



UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE CIENCIAS EXACTAS
DEPARTAMENTO DE MATEMÁTICA

TESIS PRESENTADA PARA OPTAR AL GRADO DE DOCTOR DE LA
FACULTAD DE CIENCIAS EXACTAS
EXISTENCIA DE SOLUCIONES DE
ECUACIONES DIFERENCIALES IMPLÍCITAS

Germán Eduardo Zorba

Director: Dra. María del Rosario Etchechoury
Codirector: Dr. Hernán Cendra

AÑO 2013

Índice general

1. Introducción	3
2. Preliminares	8
2.1. Fibrados, distribuciones y distribuciones de rango constante.	8
2.2. Ecuaciones Diferenciales Implícitas Cuasilineales	17
3. Teoría de ligaduras de Dirac	20
3.1. Método de Dirac	21
3.2. Algoritmo de Gotay-Nester	24
3.3. Sistemas dinámicos de Dirac y algoritmo de ligaduras.	27
3.4. Ejemplos	32
4. Sistemas con fricción	37
4.1. Principio D'Alembert generalizado	37
4.2. Algoritmo de ligaduras para formas que no son antisimétricas	39
4.3. Estructuras de Dirac generalizadas	41
4.3.1. Estructuras de Dirac generalizadas sobre un espacio vectorial	42
4.3.2. Estructuras de Dirac generalizadas sobre una variedad	45
4.3.3. Sistemas de Dirac generalizados y gCAD	54
4.4. Ejemplos	58
5. Algoritmos de restricciones	65
5.1. Algoritmo de restricciones regular	66
5.2. Algoritmo de restricciones singular	67
5.3. Comparación entre ambos algoritmos	74
6. Singularidades de Ecuaciones Diferenciales Implícitas Cuasi- lineales	76
6.1. Definición de puntos de cruce y puntos de impasse	76
6.2. Reparametrización	81

6.3. Algunos resultados sobre existencia de curvas invariantes . . .	85
6.4. Existencia de soluciones de cruce y de impasse	89
7. Conclusiones	105
A. Demostración del teorema de la variedad estable	108

Capítulo 1

Introducción

Las *ecuaciones diferenciales implícitas* -EDIs- aparecen frecuentemente en diferentes ciencias. Una EDI definida sobre una variedad M puede representarse de manera general así:

$$\varphi(x, \dot{x}) = 0, \quad (1.1)$$

siendo una *ecuación diferencial ordinaria* -EDO- $\dot{x} = f(x)$ el caso particular más simple.

Las ecuaciones de Euler-Lagrange para un Lagrangiano dado [17], las ecuaciones de Lagrange-D'Alembert para un sistema no-holónomo y su versión reducida: las ecuaciones de Lagrange-D'Alembert-Poincaré [18], son algunos ejemplos de EDIs que provienen de la mecánica.

Un gran número de trabajos sobre EDIs han sido también motivados por aplicaciones en teoría de circuitos [19]. En este contexto, como en otros, las EDIs se conocen como *ecuaciones diferenciales algebraicas* -EDAs-. La forma diferencial-algebraica de las ecuaciones de un circuito se debe de manera natural a la combinación de ecuaciones diferenciales con relaciones algebraicas, donde estas últimas modelan las leyes de Kirchoff.

La teoría de control ha considerado también la literatura de las EDAs. En este contexto las EDAs se conocen como *sistemas descriptores*. En las últimas tres décadas los sistemas de control lineales y no-lineales, así como también los variantes en el tiempo, han sido un foco de interés desde la perspectiva de las EDAs, jugando un rol importante los problemas de control óptimo [20].

En los años '70 un nuevo enfoque matemático de las EDIs ha comenzado a desarrollarse, de algún modo independiente de un campo de aplicación específico. En este nuevo enfoque se pone énfasis en los aspectos numéricos de las EDIs, [21].

Una *solución de (1.1) en un punto x* es un vector $(x, v) \in T_x M$ que satisface (1.1). Una *curva solución de (1.1)* $x : I \rightarrow M$, con

I intervalo abierto, es una función diferenciable en I y tal que $(x(t), \dot{x}(t))$ es una solución en el punto $x(t)$, para todo t en I .

Cuestiones básicas tales como existencia, unicidad o extensión de soluciones para una condición inicial dada no han sido aún completamente resueltas, aunque varios resultados parciales se han establecido para ciertas clases de EDIs [12, 2, 30]. Un camino a seguir es usar un *algoritmo de restricciones*, que consiste en construir una secuencia decreciente de *variedades de restricción* $M \supseteq M_1 \supseteq \dots \supseteq M_q$ definida así,

$$M_{k+1} := \{x \in M_k \mid \text{existe } (x, v) \in TM_k \text{ tal que } \varphi(x, v) = 0\},$$

con $M_0 := M$, con el objetivo de obtener una ecuación equivalente sobre la *variedad final* M_q de modo tal de poder asegurar existencia de solución para cada condición inicial en M_q . Este algoritmo, que usa sólo la estructura diferenciable, independientemente de cualquier otra estructura que pudiera estar presente, como por ejemplo una estructura presimpléctica o de Poisson o de Dirac, representa un aspecto geométrico-diferencial subyacente en los algoritmos de Gotay-Nester, Dirac ó CAD. Estos tres algoritmos, que pueden considerarse tipos particulares de algoritmos de restricciones y que se aplican a EDIs con determinadas estructuras, serán desarrollados en el Capítulo 3. Para asegurar, por un lado, que en cada paso del algoritmo el correspondiente conjunto M_k resulta una subvariedad, de modo tal de poder continuar con el proceso, y por otro lado, que el algoritmo se detiene luego de un número finito de pasos, se puede elegir asumir ciertas hipótesis del tipo condiciones de “*rango localmente constante*.” Luego, la EDI puede reducirse a una EDO equivalente, dependiendo de parámetros, sobre la *variedad final de restricciones* M_q . En efecto, por construcción, M_q se caracteriza por la propiedad de ser la menor subvariedad que contiene todas las curvas solución de la EDI dada. Luego, si M_q es vacío, no existen curvas solución. Otra manera de definir la variedad final M_q , sin necesidad de calcular las variedades intermedias, es la siguiente:

$$M_q = \{x(t) : x(s), s \in (a, b), \text{ es una curva solución de } \varphi(x, \dot{x}) = 0, t \in (a, b)\}.$$

A través de las siguientes referencias se puede ver cómo algún tipo de algoritmo de restricciones se aplica en diferentes contextos para ciertas clases de EDIs: en [10] se trabaja con variedades presimplécticas; en [12] se suponen ciertas hipótesis de diferenciabilidad sobre los datos; en [11] se trabaja con variedades algebraicas complejas; mientras que en [2] se plantea el algoritmo en el contexto analítico, admitiendo en este caso que en algún paso del algoritmo el correspondiente conjunto M_k puede tener singularidades para luego

aplicar una técnica de desingularización.

Desde el punto de vista geométrico la idea básica del algoritmo de restricciones, en sus diferentes variantes, es muy simple y natural y se encuentra ya presente en la teoría de ligaduras de Dirac [7, 8, 9].

Presentaremos ahora algunos ejemplos simples de EDIs, con la intención de ilustrar que tipo de situaciones se pueden presentar en lo que se refiere a existencia, unicidad y extensión de soluciones.

Ejemplo 1.1. Sea $F : T\mathbb{R}^3 \rightarrow \mathbb{R}^3$, con

$$F(x, \dot{x}) \equiv \begin{bmatrix} x_1 - \cos x_2 \\ \dot{x}_1 - x_3 \\ \dot{x}_2 - 1 \end{bmatrix} = 0. \quad (1.2)$$

Observar que $DF(x, \dot{x})$ tiene rango 3 cualquiera sea $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, luego $M = F^{-1}(0)$ es una subvariedad de $T\mathbb{R}^3$ de dimensión 3. Si $x = x(t)$ es cualquier solución \mathcal{C}^1 de (1.2), entonces derivando la ecuación algebraica $x_1(t) - \cos x_2(t) = 0$ y reemplazando en las ecuaciones diferenciales se puede ver que esta solución debe satisfacer también la ecuación $x_3(t) + \sin x_2(t) = 0$. En otras palabras cualquier solución debe estar contenida en el conjunto $Y = g^{-1}(0) \subset \mathbb{R}^3$, con $g(x) = (x_1 - \cos x_2, x_3 + \sin x_2)^T$, $x \in \mathbb{R}^3$. Como $Dg(x)$ tiene rango 2 para todo x , se concluye que Y es una subvariedad de \mathbb{R}^3 de dimensión 1. Luego, si $x = x(t)$ es una solución de (1.2), entonces $(x(t), \dot{x}(t)) \in TY \cap F^{-1}(0)$, y recíprocamente.

Ejemplo 1.2. Consideremos ahora la EDI definida sobre \mathbb{R}^2 así,

$$\dot{x}_1 + \dot{x}_2 = 1 \quad (1.3)$$

$$x_1 x_2 = 0. \quad (1.4)$$

Las curvas $c_1(t) = (t, 0)$ y $c_2(t) = (0, t)$ tienen el mismo valor inicial en $t = 0$ y son ambas solución de (1.3)-(1.4). Además, $Dc_1(0) \neq Dc_2(0)$, esto se debe al hecho que $x_1 x_2 = 0$ no es una subvariedad.

Ejemplo 1.3. Sea la siguiente EDI definida sobre \mathbb{R}^2 ,

$$x_1^2 + x_2 = 0 \quad (1.5)$$

$$\dot{x}_2 = 1, \quad (1.6)$$

con condición inicial $x(0) = (1, -1)$ (y entonces $\dot{x}(0) = (-\frac{1}{2}, 1)$). Este problema de valores iniciales tiene una única solución $x(t) = ((1-t)^{\frac{1}{2}}, t-1)$, que no puede continuarse más allá de $t = 1$, a pesar del hecho de que $x(1) = (0, 0)$ existe y $\lim_{t \rightarrow 1^-} x(t) = x(1)$. El punto $(0, 0)$, donde la solución $x(t) = ((1-t)^{\frac{1}{2}}, t-1)$ de (1.5)-(1.6) termina en tiempo finito $t = 1$, es un *punto de impasse* [26], como suele llamarse en la literatura ingenieril, especialmente en conexión con circuitos RLC no lineales [27]. Por último observar que la solución está definida en $(-\infty, 1]$, pero a diferencia de lo que ocurre cuando se trabaja con EDOs, no se puede continuar.

