H \subset G$ es un subgrupo regular, entonces H es cerrado.*

Juntando el teorema del subgrupo cerrado y la anterior proposición, tenemos lo siguiente.

Corolario I.46. *Un subgrupo $H \subset G$ es cerrado si y sólo si la subvariedad H es regular.*

I.10. Álgebras de Lie Solubles

Denotaremos por $D\mathbf{g} = [\mathbf{g}, \mathbf{g}]$ para ser el espacio lineal generado por elementos de la forma $[X, Y]$ con $X, Y \in \mathbf{g}$. $D\mathbf{g}$ es una subálgebra de Lie de $\mathbf{g}$ y es llamado *el álgebra derivado de* $\mathbf{g}$. De manera inductiva, definamos $D^p\mathbf{g}$ donde $p \geq 1$ por

$$D^0\mathbf{g} = \mathbf{g}, \qquad D^p\mathbf{g} = [D^{p-1}\mathbf{g}, D^{p-1}\mathbf{g}].$$

Si $\mathbf{a}$ es una subálgebra de $\mathbf{g}$, entonces $D\mathbf{a} = [\mathbf{a}, \mathbf{a}]$ es también una subálgebra de $\mathbf{g}$ y está claro por la definición de encima que está bien definida la sucesión

$$D^0\mathbf{g} \supseteq D^1\mathbf{g} \supseteq ...$$

de subálgebras de $\mathbf{g}$.

$D^p\mathbf{g}$ es llamada la p-ésima algebra derivada de $\mathbf{g}$.

Definición I.47. *Un álgebra de Lie $\mathbf{g}$ es llamada soluble, si existe $p \geq 1$ tal que $D^p\mathbf{g} = 0$.*

Si V es un espacio vectorial, $\{v_1, ..., v_n\}$ una base para V. Sea $\mathbf{g}$ el álgebra de Lie de todos los endomorfismos X de V cuyas matrices en la base $\{v_1, ..., v_n\}$ tienen la forma

$$\begin{pmatrix} \lambda_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ \mathbf{0} & & & \lambda_n \end{pmatrix}.$$

Entonces $\mathbf{g}$ es soluble. Pues $D\mathbf{g}$ está contenida en el álgebra de Lie nilpotente de todos los endomorfismos de V cuyas matrices tienen ceros por debajo de la diagonal; luego $D\mathbf{g}$ es soluble y así $\mathbf{g}$ lo es. Cualquier subálgebra de $\mathbf{g}$ es también soluble. El resultado básico en la teoría de álgebras de Lie solubles es el teorema de Lie que afirma que cualquier álgebra de Lie soluble de matrices puede ser obtenida de la manera explicada encima.

Teorema I.48. *Sea $\mathbf{g}$ un álgebra de Lie soluble y ρ una representación de $\mathbf{g}$ en el espacio vectorial V de dimensión finita n. Entonces existen $\lambda_i \in \mathbf{g}^*$ $(1 \leq i \leq n)$ y una base $\{v_1, ..., v_n\}$ de V tal que para cada $X \in \mathbf{g}$, la matriz de $\rho(X)$ en la base tiene la forma*

$$\begin{pmatrix} \lambda_1(X) & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ \mathbf{0} & & & \lambda_n(X) \end{pmatrix}.$$

En particular, para todo $X \in \mathbf{g}$,

$$\rho(X)v_1 = \lambda_1(X)v_1.$$

La demostración de este teorema puede ser encontrada en [16].

Remarca I.49. *Si $\mathbf{g}$ es un álgebra de Lie con $\dim \mathbf{g} = 2$ no abeliana entonces es soluble. En efecto, por el ejemplo (I.10), existe una base $\{X, Y\}$ tal que $[X, Y] = Y$, luego $D\mathbf{g} = [\mathbf{g}, \mathbf{g}] = \{\lambda Y : \lambda \in \mathbf{R}\}$ y claro $D^2\mathbf{g} = 0$. Por otro lado, en el ejemplo (I.30) exhibimos una representación fiel de $\mathbf{g}$, esto* *es, básicamente el álgebra de Lie es*

$$\mathbf{g} = \{ \begin{pmatrix} a & b \\ 0 & -a \end{pmatrix} : a, b \in \mathbf{R}\}.$$

CAPÍTULO II

Sistemas de Control

En este capítulo se desarrollaran los conceptos, propiedades elementales de la teoría de los sistemas de control, para la comprensión y el seguimiento de este capítulo, incluyendo las demostraciones que no se encuentran explícitas, se recomienda [10], sin embargo también se pueden consultar [6] y [13].

El desarrollo de este capítulo está dividido en tres secciones. En la primera sección definiremos un sistema de control y desarrollaremos las ideas principales de esta teoría. Comenzaremos con el tratamiento y estudio de una familia de campos de vectores definidos sobre una variedad diferenciable, posteriormente definiremos los conjuntos accesibles desde un punto del espacio de estados por dicha familia de campos vectoriales, enunciaremos las propiedades básicas que estos conjuntos poseen cuando la familia de campos vectoriales tienen una propiedad adicional (sistemas Lie-determinados), hablaremos acerca la noción y delicado problema de controlabilidad de un sistema de control y finalmente mostraremos un resultado abstracto sobre la controlabilidad de un sistema.

En la segunda sección realizaremos una descripción de los sistemas lineales y bilineales en $\mathbf{R}^n$, mencionando resultados que han sido obtenidos con respecto a la controlabilidad de dichos sistemas, el más importante sin lugar a duda el de Kalmann para sistemas lineales. Mostraremos con un ejemplo como se puede convertir un sistema lineal no controlable en un sistema bilineal (adecuado) controlable. En la última sección haremos un tratamiento a los sistemas de control sobre grupos de Lie como espacio de estados, y destacaremos las propiedades que se obtienen por considerar dicha clase de espacio estado.

II.1. Nociones de Sistemas de Control

Un sistema de control puede ser visto como un sistema dinámico, cuya dinámica no se mantiene enteramente fija como podemos encontrar en los problemas de física clásica, pero que dependen de parámetros llamados controles, que pueden variar y que serán llamados los controles del sistema.

Es una cuestión muy natural suponer que el espacio de todas las configuraciones del sistema sea una variedad n-dimensional M y que la dinámica del sistema se encuentre descrita por campos vectoriales que (en contraste a la situación clásica) dependen de controles como parámetros. La noción de que el sistema en punto de M pueda seguir un número de direcciones tangentes depende de la elección del control.

Definición II.1. *Un sistema de control esta dado por:*

1. *El espacio estado M, una variedad diferenciable de dimensión n.*

2. *El rango de control $U \subset \mathbf{R}^m$ y el conjunto de funciones de control admisibles*

$$\mathcal{U} = \{u : \mathbf{R} \to U, \textit{localmente integrable}\}.$$

   Esto requiere que los controles son integrables sobre cualquier intervalo acotado; alternativamente, admitiremos solo controles constantes por pedazos u (es decir, $\mathbf{R}$ es descompuesto en subintervalos de longitud acotada por un número positivo tal que u es constante sobre cada subintervalo).

3. *La dinámica está descrita por una función $F : M \times \mathcal{U} \to TM$ tal que para cada $u \in \mathcal{U}$, $F_u : M \to TM$ definida por $F_u(x) = F(x, u)$ para $x \in M$ es un campo vectorial diferenciable.*

Las funciones de control pueden ser de muchas maneras distintas. Un control u se llama *controles lazo cerrado*, si $u : M \to U$. Cuando u es una función diferenciable y cuando F es diferenciable, el correspondiente campo vectorial $x \to F(x, u(x))$ es un campo vectorial diferenciable. Cualquier curva integral de este campo vectorial es llamada una trayectoria lazo cerrado.

Un control u es llamado *cotrol a lazo abierto* si $u : \mathbf{R} \to U$. Controles a lazo abierto son curvas con valores en U. Sus trayectorias son curvas integrales del campo vectorial que varia con el tiempo $x \to F(x, u(t))$. Finalmente, un control puede ser una combinación de ambos tipos, esto es, una función $u : M \times \mathbf{R} \to U$. Las trayectorias correspondientes a tal elección de control son soluciones del sistema diferencial que varia con el tiempo

$$\frac{dx}{dt} = F(x(t), u(x, t)).$$

Cada trayectoria $x(t)$ del sistema diferencial de encima corresponde a una trayectoria generada por un control a lazo abierto $v(t)$, definida por $v(t) = u(x(t), t)$. Además, todos los estados que pueden ser alcanzados por controles a lazo cerrado también pueden ser alcanzados por controles a lazo abierto. Esta es la razón por la que inicialmente se admiten solo controles a lazo abierto.

El interés primario para el estudio de los sistemas de control, se encuentra en los siguientes tópicos:

1. La existencia de una función (control) que pueda transferir el sistema de una configuración inicial dada a una configuarción terminal dada.
2. La longitud de tiempo requerido para alcanzar un estado terminal, y
3. La existencia y propiedades de un control optimal.

Para la primera parte de esta exposición es suficiente considerar solo controles constantes por pedazos que tengan discontinuidades en un subconjunto discreto de $\mathbf{R}$. Esta suposición permite establecer la base geométrica con todas las consideraciones teóricas de medida.

Ejemplo II.2 (Carro Cohete)**.** *Este ejemplo sirve para ilustrar, motivar los conceptos y las propiedades de los sistemas de control. El carro cohete se encuentra sobre una riel y se encuentra equipado con dos motores de cohete uno a cada lado. El problema consiste en mover el carro de* *una ubicación dada a un destino fijo preasignado. Por simplicidad, supondremos que el punto de partida es el origen y denotaremos la posición del centro del carro en un tiempo $t > 0$, por $p(t)$. Si el carro está en una posición p_0 en un tiempo $t = 0$, con velocidad v_0, deseamos encender los dos motores de acuerdo a alguna necesidad (patron, programa) que nos lleve desde p_0 en reposo (con velocidad cero) hasta algún instante $t_1 > 0$. Podemos expresar como un sistema que representa al carro (su posición) y su velocidad; entonces como estado del sistema podemos tomar el vector*

$$x(t) = (p(t), \dot{p}(t));$$

el estado inicial (p_0, v_0) lo asumiremos que ya se encuentra fijado. El estado objetivo (donde queremos llegar) es $(0, 0)$.

Un control es una función real, que para nosotros representará la fuerza ejercida sobre el carro debido a alguno de los motores en el tiempo t. Si encendemos el motor de la derecha en un tiempo t^*, diremos que la fuerza es negativa, si usamos el motor de la izquierda tenemos fuerza positiva. Luego la dinámica del sistema está dada por la Ley de Newton

$$F = ma,$$

entonces podemos escribir $\ddot{p}(t) = u(t)$. En forma vectorial, podemos presentar al sistema descrito anteriormente por

$$x(t) = \begin{pmatrix} p(t) \\ \dot{p}(t) \end{pmatrix}, \qquad \dot{x}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + u(t) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Ahora de manera natural, existen restricciones sobre la magnitud de $u(t)$, basada sobre el tamaño del motor del cohete y el aumento de la aceleración en el carro.

Una suposición matemática muy razonable es pedir que la función $u(t)$ sea medible y acotada, luego por simplicidad podemos suponer $|u(t)| \leq 1$.

Como es de nuestro conocimiento, las funciones medibles pueden ser muy abstractas, entonces buscamos una clase mas simple de funciones que son básicamente razonables, estos son los controles *constantes por pedazos.*

Recordemos que en nuestro análisis, una función de control $u(t)$ es un motor. Por ejemplo, el control

$$u(t) = \begin{cases} +1, & 0 \leq t \leq 1 \\ -\frac{1}{2}, & 1 < t \leq 3, \end{cases}$$

que para nosotros representa lo siguiente: primero encendemos el motor de la izquierda con una fuerza completa en una unidad de tiempo y luego encendemos el motor de la derecha con la mitad de fuerza por dos unidades de tiempo.

Si $x_0 = (p_0, v_0)$ es la posición del carro en $t = 0$, integrando dos veces la ecuación diferencial y luego haciendo integración por partes tenemos:

$$\dot{p}(t) = v_0 + \int_0^t u(r)dr, \qquad p(t) = p_0 + v_0 + \int_0^t (t-r)u(r)dr.$$

Entonces la elección de un control $u(t)$ genera una solución

$$x[t] = x(t; x_0, u(t)).$$

Si la solución $x[t]$ alcanza el objetivo $(0,0)$ en algún $t_1 > 0$, entonces diremos que el control $u(t)$ es exitoso. Ahora, estos controles pueden existir o no. Cuando existen muchos de estos controles, la elección de uno u otro es realizada de acuerdo a la practicabilidad y/o por un costo o un funcionamiento. Por ejemplo, podemos considerar los criterios, (a) tiempo mínimo, (b) menor energía usada, (c) menor uso de gasolina. Luego el problema de control se convierte en un problema de control óptimo.

Usaremos la notación $(p(t), q(t))$ para denotar la posición y la velocidad en un instante t del carro cohete, luego la dinámica del sistema está descrita por

$$\dot{p}(t) = q(t), \qquad \dot{q}(t) = u(t), \qquad -1 \leq u(t) \leq 1,$$

con $u(t)$ medible, con (p_0, q_0) en $t = 0$. El estado objetivo es $(0,0)$, esto es, $(p(t_1), q(t_1)) = (0,0)$

para algún $t_1 > 0$.

Si $u(t) = 1$ en algún intervalo de tiempo partiendo de $t = 0$, entonces

$$(+) \qquad \dot{p} = q \qquad \dot{q} = 1 \Rightarrow q\dot{q} = p \Rightarrow (q(t))^2 - q_0^2 = 2(p(t) - p_0);$$

si $u(t) = -1$ sobre tal intervalo, entonces

$$(-) \qquad \dot{p} = q \qquad \dot{q} = -1 \Rightarrow q\dot{q} = -p \Rightarrow (q(t))^2 - q_0^2 = -2(p(t) - p_0).$$

Por lo tanto las soluciones de nuestro sistema son parábolas

$$\pm 2p = q^2 + \alpha$$

donde α es constante.

Para describir el conjunto accesible $\mathcal{A}(t_1, x_0)$, con t_1 fijo y tomado $x_0 = (p_0, 0)$ por simplicidad. Tenemos dos trayectorias correspondientes a la elección $u(t) = -1$ sobre $[0, t_1]$, y $u(t) = 1$ sobre $[0, t_1]$.

La descripción del conjunto controlable $\mathcal{C}$, para el carro cohete, es decir, aquellos estados iniciales que pueden ser trasladados al estado cero por controles constantes por pedazos bang-bang

$$\mathcal{C} = \{x_0 / \exists u(\cdot), t_1 > 0 : x(t_1; x_0, u(\cdot)) = 0\}.$$

Tenemos que $\mathcal{C} = \mathbf{R}^2$. Primero observemos la trayectoria $q^2 = 2p$ con $u(t) = 1$ para $t \in [0, \infty)$ que pasa por $(0, 0)$; similarmente la trayectoria $q^2 = -2p$ con $u(t) = -1$ para $t \in [0, \infty)$. Esas dos trayectorias (+),(-) respectivamente que pasa por $(0, 0)$, nos conducen desde cualquier estado inicial $x_0 = (p_0, v_0)$ hacia el origen, decimos entonces que el sistema es controlable.

El siguiente ejemplo nos da importantes aplicaciones y muestra el contexto de propiedades geométricas y consideraciones teóricas.

Ejemplo II.3 (Control de la ecuación de Liénard's). *Un oscilador no lineal general con una fuerza* *externa u está descrita por la ecuación*

$$\frac{d^2q}{dt^2} + f(q)\frac{dq}{qt} + g(q) = u(q, \frac{dq}{dt}, t).$$

Como una costumbre de la literatura, esta ecuación puede ser expresada como un sistema en el plano introduciendo las coordenadas

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

definido por $x_1 = q$ *y* $x_2 = \frac{dq}{dt}$. *Entonces* $\frac{dx_1}{dt} = x_2$, *y* $\frac{dx_2}{dt} = \frac{d^2q}{dt^2} = -f(x_1)x_2 - g(x_1) + u$. *Si* X *denota el campo vectorial con coordenadas*

$$\begin{pmatrix} x_2 \\ -f(x_1)x_2 - g(x_1) \end{pmatrix},$$

e Y *denota el campo vectorial*

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

entonces el sistema precedente será

$$\frac{dx}{dt} = X(x) + u(x,t)Y(x)$$

con

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Y podemos asumir que la fuerza externa u juega el papel de un control.

Ejemplo II.4 (Sistema mecánico)**.** *Consideremos ahora el problema de controlabilidad de un sistema mecánico*

$$\frac{d^2q}{dt^2} + u(\frac{dq}{dt}) + g(q) = 0$$

por una función de control u. El sistema equivalente de primer orden está dado por

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -g(x) - ux_2. \tag{II.1}$$

Por simplicidad, asumiremos que $g(x) = kx$ para alguna constante k. Entonces el sistema mencionado puede ser prescrito como

$$\frac{dx}{dt} = (A + uB)x$$

para ciertas matrices

$$A = \begin{pmatrix} 0 & 1 \\ -k & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix},$$

con

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Para cualquier función de control constante u, el sistema resultante es lineal. El polinomio característico de $A + uB$ es igual a

$$\lambda^2 + \lambda u + k = 0.$$

Caso 1 si $k > 0$. Las curvas integrales que corresponden a $u = 0$ son elipses concéntricas centradas en el origen. Sea u_1 un número tal que los autovalores son negativos, y u_2 un número para el cual los autovalores sean positivos. El sistema lineal $A + u_1B$ es una pila o fregadero, mientras que $A + u_2B$ es una fuente. Evidentemente cualquier estado inicial no nulo a puede ser transferido a un estado b no nulo por un control constante a trozos que tome solo los valores $0, u_1$ y u_2.
Caso 2 $k \leq 0$. Sea $x(t)$ una trayectoria de II.1 generado por un control $u(t)$. Denotemos por $\phi(t)$ el producto $\phi(t) = x_1(t)x_2(t)$. Entonces

$$(\frac{d}{dt})\phi + u\phi = x_2^2 - kx_1^2.$$

Puesto que $k \leq 0$, el lado derecho del sistema de encima es siempre no negativo. Por tanto $\phi(t) \geq 0$ para todo $t \geq 0$ siempre que $\phi(0) \geq 0$, como

$$\phi(t) = e^{-\int_0^t u}(\phi(0) + \int_0^t e^{\int_0^t u}(x_2^2 - kx_1^2))ds.$$

Luego, el primer cuadrante $x_1 \geq 0, x_2 \geq 0$ es un conjunto invariante que no importa en la elección de la función de control u. Este ejemplo muestra que existen estados que no pueden ser alcanzados de un estado inicial a, incluso cuando se intente forzar una función de control.

II.1.1. Familias de campos vectoriales y sistemas de control

Sea $F : M \times \mathcal{U} \to TM$ un sistema de control, como fue descrito en la sección anterior. Para todo $u \in \mathcal{U}$, F_u es el campo vectorial correspondiente. Usaremos la notación $\mathcal{F}$ para denotar la familia de campos vectoriales $\mathcal{F} = \{F_u : u \in \mathcal{U}\}$ generado por F. Lo primero que destacaremos es relativo a las trayectorias correspondientes a controles constantes por pedazos con valores en U de las curvas integrales de elementos de $\mathcal{F}$.