Notar que todos los problemas a tratar a lo largo de este trabajo, relacionados a existencia de soluciones, serán en general de naturaleza local, es decir, en un entorno del punto de la variedad donde se esté trabajando, aunque algunos resultados valdrán globalmente.

Si bien el tema general de esta tesis es el estudio de soluciones de EDIs, es importante remarcar que hay dos líneas bien diferenciadas de las que nos ocuparemos y que se describen más abajo.

En el Capítulo 2 se establecerán algunas definiciones y algunos resultados básicos que utilizaremos a lo largo de esta tesis.

En los Capítulos 3 y 4 trabajaremos con EDIs que modelan sistemas dinámicos que provienen de la mecánica. Más concretamente, en el Capítulo 3 describiremos las ideas presentadas por Paul Dirac para el estudio de sistemas mecánicos con Lagrangianos degenerados y desarrollos posteriores en el mismo sentido, que incluyen en el mismo contexto a los sistemas Lagrangianos con vínculos no-holónomos. En este Capítulo se repasan algunos algoritmos ya conocidos, como el propio algoritmo de Dirac y el algoritmo de Gotay-Nester, y se detalla también como es posible utilizar el concepto de estructura de Dirac para generalizar estos algoritmos a través de un nuevo algoritmo llamado CAD, de modo tal de poder tratar con sistemas más generales, que incluyen ejemplos importantes, tales como los circuitos LC y los sistemas Lagrangianos con vínculos no-holónomos, en todos los casos se trata de *sistemas conservativos*. En el Capítulo 4 se presentarán algunos de los resultados originales de la tesis y es de alguna manera una generalización del Capítulo anterior. La idea fundamental es generalizar el concepto de estructura de Dirac de modo tal de poder plantear un nuevo algoritmo, también más general, que sea aplicable a sistemas donde no haya necesariamente conservación de la energía. Los ejemplos que se tratan con esta nueva formulación son los circuitos eléctricos RLC, donde además de capacitores e inductores hay resistencias, y los sistemas mecánicos con fricción.

En los Capítulos 5 y 6 trabajaremos con EDIs más generales, llamadas *cua-*

silineales. En el Capítulo 5 se detallarán dos algoritmos de restricciones ya conocidos, que llamaremos en este trabajo *regular* y *singular*. En ambos casos se busca hallar una EDI equivalente a la dada y un dominio donde pueda garantizarse existencia de solución. En el caso del algoritmo regular se pedirán ciertas hipótesis de regularidad, que no serán necesarias en el caso del algoritmo singular, donde se utilizará una técnica llamada de *desingularización* que permitirá aplicar el algoritmo aún en los pasos donde las hipótesis de regularidad no se cumplan. Si bien el algoritmo singular es una generalización del algoritmo regular, no da respuesta a ciertos problemas de valores iniciales, que se vinculan con existencia y extensión de soluciones por ciertos puntos que llamaremos *singulares*. El Capítulo 6, que contiene otros de los aportes originales de esta tesis, da respuesta a algunos de estos problemas. En efecto, en ese Capítulo nos ocuparemos de estudiar existencia y extensión de soluciones de cierta clase de EDIs cuasilineales, por algunos puntos singulares que llamaremos *puntos de impasse* y *puntos de cruce*.

Capítulo 2

Preliminares

En este capítulo recordaremos algunas definiciones y algunos resultados que serán necesarios en el desarrollo de esta tesis. Básicamente, en los Capítulos 3 y 4, utilizaremos las nociones de *fibrados*, *distribuciones en general* y *distribuciones de rango constante*; mientras que para desarrollar los Capítulos 5 y 6 necesitaremos definir una *ecuación diferencial implícita cuasilineal* -EDICL- y establecer cuándo este tipo de ecuación se dice de *rango localmente constante*.

2.1. Fibrados, distribuciones y distribuciones de rango constante.

Gran parte de esta sección está formada por resultados conocidos que se mencionan aquí por completitud, de modo que el lector que conozca el tema puede saltar gran parte de esta sección. La excepción es la definición de distribución (definición 2.12), que es diferente de la estándar.

Variedades. Las variedades son espacios que, localmente, son muy parecidas a \mathbb{R}^n y son necesarias para poder dar una visión global de muchos problemas.

Definición 2.1. Una *variedad* es un espacio topológico Hausdorff M 2-numerable, paracompacto y localmente homeomorfo a un espacio vectorial de dimensión finita. Si M es conexo, entonces la dimensión será constante, en cambio, si M tiene más de una componente conexa, cada una de ellas podrá tener una dimensión diferente.

Más explícitamente, que sea localmente homeomorfo a un espacio vectorial significa que existe un cubrimiento de M por abiertos (U_i) y homeomor-

fismos $\phi_i : U_i \rightarrow V_i$ donde V_i es un abierto de \mathbb{R}^{n_i} . Cada uno de estos pares (U_i, ϕ_i) se llama *carta local* o *sistema de coordenadas locales*. Un conjunto de cartas locales tales que $\bigcup U_i = M$ e denomina *atlas*.

Las aplicaciones $\phi_{i_1 i_2} = \phi_{i_2} \circ \phi_{i_1}^{-1} : \phi_{i_1}(U_{i_1} \cap U_{i_2}) \rightarrow \phi_{i_2}(U_{i_1} \cap U_{i_2})$, donde $U_{i_1} \cap U_{i_2} \neq \emptyset$, se llaman *cambios de coordenadas*.

$$\begin{array}{ccc}
 & U_{i_1} \cap U_{i_2} \subset M & \\
 \swarrow \phi_{i_1} & & \searrow \phi_{i_2} \\
 \phi_{i_1}(U_{i_1} \cap U_{i_2}) & \xrightarrow{\phi_{i_1 i_2}} & \phi_{i_2}(U_{i_1} \cap U_{i_2})
 \end{array}$$

Una *variedad C^k* es una variedad provista de un atlas tal que todos los cambios de coordenadas son difeomorfismos C^k (i.e. funciones C^k con inversa también C^k). De modo similar definimos variedades C^∞ y C^ω , pidiendo que los mapas de transición sean difeomorfismos suaves y analíticos, respectivamente. Dos atlas C^k son equivalentes si su union es también un atlas C^k , debido a esto puede pensarse que una variedad C^k es una variedad con un atlas C^k maximal, aunque para definirla alcanza con dar un atlas no maximal.

Utilizaremos la frase *variedad diferencial* cuando $k = 1, \dots, \infty, \omega$ pueda quedar implícita.

Definición 2.2. Una *subvariedad* de una variedad M es un subconjunto N tal que para cada $x \in N$ existen sistemas de coordenadas locales $\phi : U \rightarrow \mathbb{R}^n$ con $x \in U \subset M$ tales que $\phi(N \cap U)$ es igual a $V \cap \phi(U)$ donde V es algún subespacio de \mathbb{R}^n .

Definición 2.3. Dadas M_1 y M_2 dos variedades C^k (respectivamente C^∞ , C^ω), una función $f : M_1 \rightarrow M_2$ y dos cartas locales $\phi_{i_1} : U_{i_1} \subset M_1 \rightarrow \mathbb{R}^{n_1}$ y $\phi_{i_2} : U_{i_2} \subset M_2 \rightarrow \mathbb{R}^{n_2}$ tales que $f(U_{i_1}) \subset U_{i_2}$, se llama *representante local de f* a la aplicación $f_{i_1 i_2} = \phi_{i_2} \circ f \circ \phi_{i_1}^{-1}$.

Se dice que una función $f : M_1 \rightarrow M_2$ es *de clase C^k* (respectivamente C^∞ , C^ω) si todos sus representantes locales lo son. También utilizaremos la palabra *suave* cuando $k = 1, \dots, \infty, \omega$ pudiera quedar implícito.

$$\begin{array}{ccc}
 M_1 & \xrightarrow{f} & M_2 \\
 \downarrow \phi_1 & & \downarrow \phi_2 \\
 \phi_1(U_1) & \xrightarrow{f_{i_1 i_2}} & \phi_2(U_2)
 \end{array}$$

Nota. Si bien pueden tomarse definiciones más amplias para variedades diferenciales, aquí decidimos restringirnos a espacios de dimensión finita y agregar las hipótesis necesarias para garantizar la existencia de particiones de la unidad subordinadas a las cartas locales. (Ver, por ejemplo, comentario en el libro de Hirsch [35]1§5).

Definición 2.4. Dada una variedad M y un cubrimiento por abiertos $\mathcal{U} = (U_i)$, una *partición de la unidad* subordinada a \mathcal{U} es una familia de funciones suaves (λ_i) tales que:

1. para todo i , λ_i se anula fuera de U_i ,
2. para todo $x \in M$ existe un entorno U_x , con $x \in U_x$, tal que sólo un número finito de las λ_i toma valores no nulos en U_x ,
3. $\sum \lambda_i(x) = 1$, para todo x (por hipótesis 2 esta suma es finita).

Teorema 2.1. *Sea M una variedad C^k con $k = 1, \dots, \infty$. Todo cubrimiento por abiertos de M tiene una partición de la unidad subordinada.*

Demostración. Ver [35] teorema 2§2.1. □

Las particiones de la unidad son útiles porque permiten construir objetos globales a partir de combinaciones lineales de objetos locales.

Fibrados. Definiremos ahora una estructura que será fundamental en los capítulos 3 y 4.

Definición 2.5. Consideremos dos variedades E, M y una aplicación $\tau : E \rightarrow M$.

Una *carta fibrada* de rango n es una función suave $\psi : \tau^{-1}(U) \rightarrow U \times \mathbb{R}^n$, donde U es un abierto de M , tal que el diagrama siguiente conmuta,

$$\begin{array}{ccc} \tau^{-1}(U) & \xrightarrow{\psi} & U \times \mathbb{R}^n \\ \tau \downarrow & \swarrow \pi_U & \\ U & & \end{array}$$

donde $\pi_U(x, v) = x$ es la proyección sobre la primera componente.

Llamamos *fibra sobre x* al conjunto $E_x = \tau^{-1}(x)$. La fibra E_x adquiere una estructura de espacio vectorial a través del homeomorfismo $\psi_x : E_x \rightarrow \mathbb{R}^n$ que resulta de restringir ψ a E_x .

$$\begin{array}{ccc}
& E_x & \\
\psi_{i_1 x} \swarrow & & \searrow \psi_{i_2 x} \\
\mathbb{R}^n & \xrightarrow{\psi_{i_1 i_2 x}} & \mathbb{R}^n
\end{array}$$

Un *atlas fibrado* de rango n es una familia de cartas fibradas $\psi_i : \tau^{-1}(U_i) \rightarrow U_i \times \mathbb{R}^n$ tales que (U_i) es un cubrimiento por abiertos de M y las aplicaciones $\psi_{i_1 i_2 x} = \psi_{i_2 x} \circ \psi_{i_1 x}^{-1}$ son aplicaciones lineales.

Un *fibrado vectorial* sobre M es una aplicación $\tau : E \rightarrow M$ con un atlas fibrado. La variedad M se denomina *base* y las aplicaciones $\psi_{i_1 i_2(\cdot)} : U_{i_1} \cap U_{i_2} \rightarrow GL(\mathbb{R}^n)$ se denominan *mapas de transición*. En ocasiones se menciona un fibrado haciendo referencia sólo al conjunto E , cuando los demás elementos puedan quedar implícitos.

Alternativamente, un fibrado vectorial puede ser definido mediante un cubrimiento $\{U_i\}$ y mapas de transición suaves $\psi_{ij} : U_i \cap U_j \rightarrow GL(\mathbb{R}^n)$ que verifiquen las siguientes hipótesis:

1. $\psi_{ij} \circ \psi_{ji} \equiv I_n$.
2. $\psi_{jk} \circ \psi_{ij} = \psi_{ik}$ allí donde la composición esté bien definida.

En este trabajo, siempre usaremos la palabra fibrado para referirnos a fibrados vectoriales.

Definición 2.6. Dado un fibrado $\tau : E \rightarrow M$, una *sección local* es una aplicación suave $v : U \rightarrow E$ donde $U \subset M$ es un abierto y tal que para todo $x \in U$, $\tau(v(x)) = x$.

Ejemplo 2.1. Dada una variedad M (que sea al menos C^1) se define el *fibrado tangente* TM formado por los elementos (x, v) , donde $x \in M$ y v es una clase de equivalencia de curvas $\gamma : (-\epsilon, \epsilon) \rightarrow M$ con $\gamma(0) = x$, donde dos curvas son equivalentes si y sólo si, al componerlas con alguna carta local $\phi : M \rightarrow \mathbb{R}^n$, coinciden hasta primer orden (la condición que la variedad sea al menos C^1 garantiza que esta relación no depende de la elección de la carta local).

Definimos $\tau : TM \rightarrow M$ por $\tau(x, v) = x$. Dado un atlas de M formado por las cartas locales ϕ_i , el atlas fibrado correspondiente de TM estará formado por las cartas fibradas $\psi_i(x, v) = (\phi_i(x), \frac{d}{dt}(\phi_i \circ \gamma)(0))$, donde γ es un representante de la clase de equivalencia v .

Los mapas de transición resultan ser $\psi_{ij} = d\phi_{ij}$.

Notar que si los cambios de coordenadas de M son de clase C^k , entonces los mapas de transición de TM serán de clase C^{k-1} , por lo tanto si M es una

variedad C^k , entonces TM será un fibrado de clase C^{k-1} , y en caso de que M fuera de clase C^∞ o C^ω , entonces TM será un fibrado de la misma clase. (Por este motivo en el Capítulo 4 nos limitaremos a trabajar en los casos C^∞ y C^ω).

La notación T_xM se usa para referirse a la fibra $(TM)_x$.

Definición 2.7. Dado un fibrado $\tau : E \rightarrow M$ y una aplicación suave $f : \tilde{M} \rightarrow M$ se define el *pullback* f^*E del siguiente modo. Se toma $f^*E = \{(x, y) \in \tilde{M} \times E | f(x) = \tau(y)\}$ y $\tilde{\tau} : f^*E \rightarrow \tilde{M}$ definido por $\tilde{\tau}(x, y) = x$. Las cartas fibradas de f^*E se definen tomando cartas fibradas de E , $\psi : \tau^{-1}(U) \rightarrow U \times \mathbb{R}^n$ y definiendo $\tilde{\psi} : \tilde{\tau}^{-1}(f^{-1}(U)) \rightarrow f^{-1}(U) \times \mathbb{R}^n$ mediante $\tilde{\psi}_x = \psi_{f(x)}$.

Dado un fibrado $\tau : E \rightarrow M$ y una subvariedad $N \subset M$ llamamos *restricción* a $E_N = i^*E$, donde $i : N \rightarrow M$ es la inclusión.

Definición 2.8. Dados dos fibrados $\tau_i : E_i \rightarrow M$, $i = 1, 2$ sobre el mismo espacio base M se define la suma de Whitney como $E_1 \oplus E_2 = \{(v_1, v_2) \in E_1 \times E_2 | \tau_1(v_1) = \tau_2(v_2)\}$.

La estructura de fibrado de $E_1 \oplus E_2$ sobre M puede obtenerse tomando la estructura de $E_1 \times E_2$ como fibrado sobre $M \times M$ y restringiendo a la diagonal $\{(x, x) | x \in M\}$ e identificando a esta diagonal con M .

Definición 2.9. Dado un fibrado $\tau : E \rightarrow M$, un *producto interno* sobre E es una familia $(\langle \cdot, \cdot \rangle_x)_{x \in M}$ de productos internos sobre E_x tales que la aplicación $(x, u, v) \mapsto (\langle u, v \rangle_x)$ es una aplicación suave de $E \oplus E \rightarrow \mathbb{R}$.

Para los casos C^k , $k = 1, \dots, \infty$ siempre se puede construir un producto interno tomando un producto interno para cada carta fibrada y combinándolos mediante una partición de la unidad.

Un producto interno sobre TM se denomina *métrica riemanniana*.

Morfismos. Los fibrados vectoriales sobre una misma base forman una categoría tomando como morfismos a los mapas lineales.

Definición 2.10. Dados dos fibrados $\tau_1 : E_1 \rightarrow M_1$ y $\tau_2 : E_2 \rightarrow M_2$, un *mapa fibrado* es una función suave $f : E_1 \rightarrow E_2$ que cubre una función suave $\tilde{f} : M_1 \rightarrow M_2$, es decir, que el siguiente diagrama conmuta

$$\begin{array}{ccc} E_1 & \xrightarrow{f} & E_2 \\ \downarrow \tau_1 & & \downarrow \tau_2 \\ M_1 & \xrightarrow{\tilde{f}} & M_2 \end{array}$$

el hecho de que el diagrama conmute implica que f envía la fibra E_{1x} en la fibra $E_{2\tilde{f}(x)}$.

Una *aplicación lineal* o *morfismo* de fibrados vectoriales es un mapa fibrado tal que la restricción de f a cada fibra resulta ser una función lineal.

Ejemplo 2.2. Dada una función suave entre dos variedades $f : M \rightarrow N$ se define el mapa $Tf : TM \rightarrow TN$ mediante la fórmula $Tf(x, [\gamma]) = (f(x), [f \circ \gamma])$, donde $[\cdot]$ representa la clase de equivalencia de la curva correspondiente.

También se usa la notación $T_x f(v) = Tf(x, v)$. Que $T_x f$ es una aplicación lineal se deduce de la linealidad de la derivación.

Con esto podemos dar una forma de definir subvariedades que es más sencilla de verificar.

Definición 2.11. Dadas dos variedades N y M y una aplicación $f : N \rightarrow M$ se dice que

- f es una inmersión si $T_x f$ es monomorfismo para todo x .
- f es una submersión si $T_x f$ es epimorfismo para todo x .
- f es un encaje si es una inmersión y además es un homeomorfismo de N en $f(N)$.

Teorema 2.2. Dada una variedad M y un subconjunto $N \subset M$, N es una subvariedad si y sólo si N es la imagen de un encaje.

Demostración. Ver [35] teorema 1§3.1 □

Construcción functorial de fibrados Una forma bastante general de construir nuevos fibrados a partir de fibrados anteriores es la siguiente:

Un functor en la categoría de espacios vectoriales de dimensión finita es una aplicación \mathcal{F} que, dados los espacios vectoriales $(V_i, W_j)_{i \in I, j \in J}$ construye el espacio vectorial $\mathcal{F}(V_i, W_j)$, y que dadas las transformaciones lineales $A_i : V_i \rightarrow \tilde{V}_i$ y $B_j : \tilde{W}_j \rightarrow W_j$ construye la transformación lineal $\mathcal{F}(A_i, B_j) : \mathcal{F}(V_i, W_j) \rightarrow \mathcal{F}(\tilde{V}_i, \tilde{W}_j)$. Se dice que el functor \mathcal{F} es covariante en las variables V_i y contravariante en las variables W_j .

Por ejemplo, el dual $W \mapsto W^*$ es un functor contravariante (aplicado a transformaciones lineales da la transpuesta), la suma directa $V_1 \oplus V_2$ es un functor covariante en las dos variables y $L(W, V)$ (espacio de transformaciones lineales de W en V) es covariante en la segunda variable y contravariante en la primera (aplicar este último functor a transformaciones lineales equivale a componer a derecha con una transformación y a izquierda con la otra).

Un functor se dice C^k $k = 1, \dots, \infty, \omega$ si al aplicarlo a transformaciones lineales resulta ser una función suficientemente suave (los conjuntos $L(V, \tilde{V})$ son espacios vectoriales de dimensión finita y por lo tanto tienen una estructura diferencial).

Teorema 2.3. *Sea \mathcal{F} un functor suave, entonces para cada variedad M existe un functor \mathcal{F}_M de fibrados vectoriales sobre M que verifica que, dados los fibrados E^i , y los morfismos de fibrados f^i ,*

1. *para todo $x \in M$ $\mathcal{F}_M(E^i)_x = \mathcal{F}(E^i_x)$;*
2. *para todo $x \in M$ $\mathcal{F}_M(f^i)_x = \mathcal{F}(f^i_x)$;*
3. *si los E^i son fibrados triviales entonces $\mathcal{F}_M(E^i)$ también es un fibrado trivial;*
4. *si $h : N \rightarrow M$ es suave entonces $\mathcal{F}_N(h^*E^i) = h^*\mathcal{F}_M(E^i)$.*

Demostración. Ver [37] teorema 3§4.1. □

En particular, esto nos permite dar una construcción alternativa para la suma de Whitney $E_1 \oplus E_3$ (ver definición 2.8) y además nos permite definir el fibrado dual E^* y ver al conjunto de todas las aplicaciones lineales $L(E_1, E_2)$ como un fibrado sobre la misma base.

Ejemplo 2.3. El dual del fibrado tangente TM se denomina *fibrado cotangente* y se simboliza T^*M .

El fibrado $TM \oplus T^*M$ jugará un rol importante en los Capítulos 3 y 4 de esta tesis.