Definición II.5. *Una curva continua $x(t)$ en M, definida sobre un intervalo $[0, T]$, es llamada una curva integral de $\mathcal{F}$ si existe una partición $0 = t_0 < t_1 < \cdots < t_m = T$ y campos vectoriales $X_1, \cdots, X_m$ en $\mathcal{F}$ tal que la restricción de $x(t)$ a cada intervalo abierto (t_{i-1}, t_i) es diferenciable, y $\frac{dx(t)}{dt} = X_i(x(t))$ para todo $i = 1, \cdots, m$.*

Los elementos de $\mathcal{F}$ son parametrizados por controles, se sigue que cada X_i es igual a F_{u_i} para algún $u_i \in U$. Luego $x(t)$ es la curva solución del campo vectorial $F(x, u(t))$ que varia con el tiempo, con $u(t)$ igual a controles constantes por pedazos, con valores constantes u_i en el intervalo $[t_{i-1}, t_i]$, y $x(t)$ puede ser visto como una curva continua que consiste en la "soldadura" de pedazos de curvas integrales de campos vectoriales correspondientes a la elección de controles diferentes.

En cualquier familia de campos vectoriales, el objeto básico de estudio son los conjuntos accesibles.

Definición II.6. *Sea M una variedad diferenciable y $\mathcal{F}$ un sistema de control definido sobre M.*

1. *Para cada $T > 0$, y cada x_0 en M, el conjunto de puntos accesibles desde x_0 en el tiempo T, denotado por $\mathcal{A}(x_0, T)$, es igual al conjunto de puntos terminales $x(T)$ de curvas integrales de $\mathcal{F}$ que empiezan en x_0.*

2. *La unión de $\mathcal{A}(x_0, T)$, para $T \geq 0$, es llamado el conjunto accesible desde x_0. A este conjunto lo denotaremos por $\mathcal{A}(x_0)$.*

También a veces es conveniente considerar, conjuntos accesibles en T unidades de tiempo o menos, a estos conjuntos los denotaremos por

$$\mathcal{A}(x_0, \leq T).$$

Los conjuntos accesibles admiten muchas descripciones geométricas que las destacamos formalmente a continuación.

Asumiendo que los campos vectoriales pertenecientes a $\mathcal{F}$ son completos, entonces cada elemento X en $\mathcal{F}$ genera un grupo uno-parámetro de difeomorfismos

$$\{\exp tX : t \in \mathbf{R}\}.$$

Sea $G_{\mathcal{F}}$ el subgrupo del grupo de difeomorfismos en M generado por la unión de

$$\{\exp tX : t \in \mathbf{R}, X \in \mathcal{F}\}.$$

Cada elemento Φ de $G_{\mathcal{F}}$ es un difeomorfismo de M de la forma

$$\Phi = (\exp t_k X_k)(\exp t_{k-1} X_{k-1}) \cdots (\exp t_1 X_1)$$

para ciertos números reales $t_1, ..., t_k$ y campos vectoriales $X_1, ..., X_k$ en $\mathcal{F}$.

Definición II.7. *El grupo del sistema $G_{\mathcal{F}}$ y el semigrupo del sistema $S_{\mathcal{F}}$ generado por la familia $\mathcal{F}$ son definidos, respectivamente, como,*

$$G_{\mathcal{F}} = \{\Phi, X_i \in \mathcal{F}, t_i \in \mathbf{R}, k \in \mathbf{N}\},$$

$$S_{\mathcal{F}} = \{\Phi, X_i \in \mathcal{F}, t_i \geq 0, k \in \mathbf{N}\},$$

donde Φ es como encima.

Entonces el conjunto accesible desde x_0 en el tiempo T consiste de todos los puntos $\Phi(x_0)$ correspondiente a elementos Φ de $G_{\mathcal{F}}$ que pueden ser expresados como

$$\Phi = (\exp t_k X_k)(\exp t_{k-1} X_{k-1}) \cdots (\exp t_1 X_1),$$

con $t_1 \geq 0$, $t_2 \geq 0$,...,$t_k \geq 0$, $t_1 + t_2 + \cdots + t_k = T$, y $X_1, X_2, ..., X_k$ en $\mathcal{F}$.

Los otros conjuntos accesibles tienen descripciones análogas. En particular, el conjunto accesible desde x_0 por $\mathcal{F}$ es igual a la órbita del semigrupo $S_{\mathcal{F}}$ desde x_0, esto es,

$$\mathcal{A}(x_0) = S_{\mathcal{F}}(x_0).$$

La órbita de $S_{\mathcal{F}}$ desde x_0, que la escribimos como $S_{\mathcal{F}}(x_0)$, es igual a $\{\Phi(x_0) : \Phi \in S_{\mathcal{F}}\}$.

Definición II.8. *Para un sistema de control con campos vectoriales en $\mathcal{F}$ sobre el espacio de estados M y $x \in M$ el conjunto*

$$\mathcal{O}_{\mathcal{F}}(x) = \{y \in M : \exists \Phi \in G_{\mathcal{F}}, y = \Phi(x)\},$$

es la órbita desde $x \in M$, y

$$\mathcal{O}^+_{\mathcal{F}}(x) = \{y \in M : \exists \Phi \in S_{\mathcal{F}}, y = \Phi(x)\},$$

$$\mathcal{O}^-_{\mathcal{F}}(x) = \{y \in M : \exists \Phi \in S_{\mathcal{F}}, x = \Phi(y)\},$$

son la órbita positiva y negativa desde $x \in M$.

Por definición, la órbita $\mathcal{O}_{\mathcal{F}}(x)$ coincide con M para algún $x \in M$ (y luego para cada $x \in M$) si y solo si el grupo del sistema $G_{\mathcal{F}}$ actúa transitivamente sobre M.

Con estas notaciones tenemos la igualdad

$$\mathcal{O}^+_{\mathcal{F}}(x) = S_{\mathcal{F}}(x_0).$$

Definición II.9 (Accesibilidad)**.** *Un sistema de control es accesible desde $x \in M$ si $\mathcal{O}^+(x)$ y*

$\mathcal{O}^{-}(x)$ tienen interior no vacío en M. El sistema es llamado localmente accesible desde $x \in M$ si

$$int\mathcal{O}^{+}_{\leq T}(x) \neq \emptyset, \quad y \quad int\mathcal{O}^{-}_{\leq T}(x) \neq \emptyset,$$

para todo $T > 0$.

Si esas propiedades son válidas para todo $x \in M$ diremos que el sistema es accesible o localmente accesible, respectivamente.

Ejemplo II.10. *Sea $M = \mathbf{R}^n$, y $\mathcal{F} = \{X, Y\}$, donde X e Y son campos vectoriales constantes. Sea $X(x) = a$, $Y(x) = b$ para todo $x \in \mathbf{R}^n$. Entonces $(\exp tX)(x) = x + ta$ y $(\exp tY)(x) = x + tb$. Denotemos por G el subgrupo del grupo de difeomorfismos generado por las exponenciales en $\mathcal{F}$. Se sigue que G es igual al grupo de traslaciones de $\mathbf{R}^n$ por los elementos en el espacio vectorial generado por a y b.*
Para cada $x \in \mathbf{R}^n$, la orbita Gx es igual a $x + V$, donde V es el espacio vectorial generado por los vectores a y b. $S_{\mathcal{F}}(x)$ es el cuadrante positivo en V dado por $x + \alpha a + \beta b$ con α y β variando en los números reales no negativos. Para cada $T > 0$, el conjunto accesible desde x en el tiempo es el segmento de recta $\{x + \alpha a + (T - \alpha)b : 0 \leq \alpha \leq T\}$.

Ejemplo II.11. *Sea $M = \mathbf{R}^n$, y $\mathcal{F} = \{X, Y\}$, donde X es un campo lineal $X(x) = Ax$ para alguna matriz A, e Y es un campo vectorial constante dado por $Yx = a$.*
Sea G el grupo generado por $\{\exp tX : t \in \mathbf{R}\} \cup \{\exp tY : t \in \mathbf{R}\}$. Se sigue que

$$((\exp tX) \circ (\exp tY) \circ (\exp -tX))(x) = e^{tA}(e^{-tA}x + sa) = x + s(e^{tA}(a)).$$

Luego G incluye el grupo de G_0 de las traslaciones en la dirección $se^{tA}a$ para cada t y s en $\mathbf{R}$. Denotemos por C el espacio vectorial generado por $\{e^{tA} : t \in \mathbf{R}\}$. Mostramos consecuentemente que C es el espacio lineal generado por $\{a, Aa, A^2a, ..., A^{n-1}a\}$. Se sigue que $G_0 = \{T : Tx = x + c, x \in \mathbf{R}^n, c \in C\}$. Cualquier elemento de G es de la forma $T \circ \exp tA$ para algún $t \in \mathbf{R}$ y $T \in G_0$. La órbita de G desde el origen es igual a C, y la órbita desde cualquier punto arbitrario $x \in \mathbf{R}^n$ es la variedad $\{\exp tAx + c : t \in \mathbf{R}, c \in C\}$.

Una propiedad básica de las familias de campos vectoriales es que sus órbitas son variedades. Este hecho es conocido como el teorema de la órbita, que es un punto de apertura para la teoría de control geométrica, la demostración de este resultado pude ser encontrado en [10].

Teorema II.12 (El teorema de la orbita)**.** *La órbita de $\mathcal{F}$ desde cualquier punto $x \in M$ es una subvariedad conexa de M.*

Para cualquier familia de campos vectoriales diferenciables $\mathcal{F}$ y para cualquier órbita N de $\mathcal{F}$, cada campo vectorial X en $\mathcal{F}$ es tangente a N. En la siguiente proposición damos una condición de tangencia mas general.

Proposición II.13. *Supongamos que X e Y son campos vectoriales cualesquiera sobre M. Entonces*

1. $(d/dt)(\exp t\alpha X) \circ (\exp t\beta Y)(x)|_{t=0} = \alpha X(x) + \beta Y(x)$ *para cualquier $x \in M$, y*
2. $(d/dt)((\exp -\sqrt{t}Y) \circ (\exp -\sqrt{t}X) \circ (\exp \sqrt{t}Y) \circ (\exp \sqrt{t}X(x)))|_{t=0} = [X, Y](x)$.

Con respecto a la tangencia a subvariedades integrales sobre M tenemos el siguiente resultado.

Proposición II.14. *Si X e Y son campos vectoriales tangentes a la subvariedad N, entonces su corchete de Lie es también tangente a N, y también cualquier combinación lineal $\alpha X + \beta Y$.*

Usando un proceso inductivo y de la proposición **??** que si $\mathcal{F}$ es tangente a N, también lo es $Lie(\mathcal{F})$, que representa el álgebra de Lie generada por $\mathcal{F}$. Está observación nos da inmediatamente el próximo teorema.

Teorema II.15. *Supongamos que $\mathcal{F}$ es tal que $Lie_x(\mathcal{F}) = T_xM$ para algún x en M. Entonces la órbita $G(x)$ de $\mathcal{F}$ desde x es abierta. Si además, $Lie_y(\mathcal{F}) = T_yM$ para cada $y \in M$, y M es conexo, entonces existe solo una órbita de $\mathcal{F}$ que es igual a M.*

Demostración. La órbita de $\mathcal{F}$ desde cada punto $x \in M$ es una subvariedad de M, y además su espacio tangente es de dimensión constante en todos sus puntos. En particular si $Lie_x(\mathcal{F}) = T_xM$, entonces el espacio tangente de $G(x)$ en x es igual a T_xM, ya que $Lie_x(\mathcal{F})$ es tangente a $G(x)$ en x. Por tanto, $\dim G(x) = \dim M$, y luego $G(x)$ es abierto en M. Si cada órbita es abierta, entonces cada órbita es también cerrado. Puesto que M es conexo, cada órbita es igual a M, con eso termina la prueba. □

En está terminología tenemos una nueva versión del teorema de Frobenius.

Definición II.16. *Una familia $\mathcal{F}$ se dice involutiva si para cualesquiera campos vectoriales $X, Y \in \mathcal{F}$, $[X, Y](x)$ está contenido en el espacio vectorial generado por los elementos de $\mathcal{F}(x)$ para cada $x \in M$. $\mathcal{F}(x)$ es la evaluación de $\mathcal{F}$ en x igual a $\{V(x) : V \in \mathcal{F}\}$.*

Teorema II.17 (Frobenius)**.** *Sea $\mathcal{F}$ una familia involutiva de campos vectoriales diferenciables para el cual la dimensión del espacio lineal generado por $\mathcal{F}(x)$ es constante para cada $x \in M$. Entonces el espacio tangente en un punto x de una orbita de $\mathcal{F}$ es igual al espacio lineal generado por $\mathcal{F}(x)$.*

Corolario II.18. *Sea $\mathcal{F}$ una familia de campos vectoriales diferenciables tal que la dimensión de cada espacio vectorial $Lie_x(\mathcal{F})$ es constante cuando x varia sobre M. Sea k la dimensión de $Lie_x(\mathcal{F})$. Entonces para cada $x \in M$, el espacio tangente a x de la órbita de $\mathcal{F}$ desde x coincide con $Lie_x(\mathcal{F})$. Consecuentemente, cada órbita de $\mathcal{F}$ es una subvariedad k-dimensional de M.*

Ahora definiremos la controlabilidad de un sistema de control.

Definición II.19 (Controlabilidad)**.** *Sea $\mathcal{F}$ un sistema de control definido sobre la variedad diferenciable M.*

1. *Diremos que $\mathcal{F}$ es fuertemente (completamente) controlable si para cualquier $T > 0$ cualquier punto de M es accesible desde cualquier otro por $\mathcal{F}$ en T o menos unidades de tiempo.*

2. *$\mathcal{F}$ es controlable si cualquier punto de M es accesible desde cualquier otro punto de M.*

Puesto que el problema de controlabilidad está íntimamente relacionada con los conjuntos accesibles, estamos interesados en buscar conjuntos accesibles de familias de campos de vectores que tengan representaciones geométricas muy interesantes. Para ello, requerimos la noción de familia de campos vectoriales Lie-determinados.

Definición II.20. *Una familia de campos vectoriales diferenciables $\mathcal{F}$ es Lie-determinado si el espacio tangente de cada punto x en una órbita de $\mathcal{F}$ coincide con la evaluación de x del álgebra de Lie generada por $\mathcal{F}$.*

La propiedad esencial de los sistemas Lie-determinados $\mathcal{F}$, que lo hace merecedor de este nombre, es que la estructura de sus órbitas está determinada por las propiedades locales de los elementos de $\mathcal{F}$ y sus derivadas de Lie.

El significado de los sistemas Lie-determinados también lo encontramos claramente en las propiedades topológicas de sus conjuntos accesibles. Para tales sistemas, los conjuntos accesibles $\mathcal{A}(x, \leq T)$ tienen interior no vacío en la topología de la orbita que es una variedad.

Para describir las propiedades topológicas de los conjuntos accesibles, usaremos la notación $int(\mathcal{A})$ para denotar el interior del conjunto $\mathcal{A}$, y usaremos $cl(\mathcal{A})$ para denotar la clausura topológica.

Proposición II.21. *Supongamos que $\mathcal{F}$ es una familia de campos de vectores diferenciables sobre M tal que $Lie_x(\mathcal{F}) = T_xM$ para algún x en M. Entonces para cada $T > 0$ y cada $\epsilon > 0$,*

1. $\mathcal{A}_{\mathcal{F}}(x, \leq T) \subset cl(int(\mathcal{A}_{\mathcal{F}}(x, \leq T)))$.
2. $int(cl(\mathcal{A}_{\mathcal{F}}(x, \leq T))) \subset int(\mathcal{A}_{\mathcal{F}}(x, \leq T + \epsilon))$.
3. $int(cl(\mathcal{A}_{\mathcal{F}}(x))) = int(\mathcal{A}_{\mathcal{F}}(x))$.

Como una consecuencia de esta proposición, tenemos el siguiente corolario.

Corolario II.22. *Sea N una órbita de una familia de campos vectoriales Lie-determinada $\mathcal{F}$. Entonces para cada $x \in N$, para cada $T > 0$, y cualquier $\epsilon > 0$,*

1. $\mathcal{A}_{\mathcal{F}}(x, \leq T) \subset cl(int(\mathcal{A}_{\mathcal{F}}(x, \leq T)))$, *y*
2. $int(cl(\mathcal{A}_{\mathcal{F}}(x, \leq T))) \subset int(\mathcal{A}_{\mathcal{F}}(x, \leq T + \epsilon))$, *con el interior y la clausura tomadas en la topología de la órbita N.*

Ejemplo II.23. *Sea M el circulo unidad S^1, y X el campo vectorial rotación $X(x_1, x_2) = (-x_2, x_1)$. Denotaremos por f la función monótona no creciente que satisface*

$$\lim_{u \to -\infty} f(u) = -\frac{1}{2} \text{ y } \lim_{u \to \infty} f(u) = \frac{1}{2}.$$

Consideremos $\mathcal{F} = \{f(u)X : u \in \mathbf{R}\}$. Entonces

$$\mathcal{A}_{\mathcal{F}}(x, \leq 1) = S^1 - \{-x\}$$

para cualquier $x \in S^1$. *Luego* $int(cl(\mathcal{A}_{\mathcal{F}}(x, \leq 1))) = S^1$, *el cual no está contenido en* $int(\mathcal{A}_{\mathcal{F}}(x, \leq 1))$.

Corolario II.24. *Para sistemas Lie-determinados, los conjuntos accesibles* $\mathcal{A}_{\mathcal{F}}(x)$ *no pueden ser densos en una órbita de* $\mathcal{F}$ *sin ser igual a la órbita entera.*

Definición II.25. *Sea* $\mathcal{F}$ *una familia de campos vectoriales Lie-determinado.*

1. *El Lie-saturado fuerte de* $\mathcal{F}$ *es el más grande subconjunto* $\mathcal{F}'$ *de* $Lie(\mathcal{F})$ *con la propiedad que* $cl(\mathcal{A}_{\mathcal{F}'}(x, \leq T)) = cl(\mathcal{A}_{\mathcal{F}}(x, \leq T))$ *para cada* $x \in M$ *y cada* $T > 0$. *Lo denotaremos por* $\mathcal{LS}_s(\mathcal{F})$.

2. *El Lie-saturado de* $\mathcal{F}$, *denotado por* $\mathcal{LS}(\mathcal{F})$, *es el más grande subconjunto* $\mathcal{F}'$ *de* $Lie(\mathcal{F})$ *tal que* $cl(\mathcal{A}_{\mathcal{F}'}(x)) = cl(\mathcal{A}_{\mathcal{F}}(x))$ *para todo* x *en* M.

El siguiente teorema nos da un criterio abstracto de controlabilidad.

Teorema II.26. *Supongamos que* $\mathcal{F}$ *es una familia de campos vectoriales Lie-determinado. Entonces* $\mathcal{F}$ *es fuertemente controlable si y sólo si el Lie-saturado fuerte es igual a* $Lie(\mathcal{F})$ *y* $Lie_x(\mathcal{F}) = T_xM$ *para cada* $x \in M$. $\mathcal{F}$ *es controlable si y sólo si el Lie-saturado es igual a* $Lie(\mathcal{F})$ *y* $Lie_x(\mathcal{F}) = T_xM$.