Subfibrados y distribuciones. Definiremos a continuación los conceptos de subfibrado, distribución y codistribución. La definición de distribución dada aquí se aleja bastante de la habitualmente hallada en la bibliografía.

Definición 2.12. Dados dos fibrados E y \tilde{E} sobre M y un morfismo $f : \tilde{E} \rightarrow E$ definimos la imagen de f como $\text{Im}(f) = \{f(v) \in E | v \in \tilde{E}\}$ y el núcleo de f como $\text{ker}(f) = \{v \in \tilde{E} | f(v) \in M \times \{0\} \subset E\}$ (esto es, tomar la imagen y el núcleo fibra a fibra).

- Una *distribución* es cualquier subconjunto de un fibrado que pueda obtenerse como imagen de un morfismo.
- Una *codistribución* es cualquier subconjunto de un fibrado que pueda obtenerse como núcleo de un morfismo.

- Un *subfibrado* es una distribución que puede obtenerse como imagen de un monomorfismo (es decir, imagen de un morfismo cuyo núcleo es la sección nula).

Una definición equivalente para subfibrados consiste en decir que un subfibrado es un subconjunto de un fibrado localmente generado por k secciones linealmente independientes.

La definición de distribución dada arriba es diferente a las definiciones halladas en la bibliografía. En algunos casos, la palabra distribución se utiliza para subfibrados del fibrado tangente. Nosotros damos una definición más general debido a que trabajaremos con mucha frecuencia con el fibrado $TM \oplus T^*M$.

En otros casos, la palabra distribución se usa para referirse a subconjuntos de TM localmente generados por secciones que no son necesariamente linealmente independientes. Esta definición es, en principio, más general que la de distribuciones de TM en el sentido de la definición 2.12, pero en los casos en que la cantidad total de secciones utilizadas para generar el subconjunto sea finita, ambas definiciones coinciden (en el caso de la definición 2.12 puede construirse un morfismo con dominio en un fibrado trivial de rango suficientemente grande).

El rango de una distribución es una función semicontinua inferiormente y el rango de una codistribución es una función semicontinua superiormente. El rango de un subfibrado es constante. Es inmediato verificar que cualquier distribución de rango constante es un subfibrado.

Las codistribuciones introducidas en la definición 2.12 son objetos duales de las distribuciones en el siguiente sentido: Dada una distribución $D \subset E$, entonces existe un morfismo $f : \tilde{E} \rightarrow E$ cuya imagen es D . El morfismo f induce otro morfismo $f^* : E^* \rightarrow \tilde{E}^*$. El núcleo de f^* es D° que resulta ser una codistribución.

Las distribuciones son, como hemos visto, objetos generados por secciones, y las codistribuciones resultan ser objetos definidos por ecuaciones lineales.

Ejemplo 2.4. Este ejemplo adelanta algunas de las discusiones sobre estructuras de Dirac generalizadas que se darán en el capítulo 4, y sirve para motivar la definición de mapas sobre distribuciones que se dará más adelante.

Dado un fibrado E sobre M , una forma bilineal es una sección ω de $L(E \oplus E, M \times \mathbb{R})$. También puede ser vista como una sección ω^b de $L(E, E^*)$, es decir como un morfismo $\omega^b : E \rightarrow E^*$ (esto es así porque para espacios vectoriales $L(V \oplus V, \mathbb{R}) \simeq L(V, V^*)$). Finalmente, también puede verse a la forma bilineal como la gráfica de ω^b , que es un subfibrado de $E \oplus E^*$.

Una forma de generalizar el concepto de formas bilineales consiste en considerar cualquier subfibrado $D \subset E \oplus E^*$ y considerarlo como una forma bilineal cuyas variables toman valores en dos distribuciones adecuadas.

Tomemos por ejemplo $E = T\mathbb{R}$ y $D \subset T\mathbb{R} \oplus T^*\mathbb{R}$ generado por la sección $x \mapsto (x, x^3, 1)$. La proyección de D sobre E es una distribución E_D que verifica $(E_D)_x = T_x\mathbb{R}$ para $x \neq 0$ y $(E_D)_0 = \{0\}$. Para $x \neq 0$ resulta que D es la gráfica de una función $E \rightarrow E^*$ dada por $(x, v) \mapsto v/x^3$. Querriamos entonces que la función $\omega : E_D \otimes E_D \rightarrow M \times \mathbb{R}$ dada por

$$\omega(x, v_1, v_2) = \begin{cases} (x, v_1 v_2 / x^3) & x \neq 0 \\ (0, 0) & (x, v_1, v_2) = (0, 0, 0) \end{cases}$$

fuera una aplicación lineal de $E_D \otimes E_D$ en $M \times \mathbb{R}$, donde $E_D \otimes E_D$ es una distribución de $E \otimes E$ tal que $(E_D \otimes E_D)_x = (E_D)_x \otimes (E_D)_x$.

En principio, ω no puede extenderse a una función continua $E \otimes E$, por eso es que hace falta una definición más sutil de aplicaciones lineales definidas sobre una distribución.

Definición 2.13. Dados dos fibrados E y F sobre la misma variedad M , dada una distribución $E_D \subset E$ y una función $f : E_D \rightarrow F$, decimos que f es un *mapa lineal* si la gráfica de f es una distribución en $E \oplus F$.

Entornos tubulares. Hay dos operaciones diferentes con fibrados que generan fibrados más pequeños: los subfibrados y la restricción a subvariedades. Esas dos operaciones se relacionan del siguiente modo: dada una variedad M y una subvariedad N , el fibrado TN es un subfibrado de $T_N M$.

Hay un resultado un poco más fuerte que esto, que vamos a usar en el capítulo 4, que si bien parece estándar no lo hemos encontrado en la bibliografía de referencia: existe U un entorno de N y un fibrado E sobre U tal que E es un subfibrado de $T_U M$ y $E_N = TN$, es decir, E es una extensión de TN como subfibrado de $T_N M$ a un entorno de N . Esto lo utilizaremos en el capítulo 4 para poder escribir ciertas ecuaciones de manera global.

El resultado en sí no es muy novedoso, y resulta de combinar dos resultados conocidos:

Definición 2.14. Sea $N \subset M$ una subvariedad. Un *entorno tubular* de N en M es un fibrado E sobre N junto con un encaje $f : E \rightarrow M$ tal

- $f|_N$ es la identidad en N (identificando N con la sección nula),
- $f(E)$ es un entorno abierto de N en M .

También se llama entorno tubular a la imagen $f(E)$.

Teorema 2.4. *Sea $N \subset M$ una subvariedad, entonces existe un entorno tubular de N en M*

Demostración. Ver [35] teorema 4§5.2. □

Lema 2.5. *Dado un fibrado E , existe una métrica riemanniana en E tal que las fibras sean ortogonales a la sección nula.*

Demostración. El resultado es trivial para fibrados triviales. Tomando trivializaciones locales, pueden construirse métricas locales tales que las fibras resulten ortogonales a la sección nula. Esas métricas locales pueden combinarse mediante una partición de la unidad para obtener una métrica como en el enunciado. □

Corolario 2.6. *Sea $N \subset M$ una subvariedad y U un entorno tubular, entonces existe F un subfibrado de $T_U M$ tal que $F_N = TN$.*

Demostración. Sea E un fibrado y $f : E \rightarrow M$ un encaje como en la definición de entornos tubulares, con $f(E) = U$.

Consideremos en TE el subfibrado F tangente a cada fibra (es decir que $F_{(x,v)} = T_{x,v} E_x$; F es un fibrado isomorfo a $E \oplus E$). Elijamos una métrica riemanniana en E como en el lema 2.5. y consideremos F^\perp el complemento ortogonal de F . Por construcción F^\perp es ortogonal a las fibras de E , y por lo tanto es tangente a la sección nula, luego $F^\perp|_N = TN$. □

2.2. Ecuaciones Diferenciales Implícitas Cuasilineales

A continuación definiremos una clase particular de EDIs, *las cuasilineales*, con las que trabajaremos en los Capítulos 5 y 6, mostraremos además como una EDI general del tipo (1.1) puede llevarse de manera simple a una cuasilineal equivalente. Definiremos también una hipótesis fundamental que se utilizará en esos capítulos, y que llamaremos hipótesis de “*rango localmente constante*.”

Ecuación diferencial implícita cuasilineal. Sea M una variedad de dimensión n y F un espacio vectorial de dimensión m . Sea $a : TM \rightarrow F$ una aplicación suave tal que $a(x, \dot{x}) \equiv a(x)\dot{x}$ es lineal en \dot{x} . Sea $f : M \rightarrow F$ una aplicación suave dada. Una *ecuación diferencial implícita cuasilineal* - EDICL- tiene la forma,

$$a(x)\dot{x} = f(x). \tag{2.1}$$

Dada una EDICL del tipo (2.1) se tiene inmediatamente un *sistema lineal algebraico*, para cada $x \in M$, que depende de manera suave sobre x , donde la incógnita del sistema lineal es el vector \dot{x} , con base en x , para cada $x \in M$.

Observar que los ejemplos presentados en la Introducción representan EDICLs. En efecto, si llamamos

$$a(x) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad f(x) = \begin{bmatrix} x_1 - \cos x_2 \\ x_3 \\ 1 \end{bmatrix},$$

podemos ver que (1.2) resulta una EDICL. Lo mismo ocurre con (1.3)-(1.4) considerando ahora

$$a(x) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad f(x) = \begin{bmatrix} 1 \\ x_1 x_2 \end{bmatrix}.$$

Por último (1.5)-(1.6) resulta una EDICL tomando

$$a(x) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad f(x) = \begin{bmatrix} x_1^2 + x_2 \\ 1 \end{bmatrix}.$$

Reducción de una EDI general a una EDICL. Es fácil ver que una EDI general del tipo (1.1) donde la aplicación $\varphi : TM \rightarrow F$ puede ser no lineal en \dot{x} , no es mucho más general que una ecuación cuasilineal del tipo (2.1).

En efecto, supongamos por simplicidad que M es un subconjunto abierto de un espacio vectorial E de dimensión finita. Una EDI del tipo (1.1) se puede reescribir en la forma (2.1) con dominio $M \times E$ y rango $E \times F$ así:

$$\dot{x} = u, \quad 0 = \varphi(x, u),$$

que tiene la forma (2.1) si consideramos

$$a(x, u) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad f(x, u) = \begin{bmatrix} u \\ \varphi(x, u) \end{bmatrix}.$$

EDICL de rango localmente constante. Consideramos la ecuación (2.1), y supongamos que el sistema lineal algebraico asociado tiene solución \dot{x} para cada $x \in M$. Entonces podemos pensar que (2.1) define una distribución afín, en general singular, sobre M . Si además, $\text{rg}[a(x)] = \text{rg}[a(x), f(x)]$ es localmente constante, es decir, es constante sobre cada componente conexa de M , entonces la ecuación (2.1) se dice que es de *rango localmente constante*. Esto es equivalente a decir que la correspondiente distribución afín tiene *rango constante* sobre cada componente conexa de M .

Curva solución de una EDICL. Recordemos que una solución de (2.1) en un punto $x \in M$ es un vector $(x, v) \in T_x M$ tal que satisface (2.1). En el Capítulo 1 introdujimos la noción de curva solución en un intervalo abierto. Estableceremos ahora el concepto de *curva solución* de una EDICL pero admitiendo que su dominio pueda ser también un intervalo semiabierto o cerrado.

Definición 2.15. Una curva $x(t)$ es solución de (2.1) si es una función continua $x: I \rightarrow M$ con $I = (t_0, t_1)$ un intervalo real (resp. $I = (t_0, t_1]$, $I = [t_0, t_1)$ o $I = [t_0, t_1]$), x diferenciable en el interior de I y $(x(t), \dot{x}(t))$ es solución de la ecuación en el punto $x(t)$, para todo t en el interior de I . Además, si $t_0 \in I$ entonces $(x(t_0), \dot{x}(t_0^+))$ es solución y si $t_1 \in I$ entonces $(x(t_1), \dot{x}(t_1^-))$ es solución; siendo $\dot{x}(t_0^+) = \lim_{t \rightarrow t_0^+} \frac{x(t) - x(t_0)}{t - t_0}$ y $\dot{x}(t_1^-) = \lim_{t \rightarrow t_1^-} \frac{x(t) - x(t_1)}{t - t_1}$.

Capítulo 3

Teoría de ligaduras de Dirac

En este capítulo estudiaremos las ideas presentadas por Paul Dirac a partir de 1950 para el estudio de sistemas mecánicos con Lagrangianos degenerados [7] y desarrollos posteriores en el mismo sentido, que incluyeron en el mismo contexto a los sistemas Lagrangianos con vínculos no-holónomos.

La formulación Lagrangiana establece la dinámica de un sistema en términos de las coordenadas q y de las velocidades \dot{q} mediante un principio variacional. En efecto, dado un espacio de configuración Q y dada una función $L : TQ \rightarrow \mathbb{R}$, a la que se denomina Lagrangiano, se considera el funcional que a cada curva $q(t)$ le asigna el valor $\int L(q(t), \dot{q}(t)) dt$. La dinámica del sistema estará dada por las curvas extremales de este funcional, es decir, por las curvas que satisfagan la ecuación

$$\delta \int L(q, \dot{q}) = 0.$$

A partir de este principio se obtienen las *ecuaciones de Euler-Lagrange*, que son un ejemplo muy importante de ecuaciones diferenciales implícitas. Las ecuaciones de Euler-Lagrange habitualmente se escriben en la forma

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) = \frac{\partial L}{\partial q}(q, \dot{q}),$$

y pueden reescribirse como

$$\frac{\partial^2 L}{\partial \dot{q}^2}(q, \dot{q}) \ddot{q} = \frac{\partial L}{\partial q}(q, \dot{q}) - \frac{\partial^2 L}{\partial q \partial \dot{q}}(q, \dot{q}) \dot{q},$$

de este modo se ve más claramente que se trata de una ecuación diferencial implícita de segundo orden.

En cambio, mediante la formulación Hamiltoniana, se obtienen ecuaciones diferenciales explícitas en términos de las coordenadas q y de los momentos

p , con $(q, p) \in T^*Q$. En efecto, dado un Hamiltoniano $H : T^*Q \rightarrow \mathbb{R}$, mediante el corchete de Poisson, la fórmula

$$\dot{f} = \{f, H\}$$

define de manera explícita una derivación, y por lo tanto define también un campo vectorial. Las *ecuaciones de Hamilton* muestran esto de manera más clara y se escriben así:

$$\dot{q} = \frac{\partial H}{\partial p} \quad , \quad \dot{p} = -\frac{\partial H}{\partial q} .$$

En el caso de contar con un Lagrangiano regular, puede pasarse de la formulación Lagrangiana a la Hamiltoniana, y recíprocamente, mediante el cambio de variables

$$p = \frac{\partial L}{\partial \dot{q}},$$

pero en el caso en que la matriz $\left(\frac{\partial^2 L}{\partial \dot{q}^2}\right)$ no sea invertible, la transformación $p = \frac{\partial L}{\partial \dot{q}}$ tampoco lo será.

El problema de estudiar sistemas con Lagrangianos degenerados fue analizado por Paul Dirac a partir del año 1950, siendo uno de sus objetivos cuantizar estos sistemas [7, 8, 9].

En lo que sigue revisaremos el método de Dirac y algunas versiones más geométricas que se desarrollaron con posterioridad.

3.1. Método de Dirac

En su trabajo de 1950 [7], Dirac propone un método para transformar un sistema con Lagrangiano degenerado en otro del tipo Hamiltoniano. Este método fue reformulado por el propio Dirac en 1958 [8], y presentado de un modo matemáticamente más adecuado.

El primer paso del método consiste en considerar las ecuaciones

$$p_i = \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q}). \quad (3.1)$$

A partir de (3.1) se trata de eliminar \dot{q} , es decir hallar la condición que debe satisfacer un punto (q, p) para que el sistema de ecuaciones (3.1) tenga al menos una solución \dot{q} . Esta condición no siempre viene dada por una relación

$$\phi_m(q, p) = 0,$$

con ϕ_m diferenciable. Sin embargo Dirac asume que sí es posible, y supone también que las otras relaciones que aparecen a lo largo del algoritmo también pueden escribirse de esta manera.

Desde un punto de vista más geométrico, el planteo Lagrangiano se desarrolla en el fibrado tangente $(q, \dot{q}) \in TQ$, siendo Q el espacio de configuración, y el planteo Hamiltoniano en el fibrado cotangente $(q, p) \in T^*Q$. La aplicación $\mathcal{L} : (q, \dot{q}) \mapsto (q, \frac{\partial}{\partial \dot{q}}L(q, \dot{q}))$, llamada transformada de Legendre, es una función entre dos fibrados vectoriales (no necesariamente lineal) y las ecuaciones $\phi_m(q, p) = 0$ que definen localmente a la imagen de \mathcal{L} se denominan *restricciones, vínculos o ligaduras* (no hay una traducción unificada, en inglés se las llama *constraints*). En el caso en que el Lagrangiano sea regular resulta que \mathcal{L} es localmente invertible, y por lo tanto no hay restricciones ϕ_m .

El segundo paso consiste en definir el Hamiltoniano, que en el caso regular resulta,

$$H(q, p) = p\dot{q} - L(q, \dot{q}),$$

donde \dot{q} es calculado usando la inversa de la transformación \mathcal{L} . Pero en el caso de un Lagrangiano degenerado, esta ecuación sólo define a H en la imagen de \mathcal{L} , y luego debe ser extendido al resto de T^*Q . Esta extensión no es única, ya que puede ser reemplazada por

$$H_T = H + \sum \lambda^m \phi_m,$$

donde los λ^m son parámetros arbitrarios. La función H_T se llama Hamiltoniano total y los multiplicadores λ^m son nuevas variables a determinar.

El tercer paso surge de observar que, como $\phi_m = 0$ a lo largo de la evolución del sistema, entonces necesariamente

$$\dot{\phi}_m = \{\phi_m, H_T\} = 0,$$

lo que da lugar a nuevas ecuaciones, que en algunos casos sirven para determinar los valores de λ^m , pero en otros casos dan lugar a nuevas restricciones para los valores de q y p , apareciendo entonces nuevas ecuaciones

$$\chi_j(q, p) = 0,$$

que deben satisfacerse a lo largo de las trayectorias del sistema, y por lo tanto nuevamente se debe cumplir

$$\{\chi_j, H_T\} = 0,$$

lo que podría implicar nuevas restricciones. El proceso se repite hasta que las nuevas ecuaciones no impongan nuevas restricciones a las variables q y p .

Al finalizar el proceso tenemos las restricciones ϕ_m y χ_j , el Hamiltoniano total $H_T = H + \sum \lambda^m \phi_m$ y las ecuaciones para los λ^m dadas por $\{\phi_m, H_T\} = 0$ y $\{\chi_j, H_T\} = 0$. Para simplificar la escritura, unificaremos la notación de las *restricciones primarias* ϕ_m y las *restricciones secundarias* χ_j . Para eso llamaremos $\chi_m = \phi_m$ para $m = 1, \dots, M$ a las restricciones primarias y χ_j para $j = M + 1, \dots, J$ a las restricciones secundarias. Aunque la notación resulta redundante, esto permite simplificar las expresiones puesto que ahora ϕ_m se refiere sólo a las restricciones primarias mientras que χ_j se refiere a todas las restricciones, tanto primarias como secundarias. Esta elección de notación difiere levemente del trabajo de Dirac [7], donde llama ϕ_m a todas las restricciones.

Con la nueva notación, las ecuaciones $\{\chi_j, H_T\} = 0$, $j = 1, \dots, J$ escritas en forma matricial resultan

$$\begin{pmatrix} \{\chi_1, \phi_1\} & \cdots & \{\chi_1, \phi_M\} \\ \vdots & & \vdots \\ \{\chi_J, \phi_1\} & \cdots & \{\chi_J, \phi_M\} \end{pmatrix} \begin{pmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^M \end{pmatrix} = - \begin{pmatrix} \{\chi_1, H\} \\ \vdots \\ \{\chi_J, H\} \end{pmatrix}. \quad (3.2)$$

Por último, para concluir con el método propuesto por Dirac, se considera a la ecuación (3.2) como un sistema lineal en el que los parámetros λ^m son las incógnitas, que dependen de p y q , y a esta ecuación se le aplica un cambio de coordenadas lineal, de manera que se anulen la mayor cantidad posible de columnas de la matriz de la izquierda en (3.2) (es decir, se realizan combinaciones lineales de las columnas para anular la mayor cantidad posible, o más explícitamente, se calcula el núcleo de la matriz). Esto puede hacerse siempre que la matriz tenga rango localmente constante. Supongamos que ese cambio ya fue realizado y se tiene entonces que las primeras K columnas de la matriz son linealmente independientes y las últimas $M - K$ columnas son todas nulas, siendo $K = 2s$. Luego, teniendo en cuenta la ecuación (3.2), los parámetros $\lambda^1, \dots, \lambda^K$ están unívocamente determinados como funciones de q y p , mientras que $\lambda^{K+1}, \dots, \lambda^M$ son funciones arbitrarias de t .