Demostración. Evidentemente, si $\mathcal{F}$ es (fuertemente) controlable entonces el Lie-saturado (fuerte) es igual a $Lie(\mathcal{F})$. Recíprocamente, si el Lie-saturado fuerte es igual a $Lie(\mathcal{F})$ y si $Lie_x(\mathcal{F}) = T_xM$, entonces $cl(\mathcal{A}_{\mathcal{F}}(x, \leq T)) = M$ para cada $x \in M$ y $T > 0$. Pero entonces, de acuerdo al teorema II.21,

$$M = int(cl(\mathcal{A}_{\mathcal{F}}(x, \leq T))) \subset \mathcal{A}_{\mathcal{F}}(x, \leq T + \epsilon).$$

Ya que $\epsilon > 0$ es arbitrario, $\mathcal{A}_{\mathcal{F}}(x, \leq T) = M$ para cada $T > 0$ y cada $x \in M$. El argumento es el mismo cuando sucede $cl(\mathcal{A}_{\mathcal{F}}(x)) = M$, esto es, cuando el Lie-saturado de $\mathcal{F}$ es igual a $Lie(\mathcal{F})$ y $Lie_x(\mathcal{F}) = T_xM$ para todo x. $\square$

II.2. Sistemas de Control Lineales y Bilineales en $\mathbf{R}^n$

II.2.1. Sistemas lineales

Si $\mathcal{L}$ es un sistema lineal de control de tiempo no acotado y conjunto de controles constantes por pedazos en $\mathbf{R}^n$, entonces $\mathcal{L}$ está determinado por la dinámica de la ecuación diferencial

$$\dot{x} = Ax + Bu,$$

donde $x \in \mathbf{R}^n$, $u \in U$ (controles constantes por pedazos), A y B son matrices de dimensiones apropiadas. Denotamos por $b_1, b_2, ..., b_m$ las columnas de B, el correspondiente sistema lineal es dado por

$$\frac{dx}{dt} = Ax + \sum_{i=1}^{m} u_i b_i. \tag{II.2}$$

El sistema anterior es obtenido también de manera mas compacta como

$$\frac{dx}{dt} = Ax + Bu.$$

con u denotando el vector columna de controles

$$\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix},$$

y B denota la función lineal de $\mathbf{R}^n$ en M, dado por

$$Bu = \sum_{i=1}^{m} u_i b_i.$$

Los sistemas lineales están estrechamente conectados con las ecuaciones diferenciales no homogéneas de orden n con coeficientes constantes

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = u(t).$$

En esta notación y es una función real y $y^{(k)}$ denota la k-ésima derivada. La ecuación anterior modela un sistema lineal eléctrico o mecánico controlado por una fuerza externa u. La ecuación anterior puede ser convertido en un sistema de primer orden en $\mathbf{R}^n$ via la transformación

$$x_1 = y, \;\; x_2 = y^{(1)}, \cdots, x_n = y^{(n-1)}.$$

Entonces

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = x_3, \cdots, \frac{dx_{n-1}}{dt} = x_n,$$

$$\frac{dx_n}{dt} = -a_1 x_n - a_2 x_{n-1} - \cdots - a_n x_1 + u.$$

Denotando por x el vector columna

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

el sistema diferencial anterior puede ser escrito como la equación (1.1), con $m = 1$, $b = e_n$, y

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ & & & 1 \\ -a_n & \cdots & & -a_1 \end{pmatrix}.$$

Recíprocamente, para cualquier solución

$$x(t) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

del sistema de control afín precedente, la primera coordenada $x_1(t)$ es una solución de la ecuación diferencial de orden n descrita encima.

Es necesario comentar que en muchas ocasiones, sistemas lineales de control son simplemente ecuaciones diferenciales de orden n sin distinguirlas, como indica el siguiente teorema.

Teorema II.27. *Sea* $\dfrac{dx}{dt} = Ax + bu$ *un sistema de control lineal para el cual* $b, Ab, ..., A^{(n-1)}b$ *forma una base para el espacio estado* M*. Entonces existe un sistema lineal de coordenadas en* M *en el cual* A *es igual a la matriz*

$$\begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ & & & 1 \\ -a_n & \cdots & & -a_1 \end{pmatrix}, \quad y \quad b = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Sea U un subconjunto compacto de $\mathbf{R}^n$, denotaremos por $\mathcal{A}(x, U, T)$ el conjunto de puntos accesibles desde x en exactamente T unidades de tiempo por las trayectorias de (II.2) generado por controles medibles sobre $[0, T]$ que toman valores en U. Con está terminología podemos enunciar el siguiente teorema:

Teorema II.28. *Si* U *es compacto entonces* $\mathcal{A}(x, U, T)$ *es compacto para cada* $x \in \mathbf{R}^n$ *y cualquier* $T > 0$.

Este teorema demuestra que, controlabilidad fuerte no es posible cuando los controles toman valores en un subconjunto compacto U, pues entonces cada uno de los conjuntos accesibles $\mathcal{A}(x, U, T)$ son compactos. Entonces el interés radica en buscar condiciones sobre A y $b_1, ..., b_m$ que garanticen la controlabilidad cuando los controles $u = (u_1, ..., u_m)$ toman valores en un conjunto compacto U.

Antes de hacer los detalles y presentar el resultado, consideraremos dos ejemplos que son típicos del resultado general.

Ejemplo II.29. *Sea*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad y \quad b = e_2.$$

Entonces $\{b, Ab\} = \{e_1, e_2\}$, *y por tanto la condición de controlabilidad se satisface. Supongamos que la magnitud del control* u *está acotada por alguna constante* $K < \infty$. *Ahora, el sistema de control está dado por*

$$\frac{dx_1}{dt} = x_1 + x_2, \quad \frac{dx_2}{dt} = u(t).$$

Supongamos que $x_2(0) = 0$. *Entonces* $|x_2(t)| \leq Kt$ *y además* $|\int_0^t e^{-s} x_2(s)ds| \leq K \int_0^t e^{-s} ds \leq K$. *Puesto que*

$$x_1(t) = e^t(x_1(0) + \int_0^t e^{-s} x_2(s)ds),$$

se sigue que $x_1(t) > 0$, *la elección del control* u *no importa siempre que* $x_1(0) > K$. *Por tanto el sistema es controlable.*

Ejemplo II.30. *Sea*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y \quad b = e_2.$$

Supongamos que $\epsilon > 0$ *es cualquier número real, y sea* $U = \{u : |u| \leq \epsilon\}$. *El sistema de control está dado por*

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} = u.$$

Si $u(t) = \pm\epsilon$, *entonces las soluciones correspondientes son* $x_1(t) = x_1(0) + x_2(0)t \pm \epsilon(t^2/2)$ *y* $x_2(t) = x_2(0) \pm \epsilon t$. *Las trayectorias correspondientes son parábolas*

$$x_1 - c_1 = \pm(1/2\epsilon)(x_2^2 - c_2^2),$$

donde $(c_1, c_2) = (x_1(0), x_2(0))$. *En este caso tenemos que dos puntos cualesquiera en* $\mathbf{R}^2$ *pueden ser conectados por esas parábolas, y por tanto el sistema es controlable.*

Teorema II.31. *Supongamos que el conjunto de control* U *es compacto. Entonces una condición necesaria para la controlabilidad es que cualquier autovalor de* A *tenga parte real igual a cero.*

Gracias a está condición espectral podemos enunciar el siguiente resultado que asegura controlabilidad.

Teorema II.32. *Supongamos que A es tal que todos sus autovalores tienen parte real igual a cero. Sea U cualquier conjunto de control que es una vecindad del origen en $\mathbf{R}^n$. Entonces el sistema de control lineal cuyos controles están en U es controlable cuando $\bigcup_{k=1}^{n-1}\{A^k b_j : j = 1, ..., m\}$ genera M.*

En el siguiente resultado daremos una condición necesaria para la controlabilidad fuerte, esto surge precisamente puesto que para la controlabilidad fuerte es muy común que se requiera un control infinito, pero está condición en la práctica a veces es difícil de verificar.

Teorema II.33. *El sistema lineal de control II.2 con u acotado no es completamente controlable si los autovalores de A tienen parte real negativa.*

Sin lugar a duda el resultado más importante para la controlabilidad de sistemas lineales sobre $\mathbf{R}^n$ está dado en el siguiente teorema

Teorema II.34 (Kalman). *Sea M un espacio vectorial, y sea A un campo vectorial lineal sobre M. Supongamos además que b es cualquier campo vectorial constante sobre M, y consideremos el siguiente sistema de control sobre M:*

$$\frac{dx}{dt} = Ax + u(t)b,$$

con $u(t)$ cualquier función de control constante a trozos. Entonces $\mathcal{A}(x, \leq T) = M$ para cualquier x en M y cualquier $T > 0$ si, y sólo si, M es igual al espacio lineal generado por $\{b, Ab, ..., A^{n-1}b\}$.

II.2.2. Sistemas Bilineales

Luego de estudiar los sistemas lineales de la forma $\dot{x} = Ax$, donde A es una matriz constante, estamos interesados en considerar sistemas lineales que varían con el tiempo, es decir, sistemas lineales de la forma

$$\dot{x} = F(t)x \qquad \text{(II.3)}$$

donde F es una función diferenciable de $\mathbf{R}^n$ en $\mathbf{R}^n$, que también en ocasiones es considerado como un campo vectorial diferenciable. La primera cosideración es suponer que $F(t)$ es una combinación lineal de matrices constantes en los cuales los coeficientes varían con el tiempo, esto es,

$$F(t) = A + u_1B_1 + u_2B_2 + \cdots + u_mB_m.$$

Luego el sistema (II.3) puede ser escrito de la forma

$$\dot{x} = (A + \sum_{i=1}^{m} u_i(t)B_i)x.$$

De una manera mas precisa y compacta damos la siguiente definición.

Un sistema bilineal de control en $\mathbf{R}^n$, es un sistema de la forma

$$\dot{x} = Ax + uBx$$

donde $A, B \in M(n, \mathbf{R})$ y las funciones de control u son controles admisibles $u : \mathbf{R} \to U$ con $U \subset \mathbf{R}^m$.

Dentro de la literatura escrita, no se tiene todavía un resultado parecido al de Kalman que nos de un condición necesaria y suficiente para la controlabilidad del sistema, con el simple cálculo del rango de una matriz. Sin embargo existen conclusiones que derivan directamente de la teoría de los sistemas lineales en los cuales se encuentran caracterizados por ejemplo la no controlabilidad de un sistema bilineal en $\mathbf{R}^n$. Para ver en más detalle los resultados sobre sistemas bilineales sugerimos [7].

El álgebra de Lie asociado a un sistema de control bilineal está dado por

$$\mathcal{LA}\{A + uB, u \in \mathcal{U}\} \subset gl(n, \mathbf{R}).$$

En general para cada $x \in \mathbf{R}^n$, tenemos

$$\mathcal{LA}\{A + uB, u \in \mathcal{U}\}(x) \subset T_x\mathbf{R}^n.$$

Definición II.35. *Diremos que el sistema de control bilineal satisface la Condición de Rango de Álgebras de Lie (LARC) en $x \in \mathbf{R}^n$ si*

$$\dim(\mathcal{LA}\{A + uB, u \in \mathcal{U}\}(x)) = \dim \mathbf{R}^n. \tag{II.4}$$

Si (II.4) vale para todo $x \in \mathbf{R}^n - \{0\}$, se dice que el sistema bilineal de control satisface LARC.

Ejemplo II.36. *Consideremos el sistema de control bilineal*

$$\begin{aligned}\dot{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + u\begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2(1+u) & 2u \\ 0 & 1+u \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.\end{aligned}$$

Este sistema de control no satisface la condición de rango de álgebras de Lie, puesto que

$$[A, B] = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \quad [A, [A, B]] = [A, B], \quad [B, [A, B]] = [A, B],$$

y por tanto

$$(A, B, [A, B])\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 & 2x_1 + 2x_2 & 2x_2 \\ x_2 & x_2 & 0 \end{pmatrix}$$

tiene rango 1 para $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Un importante hecho que justifica el estudio de los sistemas bilineales de control es que existen sistemas lineales que son difíciles de verificar que sean controlables pero que al extenderlo a un sistema bilineal adecuado podemos obtener la controlabilidad del sistema, tal y como se especifica en el siguiente ejemplo.

De hecho este proceso va en dirección contraria a la teoría clásica de resolución de sistemas de ecuaciones diferenciales, es decir, cuando nos encontramos con la necesidad de resolver una ecuación diferencial no lineal, el proceso que se sigue es linealizarlo y luego estudiar su parte lineal.

Ejemplo II.37. *Consideremos el sistema lineal en $\mathbf{R}^2$ dado por la ecuación diferencial*

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -2x_1 - x_2 + u, \tag{II.5}$$

o de manera mas compacta

$$\frac{dx}{dt} = Ax + ub$$

donde $A = \begin{pmatrix} 0 & 1 \\ -2 & -1 \end{pmatrix}$ *y* $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. *Si denotamos por E el espacio lineal generado por $\{b, Ab\}$ como en el teorema de Kalman. Primero si consideramos un control no acotado $u(t)$, se puede ver que el rango de E es 2, luego el sistema es completamente controlable si y solo si el control u es no acotado. Si el control es acotado, digamos $|u(t)| \leq 1$ entonces tenemos que el sistema no es controlable.*

En un sistema bilineal de control adecuado a partir de II.5, puede ser completamente controlable. Por ejemplo, consideremos el sistema

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -2x_1 - x_2 + u + x_1u + 2x_2u, \tag{II.6}$$

con $|u(t)| \leq 1$. Para $u = +1$, el sistema exhibe una característica de foco inestable con punto de equilibrio en $(1, 0)$ sobre el retrato fase, y para $u = -1$ un foco estable con punto de equilibrio en $(\frac{1}{3}, 0)$. Los correspondientes autovalores son $\lambda_1, \lambda_2 = \frac{1}{2}(1 \pm i\sqrt{3})$ y $\lambda_1, \lambda_2 = -\frac{1}{2}(3 \pm i\sqrt{3})$, respectivamente. Al realizar la combinación de estos focos estables e inestables nos dan controlabilidad completa.

II.3. Sistemas de Control sobre Grupos de Lie

Los sistemas de control que consideraremos en está sección son descritos por un sistema evolucionado en un grupo de Lie G de la forma

$$\frac{dx}{dt}(t) = X_0(x(t)) + \sum_{i=1}^{m} u_i(t)X_i(x(t)) \tag{II.7}$$

donde $X_0, ..., X_m$ son campos de vectores invariantes a derecha sobre G. Sistemas descritos de está forma son llamados invariantes a derecha.

II.3.1. Órbitas de campos de vectores invariantes

En esta sección consideraremos las propiedades básicas de los flujos generados por los campos de vectores invariantes a derecha sobre un grupo de Lie G. Se tiene que los campos de vectores invariantes a izquierda sobre G, satisfacen propiedades análogas.

Sea Φ el flujo correspondiente al campo vectorial invariante a derecha X. Sea

$$e(t) = \Phi(t, e)$$

la curva integral en G que pasa por la identidad del grupo G. Para cualquier $g \in G$, la curva $g(t)$ definida por

$$g(t) = e(t)g = R_g(e(t)),$$

satisface

$$\frac{d}{dt}g(t) = (R_g)_* \frac{d}{dt}e(t) = (R_g)_* X \circ e(t) = X(e(t)g) = X(g(t)).$$

Así $g(t)$ es la curva integral de X que pasa por g, que satisface la propiedad

$$\Phi(t, g) = \Phi(t, e)g$$

para cualquier $g \in G$ y todo $t \in \mathbf{R}$. Esta igualdad implica las siguientes propiedades básicas

1. La curva integral $e(t)$ se encuentra definida para todo $t \in \mathbf{R}$.
   Tenemos que $H = \{e(t) : t \in \mathbf{R}\}$ es un subgrupo abeliano de G. La prueba de esta afirmación es muy sencilla,

   $$\Phi(t, \Phi(s, e)) = \Phi(t + s, e) = \Phi(s + t, e) = \Phi(s, \Phi(t, e)).$$

   La relación de los flujos escrita encima puede ser transcrita en términos de las curvas integrales, de la siguiente manera

   $$e(t)e(s) = e(t + s) = e(s)e(t),$$

   de esto se sigue que H es un grupo abeliano. El hecho que $e(t)$ esté definida para todo t, se sigue también de la anterior igualdad: si la curva e es definida para un valor particular de t, entonces e también se encuentra definido para $t + \epsilon$, donde ϵ es independiente de t, ya que $e(t + \epsilon) = e(t)e(\epsilon)$. Por lo tanto, $e(t)$ esta definida para todo t.

2. X es un campo vectorial completo.
   Basta que observemos la igualdad

   $$\Phi(t, g) = \Phi(t, e)g,$$

   y por lo tanto $\Phi(t, g)$ está definida para todo número real t.

3. Si X es un campo vectorial invariante a izquierda, entonces su flujo Φ satisface la igualdad $\Phi(t, g) = g\Phi(t, e)$. Se sigue que las afirmaciones (1) y (2) son válidas para campos de vectores invariantes a izquierda.

Definición II.38. *Para cualquier campo de vectores X invariante a derecha sobre G,*

$$e^{tA}$$

denotará la curva integral de X que pasa por la identidad con $A = X(e)$. En términos de esta *notación,*

$$(\exp tX)(g) = e^{tA}g$$

El teorema de la órbita aplicado a una familia de campos de vectores invariantes a derecha sobre un grupo de Lie G es muy efectivo para obtener algunos resultados de la teoría clásica de los grupos de Lie. Para tener una ilustración de esto, presentamos los siguientes hechos:

Sea $\mathcal{F}$ una familia de campos de vectores invariantes a derecha sobre un grupo de Lie G, y sea

$$\Gamma = \{X(e) : X \in \mathcal{F}\}.$$

Entonces evidentemente se tiene $\Gamma \subset L(G)$, donde $L(G)$ es el álgebra de Lie de G. Denotemos por H la órbita de $\mathcal{F}$ en la identidad. Tenemos por el teorema de la órbita que H es una subvariedad analítica de G. Pero esta órbita es también un subgrupo de G generado por $\{e^{tA} : t \in \mathbf{R}, A \in \Gamma\}$.

Así H es un grupo de Lie conexo, y su álgebra de Lie es igual a la subálgebra de Lie de $L(G)$ generado por Γ. Luego tenemos la demostración del clásico teorema de grupos de Lie, que dice que a cualquier subálgebra $\mathcal{L}$ de $L(G)$ le corresponde exactamente un único subgrupo de Lie conexo de G. En este contexto, $Lie(\Gamma)$ es el álgebra de Lie $\mathcal{L}$ dada, y H es el correspondiente subgrupo de Lie conexo de G. Pero el teorema de la órbita también demuestra que dado cualquier subconjunto Γ tal que $Lie(\Gamma) = \mathcal{L}$, entonces cualquier elemento de H puede ser escrito de la forma

$$e^{t_1A_1}e^{t_2A_2}\dots e^{t_pA_p}$$

para ciertos elementos $A_1, A_2, ..., A_p$ en Γ.

A manera de discusión también tenemos el famoso teorema de grupos de Lie: Un subgrupo cerrado H de un grupo de Lie G es por si mismo un grupo de Lie. Su prueba la exhibimos en las siguientes líneas. Sea

$$\Gamma = \{A : e^{tA} \in H\}$$

para todo t. Se puede demostrar que Γ es un subálgebra de Lie de $L(G)$. Si definimos $\mathcal{F}$ como la familia de campos de vectores invariantes a derecha cuyos valores en la identidad están en Γ, entonces la órbita de $\mathcal{F}$ desde la identidad es un grupo de Lie. Este es igual a H, cuando H es conexo.