Dirac llama *funciones de primera clase* a aquellas funciones f que verifiquen $\{f, \chi_j\} = 0$ sobre la variedad de restricciones final para todo j , mientras que las restantes se llaman *funciones de segunda clase*. Por ejemplo las restricciones ϕ_m , $m = K + 1, \dots, M$ son funciones de primera clase. Esta distinción es importante porque los valores $\lambda^{K+1}, \dots, \lambda^M$ son variables dinámicas arbitrarias, y por lo tanto no deberían tener significado físico.

3.2. Algoritmo de Gotay-Nester

El método de Dirac desarrollado arriba es local, en el sentido que las restricciones ϕ_m y χ_j definen localmente el conjunto de valores posibles para q y p . Un paso importante es dar a este método una formulación global, que permita una visión más geométrica del problema.

Ese paso fue dado por Gotay, Nester y Hinds en 1978 [10]. La ecuación estudiada por ellos toma como datos una variedad presimpléctica (M_0, ω) y una 1-forma α . Esta ecuación

$$\omega_x(\dot{x}, \cdot) = \alpha_x \quad (3.3)$$

puede ser vista como un caso particular de la ecuación que estudia Dirac, si tomamos M_0 igual a la subvariedad de restricciones primarias, y ω igual a la restricción de la 2-forma canónica Ω . En efecto, la ecuación estudiada por Dirac

$$\dot{f} = \{f, H + \sum \lambda^m \phi_m\} = 0$$

puede ser reescrita como

$$\Omega(\dot{x}, \cdot) = -dH + \sum \lambda^m d\phi_m \quad (3.4)$$

y como los $d\phi_m$ son un sistema de generadores del anulador de TM_0 , la existencia de parámetros λ tales que se satisfaga (3.4) es equivalente a

$$\Omega(\dot{x}, \cdot)|_{TM_0} = -dH|_{TM_0}$$

que, junto con la condición de que la imagen de x esté contenida en M_0 , resulta equivalente a (3.3).

En [10] desarrollan un algoritmo similar al de Dirac para la ecuación (3.3), y luego demuestran que en el caso en que M_0 sea la subvariedad de restricciones primarias, el algoritmo de Gotay-Nester reproduce exactamente los mismos pasos que el algoritmo de Dirac.

A diferencia del camino utilizado por estos autores para la presentación del método, seguiremos el camino inverso, es decir partiremos del algoritmo de Dirac y lo reescribiremos aprovechando la estructura presimpléctica de M_0 , para llegar al algoritmo de Gotay-Nester.

Recordemos que, luego de haber calculado las restricciones primarias ϕ_m , cada paso del método de Dirac consiste en considerar las restricciones obtenidas en los pasos anteriores, a las que llamaremos χ_j (en el primer paso, las χ_j serán exactamente las ϕ_m , en los otros pasos serán las ϕ_m junto a

las restricciones calculadas en los pasos anteriores), considerar las ecuaciones provenientes del hecho de que $\dot{\chi}_j = 0$

$$\{\chi_j, H + \sum \lambda^m \phi_m\} = 0 \quad (3.5)$$

y buscar los valores de p y q para los cuales (3.5) tiene solución (λ^m). Las nuevas restricciones secundarias resultan las ecuaciones que definen el subconjunto donde el sistema tiene solución.

Llamaremos M_0 al conjunto definido por las restricciones primarias

$$\phi_m = 0;$$

llamaremos $M_1 \subset M_0$ al subconjunto que surge de agregar las restricciones secundarias que aparecen en el primer paso, es decir que

$$M_1 = \{(q, p) \in M_0 \mid \text{el sistema} \{\chi_j, H + \sum \lambda^m \phi_m\} = 0, j = 1, \dots, M, \\ \text{tiene solución}\},$$

(recordar que $\chi_j = \phi_j = 0, j = 1, \dots, M$, son las ecuaciones que definen M_0); las ecuaciones que definen M_1 serán $\chi_j = 0, j = 1, \dots, J_1$; luego de aplicar el segundo paso resulta la restricción $M_2 \subset M_1$

$$M_2 = \{(q, p) \in M_1 \mid \text{el sistema} \{\chi_j, H + \sum \lambda^m \phi_m\} = 0, j = 1, \dots, J_1, \\ \text{tiene solución}\},$$

y así sucesivamente. También supondremos que cada uno de estos subconjuntos es una subvariedad, definida regularmente por ecuaciones (al menos localmente).

Supongamos que ya se han realizado k pasos del algoritmo y queremos describir el paso $k + 1$ de un modo más geométrico.

El sistema (3.5) puede reescribirse en términos de la estructura simpléctica de $M = T^*Q$ como

$$\left\langle dH + \sum \lambda^m d\phi_m, (d\chi_j)^\sharp \right\rangle = 0, \quad (3.6)$$

donde $(d\chi_j)^\sharp$ es el único campo vectorial v tal que $d\chi_j = \Omega(v, \cdot)$, siendo Ω la 2-forma simpléctica canónica.

Dado que M_0 está definido por las ecuaciones $\phi_m = 0$ y que M_k está definido por las ecuaciones $\chi_j = 0, j = 1, \dots, J_k$, y como estamos suponiendo que M_k es subvariedad de M , para todo k , resulta entonces que los covectores $d\phi_m$ forman un conjunto de generadores del anulador $(TM_0)^\circ \subset T^*M$ y los

covectores $d\chi_j$ forman un conjunto de generadores de $(TM_k)^\circ \subset T^*M$. Se tiene entonces que los vectores $(d\chi_j)^\sharp$ generan $(TM_k)^\Omega$, donde el superíndice Ω significa ortogonal respecto a la 2-forma Ω , es decir, dado un vector $V_x \in T_xM$ se cumple $V_x \in (TM_k)^\Omega$ si y sólo si $\Omega(V_x, W_x) = 0$ para todo $W_x \in T_xM_k$.

Luego, podemos reescribir el sistema (3.6) como

$$\left\langle dH + \sum \lambda^m d\phi_m, (TM_k)^\Omega \right\rangle = 0$$

donde se han reemplazando las funciones χ_j , que definen localmente a M_k , por $(TM_k)^\Omega$, que es un objeto geométrico definido globalmente.

Observar que aún aparecen las ϕ_m en la ecuación. Dado que las $d\phi_m$ anulan a TM_0 , podemos descomponer el sistema en la parte que está en TM_0 y la que está en su complemento,

$$\langle dH, (TM_k)^\Omega \cap TM_0 \rangle = 0, \quad (3.7)$$

$$\left\langle dH + \sum \lambda^m d\phi_m, V_{M_k} \right\rangle = 0, \quad (3.8)$$

siendo V_{M_k} un fibrado vectorial tal que $(TM_k)^\Omega = ((TM_k)^\Omega \cap TM_0) \oplus V_{M_k}$. En el caso en que $(TM_k)^\Omega \cap TM_0$ tenga rango constante, V_{M_k} puede ser tomado, por ejemplo, como el complemento ortogonal de $(TM_k)^\Omega \cap TM_0$ respecto a alguna métrica.

La ecuación (3.8) contiene elementos definidos localmente, sin embargo siempre tiene solución para λ^m (como mostraremos a continuación) y por eso no interviene en la definición de M_{k+1} . Para probar que (3.8) siempre tiene solución consideramos el subespacio generado por los $d\phi_m$, $m = 1, \dots, M$; el anulador de este subespacio es TM_0 , que contiene al complemento de V_{M_k} , entonces los $d\phi_m|_{V_{M_k}}$ generan al dual $(V_{M_k})^*$; en particular, hay una combinación lineal de los $d\phi_m|_{V_{M_k}}$ que es igual a $-dH|_{V_{M_k}}$, y los coeficientes de esa combinación lineal son una solución de (3.8). En general esa solución no es única. Sin embargo lo que en este momento nos interesa es el hecho que para que el sistema completo (3.7)-(3.8) tenga solución es necesario y suficiente que se satisfaga la ecuación (3.7).

Otra cuestión interesante es que, dado que los vectores que aparecen en (3.7) están todos en TM_0 , entonces sólo son importantes los valores de H en M_0 ; de este modo se salva la ambigüedad que surgía en el método de Dirac por las múltiples posibles formas de extender el hamiltoniano (originalmente definido en M_0) a todo $M = T^*Q$. Además, restringiéndonos a M_0 podemos simplificar más la notación de la ecuación del siguiente modo: consideramos la inclusión $i: M_0 \rightarrow M$ y el pullback $\omega = i^*\Omega$, de modo que $(TM_k)^\Omega \cap TM_0 = (TM_k)^\omega$.

Con esto hemos llegado a la expresión para M_k que se define en el algoritmo desarrollado por Gotay y Nester,

$$M_{k+1} = \{x \in M_k | \langle dH, (TM_k)^\omega \rangle = 0\}. \quad (3.9)$$

De este modo la subvariedad M_{k+1} queda definida en términos de la 1-forma cerrada dH (sólo nos interesan los valores de H a lo largo de M_0) y de la estructura presimpléctica de M_0 . Además de la ventaja de que todo el algoritmo se desarrolle dentro de M_0 , también es importante remarcar que ahora el proceso está planteado en términos geométricos globales, que no dependen de ninguna elección de coordenadas.

El enfoque que hemos utilizado para presentar el método sigue un orden opuesto al seguido por Gotay–Nester–Hinds. En efecto, ellos parten de una variedad presimpléctica (M_0, ω) , una 1-forma cerrada dH y la ecuación

$$\iota(\dot{x})\omega = dH \quad (3.10)$$

(donde $\iota(\dot{x})\omega$ es la 1-forma $\omega(\dot{x}, \cdot)$). Muestran luego que la ecuación

$$\langle dH, (TM_k)^\omega \rangle = 0$$

es una condición necesaria para la existencia de soluciones tangentes a la subvariedad M_k y utilizan esta ecuación para establecer la sucesión de subvariedades M_1, M_2, \dots que construimos más arriba; muestran además que el método de Dirac es un caso particular del algoritmo planteado por ellos.

La ventaja de trabajar directamente con una variedad presimpléctica (sin hacer referencia a la variedad simpléctica de fondo) es que es posible aplicar otros métodos para simplificar las ecuaciones, como por ejemplo reducción por un grupo de simetría actuando en M_0 por acciones que preserven ω .

3.3. Sistemas dinámicos de Dirac y algoritmo de ligaduras.

La técnica para trabajar con ecuaciones diferenciales implícitas utilizada en las secciones anteriores puede generalizarse aún más, incluso cuando se trata de una ecuación diferencial implícita general. Nos ocuparemos de esta generalización en otro capítulo.

Ahora nos centraremos en la generalización necesaria para incluir el principio de D'Alembert dentro del mismo formalismo.

La ecuación de Euler–Lagrange con una fuerza externa F toma la forma

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) - \frac{\partial L}{\partial q}(q, \dot{q}) = F.$$

Supongamos que el sistema en consideración tiene las velocidades restringidas por una distribución lineal $\Delta \subset TQ$, tenemos entonces la ecuación

$$\dot{q} \in \Delta,$$

esto significa que localmente las velocidades están restringidas por ecuaciones lineales. Un ejemplo de esto sería una rueda que gira sin deslizarse, en este caso la velocidad de avance es proporcional a la velocidad de giro.

Si suponemos además que las fuerzas que realizan el vínculo no hacen trabajo sobre las velocidades, es decir, que se conserva la energía, entonces tenemos la ecuación

$$F \in \Delta^\circ.$$

Se obtienen así las ecuaciones de Lagrange–D'Alembert

$$\begin{cases} \dot{q} \in \Delta \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) - \frac{\partial L}{\partial q}(q, \dot{q}) \in \Delta^\circ \end{cases} \quad (3.11)$$

que son otro ejemplo importante de ecuaciones diferenciales implícitas. Estas ecuaciones pueden deducirse del mismo principio variacional de Lagrange,

$$\begin{cases} \dot{q} \in \Delta \\ \delta \int L(q, \dot{q}) \in \Delta^\circ \end{cases}$$

restringiendo las velocidades y las variaciones a Δ . Esto se conoce como principio de D'Alembert.

Luego de escribir estas ecuaciones en el contexto Hamiltoniano se obtienen las ecuaciones

$$\begin{cases} \dot{q} \in \Delta \\ \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} - \frac{\partial H}{\partial q} \in \Delta^\circ. \end{cases} \quad (3.12)$$

Estas ecuaciones ya no corresponden a una función que relaciona dH con (\dot{q}, \dot{p}) , sino que es un sistema de ecuaciones lineales donde aparecen (\dot{q}, \dot{p}) y dH pero de donde quizá no sea posible despejar \dot{p} .

Las ecuaciones de Hamilton–Jacobi se escriben de manera intrínseca mediante el uso de una estructura simpléctica. Para lograr algo similar con ecuaciones más generales como (3.12) utilizaremos una estructura más general, llamada *estructura de Dirac*.

Estructuras de Dirac. Las estructuras y variedades de Dirac fueron definidas por Theodore Courant en 1990 [6], como una generalización de las variedades simplécticas y de las variedades de Poisson y desde entonces se utilizaron en diferentes contextos . Así por ejemplo Yoshimura y Marsden utilizaron estructuras de Dirac con el objeto de transformar sistemas Lagrangianos con vínculos no holónomos a una forma similar a la Hamiltoniana, [14]. A continuación definimos estructura de Dirac y damos algunas propiedades que usaremos más adelante.

Definición 3.1. Una estructura de Dirac lineal sobre un espacio vectorial n -dimensional V es un subespacio $D \subset V \oplus V^*$ con $\dim(D) = n$ y tal que si $(v_1, \alpha_1), (v_2, \alpha_2) \in D$ se cumple que

$$\langle \alpha_1, v_2 \rangle + \langle \alpha_2, v_1 \rangle = 0. \quad (3.13)$$

Notar que en particular esta ecuación implica que $\langle \alpha, v \rangle = 0$ para todo $(v, \alpha) \in D$.

La condición que $\dim(D) = n$ es equivalente a la condición de maximalidad respecto a la condición (3.13). Es decir que $(v_1, \alpha_1) \in D$ si y sólo si la condición (3.13) se cumple para todo $(v_2, \alpha_2) \in D$.

Dada una estructura de Dirac D , llamamos E_D al subespacio

$$E_D = \{v \in V \mid \exists \alpha \in V^*, \text{ tal que } (v, \alpha) \in D\}.$$

En E_D puede definirse el producto

$$\omega_D(v, w) = \alpha w, \text{ con } (v, \alpha) \in D.$$

A partir de la igualdad (3.13) resulta que el producto ω_D está bien definido y es antisimétrico. Más aún, una estructura de Dirac cualquiera queda unívocamente determinada por E_D y ω_D .

Dado un subespacio $W \subset V$ definimos

$$D^b(W) = \{\alpha \in V^* \mid \exists v \in W, \text{ tal que } (v, \alpha) \in D\},$$

y definimos el ortogonal a W con respecto a D como

$$W^D = (D^b(W))^\circ.$$

Notar que esta definición es equivalente a $W^D = (W \cap E_D)^{\omega_D}$, que es una forma más práctica para calcular el ortogonal cuando E_D y ω_D son conocidos.

A partir de la definición de ortogonalidad con respecto a D obtenemos la fórmula

$$D^b(W) = (W^D)^\circ, \quad (3.14)$$

que nos da una forma más simple para calcular $D^b(W)$.

Definición 3.2. Una estructura de Dirac sobre una variedad M es un subfibrado de la suma de Whitney $D \subset TM \oplus T^*M$, tal que $D_x \subset T_xM \oplus T_x^*M$ es una estructura de Dirac lineal sobre T_xM , para cada $x \in M$. En este caso E_D es la distribución que a cada punto $x \in M$ le asigna el subespacio E_{D_x} .

Sistemas dinámicos de Dirac

Definición 3.3. Un *sistema dinámico de Dirac* sobre una variedad M es una ecuación de la forma

$$\dot{x} \oplus \mu \in D, \quad (3.15)$$

donde μ es una 1-forma sobre M y D es una estructura de Dirac sobre M .

Ejemplo 3.1. Las ecuaciones (3.12) pueden reescribirse como

$$(\dot{q}, \dot{p}) \oplus dH \in D_{(\Omega, \Delta)}, \quad (3.16)$$

donde $D_{(\Omega, \Delta)}$ es la estructura de Dirac sobre $M = T^*Q$ definida por

$$\begin{aligned} D_{(\Omega, \Delta)} &= \{(\dot{q}, \dot{p}) \oplus (\alpha dq + \beta dp) \in TM \oplus T^*M \mid \\ &\quad \dot{q} \in \Delta, \dot{q} - \beta = 0, \dot{p} + \alpha \in \Delta^\circ\} \\ &= \{(\dot{q}, \dot{p}) \oplus \nu \in TM \oplus T^*M \mid \dot{q} \in \Delta, (\dot{q}, \dot{p})^\flat - \nu \in \Delta^\circ \oplus 0\}. \end{aligned}$$

Mediante cálculos sencillos se puede probar que la ecuación (3.16) es equivalente a (3.12) y que $D_{(\Omega, \Delta)}$ así definida es una estructura de Dirac. Notar que para que la estructura anterior sea una estructura de Dirac no es necesario que Ω sea no-degenerada sino que sólo hace falta que sea antisimétrica, esto es importante para poder incluir el caso de sistemas definidos con formas antisimétricas degeneradas.

Antes de seguir avanzando, remarquemos que la definición de $D_{(\Omega, \Delta)}$ en el ejemplo anterior es independiente del sistema de coordenadas. En efecto, si llamamos τ_M a la proyección canónica $\tau_M : M = T^*Q \rightarrow Q$, entonces mediante la aplicación tangente $T\tau_M : TM \rightarrow TQ$ podemos redefinir $D_{(\Omega, \Delta)}$ como

$$D_{(\Omega, \Delta)} = \{\dot{x} \oplus \mu \in TM \oplus T^*M \mid \dot{x} \in ((T\tau_M)^{-1}\Delta), \dot{x}^\flat - \mu \in ((T\tau_M)^{-1}\Delta)^\circ\}.$$

Ejemplo 3.2. Vamos a escribir el sistema Lagrangiano con vínculos no-holónomos (3.11) utilizando estructuras de Dirac como en el ejemplo anterior, pero ahora incluyendo el caso de Lagrangianos degenerados. Esto ya fue hecho en [14], pero aquí vamos a recorrer un camino diferente considerando $M = TQ \oplus T^*Q$, ver [15], en lugar de $M = T^*Q$, como en [14].

Para poder definir la función de energía $\mathcal{E} = pv - L(q, v)$ trabajaremos en coordenadas $(q, v, p) \in TQ \oplus T^*Q$. Al costo de agregar nuevas variables hemos simplificado el problema que aparece cuando la transformada de Legendre no es invertible (y por lo tanto las coordenadas v no son funciones de las coordenadas p).

Supongamos dados un Lagrangiano $L : TQ \rightarrow \mathbb{R}$ y una distribución $\Delta \subset TQ$. Consideramos $M = TQ \oplus T^*Q$, y sean $\tau_M : M \rightarrow Q$ y $\pi_{T^*Q} : M \rightarrow T^*Q$ las proyecciones canónicas, Ω la 2-forma canónica sobre T^*Q , $\omega = \pi_{T^*Q}^* \Omega$ y $\tilde{\Delta} = (T\tau_M)^{-1} \Delta$.

Definimos

$$D_{(\omega, \tilde{\Delta})} = \{\dot{x} \oplus \mu \in TM \oplus T^*M \mid \dot{x} \in \tilde{\Delta}, \omega^\flat(\dot{x}) - \mu \in \tilde{\Delta}^\circ\}, \quad (3.17)$$

y el sistema dinámico de Dirac asociado

$$\dot{x} \oplus d\mathcal{E} \in D_{(\omega, \tilde{\Delta})}, \quad (3.18)$$

que resulta equivalente a (3.11). En efecto, en coordenadas tenemos $\dot{x} = (\dot{q}, \dot{v}, \dot{p})$ y $d\mathcal{E} = -\frac{\partial}{\partial q} L dq + (p - \frac{\partial}{\partial v} L) dv + v dp$, además $\omega = dq \wedge dp$ de donde $\omega^\flat(\dot{x}) = -\dot{p} dq + \dot{q} dp$ y finalmente $\tilde{\Delta}_x = \Delta_x \oplus \mathbb{R}^n \oplus \mathbb{R}^n$ y $\tilde{\Delta}^\circ = \Delta^\circ \oplus 0 \oplus 0$.