II.3.2. $GL(n, \mathbf{R})$ y sus subgrupos

Para cada $A \in M(n, \mathbf{R})$, $\exp tA$ es una curva en $GL(n, \mathbf{R})$ cuyo vector tangente en la identidad es igual a A. Luego $M(n, \mathbf{R})$ está contenido en el espacio tangente en la identidad de $GL(n, \mathbf{R})$, y ya que $GL(n, \mathbf{R}) \subset M(n, \mathbf{R})$, tenemos que son iguales. Como ya fue visto en el capítulo 1 las traslaciones a derecha (resp. a izquierda) en $GL(n, \mathbf{R})$ son multiplicación de matrices por la derecha (resp. a izquierda). Luego los campos de vectores invariantes a derecha son de la forma

$$X(g) = Ag, \qquad A \in M(n, \mathbf{R}), g \in GL(n, \mathbf{R}).$$

A es el valor de X en la identidad. Los campos de vectores invariantes a izquierda son de la forma $X(g) = gA$. Si X, Y son campos de vectores invariantes a derecha entonces

$$[X, Y]g = (AB - BA)g = [A, B]g$$

para todo $g \in GL(n, \mathbf{R})$. Por tanto el álgebra de Lie de $GL(n, \mathbf{R})$ es el espacio vectorial $M(n, \mathbf{R})$ munido con el conmutador de matrices como su corchete de Lie.

Un subgrupo de Lie de $GL(n, \mathbf{R})$ es llamado un *grupo lineal*. El álgebra de Lie de cualquiera de tales subgrupos es una subálgebra de $M(n, \mathbf{R})$. Por ejemplo el conjunto de las matrices antisimétricas es el álgebra de Lie del grupo ortogonal $O(n, \mathbf{R})$. El conjunto de todas las matrices con traza igual a cero es el álgebra de Lie del grupo especial lineal $SL(n, \mathbf{R})$.

Cualquier colección de matrices Γ determina una familia de campos de vectores invariantes a derecha o izquierda $\mathcal{F}_\Gamma$. La orbita de $\mathcal{F}_\Gamma$ desde la identidad de $GL(n, \mathbf{R})$ es un subgrupo lineal G consistiendo de todos los productos de matrices de la forma

$$e^{t_1 A_1} \cdots e^{t_p A_p}$$

para ciertas matrices $A_1, ..., A_p \in \Gamma$ y números reales $t_1, ..., t_p$. La órbita de $\mathcal{F}_\Gamma$ desde cualquier otro punto g en $GL(n,\mathbf{R})$ es una clase a derecha o izquierda de G (dependiendo si $\mathcal{F}_\Gamma$ son campos invariantes a izquierda o derecha). $G = GL^+(n,\mathbf{R})$ si y sólo si $Lie(\Gamma) = M(n,\mathbf{R})$, donde $Lie(\Gamma)$ denota el álgebra de Lie generada por Γ.

II.3.3. Espacios Homogéneos

Definición II.39. *Sea G un grupo de Lie y M una variedad diferenciable. Diremos que G actúa sobre M si existe una función diferenciable*

$$\theta : G \times M \to M$$

que satisface

1. $\theta(g_1g_2, x) = \theta(g_1, \theta(g_2, x))$ *para todo* $g_1, g_2 \in G$ *y todo* $x \in M$, *y*

2. $\theta(e, x) = x$ *para todo* $x \in M$.

Para cada $g \in G$, sea $\theta_g : M \to M$ definida por $\theta_g(x) = \theta(g, x)$. La función $g \to \theta_g$ es llamada una acción de G sobre M. Cualquier tal acción es un homomorfismo de grupos de G en el grupo de difeomorfismos sobre M. En particular, si A es un elemento de $L(G)$, entonces $\{\theta_{\exp tA}\}$ es un grupo a 1-parámetro de difeomorfismos sobre M. Sea X_A su generador infinitesimal. La correspondencia $A \to X_A$ es un homomorfismo de álgebras de Lie de $L(G)$ en el álgebra de Lie de campos de vectores sobre M. Por tanto, la familia $\mathcal{F} = \{X_A : A \in L(G)\}$ es un álgebra de Lie de dimensión finita de campos vectoriales diferenciables sobre M. Nos referiremos a sus elementos como los campos vectoriales subordinados por G. Los elementos de $\mathcal{F}$ son necesariamente campos vectoriales completos sobre M.

Diremos que G actúa *transitivamente* sobre M si existe $x_0 \in M$ tal que la órbita

$$Gx = \{\theta(g, x) : g \in G\} = M.$$

Se sigue que para todo $x \in M$, la órbita $\{\theta(g, x) : g \in G\}$ es igual a M. En efecto, existe $g \in G$

tal que $gx_0 = x$ entonces

$$Gx = Ggx_0 = Gx_0 = M.$$

Si G actúa transitivamente sobre M, entonces tenemos la siguiente versión del teorema de la órbita:

Teorema II.40. *Sea Γ un subconjunto arbitrario de $L(G)$, y sea $\mathcal{F} = \{X_A : A \in \Gamma\}$. Denotemos por H el subgrupo de G generado por $\{\exp tA : A \in \Gamma, t \in \mathbf{R}\}$. Entonces la órbita de $\mathcal{F}$ desde cualquier punto $x \in M$ es dado por la acción de H sobre x, esto es,*

$$G(\mathcal{F})(x) = \{\theta_h(x) : h \in H\}.$$

En particular, si Γ genera el álgebra de Lie de G, entonces $G(\mathcal{F})(x) = M$ para cada $x \in M$.

Variedades diferenciables que admiten acciones transitivas de grupos de Lie son llamados *espacios homogéneos.*

II.3.4. Controlabilidad

Una condición necesaria para la controlabilidad de un sistema de control sobre un grupo de Lie G de la forma II.7 es que el conjunto $\mathcal{A}(e)$ de puntos accesibles desde la identidad de G sea un subgrupo de G. Luego, el problema de controlabilidad se reduce a lo siguiente:

1. Cuando $\mathcal{A}(e)$ es un subgrupo? y
2. Si $\mathcal{A}(e)$ es un subgrupo, cuando $\mathcal{A}(e) = G$?

La pregunta (2) es más fácil de responder que (1). De hecho en un teorema que enunciaremos posteriormente se prueba que si $\mathcal{A}(e)$ es un subgrupo de G entonces necesariamente, este subgrupo es un grupo de Lie conexo sobre S de G cuya álgebra de Lie es el subálgebra L generada por $X_0, ..., X_m$. De esto se sigue que el sistema II.7 es controlable si y sólo si

1. $\mathcal{A}(e)$ es un subgrupo,

2. G es conexo, y

3. L es el álgebra de Lie de G.

Esto demuestra que la importante cuestión está en determinar cuando $\mathcal{A}(e)$ es un subgrupo.

La razón por la cual solo se consideran los conjuntos accesibles desde la identidad esta explicada en la sección II.3.1, que se traduce de la siguiente manera

$$\mathcal{A}(g,T) = \mathcal{A}(e,T)g \quad y \quad \mathcal{A}(g) = \mathcal{A}(e)g.$$

Para darle simplicidad a la escritura denotaremos el sistema de control invariante a derecha II.7 de la forma (X, U) donde U es el conjunto de control.

Sin entrar en los detalles de esta sección, a continuación enunciaremos los resultados obtenidos para la controlabilidad de un sistema invariante a derecha.

Teorema II.41. *Una condición necesaria para que (X, U) sea controlable es que G sea conexo y que $L = L(G)$. Si G es compacto o si el sistema es homogéneo ($X_0 = 0$) la condición es también suficiente.*

Teorema II.42. *Sea G compacto y (X, U) controlable. Entonces el sistema es completamente controlable. Esto es, existe $T > 0$ tal que, para cualesquiera $g, g' \in G$ existe un control que me traslada de g hasta g' en T unidades de tiempo o menos.*

CAPÍTULO III

Semigrupos de $GL(2)$ transitivos sobre $\mathbf{R}^2 - \{0\}$

Como fue mencionado en la introducción de este trabajo, es de nuestro interés obtener condiciones para que el semigrupo S_Σ generado por $\Sigma = \{A, \pm B\}$ del sistema bilineal

$$\dot{x} = Ax + uBx$$

actúe transitivamente sobre $\mathbf{R}^2 - \{0\}$.

Con este propósito, en este capítulo desarrollaremos a detalle y encontraremos los grupos lineales (grupos de Lie de matrices) que actúan transitivamente sobre $\mathbf{R}^2 - \{0\}$.

Dividiremos el capítulo en dos secciones. En la primera sección presentaremos los subgrupos de Lie de $GL(2)$ que actuán transitivamente sobre $\mathbf{R}^2 - \{0\}$, mostrando que son únicos. Sin embargo al ser esto una tarea complicada y teniendo en cuenta la relación uno a uno de subgrupos de Lie conexos con las subálgebras de Lie, con la ayuda de unos lemas mostraremos que las únicas subálgebras de Lie lineales que actúan transitivamente sobre $\mathbf{R}^2$ son $\mathbf{gl}(2)$, $\mathbf{sl}(2)$ y $\mathbf{so}(2) \oplus \mathbf{R}1$, que son precisamente las subálgebras de Lie de esos grupos.

En la sección dos buscaremos dentro de $SL(2)$, semigrupos que actúen transitivamente sobre el espacio proyectivo unidimensional $\mathbf{P}^1$, esto se hace precisamente pues se tiene que la transitividad sobre el espacio proyectivo implica la transitividad sobre $\mathbf{R}^2 - \{0\}$, este es uno de los resultados fundamentales de este capítulo. De manera mas precisa el resultado dice que la transitividad de un semigrupo $S \subset SL(2)$ con interior no vacío sobre el proyectivo implica que $S = SL(2)$. Luego puesto que $SL(2)$ actúa transitivamente sobre $\mathbf{R}^2 - \{0\}$ (como será probado en la primera sección) se tiene que la transitividad sobre el proyectivo es equivalente a la transitividad sobre $\mathbf{R}^2 - \{0\}$.

III.1. Grupos Transitivos

Sea G un subgrupo de Lie conexo de $GL(2)$, el grupo lineal de matrices inversibles, y denotemos por $\mathbf{g}$ el álgebra de Lie de G. Como fue descrito en el capítulo 1, está es el álgebra de Lie compuesta por aquellas matrices X tal que $\exp tX \in G$ para todo $t \in \mathbf{R}$.

Puesto que $\mathbf{R}^2 - \{0\}$ es conexo, una condición necesaria y suficiente para que G actúe transitivamente sobre $\mathbf{R}^2 - \{0\}$ es que la órbita

$$Gx = \{gx : g \in G\}$$

es abierto para cualquier $x \in \mathbf{R}^2 - \{0\}$. El espacio tangente a Gx en x es dado por

$$\mathbf{g}x = \{Ax : A \in \mathbf{g}\}.$$

Luego, Gx es abierto si y sólo si $\mathbf{g}x = \mathbf{R}^2$. Puesto que consideramos G conexo tenemos la equivalencia, G es transitiva si y sólo si su álgebra de Lie $\mathbf{g}$ lo es.

El propósito de esta sección es demostrar que las únicas álgebras de Lie transitivas y los grupos de Lie correspondientes son los siguientes:

1. $\mathbf{gl}(2)$, el álgebra de Lie de todas las matrices reales de orden 2, que es el álgebra de Lie del grupo conexo $GL^+(2) = \{g : \det g > 0\}$.

2. $\mathbf{sl}(2)$, el subálgebra de Lie de $\mathbf{gl}(2)$ consistiendo de las matrices de traza cero de orden 2. Este es el álgebra de Lie de $SL(2)$, el grupo de las matrices de determinante uno.

3. $\mathbf{so}(2) \oplus \mathbf{R}1$ donde $\mathbf{so}(2)$ que es el álgebra unidimensional de las matrices antisimétricas y 1 denota la matriz identidad. Este es álgebra del grupo $SO(2) \times (\mathbf{R}^{+}1)$ que consiste de una rotación (elemento de $SO(2)$) seguida de una homotecia (matriz en $\mathbf{R}^{+}1$). Los elementos de este grupo son de la forma
$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$
que están en biyección con los números complejos no nulos $a + ib$. Esto es, $SO(2) \times (\mathbf{R}^{+}1)$ es isomorfo a $\mathbf{C}^{*}$.

Es una tarea sencilla mostrar que los grupos de estas álgebras de Lie mencionadas actúan transitivamente sobre $\mathbf{R}^2 - \{0\}$. Veamos como hacemos esto, se tiene
$$\begin{pmatrix} x & 0 \\ y & \frac{1}{x} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{para } x \neq 0$$
y
$$\begin{pmatrix} 0 & -\frac{1}{y} \\ y & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \quad \text{para } y \neq 0$$
esto es, la órbita de $(1, 0)$ bajo $SL(2)$ es $\mathbf{R}^2 - \{0\}$. Consecuentemente este grupo y $GL^{+}(2) \supset SL(2)$ son transitivos sobre $\mathbf{R}^2 - \{0\}$. La transitividad de $SO(2) \times (\mathbf{R}^{+}1)$ se sigue de
$$|v| \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = v$$
para $v \in \mathbf{R}^2 - \{0\}$ donde θ es el ángulo entre v y $(1, 0)$.

La prueba que estos son los únicos grupos transitivos no se puede realizar de manera directa, lo que haremos es estudiar las subálgebras de Lie de $\mathbf{gl}(2)$ que actúan transitivamente sobre el plano. Mostraremos que las únicas subálgebras con esta propiedad, son las subálgebras de los grupos de encima, pero esto requiere algunos lemas que los desarrollaremos a continuación.

Comenzaremos con los siguientes que nos dan resultados acerca de subálgebras de dimensión dos que se encuentran dentro de $\mathbf{sl}(2)$.

Lema III.1. *Sea* $\mathbf{g} \subset \mathbf{sl}(2)$ *un subálgebra con* $\dim \mathbf{g} = 2$*. Entonces existe una base* β *de* $\mathbf{R}^2$ *tal que las matrices de las aplicaciones lineales en* $\mathbf{g}$ *con respecto a* β *son de la forma*

$$\begin{pmatrix} a & b \\ 0 & -a \end{pmatrix}.$$

Demostración. Primero notemos que si la $\dim \mathbf{g} = 2$, entonces tenemos que $\mathbf{g}$ es soluble o es abeliana (vea ejemplo 1.10). Mostraremos que $\mathbf{g}$ no puede ser abeliana. En efecto, consideremos la base de $\mathbf{sl}(2)$ formada por las matrices

$$\left\{ H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, P = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Q = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \tag{III.1}$$

si realizamos los corchetes entre estos elementos, se tiene las siguientes relaciones

$$[H, P] = 2P, \quad [H, Q] = -2Q, \quad [P, Q] = H.$$

Ahora sean X, X' dos matrices en $\mathbf{sl}(2)$ tal que $[X, X'] = 0$, escribiendo estas en términos de la base, tenemos que existen constantes a, b, c, a', b', c' no todas nulas tales que $X = aH + bP + cQ$ y $X' = a'H + b'P + c'Q$. Entonces

$$[X, X'] = (bc' - cb')H + 2(ab' - ba')P - 2(ca' - ac')Q$$

puesto que $[X, X'] = 0$ se tiene que $bc' = cb'$, $ab' = ba'$ y $ca' = ac'$. Es decir, X y X' no pueden ser linealmente independientes. En efecto, supongamos que $b \neq 0$, entonces de la primera igualdad se tiene $c' = (\frac{b'}{b})c$, de la segunda $a' = (\frac{b'}{b})a$ y finalmente $b' = (\frac{b'}{b})b$, luego X' es un múltiplo de X. Esto demuestra que $\mathbf{g}$ no es abeliana. Pues en caso contrario, suponiendo que $\mathbf{g}$ fuese abeliana, como $\dim \mathbf{g} = 2$ tomemos una base $\{X, X'\}$, entonces $[X, X'] = 0$. Pero en este caso, se tiene que X y X' son linealmente dependientes, obteniendo así una contradicción.

Por lo tanto hemos mostrado que dentro de $\mathbf{sl}(2)$ no pueden existir subálgebras abelianas de dimensión dos. Así tenemos que $\mathbf{g}$ es soluble, esto es, existe una base $\{X, Y\}$ de $\mathbf{g}$ tal que $[X, Y] = Y$.

Ahora por el Teorema de Lie sobre álgebras de Lie Solubles (teorema 1.44), existe una base de $\mathbf{R}^2$ tal que los elementos de $\mathbf{g}$ son escritos en está base de la forma

$$\begin{pmatrix} \lambda & \mu \\ 0 & -\lambda \end{pmatrix}.$$

En esta representación los elementos de la diagonal de Y son cero pues $[X, Y] = Y$. Por tanto, $Y^2 = 0$ (nilpotente) luego por Jordan existe una base β de $\mathbf{R}^2$ tal que con respecto a β, la matriz Y es triángular superior con ceros en la diagonal. De la misma manera tenemos también del hecho $[X, Y] = Y$ que X es triangular superior con respecto a β probando así el lema. $\square$

Como una consecuencia directa de este Lema podemos demostrar el siguiente corolario, que nos brinda un resultado contundente para el estudio de subálgebras de Lie contenidas en $\mathbf{sl}(2)$.

Corolario III.2. $\mathbf{sl}(2)$ *es la única de sus subálgebras que es transitiva.*

Demostración. Si $\dim \mathbf{g} = 1$ entonces el álgebra de Lie es abeliana, luego el producto de dos elementos cualesquiera es nulo, así no puede ser transitiva. Si $\dim \mathbf{g} = 3$ entonces al estar contenida en $\mathbf{sl}(2)$ que tiene dimensión tres, estos espacios deben coincidir. Finalmente si $\mathbf{g}$ es como en el lema anterior, tenemos que la recta generada por el primer elemento encontrado de la base en el lema anterior es invariante bajo $\mathbf{g}$, esto es, $[\mathbf{g}, Y] \subset \{\alpha Y : \alpha \in \mathbf{R}\}$. Se tiene entonces que esta subálgebra no puede ser transitiva, pues al hacer el corchete de más de dos elementos entonces este se anula. Con eso termina la prueba. $\square$

Con el corolario anterior, tenemos demostrado que dentro de $\mathbf{sl}(2)$ no existe ninguna subálgebra de Lie que actúa transitivamente sobre el plano.

Ahora realizaremos un análisis de aquellas subálgebras transitivas que no están enteramente contenidas en $\mathbf{sl}(2)$. Para ello notemos primero que si tomamos cualquier matriz A de 2×2, esta siempre se puede escribir de la forma

$$A = \left(A - \frac{trA}{2}1\right) + \frac{trA}{2}1.$$

Luego tenemos que $\mathbf{gl}(2)$ se puede descomponer como

$$\mathbf{gl}(2) = \mathbf{sl}(2) \oplus \mathbf{R}1. \tag{III.2}$$

Usaremos la notación $\mathbf{z} = \mathbf{R}1$ para indicar el centro de $\mathbf{gl}(2)$. Además de esta descomposición se tiene inmediatamente que cualquier subespacio $V \subset \mathbf{gl}(2)$ conteniendo $\mathbf{z}$ se descompone como $V = \mathbf{z} \oplus (\mathbf{sl}(2) \cap V)$. Con estas observaciones podemos enunciar el siguiente resultado que caracteriza las subálgebras de dimensión dos que contienen a $\mathbf{z}$.

Lema III.3. *Sea $X \in \mathbf{sl}(2)$ y supongamos que $\mathbf{g} = \mathbf{RX} \oplus \mathbf{z}$ es transitivo. Entonces $\mathbf{g}$ es isomorfo a $\mathbf{so}(2) \oplus \mathbf{z}$.*

Demostración. Asumiendo que $\mathbf{g}$ es transitivo, eso implica inmediatamente que $X \neq 0$ y como $trX = 0$, su forma canónica de Jordan es una de las siguientes

$$\begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}.$$

En los dos primeros casos las matrices en $\mathbf{g}$ son triangulares superiores y aplicando el argumento del corolario (III.2), en una base adecuada de $\mathbf{R}^2$ se tiene que $\mathbf{g}$ no puede ser transitiva. Entonces solo nos quedamos con el tercer caso, pero si es así se tiene que $\mathbf{g}$ es isomorfo a $\mathbf{so}(2) \oplus \mathbf{z}$ como queríamos. $\square$

Ciertamente este resultado nos brinda un criterio que caracteriza las subálgebras transitivas de dimensión dos, de hecho si la subálgebra se encuentra dentro de $\mathbf{sl}(2)$, no hay nada que hacer, luego el trabajo está precisamente en pensar que $\mathbf{g}$ no está dentro de $\mathbf{sl}(2)$, eso lo registramos en el siguiente corolario.

Corolario III.4. *Las subálgebras transitivas de dimensión dos son isomorfas a $\mathbf{so}(2) \oplus \mathbf{z}$.*

Demostración. Sea $\mathbf{g}$ una subálgebra transitiva con $\dim \mathbf{g} = 2$. Notemos que si $\mathbf{g} \cap \mathbf{z} \neq \{0\}$, entonces nos encontramos en el caso de encima pues $\dim \mathbf{z} = 1$, y por tanto el resultado es directo. En el otro caso, $\mathbf{g} \cap \mathbf{z} = \{0\}$ consideremos la proyeccón

$$\pi : \mathbf{gl}(2) \to \mathbf{sl}(2),$$

dada por $\pi(A) = A - \frac{tr(A)}{2}1$ de acuerdo con la descomposición (III.2). Entonces $\dim \pi(\mathbf{g}) = 2$, luego los elementos de $\pi(\mathbf{g})$ son triangulares superiores en alguna base adecuada por el lema (III.1). Esto implica que los elementos de $\mathbf{g}$ son también triangulares superiores así no puede ser transitiva. $\square$

Esto cierra el caso de estudio para las subálgebras transitivas de dimensión dos.

Con respecto a subálgebras de dimensión tres tenemos el siguiente resultado.

Lema III.5. *Supongamos que* $\mathbf{g} = \mathbf{h} \oplus \mathbf{z}$ *donde* $\mathbf{h}$ *es un subálgebra de dimensión 2 y* $\mathbf{z}$ *es como encima. Entonces* $\mathbf{g}$ *no es transitiva.*

Demostración. Por la forma que tiene $\mathbf{g}$ y recordando la relación III.2, se tiene necesariamente que $\mathbf{h} \subset \mathbf{sl}(2)$. Por el lema III.1 existe una base tal que los elementos de $\mathbf{h}$ son triangulares superiores. En esta base, los elementos de $\mathbf{g}$ son también triangulares superiores, luego está álgebra no puede ser transitiva. $\square$

De manera más general, consideremos $\mathbf{g}$ un álgebra transitiva con $\dim \mathbf{g} = 3$. Por el lema de encima, $\mathbf{z}$ no esta contenida en $\mathbf{g}$ (si lo estuviera no sería transitiva) y puesto que $\dim \mathbf{z} = 1$ si tenemos que $\mathbf{g} \cap \mathbf{z} = \{0\}$ esto implica que

$$\mathbf{gl}(2) = \mathbf{g} \oplus \mathbf{z}.$$

No obstante, $\mathbf{sl}(2)$ es el único subálgebra que complementa $\mathbf{z}$. En efecto, tomando la base $\{H, P, Q\}$ como en (III.1), el hecho que $\mathbf{g}$ complementa $\mathbf{z}$ asegura la existencia de números reales h, p y q tal que $H' = H + h1$, $P' = P + p1$ y $Q' = Q + q1$ están en $\mathbf{g}$. Los corchetes entre estas matrices son

$$[H', P'] = 2P, \quad [H', Q'] = -2Q, \quad [P', Q'] = H.$$

esto es, $\mathbf{sl}(2) \subset \mathbf{g}$ y ya que tienen la misma dimensión tenemos que esas subálgebras coinciden. Así hemos demostrado el

Lema III.6. $\mathbf{sl}(2)$ *es el única subálgebra de Lie transitiva de dimensión tres.*

Con este lema se concluye la prueba que $\mathbf{gl}(2)$, $\mathbf{sl}(2)$ y $\mathbf{so}(2) \oplus \mathbf{R}1$ son las únicas subálgebras transitivas. En efecto, si $\mathbf{g}$ es transitiva entonces $\dim \mathbf{g} \geq 2$. Si $\dim \mathbf{g} = 2$ entonces $\mathbf{g}$ es (isomorfo a) $\mathbf{so}(2) \oplus \mathbf{R}1$ por el corolario (III.4). El lema anterior muestra que $\mathbf{g} = \mathbf{sl}(2)$ si $\dim \mathbf{g} = 3$, y nos queda solamente $\mathbf{gl}(2)$ que es de dimensión cuatro.

III.2. Semigrupos en $SL(2)$

Como fue mencionado en la introducción, la controlabilidad del sistema de control es equivalente a la transitividad del semigrupo S_Σ que tiene interior no vacío en el subgrupo de Lie conexo G_Σ. Recordemos que pedimos $int(S_\Sigma) \neq \emptyset$ para que el sistema sea localmente accesible.

En esta sección estudiaremos la transitividad sobre $\mathbf{R}^2 - \{0\}$ de un semigrupo S con interior no vacío en $SL(2)$, que es uno de los grupos transitivos del capítulo anterior.

Denotaremos por $\mathbf{P}^1$ el espacio proyectivo real uno-dimensional, que es el conjunto de subespacios unidimensionales de $\mathbf{R}^2$, y para $v \in \mathbf{R}^2 - \{0\}$ sea $[v]$ el subespacio que genera. Recordemos que el espacio proyectivo real $\mathbf{P}^1$ es un espacio compacto, para ver las otras propiedades del espacio proyectivo vea [12].

Remarca III.7. *Tenemos que $SL(2)$ actúa transitivamente sobre $\mathbf{P}^1$ por la acción definida por $g[v] = [gv]$, para $g \in SL(2)$, $v \in \mathbf{R}^2 - \{0\}$. En efecto, sean $[u], [v] \in \mathbf{P}^1$ arbitrarios queremos mostrar la existencia de algún $g \in SL(2)$ tal que $g[u] = [v]$; esto sucede pues al tener $u, v \in \mathbf{R}^2 - \{0\}$ y $SL(2)$ actúa transitivamente sobre $\mathbf{R}^2 - \{0\}$, existe $g \in SL(2)$ tal que $gu = v$, luego*

$$g[u] = [gu] = [v].$$

Definición III.8. *Diremos que un subsemigrupo S de $SL(2)$ es transitivo sobre $\mathbf{P}^1$ si para todo $u, v \in \mathbf{R}^2 - \{0\}$ existe $g \in S$ satisfaciendo $g[u] = [v]$.*

En este capítulo presentamos uno de los resultados centrales de la tesis, este resultado básicamente dice que: si $S \subset SL(2)$ es un subsemigrupo con puntos interiores y transitivo sobre $\mathbf{P}^1$ entonces este es todo $SL(2)$.

Primero probaremos el siguiente lema en donde se demuestra que un semigrupo propio de $SL(2)$ no puede contener rotaciones en su interior.

Lema III.9. *Sea $S \subset SL(2)$ un semigrupo y supongamos que existe $X \in \mathbf{sl}(2)$ con autovalores imaginarios puros, y tal que $\exp X$ está en $int(S)$. Entonces bajo estas condiciones tenemos que $S = SL(2)$.*

Demostración. El hecho que los autovalores de X son imaginarios puros y que tiene traza 0 implica que su forma canónica de Jordan es

$$\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \qquad a \neq 0$$

luego

$$\exp(tX) = \begin{pmatrix} \cos(ta) & -\operatorname{sen}(ta) \\ \operatorname{sen}(ta) & \cos(ta) \end{pmatrix}.$$

Ya que $\exp X \in int(S)$ entonces existe un $t > 0$ tal que $\exp(tX) \in int(S)$ y ta es irracional (digamos que $\pi < a$ dependiendo la longitud de a podemos elegir $t = \frac{\pi}{a}$). Por otro lado, si escribimos $h = \exp(tX)$, entonces $h^n \in int(S)$ para todo estero positivo n. Se sigue que existe algún n tal que $h^n = 1$. Luego $1 \in int(S)$. De donde tenemos que $S = SL(2)$, esto sucede por que el semigrupo generado por una vecindad de la identidad en un grupo de Lie conexo es todo el grupo. □

Ahora probaremos que los semigrupos $SL(2)$ no contiene otros elementos nilpotentes.

Lema III.10. *Sea S un subsemigrupo de $SL(2)$. Supongamos que existe un elemento nilpotente $X \in \mathbf{sl}(2)$ tal que $g = \exp X \in int(S)$. Entonces $S = SL(2)$.*

Demostración. Si $X = 0$ entonces $g = \exp 0 = 1 \in int(S)$ y tenemos el resultado por el argumento final del lema anterior. En otro caso supongamos que $X \neq 0$, y que existe $n \in \mathbf{N}$ tal que $X^n = 0$, tenemos que $\det X = 0$, entonces la forma canónica de Jordan para X es

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Ya que $g \in int(S)$ y la aplicación exponencial es continua existe un entorno U de X luego un $\epsilon > 0$ suficientemente pequeño tal que la matriz

$$Y = \begin{pmatrix} 0 & 1 \\ -\epsilon & 0 \end{pmatrix} \in U$$

y por la continuidad de la exponencial

$$\exp Y = \exp \begin{pmatrix} 0 & 1 \\ -\epsilon & 0 \end{pmatrix} \in int(S).$$

Los autovalores de Y son $\pm i\sqrt{\epsilon}$ (imaginarios puros), así por el lema anterior, concluimos el resultado. □

Lema III.11. *Sea S un subsemigrupo de $SL(2)$ con $int(S) \neq \emptyset$ y asuma que S es transitivo sobre $\mathbf{P}^1$. Tomemos $v \in \mathbf{R}^2 - \{0\}$. Entonces existe $w \in \mathbf{R}^2$ y $h \in int(S)$ tal que $\{v, w\}$ es una base y en esa base h puede ser escrito como*

$$h = \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix}$$

para algún $\mu > 0$.

Demostración. Tomemos $g \in int(S)$ tal que $g[v] = [v]$. Para demostrar la existencia de este elemento, al ser $int(S)$ no vacío existe $h_1 \in int(S)$, tal que $h_1[v] = [u]$, pero como S actúa transitivamente sobre $\mathbf{P}^1$, existe $h_2 \in S$ tal que $h_2[u] = [v]$, entonces

$$h_2 h_1[v] = h_2[u] = [v],$$

luego basta tomar $g = h_2 h_1$ pertenece a $int(S)$ y deja fijo a $[v]$. Ahora sea $u \in \mathbf{R}^2$ tal que $\{v, u\}$

es una base. Para encontrar la matriz de g con relación a esta base, notemos que al ser v un autovector para g, se satisfacen las siguientes relaciones

$$gv = av + 0u, \qquad gu = bv + a^{-1}u$$

luego la matriz de g con respecto a esta base será

$$g = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}$$

con $a, b \in \mathbf{R}$ y $a \neq 0$. Claramente, $g^2 \in int(S)$ y su matriz en la base $\{v, u\}$ es

$$g^2 = \begin{pmatrix} \mu & c \\ 0 & \mu^{-1} \end{pmatrix}$$

con $c \in \mathbf{R}$ y $\mu > 0$. Estrictamente hablando se tiene que $\mu = a^2$. Si $\mu = 1$ entonces $g^2 = \exp Y = I + Y$ con

$$Y = \begin{pmatrix} 0 & c \\ 0 & 0 \end{pmatrix},$$

(pues $Y^n = 0$ para $n \geq 2$). Entonces hemos encontrado un elemento nilpotente Y tal que $\exp Y \in int(S)$ luego por el lema anterior, tenemos que $S = SL(2)$, y claramente S contiene el elemento buscado. En el otro caso, siendo el caso mas general, supongamos que $\mu \neq 1$, entonces escribiendo

$$w = u + \frac{c}{\mu^{-1} - \mu} v,$$

es tal que, el conjunto $\{v, w\}$ es una base y en esta base la matriz de $h = g^2$ tiene la forma deseada. Para mostrar esto debemos encontrar la matriz de paso de g^2 de la base $\{v, u\}$ a la base $\{v, w\}$. Para ello primero tenemos que la matriz de g^2 en la base $\{v, u\}$ es

$$g^2 = \begin{pmatrix} \mu & c \\ 0 & \mu^{-1} \end{pmatrix}.$$

Ahora queremos encontrar la matriz de paso de base a base, escribimos las siguientes relaciones

$$v = 1v + 0u, \qquad w = \frac{c}{\mu^{-1} - \mu} v + 1v.$$

Así la matriz de paso es

$$P = \begin{pmatrix} 1 & \frac{c}{\mu^{-1}-\mu} \\ 0 & 1 \end{pmatrix},$$

cuya inversa será

$$P^{-1} = \begin{pmatrix} 1 & \frac{c}{\mu-\mu^{-1}} \\ 0 & 1 \end{pmatrix}$$

por tanto la matriz de g^2 en la base $\{v, w\}$ esta representada por el producto $P^{-1}g^2P$ que la denotaremos por g_0^2 y que está dentro de $int(S)$ es igual a

$$g_0^2 = \begin{pmatrix} 1 & \frac{c}{\mu-\mu^{-1}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu & c \\ 0 & \mu^{-1} \end{pmatrix} \begin{pmatrix} 1 & \frac{c}{\mu^{-1}-\mu} \\ 0 & 1 \end{pmatrix}.$$

Desarrollando el producto del lado derecho tenemos que

$$g_0^2 = \begin{pmatrix} 1 & \frac{c}{\mu-\mu^{-1}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu & \frac{\mu^{-1}c}{\mu^{-1}-\mu} \\ 0 & \mu^{-1} \end{pmatrix} = \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix}$$

como queríamos. □

En el siguiente lema mostramos que el resultado del lema anterior puede ser mejorado en el sentido que la base puede ser encontrada de modo que $\mu > 1$. Los detalles son los siguientes.

Lema III.12. *Con las notaciones y suposiciones del lema de encima, tomando $v \in \mathbf{R}^2 - \{0\}$. Entonces existe $w \neq 0$ y $h \in int(S)$ tal que $\beta = \{v, w\}$ es una base y la matriz de h con respecto a β es*

$$\begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix}$$

con $\mu > 1$.

Demostración. Sea $\{v, w\}$ la base encontrada por el lema de encima. Todas las matrices que consideremos serán referidas a esta base.

Ahora antes de realizar la demostración, primero observemos la siguiente descomposición (descomposicón de Iwasawa) de $g_1 \in SL(2)$ que satisface $g_1[v] = [w]$. Tenemos que existen

$$h_1 = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \quad y \quad n_1 = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

tal que

$$g_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} h_1 n_1.$$

En efecto, tenemos la siguiente relación

$$g_1 v = 0v + aw, \qquad g_1 w = -a^{-1} v + bw.$$

Esto es, la matriz de g_1 en relación a esta base es de la forma

$$g_1 = \begin{pmatrix} 0 & -a^{-1} \\ a & b \end{pmatrix}.$$

Luego tenemos de manera estricta, $* = \frac{b}{a}$. De manera análoga se tiene que, si $g_2 \in SL(2)$ y $g_2[w] = [v]$ tenemos que g_2 se descompone por Iwasawa como

$$g_2 = h_2 n_2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

siendo h_2 y n_2 de la misma forma que h_1 y n_1 respectivamente.
Ahora demostraremos el lema.
Sea $h \in int(S)$ encontrado en el lema precedente. Si $\mu > 1$ entonces no hay nada que demostrar. Por lo tanto supongamos que $0 < \mu < 1$.
Puesto que estamos suponiendo que S actúa transitivamente sobre $\mathbf{P}^1$, tenemos que existen $g_1, g_2 \in$

S tal que,

$$g_1[v] = [w] \quad y \quad g_2[w] = [v],$$

y claro se tiene que

$$g_2hg_1 \in int(S)$$

pues $h \in int(S)$. Usando la descomposición de Iwasawa de g_1 y g_2 como lo realizado encima,

$$g_2hg_1 = h_2n_2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} h_1n_1,$$

y realizando el producto de las tres matrices del centro,

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix}.$$

De donde tenemos que

$$g_2hg_1 = h_2n_2 \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} h_1n_1.$$

Luego esta igualdad puede ser escrita nuevamente de la forma

$$g_2hg_1 = \bar{h}\bar{n}$$

donde

$$\bar{h} = h_2 \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} h_1 \quad y \quad \bar{n} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

En efecto, si a es una matriz diagonal y b tiene la misma forma que $\bar{n}$ encima, entonces $a^{-1}ba$ tiene la misma forma que b. Luego podemos invertir el orden del producto de la siguiente manera $ba = a(a^{-1}ba)$. Para ver esto, sean

$$a = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \quad y \quad b = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

con $x \neq 0$ e $y \neq 0$. Entonces se tiene que

$$a^{-1} = \begin{pmatrix} x^{-1} & 0 \\ 0 & y^{-1} \end{pmatrix}$$

y

$$a^{-1}ba = \begin{pmatrix} 1 & ** \\ 0 & 1 \end{pmatrix} \qquad (** = yx^{-1}*).$$

Regresando a la demostración del lema, y usando el hecho registrado tenemos lo siguiente,

$$g_2hg_1 = h_2n_2 \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} h_1n_1.$$

Luego la matriz diagonal $\begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix}$ 'conmuta' si vale el término con n_2, esto es,

$$g_2hg_1 = h_2 \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} n_2h_1n_1.$$

De la misma manera la matriz diagonal h_1 'conmuta' con n_2, luego tenemos

$$g_2hg_1 = h_2 \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} h_1n_2n_1,$$

como queríamos.

Finalmente, notemos que

$$\bar{h} = h_2 \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} h_1$$

es el producto de tres matrices diagonales, luego es una matriz diagonal. Más precisamente se tiene que

$$\bar{h} = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} \mu^{-1} & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} = \begin{pmatrix} a^2\mu^{-1} & 0 \\ 0 & a^{-2}\mu \end{pmatrix}.$$

Sea $\lambda = \frac{a^2}{\mu}$, entonces si μ es suficientemente pequeño tenemos que $\lambda > 1$. Luego,

$$g_2 h g_1 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \bar{n} = \begin{pmatrix} \lambda & * \\ 0 & \lambda^{-1} \end{pmatrix} \in int(S)$$

con $\lambda > 1$. Hacemos un cambio de base como en el lema precedente tenemos la matriz diagonal buscada. □

Este lema es fundamental para realizar la demostración del teorema central de este capítulo. Como un comentario antes de realizar la demostración del teorema mencionado, observaremos los siguientes dos hechos acerca de números reales:

1. Sea $(a, b) \subset \mathbf{R}$ un intervalo con $b > a > 1$. Entonces existe un $T_0 > 0$ tal que $(T_0, \infty) \subset \cup_{n \geq 1}(a^n, b^n)$.
2. Sea (c, d) un intervalo con $0 < c < d < 1$. Entonces existe $\epsilon > 0$ tal que $(0, \epsilon) \subset \cup_{n \geq 1}(c^n, d^n)$.

La demostración de la primera afirmación se sigue directamente del siguiente hecho. Existe un $n_0 \in \mathbf{N}$ tal que $a^{n_0+1} < b^{n_0}$. En efecto, supongamos que para todo $n \in \mathbf{N}$ se tenga

$$a^{n+1} > b^n,$$

pero en este caso tendríamos que $a > (\frac{b}{a})^n$, luego la sucesión $\{(\frac{b}{a})^n\}$ está acotada, pero eso es una contradiccón pues $\frac{b}{a} > 1$. Luego el número buscado es $T_0 = a^{n_0}$, es tal que

$$(T_0, \infty) = \cup_{n \geq n_0}(a^n, b^n)$$

agregando los intervalos que tienen subíndices $< n_0$ se tiene la inclusión buscada.

La segunda afirmación se sigue de la primera, basta notar que existe n_0 tal que

$$(\frac{1}{d})^{n_0+1} < (\frac{1}{c})^{n_0}$$

de aquí tenemos que $c^{n_0} < d^{n_0+1}$ luego el número buscado es $\epsilon = d^{n_0}$.

Teorema III.13. *Sea $S \subset SL(2)$ un subsemigrupo con puntos interiores. Supongamos que S es transitivo sobre $\mathbf{P}^1$. Entonces $S = SL(2)$.*

Demostración. Por el lema (III.12) existe una base $\{v, w\}$ de $\mathbf{R}^2$ y $h \in int(S)$ que es escrito en esta base como

$$h = \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix}$$

con $\mu > 1$. También tenemos que existe $g \in int(S)$ tal que $gw = \lambda w$ con $\lambda > 1$. En esta base $\{v, w\}$, tenemos las siguientes relaciones

$$gv = \lambda^{-1} v + *w, \qquad gw = 0v + \lambda w,$$

luego la matriz de g en esta base será

$$g = \begin{pmatrix} \lambda^{-1} & 0 \\ * & \lambda \end{pmatrix}.$$

Como $h, g \in int(S)$, sus potencias y potencias de sus vecindades siguen perteneciendo a $int(S)$. Por (1) encima existe $T_0 > 0$ tal que, para cada $t > T_0$

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \in int(S),$$

y también existe $*$ tal que

$$\begin{pmatrix} t^{-1} & 0 \\ * & t \end{pmatrix} \in int(S).$$

Luego en el interior de S existen elementos que pueden ser escritos como

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \begin{pmatrix} t^{-1} & 0 \\ * & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} = I + \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix} = \exp \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix}.$$

Puesto que la matriz $Y = \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix}$ es nilpotente y está en el interior de S, entonces por el lema (III.10), se tiene que $S = SL(2)$. □

En el análisis de sistemas bilineales probamos controlabilidad, probando la no existencia de un subconjunto invariante y compacto sobre $\mathbf{P}^1$ que garantiza $S = SL(2)$ de acuerdo con el siguiente corolario.

Corolario III.14. *Sea S un subsemigrupo de $SL(2)$ con $int(S) \neq \emptyset$. Entonces $S \neq SL(2)$ si y sólo si existe un subconjunto propio compacto $C \subset \mathbf{P}^1$, con $int(C) \neq \emptyset$, que es S-invariante, esto es, $gC \subset C$ para todo $g \in S$.*

Demostración. Para $x \in \mathbf{P}^1$ consideremos el subconjunto compacto $cl(Sx)$ de $\mathbf{P}^1$, (es compacto pues es cerrado y $\mathbf{P}^1$ es compacto). Tenemos que $cl(Sx)$ es S-invariante, por tanto, existe un subconjunto compacto propio e invariante como en la afirmación si y sólo si $cl(Sx)$ es propio para algún x. Supongamos que $cl(Sx) = \mathbf{P}^1$ para todo $x \in \mathbf{P}^1$. Entonces S actúa transitivamente sobre $\mathbf{P}^1$. En efecto, consideremos el subconjunto

$$S^{-1} = \{g^{-1} : g \in S\},$$

que es también un semigrupo con interior no vacío en $SL(2)$. Luego para cada $y \in \mathbf{P}^1$ el subconjunto

$$S^{-1}y = \{g^{-1}y : g \in S\},$$

tiene interior no vacío en $\mathbf{P}^1$. Como Sx es denso en $\mathbf{P}^1$ (pues $cl(Sx) = \mathbf{P}^1$) tenemos que

$$S^{-1}y \cap Sx \neq \emptyset.$$

Tomando $z \in S^{-1}y \cap Sx$ existen $h_1, h_2 \in S$ tal que,

$$z = h_1^{-1}y, \qquad z = h_2 x$$

esto es, $h_1^{-1}y = h_2 x$ de donde, $y = h_1 h_2 x$ con $h_1 h_2 \in S$. En resumen, hemos mostrado que para todo $x, y \in \mathbf{P}^1$ existe $h = h_1 h_2 \in S$ tal que $y = hx$, luego S es transitivo sobre $\mathbf{P}^1$.

Por tanto, no existe un subconjunto compacto propio e invariante, si y sólo si, S no es transitivo sobre $\mathbf{P}^1$. Por el teorema (III.13), la condición anterior es equivalente a $S = Sl(2)$. $\square$

Remarca III.15. *Como consecuencia del teorema 1.5 es que un semigrupo $S \subset SL(2)$ con $int(S) \neq \emptyset$, es transitivo sobre $\mathbf{R}^2 - \{0\}$ si y solo si $S = SL(2)$.*

En efecto, tenemos que la transitividad sobre $\mathbf{R}^2 - \{0\}$ implica transitividad sobre $\mathbf{P}^1$, luego por el teorema (III.13) tenemos que $S = SL(2)$. Recíprocamente, sabemos que $SL(2)$ actúa transitivamente sobre $\mathbf{R}^2 - \{0\}$. Con eso tenemos el resultado buscado.

CAPÍTULO IV

Controlabilidad en el caso $G_{\Sigma} = SL(2)$

En el presente capítulo se desarrollará de manera muy explícita y estricta la noción de controlabilidad de un sistema de control bilineal en el plano. Considerando solo el caso en que el sistema bilineal de control

$$\dot{x} = Ax + uBx$$

está formado por las matrices $A, B \in M(2, \mathbf{R})$ que tienen traza cero. Más precisamente, este capítulo lo dividiremos en dos secciones. La primera sección esta dedicada exclusivamente a mostrar condiciones necesarias y suficientes para que el álgebra de Lie generada por $\Sigma = \{A, B\}$, sea todo el espacio $\mathbf{sl}(2)$, donde A y B son matrices de traza cero y tamaño 2×2 del sistema de control bilineal como está descrito encima.

En la sección dos trataremos el delicado aspecto de controlabilidad del sistema bilineal descrito encima y daremos condiciones necesarias y suficientes para que este sistema de control sea controlable en el caso en que la traza de A y B es cero. De manera más precisa como fue mencionado en la introducción de este trabajo, el resultado fundamental que expondremos, el sistema de control bilineal en el plano descrito encima, es controlable si y sólo si $\det[A, B] < 0$.

IV.1. Espacio Generado por $\Sigma = \{A, B\}$

Puesto que tenemos que $int(S_\Sigma)$ es no vacío en G_Σ. En el caso de sistemas bilineales de dimensión 2, $G_\Sigma = SL(2)$ si y sólo si el álgebra de Lie generado por $\Sigma = \{A, B\}$ es $\mathbf{sl}(2)$.

Como fue mencionado encima, en está sección buscaremos condiciones necesarias y suficientes para que el álgebra de Lie generado por A y B sea $\mathbf{sl}(2)$. Para ello, enunciaremos y demostraremos un lema que nos dará un resultado contundente para nuestros propósitos.

Lema IV.1. *Sean* $A, B \in \mathbf{sl}(2)$. *Entonces*

1. *Supongamos que* $\det B > 0$. *Entonces* $\det[A, B] \leq 0$ *y la igualdad se cumple si y solo si* A *es un múltiplo de* B. *Por lo tanto,* $\{A, B, [A, B]\}$ *es linealmente independiente si y solo si* $\det[A, B] < 0$.

2. *Asuma que* $\det B = 0$. *Entonces* $\det[A, B] \leq 0$*, la igualdad se da si y sólo si* $[A, B] = \lambda B$. *Además el conjunto* $\{A, B, [A, B]\}$ *es linealmente independiente si y sólo si* $\det[A, B] < 0$.

3. A *y* B *generan* $\mathbf{sl}(2)$ *si y sólo si* $\det[A, B] \neq 0$.

Demostración. 1. Si $\det B > 0$ y como tiene traza cero, entonces en una base adecuada de $\mathbf{R}^2$

$$B = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \qquad a \neq 0$$

luego B es antisimétrica. Por otro lado, recordemos que cualquier matriz A puede ser escrita como la suma de una matriz antisimétrica y simétrica, era suficiente escribir

$$A = \frac{A - A^t}{2} + \frac{A + A^t}{2}.$$

Entonces, en particular para la matriz A de este caso tenemos

$$A = A_1 + A_2$$

con A_1 antisimétrica y A_2 simétrica. Se sigue que A_1 es un múltiplo de B pues el espacio de las matrices antisimétricas tiene dimensión uno, luego tenemos

$$\begin{aligned}
[A, B] &= [A_1 + A_2, B] \\
&= [A_1, B] + [A_2, B] \\
&= [\alpha B, B] + [A_2, B] \\
&= \alpha[B, B] + [A_2, B] \\
&= [A_2, B].
\end{aligned}$$

Resulta que el corchete de una matriz simétrica y una matriz antisimétrica $[A_2, B]$, es simétrica. En efecto, ya que A_2 es simétrica y B es antisimétrica, entonces su corchete

$$\begin{aligned}
[A_2, B]^t &= (A_2 B - B A_2)^t \\
&= (A_2 B)^t - (B A_2)^t \\
&= B^t A_2^t - A_2^t B^t \\
&= (-B) A_2 - A_2(-B) \\
&= A_2 B - B A_2 \\
&= [A_2, B].
\end{aligned}$$

es simétrico, de donde el corchete $[A, B] = [A_2, B]$ de A, B es una matriz simétrica. Además, puesto que A, B son matrices de traza cero se tiene que la traza de $[A, B]$ es cero (pues $\mathbf{sl}(2)$ es un álgebra de Lie). Suponiendo $A_2 = \begin{pmatrix} u & v \\ v & -u \end{pmatrix}$ calculando el corchete de A_2 y B tenemos

$$[A, B] = [A_2, B] = \begin{pmatrix} 2va & -2ua \\ -2ua & -2va \end{pmatrix}.$$

De aquí se sigue que $\det[A, B] = -4a^2(v^2 + u^2) \leq 0$.

La igualdad $\det[A, B] = 0$ se cumple si y solo si

$$-4a^2(v^2 + u^2) = 0$$

como $a \neq 0$ se tiene que $v^2 + u^2 = 0$, pero esto es válido si y solo si $v = u = 0$. Esto es,

$$A_2 = \begin{pmatrix} u & v \\ v & -u \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Por lo tanto,

$$\begin{aligned} \det[A, B] = 0 &\Leftrightarrow A_2 = 0 \\ &\Leftrightarrow A = A_1, \ \text{(antisimétrica)} \\ &\Leftrightarrow A \ \text{es un múltiplo de} \ B. \end{aligned}$$

Con esto demostramos la primera parte de la afirmación.

Para concluir que el conjunto $\{A, B, [A, B]\}$ es linealmente independiente, notemos que al negar la última equivalencia tenemos

$$\det[A, B] < 0 \Leftrightarrow \{A, B\} \ \text{es linealmente independiente.}$$

Como $[A, B] = [A_2, B]$ es simétrica y $B \neq 0$ es antisimétrica, tenemos que el conjunto $\{B, [A, B]\}$ es linealmente independiente. Por otro lado tenemos también que el conjunto $\{A, [A, B]\}$ es linealmente independiente pues al ser $[A, B]$ simétrica, la única opción es que la matriz A sea simétrica, es decir, $A = A_2$, pero $A = A_2$ es un múltiplo de $[A, B] = [A_2, B]$ solo cuando $A_2 = 0$.

2. Si $\det B = 0$ podemos asumir que

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad y \quad A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

Entonces

$$[A, B] = \begin{pmatrix} -c & 2a \\ 0 & c \end{pmatrix}.$$

Por tanto, $\det[A, B] = -c^2 \leq 0$. Supongamos que $c = 0$, esto es, la igualdad se da si y solo si $[A, B] = 2aB$. De donde tenemos que $\{B, [A, B]\}$ es linealmente independiente si y solo si $\det[A, B] < 0$.

Por otro lado, $\{A, B\}$ es linealmente independiente si $\det[A, B] \neq 0$ (su prueba es un argumento usado en la sección (1) del capítulo 3). La oportunidad para que la matriz A sea un múltiplo de $[A, B]$, es que $c = 0$. Luego $\{A, [A, B]\}$ es un subconjunto linealmente independiente si $\det[A, B] < 0$. Por lo tanto, hemos demostrado que el subconjunto $\{A, B, [A, B]\}$ es linealmente independiente si y solo si $\det[A, B] < 0$ como queríamos.

3. En vista de los items anteriores, basta que consideremos el caso en que $\det A$ y $\det B$ son < 0. Luego en bases adecuadas podemos escribir

$$A = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} \quad y \quad B = \begin{pmatrix} x & y \\ z & -x \end{pmatrix},$$

con $-x^2 - zy < 0$, $a \neq 0$, luego

$$[A, B] = \begin{pmatrix} 0 & 2ay \\ -2az & 0 \end{pmatrix}.$$

Luego $\det[A, B] \neq 0$ si y solo si $yz \neq 0$. Por otro lado notemos que si $yz = 0$ entonces las matrices A, B y $[A, B]$ son triángulares luego no pueden ser linealmente independientes. Se sigue que el subconjunto $\{A, B, [A, B]\}$ es linealmente independientes si y solo si $\det[A, B] \neq 0$. Con eso termina la demostración del lema.

□

Como una consecuencia inmediata de este lema, tenemos el siguiente criterio para determinar si el álgebra de Lie generado por la familia de campos vectoriales $\Sigma = \{A, B\}$ es el álgebra de Lie $\mathbf{sl}(2)$.

Corolario IV.2. *Sean $A, B \in \mathbf{sl}(2)$. Las siguientes condiciones son equivalentes*

1. *El álgebra de Lie generada por $\Sigma = \{A, B\}$ es todo $\mathbf{sl}(2)$.*

2. *A, B y $[A, B]$ son linealmente independientes.*

3. $\det[A, B] \neq 0$.

Para ver la demostración, basta recordar que la dimensión de $\mathbf{sl}(2)$ es igual a 3.

IV.2. Controlabilidad

En esta sección estudiaremos la controlabilidad de sistemas bilineales de control de la forma

$$\dot{x} = Ax + uBx,$$

donde los controles son constantes por pedazos.

Nuestro interés está en tener condiciones sobre las matrices A, B de modo que el álgebra de Lie generada por estas sean todo $\mathbf{sl}(2)$. Por el corolario IV.2, el sistema es controlable si y solo si este semigrupo no tiene subconjuntos compactos propios e invariantes del espacio proyectivo $\mathbf{P}^1$. La existencia o no de estos subconjuntos invariantes es detectada observando las trayectorias del correspondiente sistema lineal. Recordemos que si C es una matriz con $tr(C) = 0$, entonces su polinomio característico es

$$x^2 + \det C$$

y los autovalores son $\pm\sqrt{-\det C}$. Luego existen tres casos,

1. Si $\det C > 0$ podemos fijar una base de modo que C es

   $$\begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix}$$

   y las trayectorias del sistema lineal $\dot{x} = Cx$ son círculos centrados en el origen. En la acción inducida al espacio proyectivo existe solo una trayectoria.

2. Si $\det C = 0$ con $C \neq 0$ tenemos en alguna base que

   $$C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

y las trayectorias del sistema lineal definido por C son rectas perpendiculares al eje-y y los puntos en el eje-x son puntos estacionarios. En la acción sobre el espacio proyectivo existen dos trayectorias: un punto fijo y una trayectoria densa que comienza y termina en el punto fijo.

3. Si $\det C < 0$ entonces tenemos que

$$C = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$$

y las trayectorias son hipérbolas. La acción inducida sobre el proyectivo consiste de dos puntos fijos y dos trayectorias uniendo los puntos fijos.

Para la controlabilidad de $\dot{x} = (A + uB)x$ distinguiremos los casos de acuerdo al $\det B$. Siempre supondremos que A y B son matrices diferentes de cero.

- Si $\det B > 0$ y $A \in \mathbf{sl}(2)$ es cualquier matriz que satisface $\det[A, B] \neq 0$. Esto es cuando las trayectorias de $\dot{x} = Bx$ son círculos y no es posible encontrar algún subconjunto compacto en el proyectivo que sea invariante bajo el semigrupo del sistema. Notemos también que por el corolario (IV.2) tenemos que el interior de S_Σ es no vacío.
- Si $\det B = 0$ y sea $\det[A, B] \neq 0$ tenemos controlabilidad ya que no es posible encontrar un subconjunto compacto invariante del espacio proyectivo.
- En el caso en que $\det B < 0$. Existen dos posibilidades:

  1. Si $\det A \geq 0$ y $\det[A, B] \neq 0$ entonces las trayectorias proyectadas de A son densas y tenemos controlabilidad.
  2. Si $\det A < 0$. En este caso la controlabilidad depende del signo de $\det[A, B]$.

     Como $\det A < 0$ entonces tenemos que la matriz A tiene dos autovalores reales distintos y claro dos autovectores. Los auto-espacios generados por estos son dos puntos en el espacio proyectivo. Por tanto, geométricamente tenemos dos casos, digamos los puntos fijos de A están en la misma trayectoria de B, en este caso tenemos controlabilidad. Si los puntos fijos están en trayectorias distintas entonces no puede haber conexión entre ellas, ya que la clausura de la trayectoria de B que contiene el punto fijo atractor de A es un subconjunto propio compacto e invariante bajo el semigrupo del sistema, en este caso no existe controlabilidad. El desarrollo algebraico de estas posibilidades se reduce a lo siguiente.

Para que los puntos fijos de A en el proyectivo $\mathbf{P}^1$ se encuentren en la misma trayectoria de B, los autovectores de A deben estar en el mismo cuadrante del plano y además deben ser distintos a los autovectores de B. La segunda cuestión es inmediata de la suposición $\det[A, B] \neq 0$, pues en este caso, el conjunto $\{A, B\}$ es linealmente independiente. Si suponemos que existe un vector u que es simultáneamente un autovector de A y de B, entonces

$$Au = \lambda u, \quad Bu = \alpha u.$$

De donde,

$$(\lambda^{-1})Au + (-\alpha^{-1})Bu = 0. \tag{IV.1}$$

Esto es, existen escalares no nulos y u no nulo tal que satisface la ecuación (IV.1), luego el conjunto $\{A, B\}$ no puede ser linealmente independiente.

Para buscar la condición que garantice que los puntos fijos de A se encuentren en la misma trayectoria de B, tenemos que eso es equivalente a buscar la condición que garantice que los autovectores de A se encuentren en el mismo cuadrante. Para esto hacemos lo siguiente.

Fijemos una base tal que

$$B = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} \ \ y \ \ A = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}.$$

Los puntos fijos de A están sobre la misma trayectoria de B si y solo si los autovectores de A están en el mismo cuadrante. Ahora, los autovalores de A son

$$\lambda_{\pm} = \pm\sqrt{-detA} = \pm\sqrt{\alpha^2 + \gamma\beta} \in \mathbf{R}.$$

Sea (x, y) un autovector de A, esto es, satisface la relación

$$A(x, y) = \lambda_{\pm}(x, y),$$

reemplazando los valores tenemos

$$\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_{\pm} x \\ \lambda_{\pm} y \end{pmatrix}.$$

Reemplazando el autovalor y haciendo los cálculos tenemos el sistema

$$(\alpha \pm \sqrt{\alpha^2 + \gamma\beta})x + \beta y = 0$$

$$\gamma x + (-\alpha \pm \sqrt{\alpha^2 + \gamma\beta})y = 0.$$

Puesto que $y \neq 0$ (si lo fuera el autovector caería a un punto fijo de B en el proyectivo) podemos asumir que $y = 1$, entonces

$$x_{\pm} = \frac{-\beta}{\alpha \pm \sqrt{\alpha^2 + \gamma\beta}}$$

son los posibles valores de la primera coordenada de los autovectores de A, es decir, los autovectores de A son

$$(x_{\pm}, 1) = \Big(\frac{-\beta}{\alpha \pm \sqrt{\alpha^2 + \gamma\beta}}, 1 \Big).$$

Estos autovectores no están en el mismo cuadrante si y solo si $x_{\pm}$ tienen signos opuestos, y esto último depende de sus denominadores. Más precisamente, los autovectores están en el mismo cuadrante si y solo si $x_{\pm}$ tiene el mismo signo, pero esto ocurre si el producto de sus denominadores

$$-\gamma\beta = (\alpha + \sqrt{\alpha^2 + \gamma\beta})(\alpha - \sqrt{\alpha^2 + \gamma\beta}) > 0,$$

luego los autovectores de A están en el mismo cuadrante si y solo si $-\gamma\beta > 0$, de aquí $\gamma\beta < 0$.

Por otro lado,

$$[A, B] = \begin{pmatrix} 0 & -2a\beta \\ 2a\gamma & 0 \end{pmatrix},$$

y claro

$$\det[A, B] = 4a^2\gamma\beta$$

tenemos que

- El sistema es controlable en el caso en que $\det[A, B] < 0$, y no controlable si $\det[A, B] > 0$.

Esto concluye el analisis de sistemas que generan $SL(2)$. En resumen tenemos el siguiente teorema que reúne las ideas de todo lo hecho encima y que es el resultado central de este trabajo:

Teorema IV.3. *Sean $A, B \in \mathbf{sl}(2)$. Entonces el semigrupo S_Σ generado por $\Sigma = \{A, B\}$ coincide con $SL(2)$ si y solo si* $\det[A, B] < 0$

Demostración. Tenemos que $S_\Sigma = SL(2)$ si y solo si Σ es controlable sobre $\mathbf{R}^2 - \{0\}$, luego observemos que en los casos analizados de encima tenemos controlabilidad si y solo si $\det[A, B] < 0$. □

CAPÍTULO V

Conclusiones y Recomendaciones

V.1. Conclusiones

Una de las características del trabajo que aquí se presenta es su transversalidad a varias ramas de la matemática, como ser la Teoría de Ecuaciones Diferenciales, Topología, Álgebra, Análisis, Geométria Diferencial y Teoría de Lie. Este trabajo además de ser una incursión en una rama relativamente joven de la matemática, como es la Teoría Geométrica de Control, constituye un ejemplo interesante de la interrelación de los elementos de la matemática, un fenómeno muy propio de los últimos tiempos.

Los resultados principales desarrollados en este trabajo fueron publicados en 1996, por los profesores Braga, Gonçalvez, Rocío del Departamento de Matemática en la Universidad Estadual de Maringa y el Prof. Luis San Martin del Instituto de Matemática en la Universidad Estadual de Campinas, en Brasil [4].

En el presente trabajo realizamos un análisis para obtener la controlabilidad de un sistema bilineal de control definido en $\mathbf{R}^2 - \{0\}$ dado por

$$\dot{x} = Ax + uBx \tag{V.1}$$

donde las matrices A, B son de traza cero y de tamaño 2×2, y las funciones de control son constantes por pedazos.

La controlabilidad del sistema bilineal (V.1) para controles constantes por pedazos, esta descrita en la sección (IV.2), sin embargo realizamos un recuerdo de dichos casos analizados. Considerando que $trA = trB = 0$ y claro ninguna de ellas la matriz nula, entonces

1. Si $\det B > 0$ y $A \in \mathbf{sl}(2)$ es cualquier matriz que satisface $\det[A, B] \neq 0$ entonces el sistema (V.1) es controlable.

2. Si $\det B = 0$ y si $\det[A, B] \neq 0$ entonces el sistema bilineal (V.1) es controlable.

3. Si $\det B < 0$, $\det A \leq 0$ y $\det[A, B] \neq 0$ entonces el sistema es controlable.

4. Si $\det B < 0$, $\det A < 0$ con $\det[A, B] \neq 0$. Entonces el sistema (V.1) es controlable si $\det[A, B] < 0$ y no controlable si $\det[A, B] > 0$.

De manera resumida tenemos el resultado central de esta tesis, el cual dice

$$\text{(V.1) es controlable} \Leftrightarrow \det[A, B] < 0.$$

V.2. Recomendaciones

Un problema abierto de los sistemas bilineales de control, es precisamente considerar el espacio estado $\mathbf{R}^3$. Para este caso es posible de intentar copiar las ideas expuestas en este trabajo, y buscar una caracterización limpia para la controlabilidad del sistema bilineal.

De hecho el inicio está en encontrar los grupos de Lie lineales que actúan transitivamente sobre $\mathbf{R}^3 - \{0\}$, el cual puede ser considerado como un trabajo de tesis de licenciatura, estos resultados pueden ser encontrados en [2] y [3].

De manera natural se considera el estudio del control óptimo, las ideas fundamentales y resultados centrales de está teoría puede ser consultada en [13] y [10].

APÉNDICE A

Campos de Vectores en Variedades Diferenciables

En este apéndice se estudia con bastante detenimiento los campos de vectores en una variedad diferenciable. Para ello se precisan los conceptos de vector tangente, fibrado tangente y diferencial de una aplicación diferenciable, en términos de las dos perspectivas más usuales: la geométrica y la analítica. Luego se generalizan las ideas principales de la teoría de ecuaciones diferenciables al contexto de las variedades diferenciables: curvas integrales, existencia y unicidad de soluciones, soluciones máximas, flujos a 1-parámetro y campos conjugados. A continuación se dota de una estructura de álgebra de Lie al conjunto de campos de vectores diferenciables en una variedad, se verá, que el corchete de está álgebra permite obtener importantes consecuencias geométricas sobre el comportamiento de las curvas integrales de campos de vectores y variedades integrales de conjuntos locales de campos de vectores (distribuciones).

Se suponen familiares los conceptos de variedad diferenciable, aplicación diferenciable (submersión, inmersión y encaje) y subvariedad. Sin embargo, por la notación la mejor referencia para consultar este material es [10].

A.1. Campos de Vectores

Definición A.1. *Sea M una variedad diferenciable. Un campo de vectores en un conjunto abierto $U \subset M$ es un levantamiento de U en TM, esto es, un aplicación $X : U \to TM$ tal que*

$$X \circ \pi = id_U.$$

Se dice que el campo de vectores es diferenciable, si $X \in C^\infty(U, TM)$. El conjunto de campos vectoriales diferenciables es denotado por $\mathcal{X}(U)$.

Si $X \in \mathcal{X}(U)$ y $m \in U$ entonces $X(m)$, que usualmente lo denotaremos por X_m, es un elemento de T_mM. Claramente, se pueden definir operaciones puntuales que hacen de $\mathcal{X}(U)$ un espacio vectorial.

Definición A.2. *Sea U un abierto en la variedad diferenciable M y sea X un campo de vectores (no necesariamente diferenciable) en U. Para toda función $f \in C^\infty(U)$ definimos la aplación $Xf : U \to R$ dado por*

$$Xf(m) = X_m f$$

para todo $m \in U$.

Proposición A.3. *Sea X un campo vectorial en M. Las siguientes afirmaciones son equivalentes:*

1. *X es diferenciable, i.e. $X \in \mathcal{X}(U)$.*
2. *Si $(U, x_1, ..., x_d)$ es un sistema de coordenadas locales en M, y si $a_1, ..., a_d$ es la colección de funciones en U definidas por*
$$X|_U = \sum_{i=1}^{d} a_i \frac{\partial}{\partial x_i},$$
entonces $a_1, ..., a_d \in C^\infty(U)$.
3. *Siempre que V es un abierto en M y $f \in C^\infty(V)$, se tiene que $Xf \in C^\infty(V)$.*

A.1.1. Curvas integrales y el teorema del flujo local

Definición A.4. *Sea M una variedad diferenciable y $X \in \mathcal{X}(M)$. Una curva suave σ en M es una* curva integral *de X si*

$$\dot{\sigma} = X(\sigma(t))$$

para todo t en el dominio de σ.

Sea $X \in \mathcal{X}(M)$ y $m \in M$. En esta subsección se responderán las interrogantes naturales: ¿existe una curva integral de X?, si la respuesta es afirmativa, ¿será única?.

Una curva suave $\gamma : (a,b) \to M$ es una curva integral del campo vectorial X si, y solamente si

$$d\gamma(\frac{d}{dr}|_t) = X(\gamma(t)), \qquad (t \in (a,b)). \tag{A.1}$$

A continuación, se interpretará la ecuación (A.1) en términos de coordenadas locales.
Sea $t_0 \in (a,b)$ y $\gamma(t_0) = m$. Sea (U, φ) un sistema de coordenadas respecto de m con funciones coordenadas $x_1, x_2, ..., x_d$. De la proposición (A.3) se tiene

$$X|_U = \sum_{i=1}^{d} f_i \frac{\partial}{\partial x_i} \tag{A.2}$$

donde los f_i son funciones C^∞ en U. Por otra parte, para cada $t \in (a,b)$ tal que $\gamma(t) \in U$ se verifica

$$d\gamma(\frac{\partial}{\partial dr}|_t) = \sum_{i=1}^{d} \frac{\partial d(x_i \circ \gamma)}{\partial dr}|_t \frac{\partial}{\partial x_i}|_{\gamma(t)}. \tag{A.3}$$

Así, tenemos que la ecuación de encima se convierte en

$$\sum_{i=1}^{d} \frac{\partial d(x_i \circ \gamma)}{\partial dr}|t \frac{\partial}{\partial x_i}|_{\gamma(t)} = \sum_{i=1}^{d} f_i(\gamma(t)) \frac{\partial}{\partial x_i}|_{\gamma(t)}. \tag{A.4}$$

Entonces γ es una curva integral de X en $\gamma^{-1}(U)$, tal que $\gamma_t(0) = m$ si, y solo si,

$$\frac{\partial d(x_i \circ \gamma)}{\partial dr}|_t = f_i(\gamma(t)), \qquad i = 1, 2, ..., d. \tag{A.5}$$

o equivalentemente

$$\frac{\gamma_i}{dr}|_t = f_i \circ \gamma^{-1}(\gamma_1(t), ..., \gamma_d(t)), \qquad i = 1, 2, ..., d; t \in \gamma^{-1}(U) \tag{A.6}$$

donde $\gamma_i = x_i \circ \gamma$. Las ecuaciones en () forman un sistema de ecuaciones diferenciales ordinarias de primer orden, para el cual existen teoremas de existencia y unicidad de soluciones. Estos teoremas implican un resultado análogo en el contexto de campos en variedades diferenciables. Algo interesante de destacar es que en este ámbito la unicidad de curvas integrales máximas dependen fuertemente de que el espacio en cuestión sea Hausdorff.

Proposición A.5. *Sea X un campo vectorial diferenciable sobre la variedad diferenciable M. Para cada punto $m \in M$ existen a_m y b_m en $\mathbf{R} \cup \{\pm\infty\}$, y una curva suave*

$$\gamma_m : (a_m, b_m) \to M \tag{A.7}$$

tal que

1. *$0 \in (a_m.b_m)$ y $\gamma_m(0) = m$.*
2. *γ_m es una curva integral de X.*
3. *Si $\mu : (c, d) \to M$ es una curva suave que satisface a y b, entonces $(c, d) \subset (a_m, b_m)$ y $\mu = \gamma_m|_{(c,d)}$.*

Definición A.6. *Sea X un campo vectorial diferenciable en la variedad M. Con la terminología de la proposición A.5, para cada $t \in \mathbf{R}$ se define la transformaciíon X_t con dominio $\mathbf{D}_t = \{m \in M : t \in (a_m, b_m)\}$ por*

$$X_t(m) = \gamma_m(t). \tag{A.8}$$

A continuación se presenta un resultado básico de varias ramas de la matemática como la teoría geométrica de control, geometría diferencial y los sistemas dinámicos.

Teorema A.7 (Flujo Local)**.** *Sea X una campo vectorial diferenciable en la variedad M.*

1. *Para cada $m \in M$, existe una vecindad abierta V de m y un $\epsilon > 0$ tal que la aplicación*

$$(t,p) \mapsto X_t(p) \tag{A.9}$$

   está definida en $(-\epsilon,\epsilon) \times V$ y es C^∞.

2. *$\mathbf{D}_t$ es abierto para cada t.*

3. *$\bigcup_{t>0} \mathbf{D}_t = M$.*

4. *$X_t : \mathbf{D}_t \to \mathbf{D}_{-t}$ es un difeomorfismo con inversa X_{-t}.*

5. *Sean s y t números reales. Entonces el dominio de $X_s \circ X_t$ esta contenido en $\mathbf{D}_{s+t}$, en general no se da la igualdad. Sin embargo, el dominio de $X_s \circ X_t$ es $\mathbf{D}_{s+t}$ cuando s y t tienen el mismo signo. Es más, sobre el dominio de $X_s \circ X_t$ tenemos*

$$X_s \circ X_t = X_{s+t} \tag{A.10}$$

Remarca A.8. *Sea $m \in M$ fijo, entonces el camino $t \mapsto X_t(m)$ es una curva integral de X además $\frac{d}{dt}X_t(m) = X(X_t(m))$.*

Remarca A.9. *Existen casos en los que $\mathcal{D}_t$ es vacío, por ejemplo para el campo vectorial $\frac{d}{dr}$ en $(0,1)$, se tiene $\mathbf{D}_t = \emptyset$ para $t > 1$.*

Definición A.10. *Un campo de vectores diferenciable X en M es completo si $\mathbf{D}_t = M$ para todo t (es decir, si el dominio de $\gamma_m(t)$ es $\mathbf{R}$ para todo $m \in M$). En este caso, las transformaciones $\{X_t\}_{t\in\mathbf{R}}$ forman un grupo de transformaciones de M denominado* grupo a 1-parámetro de X, *además la aplicación () definida en $\mathbf{R} \times M$ se denomina* flujo generado por X. *Si X no es completo, las transformaciones X_t no forman un grupo porque sus dominios dependen de t. En este caso, la colección de transformaciones X_t se denomina el* grupo local a 1-parámetro de X.

Definición A.11. *Sea X un campo de vectores diferenciable en la variedad M, x es un* punto singular *de X si $X_x = 0$. Los puntos que no son singulares se denominan* puntos regulares *del campo X.*

La unicidad de las curvas integrales máximas implica el siguiente resultado.

Proposición A.12. *Sea X un campo diferenciable. Si p es un punto singular de X, entonces $X_t(p) = p$ para todo $t \in \mathbf{R}$.*

En un curso de Ecuaciones Diferenciales Ordinarias se prueba que las soluciones máximas de una ecuación definida en un abierto de $\mathbf{R}^d$ no permanece en ningún subconjunto compacto, ver teorema 3 de la página 17 de [43]. Existe un resultado similar para las curvas integrales de campos de vectores en variedades diferenciables.

Proposición A.13. *Sean X un campo vectorial diferenciable en la variedad M y $m \in M$. Sea (a_m, b_m) como en la proposición A.5. Supongamos que $b_m < \infty$ y $\{t_n\} \subset (a_m, b_m)$ es una sucesión creciente convergente a b_m. Entonces $X_{t_n}(m)$ no puede estar contenida en ningún conjunto compacto. En particular, la sucesión X_{t_n} no puede converger a un limite en M. Existe una propiedad similar para una sucesión decreciente convergente a a_m.*

Corolario A.14. *Sea X un campo vectorial diferenciable en la variedad M. Si una curva integral máxima de X esta definida en un intervalo acotado, entonces la curva integral es un subconjunto cerrado de M.*

Corolario A.15. *Si el campo vectorial diferenciable X se anula fuera de un conjunto compacto de M, entonces X es completo.*

Corolario A.16. *Si M es una variedad compacta, entonces todo campo de vectores C^∞ es completo.*

El siguiente resultado es muy útil para probar que un campo es completo. Además, permite dar otra demostración del corolario.

Proposición A.17. *Sea X un campo de vectores C^∞ en M. Si existe $\epsilon > 0$ tal que $X_t(m)$ esta definida para todo $(t, m) \in (-\epsilon, \epsilon) \times M$, entonces X es un campo de vectores completo.*

A.2. El álgebra de Lie de campos de vectores

Sea M una variedad diferenciable. En la sección anterior se hizo notar que $\mathcal{X}(M)$ tiene estructura de espacio vectorial, a continuación se dota a este espacio de estructura de álgebra de Lie y se estudian algunas de sus propiedades.

Definición A.18 (Corchete de Lie). *Sean X e Y campos de vectores diferenciables en M, se define el campo de vectores $[X,Y]$, denominado* corchete de Lie *de X e Y por*

$$[X,Y]_m(f) = X_m(Yf) - Y_m(Xf),$$

donde $m \in M$ y f es cualquier función diferenciable en una vecindad de m.

Proposición A.19. *Si $X, Y, Z \in \mathcal{X}(M)$ entonces*

1. *$[X,Y] \in \mathcal{X}(M)$.*
2. *Si $c \in \mathbf{R}$ entonces $[c(X+Y),Z] = c[X,Z] + [Y,Z]$.*
3. *$[X,Y] = -[Y,X]$.*
4. *$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$.*

Demostración. La segunda y tercera afirmación son inmediatas a partir de la definición. Sea $m \in M$ y f una función C^∞ en una vecindad de m, el campo vectorial $[X,Y]$ induce la aplicación $[X,Y]f = X(Yf) - Y(Xf)$, claramente diferenciable. Entonces la proposición A.3 prueba la primera afirmación. Por otra parte,

$$\begin{aligned}
[[X,Y],Z]_m(f) &= [X,Y]_m(Zf) - Z_m([X,Y]f) \\
&= X_m(Y(Zf)) - Y_m(X(Zf)) - Z_m((X(Yf)) - Y(Xf)) \qquad \text{(A.11)} \\
&= X_m(Y(Zf)) - Y_m(X(Zf)) - Z_m(X(Yf)) + Z_mY(Xf)
\end{aligned}$$

Análogamente,

$$[[Y,Z],X]_m(f) = Y_m(Z(Xf)) - Z_m(Y(Xf)) - X_m(Y(Zf)) + X_m(Z(Yf)) \qquad \text{(A.12)}$$

$$[[Z,X],Y]_m(f) = Z_m(X(Yf)) - X_m(Z(Yf)) - Y_m(Z(Xf)) = Y_m(X(Zf)) \tag{A.13}$$

Sumando las ecuaciones (A.11),(A.12) y (A.13) se tiene

$$[[X,Y],Z]_m(f) + [[Y,Z],X]_m(f) + [[Z,X],Y]_m(f) = 0.$$

Probando así la última afirmación. □

Corolario A.20. *Sea M una variedad diferenciable, entonces el corchete de Lie*

$$[\cdot,\cdot] : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$$

induce una estructura de álgebra de Lie en $\mathcal{X}(M)$.

A.2.1. Campos de vectores en $\mathbf{R}^d$

En la teoría de ecuaciones diferenciales, un campo vectorial en $\mathbf{R}^d$ es una aplicación $F : \mathbf{R}^d \to \mathbf{R}^d$ diferenciable. La aplicación F induce un campo de vectores X^F definido por

$$X^F = \sum_{i=1}^{d} r_i \circ F \frac{\partial}{\partial r_i},$$

donde $r_1, ..., r_d$ son las coordenadas canónicas de $\mathbf{R}^d$. Ambas nociones coinciden, en el sentido de que las curvas integrales lo hacen. Recíprocamente, todo campo de vectores diferenciable induce una aplicación diferenciable F proposición A.3. Entonces los campos de vectores están identificados con los *operadores diferenciales parciales de primer orden lineales*, es más, existe una relación algebraica que se precisa en la siguiente proposición.

Proposición A.21. *Sean X, Y campos de vectores diferenciables en $\mathbf{R}^d$. Supongamos que*

$$X = \sum_{i=1}^{d} r_i \circ F \frac{\partial}{\partial r_i}, \qquad Y = \sum_{i=1}^{d} r_i \circ G \frac{\partial}{\partial r_i}$$

donde $r_1, ..., r_d$ *denotan las coordenadas canónicas de* $\mathbf{R}^d$*. Entonces*

$$[X,Y] = (\sum_{i=1}^{d} r_i \circ F \frac{\partial}{\partial r_i})(\sum_{i=1}^{d} r_i \circ G \frac{\partial}{\partial r_i}) - (\sum_{i=1}^{d} r_i \circ G \frac{\partial}{\partial r_i})(\sum_{i=1}^{d} r_i \circ F \frac{\partial}{\partial r_i})$$

donde las operaciones en el lado derecho corresponden a las del álgebra de los operadores diferenciales parciales.

Corolario A.22. *Sean* X, Y *campos de vectores diferenciables en* $\mathbf{R}^d$*. Sean* $F, G : \mathbf{R}^d \to \mathbf{R}^d$ *las aplicaciones diferenciables inducidas, es decir,* $Y = Y^G$ *y* $X = X^F$*. Entonces* $DG(\cdot)F(\cdot) - DF(\cdot)G(\cdot)$ *induce el campo* $[X, Y]$*, es decir*

$$[X,Y](x) = \sum_{i=1}^{d} r_i (DG(x)F(x) - DF(x)G(x)) \frac{\partial}{\partial r_i}|_x$$

para todo $x \in \mathbf{R}^d$*.*

Ejemplo A.23. *Sean* X, Y *los campos en* $\mathbf{R}^3$ *definidos por*

$$X(x,y,z) = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}, \qquad Y(x,y,z) = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}.$$

Entonces

$$\begin{aligned}[X,Y](x,y,z) &= (y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y})(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}) - (z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z})(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}) \\ &= -z\frac{\partial}{\partial x} + x\frac{\partial}{\partial z}.\end{aligned}$$

Por otra parte, las aplicaciones inducidas por X *e* Y *son respectivamente,* $F(x,y,z) = (y, -x, 0)$

y $G(x,y,z)=(0,z,-y)$. *Entonces, denotando* $x=(x,y,z)$ *tenemos*

$$
\begin{aligned}
DG(x)F(x)-DF(x)G(x) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}\begin{pmatrix} y \\ -x \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ z \\ -y \end{pmatrix} \\
&= \begin{pmatrix} -z \\ 0 \\ x \end{pmatrix}.
\end{aligned}
$$

Ejemplo A.24 (Flujo de campos lineales)**.** *Sea* F *el campo vectorial diferenciable en* $\mathbf{R}^d$ *definido por* $F(x)=Ax+b$*, donde* A *es una matriz de* $d\times d$ *y* $b\in\mathbf{R}^d$*. Las curvas integrales del campo de vectores inducido* X^F *se obtiene considerando el sistema de ecuaciones diferenciales*

$$\gamma'(t)=A\gamma(t)+b.$$

Entonces el flujo estará dado por

$$X_t^F(x)=e^{At}(x+\int_0^t e^{-\tau A}bdr).$$

A.3. Campos de vectores φ-relacionados

Cuando se estudian objetos matemáticos siempre se estudian sus interrelaciones, generalmente, a través de aplicaciones que conservan la estructura de los objetos considerados. Por ejemplo, los grupos abstractos, espacios vectoriales, espacios topológicos están interrelacionados con sus similares a través de homomorfismos de grupos, transformaciones lineales, aplicaciones continuas respectivamente. Es de interés estudiar los objetos que sean esencialmente iguales, entonces surgen los conceptos de isomorfismo de grupos, isomorfismo lineal, homeomorfismo, etc...

En esta se sección se introduce el concepto de campos por una aplicación[1], lo que permitirá precisar

[1]Esta terminología es usual en la Geometría diferencial pues en la Teoría de Sistemas Dinámicos se dicen campos conjugados.

que significa que dos campos sean equivalentes. Intuitivamente es de esperarse que si $X \in \mathcal{X}(M)$, $Y \in \mathcal{X}(N)$ son equivalentes, entonces por lo menos los espacios de estado deben ser equivalentes, es decir, existe un difeomorfismo $\varphi : M \to N$. Es inmediato notar que la existencia de tal difeomorfismo no es suficiente, es de esperarse que además relacione las curvas integrales.

Definición A.25. *Sea $\varphi : M \longrightarrow N$ una aplicación diferenciable. Dos campos de vectores $X \in \mathcal{X}(M)$ e $Y \in \mathcal{X}(N)$ se dicen φ-relacionados si*

$$d\varphi \circ X = Y \circ \varphi.$$

Definición A.26. *Sea $\varphi : M \to N$ un difeomorfismo. Para cualquier $X \in \mathcal{X}(M)$ se define $\varphi^* X \in \mathcal{X}(M)$ por*

$$\varphi^* X = d\varphi^{-1} \circ Y \circ \varphi.$$

Remarca A.27. *$\varphi^* X$ y X son φ-relacionados.*

Proposición A.28. *Sean $X \in \mathcal{X}(M)$ e $Y \in \mathcal{X}(N)$ dos campos φ-relacionados. Si $\gamma : (a, b) \to M$ es una curva integral de X, entonces $\varphi \circ \gamma : (a, b) \to N$ es una curva integral de Y.*

Demostración. Existe $\varphi : M \to N$ tal que $d\varphi \circ X = Y \circ \varphi$, además $X(\lambda(t)) = \dot{\lambda}(t)$. Entonces $Y((\varphi \circ \lambda)(t)) = (Y \circ \varphi)(\lambda(t)) = (d\varphi \circ X)(\lambda(t)) = d\varphi(\dot{\lambda}(t)) = (\varphi \dot{\circ} \lambda)(t)$. $\square$

Proposición A.29. *Sean $X \in \mathcal{X}(M)$, $Y \in \mathcal{X}(N)$ campos de vectores φ-relacionados. Entonces $\varphi \circ X_t = Y_t \circ \varphi$ cuando ambos lados están definidos. En particular, si φ es un difeomorfismo entonces $(\varphi^* Y)_t = \varphi \circ Y_t \circ \varphi$.*

Corolario A.30. *Un campo de vectores diferenciable en M es invariante por el difeomorfismo φ, es decir $\varphi^* X = X$ si, y sólo si, φ conmuta con X_t para todo $t \in \mathbf{R}$. En particular $X_t^* X = X$.*

Hasta el momento, se ha probado que la aplicación que relaciona dos campos preserva las curvas integrales. A continuación se mostrará que la relación de dos campos por una relación se preserva por el corchete, como consecuencia, se verá una forma de obtener homomorfismos entre las álgebras de Lie de campos de vectores.

Proposición A.31. *Sean $c \in \mathbf{R}$, $X^1, X^2 \in \mathcal{X}(M)$, $Y^1, Y^2 \in \mathcal{X}(N)$ y $\varphi : M \to N$ una aplicación diferenciable. Si X^i e Y^i están φ-relacionados para $i = 1, 2$. entonces $cX^1 + X^2$ y $cY^1 + Y^2$ están φ-relacionados, también $[X^1, X^2]$ y $[Y^1, Y^2]$ están φ-relacionados.*

Corolario A.32. *Sea $\varphi : M \to N$ un difeomorfismo local. Si $Y \in \mathcal{X}(N)$ se define $\varphi^*Y(x) = (df(x))^{-1}Y(\varphi(x))$. Entonces la aplicación $\varphi^* : \mathcal{X}(N) \to \mathcal{X}(M)$ es un homomorfismo de álgebras de Lie, es decir,*

$$\varphi^*[Y^1, Y^2] = [\varphi^*Y^1, \varphi^*Y^2].$$

A.4. Geometría del corchete de Lie de campos de vectores

En esta sección se estudian propiedades del corchete de Lie de campos vectoriales de manera que se aclare su significado geométrico, en particular, se muestran distintas formas de calcularlos.

Sea $X \in \mathcal{X}(M)$ y $f : M \to \mathbf{R}$ una función diferenciable. A continuación veremos como cambia f respecto del flujo generado por X, es decir, se estudia $f(X_t(m))$ cuando t varia. Por la regla de la cadena siempre que $X_t(m)$ este definido se tiene

$$\frac{d}{dt}f(X_t(m)) = df(\frac{d}{dt}X_t(m)) = df(X(X_t(m))) = (Xf)(X_t(m)). \tag{A.14}$$

En particular

$$\frac{d}{dt}|_{t=0}f(X_t(m)) = Xf(m) = X_m f. \tag{A.15}$$

Definición A.33 (La derivada de Lie de funciones)**.** *Para $X \in \mathcal{X}(M)$ y $f \in C^\infty(M)$ se define $\mathcal{L}_X f : M \to \mathbf{R}$ por*

$$\mathcal{L}_X f(m) = \frac{d}{dt}|_{t=0}f(X_t(m)) = \lim_{t\to 0} \frac{f(X_t(m)) - f(m)}{t}$$

Obsérvese que para cualquier m, si t es suficientemente pequeño, $X_t(m)$ esta definido. La aplicación $\mathcal{L}_X f$ se conoce como *la derivada de Lie de f en la dirección de X*.

A la luz de la igualdad de encima se habría probado lo siguiente.

Proposición A.34. *Si $X \in \mathcal{X}(M)$ y $f \in C^\infty(M)$, entonces $\mathcal{L}_X f = Xf = df \circ X$. En particular $\mathcal{L}_X f \in C^\infty(M)$.*

Ahora se puede continuar con el proceso de diferenciación, de () se obtiene

$$\frac{d^2}{dt^2} f(X_t(m)) = \frac{d}{dt} Xf(X_t(m)) = X(Xf)(X_t(m)).$$

El teorema de Taylor implica el siguiente resultado.

Proposición A.35. *Sean $X \in \mathcal{X}(M)$ y $f \in C^\infty(M)$. Entonces*

$$f(X_t(m)) = f(x) + t(Xf)(m) + \frac{t^2}{2}(X^2 f)(m) + ... + \frac{t^k}{k!}(X^k f)(m) + O(t^{k+1}), \tag{A.16}$$

donde $k \in \mathbf{N}$ y $X^2 f = X(Xf)$, $X^3 f = X(X^2 f)$, etc.

Definición A.36 (La derivada de Lie de campos de vectores)**.** *Sean $X, Y \in \mathcal{X}(M)$, se define el campo de vectores $\mathcal{L}_X Y$ por*

$$(\mathcal{L}_X Y)_m = \lim_{t \to 0} \frac{dX_{-t}(Y_{X_t(m)}) - Y(m)}{t} = \frac{d}{dt}|_{t=0}(dX_{-t}(Y_{X_t(m)})).$$

Remarca A.37. *Es posible definir la* derivada direccional de una forma diferencial en la dirección de un campo X. *Naturalmente, coincide con los dos casos particulares que se presentaron en esta sección. Es importante recalcar que $T_m M$ es un espacio vectorial de dimensión finita, por tanto tiene asociado todas las nociones del cálculo de espacios euclídeos (por tanto la noción de variedad diferenciable), el límite de la definición anterior puede entenderse en ese contexto o equivalentemente como un vector tangente de un camino en TM.*

A continuación se muestra que la derivada de Lie de un campo no es más que el corchete de Lie.

Teorema A.38. *Si $X, Y \in \mathcal{X}(M)$ entonces $\mathcal{L}_X Y = [X, Y]$, en particular $\mathcal{L}_X Y \in \mathcal{X}(M)$.*

Remarca A.39. *La conclusión del teorema anterior suele escribirse:*

$$[X, Y] = \lim_{t \to 0} \frac{1}{t}(X_t^* Y - Y).$$

Corolario A.40. *Sean $X, Y \in \mathcal{X}(M)$. Entonces*

$$X_t^*[X, Y] = \lim_{t \to 0} \frac{1}{t}(X_{s+t}^* Y - X_s^* Y).$$

Remarca A.41. *La conclusión del anterior corolario suele escribirse:*

$$\frac{d(X_t^* Y)}{dt}|_{t=s} = X_s^*[X, Y].$$

Corolario A.42. *Sean $X, Y \in \mathcal{X}(M)$. Entonces $X_t \circ Y_s = Y_s \circ X_t$, para cualesquiera s, t tales que las expresiones están definidas si y solo si $[X, Y] = 0$.*

Remarca A.43. *Sea $m \in M$. Si s, t son suficientemente pequeños entonces $X_t \circ Y_s(m)$ y $Y_s \circ X_t(m)$ están definidos.*

Existe una caracterización mas geométrica del corchete de Lie de dos campos $X, Y \in \mathcal{X}(M)$. Se verá que el corchete de Lie de X e Y es el çonmutador infinitesimal" de los dos grupos a 1-parámetro X_t e Y_t.

Teorema A.44. *Sean $X, Y \in \mathcal{X}(M)$ y $m \in M$. Sea*

$$\lambda(t) = Y_{-\sqrt{t}} X_{-\sqrt{t}} Y_{\sqrt{t}} X_{\sqrt{t}}(m),$$

definido para t suficientemente pequeño y positivo. Entonces

$$[X, Y]_m f = \lim_{t \to 0^+} \frac{f(\lambda(t)) - f(m)}{t} = \frac{d}{dt}|_{t=0^+} \lambda(t).$$

A.5. Distribuciones y el teorema de Frobenius

Definición A.45. *Sea M una variedad diferenciable. Una* distribución *Δ en M es una aplicación que a cada $m \in M$ asocia un subespacio $\Delta(m) \subset M_m$, del espacio tangente a M en m. Una distribución Δ se dice* regular, *o* no singular, *si la dimensión de $\Delta(m)$ permanece constante como función de m, en caso contrario la distribución se dice* singular. *La dimensión constante de $\Delta(m)$, en el caso de una distribuión regular, es la dimensión de la distribución Δ.*

Definición A.46. *Una distribución Δ en M es diferenciable en un punto $m \in M$, si existen campos de vectores diferenciables $X^1, ..., X^l$ en un abierto U que contenga a m, tal que $Span\{X^1(x), ..., X^l(x)\} = \Delta(x)$, para cada $x \in U$. Una distribución es diferenciable si es diferenciable en cada punto.*

Definición A.47. *Una subvariedad (N, φ) de M es una variedad integral de una distribución Δ en M si $d\varphi(N_m) = \delta(\varphi(n))$ para cada $n \in \mathbf{N}$. Una distribución Δ en M se dice integrable en m si existe una variedad integrable de Δ que contenga a m. Una distribución es integrable en M si es integrable en cada punto m.*

Definición A.48. *Un campo vectorial X en M se dice que pertenece a la distribución Δ en M, $X \in \Delta$, si $X_m \in \delta(m)$ para todo $m \in M$. Una distribución se dice* involutiva *si $[X, Y]$ pertenece a Δ cuando X, Y son campos diferenciables que pertenecen a Δ.*

Proposición A.49. *Si Δ es una distribución integrable en M entonces Δ es involutiva.*

Se ha probado que la involutividad es una condición necesaria para la integrabilidad de una distribución, resulta que también es una condición suficiente. Se pueden encontrar demostraciones del siguiente resultado en [8], [11], [27], [45] y [46], casi todas ellas usan la rectificación de campos de vectores, [26, pp. 25].

Teorema A.50 (Frobenius)**.** *Sea Δ una distribución regular e involutiva. Si $m \in M^d$ entonces existe una variedad integral de Δ por m. Es más, si c es la dimensión de la distribución Δ, entonces existe un sistema cúbico coordenado (N, φ) centrado en m, con funciones coordenadas $x_1, ..., x_d$ tal que los conjuntos de nivel ("slices")*

$$x_i = constante \qquad i = c+1, ..., d$$

son variedades integrales de la distribución Δ; y si (N, φ) es una variedad integral conexa de Δ tal que $\psi(N) \subset U$, entonces $\psi(N)$ está contenido en uno de estos conjuntos de nivel.

Teorema A.51. *Sean $\psi : N \to M^d$ una aplicación diferenciable y (P^c, φ) una variedad integral de la distribución Δ en M. Supongamos que ψ se factoriza por (P, φ), es decir, $\psi(N) \subset \varphi(P)$. Sea $\psi_0 : N \to P$ la única aplicación tal que $\varphi \circ \psi_0 = \psi$. Entonces ψ_0 es diferenciable.*

Definición A.52. *Una* variedad integral conexa maximal *de una distribución* Δ *en* M *es una variedad integral conexa de* Δ *cuya imagen en* M *no es un subconjunto propio de toda variedad integral conexa de* Δ*.*

Teorema A.53. *Sea* Δ *una distribución regular diferenciable e involutiva en* M*. Si* $m \in M$ *entonces existe una única variedad integral maximal conexa de* Δ *por* m *que contiene cualquier otra variedad integral conexa de* Δ *por* m*.*

Bibliografía

[1] Ayala, V and Tirao, J., *Linear control systems on Lie groups and controllability.* Proceedings of Symposia in Pure Mathematics, 64 (1999), 47-64.

[2] Boothby, W., *A transitive problem from control theory.*, to appear in J. Differential Equ. 17, (1975), 296-307.

[3] Boothby, W. and Wilson, E., *Determination of transitivity of bilinear systems.*, SIAM Journal on Control and Optimization. 17, (1979), 212-221.

[4] Braga C., Goncalves J., Gemano O. y San Martin L., *Controllability of two-dimensinal bilinear systems.* Proyecciones Vol. 15, No 2, pp. 111-139, December 1996.

[5] Brockett, R., *System theory on group manifolds and coset spaces.* SIAM J. Control, 10 (1972), pp. 265-284.

[6] Colonius, F. and Kliemann, W., *The Dynamics of Control*, Systems and Control, 2000.

[7] Elliott, David., *Bilinear Control Systems.*, Aplied Mathematical Sciences 169 (2009).

[8] Jurdjevic, V. and Sussmann, J., *Controllability of nonlinear systems*, to appear in J. Differential Eq.

[9] Jurdjevic, V. and Sussmann, J., *Control system on Lie Groups.*, to appear in J. Differential Equ. 12 (1972), 313-329.

[10] Jurdjevic, Velimir, *Geometric Control Theory.*, Cambridge Studies In Advanced Mathematics 52 (1997).

[11] Jurdjevic, V. and Kupka, I., *Control system on semi-simple Lie Groups and their homogeneous spaces.*, Ann. Inst. Fourier, Grenoble 31 (1981), 151-179.

[12] Lima, Elon Lages., *Variedades Diferenciáveis*, Publicações Matemáticas, IMPA, 2007.

[13] Macki, J. and Strauss, A., *Introducction to Optimal Control Theory*, Springer-Verlag, 1982.

[14] Sagle,A. and Walde, R. *Introduction to Lie Grupos and Lie Algebras*, vol. 51 of Pure and Aplied Mathematics, Academic Press, 1973.

[15] Spivak, M., *A Comprehensive Introducction to Differential Geometry.*, Vol. 1, Publish or Perish Press (1970).

[16] Varadajan, V., *Lie Groups, Lie Algebras and Their Representations*, vol. 102 of Graduate Texts in Mathematics, Springer-Verlag, 1984.

[17] Warner, F., *Foundations of Differentiable Manifolds and Lie Groups*, vol. 92 of Graduate Texts in Mathematics, Springer-Verlag, 1983.