Luego, la ecuación $\dot{x} \in \tilde{\Delta}$ se reduce a

$$\dot{q} \in \Delta;$$

mientras que la ecuación $\omega^\flat(\dot{x}) - d\mathcal{E} \in \tilde{\Delta}^\circ$, en coordenadas (q, v, p) , corresponde a las tres ecuaciones

$$\begin{aligned} \frac{\partial L}{\partial q} - \dot{p} &\in \Delta^\circ \\ \frac{\partial L}{\partial v} - p &= 0 \\ \dot{q} - v &= 0. \end{aligned}$$

Es inmediato verificar que estas cuatro ecuaciones son equivalentes a (3.11).

Observar que las estructuras de Dirac definidas en esta subsección generalizan a las formas presimplécticas, en el sentido de que si ω es una forma presimpléctica sobre M entonces $\dot{x} \oplus \mu \in D_{(\omega, TM)}$ si y sólo si $\omega^b(\dot{x}) = \iota(\dot{x})\omega = \mu$. En consecuencia, las ecuaciones de la forma

$$\dot{x} \oplus \mu \in D,$$

donde D es una estructura de Dirac y μ es una 1-forma, generalizan la ecuación de Gotay–Nester (3.10).

Algoritmo de ligaduras para estructuras de Dirac . Presentaremos un algoritmo que generaliza el algoritmo de Gotay–Nester (3.9) para sistemas de la forma (3.15), al que llamaremos algoritmo CAD (Constraint Algorithm for Dirac Dynamical Systems), [15].

Tomaremos $M_0 = M$ y definiremos M_{k+1} como el conjunto de puntos x donde la ecuación lineal $\dot{x} \oplus \mu_x \in D_x$ tiene solución $\dot{x} \in T_x M_k$, es decir, los puntos x para los que $\mu_x \in D_x^b(T_x M_k)$. Utilizando la formula (3.14) podemos escribir,

$$M_{k+1} = \{x \in M_k \mid \langle \mu, (TM_k)^D \rangle = 0\}. \quad (3.19)$$

Luego, utilizando la forma en que hemos definido el concepto de ortogonalidad respecto a una estructura de Dirac, obtenemos un algoritmo que formalmente es muy similar al algoritmo de Gotay–Nester, y que bajo las mismas hipótesis, es decir suponiendo que en cada paso se obtiene una subvariedad, se termina construyendo una subvariedad final donde hay existencia de soluciones (aunque en general no hay unicidad).

3.4. Ejemplos

En esta subsección daremos algunos ejemplos de sistemas que pueden describirse mediante estructuras de Dirac y obtendremos en cada caso el resultado de aplicar el algoritmo CAD.

Ejemplo 3.3 (Algoritmo de Gotay–Nester y algoritmo CAD). Utilizando la estructura de Dirac correspondiente a la forma presimpléctica ω :

$$D_\omega = \{\dot{x} \oplus \mu \in TM \oplus T^*M \mid \mu = \omega(\dot{x}, \cdot)\},$$

resulta que las ecuaciones (3.10) y (3.15) son equivalentes. Además $(TM_k)^D = (TM_k)^\omega$. Por lo tanto en este caso el algoritmo CAD coincide exactamente con el algoritmo de Gotay–Nester.

En [15] se prueba que en el caso en que la distribución E_D es integrable, entonces el algoritmo CAD restringido a cada hoja integral de E_D es equivalente a un algoritmo de Gotay–Nester. El ejemplo anterior es un caso particular de esto.

Ejemplo 3.4 (Circuitos LC). Un circuito LC es un circuito eléctrico compuesto por inductores y capacitores (en la figura 3.1, L es un inductor y C_1, C_2 son capacitores).

Hay diferentes maneras de describir un circuito LC, en este trabajo tomaremos la utilizada en [14, 15], que parten del espacio de cargas Q , que es un espacio vectorial de dimensión igual a la cantidad de componentes del circuito. El fibrado tangente TQ es el espacio de corrientes y el fibrado cotangente T^*Q es el espacio de *flux linkage*.

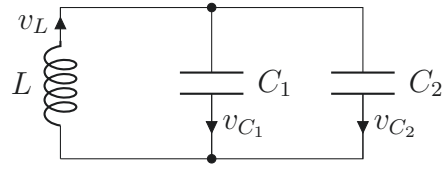


Figura 3.1: circuito LC

La ley de Kirchoff para las corrientes (LKC) indica que las corrientes que entran a un nodo en un circuito deben ser iguales a las que salen ($v_L = v_{C_1} + v_{C_2}$ en el circuito del gráfico). Las ecuaciones de la LKC definen una distribución constante $\Delta \subset TQ$.

La ley de Kirchoff para los voltajes (LKV) indica que la suma de los voltajes en un circuito cerrado debe ser cero (en el gráfico $v_L + v_{C_1} = 0$ y $v_{C_1} - v_{C_2} = 0$). La distribución constante definida por la LKV es $\Delta^\circ \subset T^*Q$.

La dinámica de un circuito LC puede ser calculada mediante el principio variacional

$$\delta \int \sum \frac{L_i}{2} (\dot{q}_{L_i})^2 - \frac{1}{2C_j} (q_{C_j})^2 = 0 \quad (3.20)$$

sometido a las restricciones $v = \dot{q} \in \Delta$ y $\delta q \in \Delta^\circ$. El hecho que el subespacio que resulta de la aplicación de la ley de Kirchoff para los voltajes sea el anulador de Δ puede demostrarse usando técnicas de teoría de grafos, y físicamente corresponde a la conservación de la potencia eléctrica.

El subespacio Δ puede calcularse de manera sencilla a partir de la topología del grafo del circuito y Δ° es su anulador.

Este planteo variacional da lugar a las ecuaciones de Lagrange–D’Alembert

$$\begin{cases} \dot{q} \in \Delta \\ \sum \frac{d}{dt} (L_i \dot{q}_{L_i}) + q_{C_j} / C_j \in \Delta^\circ. \end{cases} \quad (3.21)$$

Estas ecuaciones pueden escribirse mediante una estructura de Dirac. En [14] construyen una estructura de Dirac sobre T^*Q , pero para describir las ecuaciones necesitan definir el *diferencial de Dirac* del Lagrangiano. En

[15] evitan esta complicación al costo de trabajar con más variables, ya que construyen la estructura de Dirac sobre $M = TQ \oplus T^*Q$. Aquí tomaremos este segundo camino. La estructura de Dirac queda entonces definida sobre M (con coordenadas (q, v, p)) como:

$$D = \{(x, \dot{x}, \mu) \in TM \oplus T^*M \mid \dot{q} \in \Delta, \dot{q} = \mu_p, \dot{p} + \mu_q = 0\}.$$

Definimos la función de energía

$$\mathcal{E} = p \cdot v - \sum \frac{L_i}{2} (v_{L_n})^2 - \frac{1}{2C_n} (q_{C_i})^2.$$

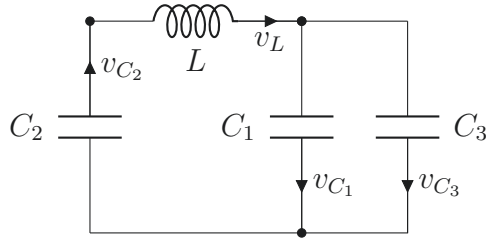
Luego, la ecuación

$$\dot{x} \oplus d\mathcal{E} \in D$$

describe la dinámica del circuito.

Como caso particular, observemos que en el circuito de la derecha la ley de Kirchoff para las corrientes está dada por

$$\Delta = \{(v_L, v_{C_1}, v_{C_2}, v_{C_3}) \mid v_L = v_{C_2} = v_{C_1} + v_{C_3}\},$$



y la función de energía resulta

$$\mathcal{E} = p \cdot v - \frac{1}{2} L v_L^2 + \frac{1}{2} \frac{q_{C_1}^2}{C_1} + \frac{1}{2} \frac{q_{C_2}^2}{C_2} + \frac{1}{2} \frac{q_{C_3}^2}{C_3}.$$

Al aplicar el algoritmo CAD a la ecuación $\dot{x} \oplus d\mathcal{E} \in D$ obtenemos,

$$M_0 = M = TQ \oplus T^*Q,$$

$$(TM_0)^D = \{(x, \dot{x}) \in TM \mid \dot{q} = 0, \dot{p}_L + \dot{p}_{C_1} + \dot{p}_{C_2} = 0, \dot{p}_L + \dot{p}_{C_2} + \dot{p}_{C_3} = 0\},$$

$$\begin{aligned} d\mathcal{E} = & \frac{q_{C_1}}{C_1} dq_{C_1} + \frac{q_{C_2}}{C_2} dq_{C_2} + \frac{q_{C_3}}{C_3} dq_{C_3} + \\ & (p_L - L v_L) dv_L + p_{C_1} dv_{C_1} + p_{C_2} dv_{C_2} + p_{C_3} dv_{C_3} + \\ & v_L dp_L + v_{C_1} dp_{C_1} + v_{C_2} dp_{C_2} + v_{C_3} dp_{C_3}. \end{aligned}$$

Entonces la expresión $\langle d\mathcal{E}, (TM_0)^D \rangle = 0$ corresponde a las ecuaciones

$$\begin{cases} p_L = L v_L \\ p_{C_1} = p_{C_2} = p_{C_3} = 0 \\ v_L = v_{C_2} = v_{C_1} + v_{C_3} \end{cases}$$

que definen M_1 .

Calculamos ahora

$$TM_1^D = \{(x, \dot{x}) \in T_{M_1}M \mid \dot{q}_L = \dot{q}_{C_1} = \dot{q}_{C_2} + \dot{q}_{C_3} = 0, \\ \dot{p}_L + \dot{p}_{C_1} + \dot{p}_{C_2} = 0, \dot{p}_L + \dot{p}_{C_2} + \dot{p}_{C_3} = 0\}.$$

Luego la ecuación $\langle d\mathcal{E}, (TM_1)^D \rangle = 0$ agrega una nueva restricción,

$$\begin{cases} p_L = Lv_L \\ p_{C_1} = p_{C_2} = p_{C_3} = 0 \\ v_L = v_{C_2} = v_{C_1} + v_{C_3} \\ \frac{q_{C_1}}{C_1} = \frac{q_{C_3}}{C_3}. \end{cases}$$

Estas ecuaciones definen M_2 .

De manera análoga al paso anterior calculamos

$$TM_2^D = \{(x, \dot{x}) \in T_{M_2}M \mid \dot{q}_L = \dot{q}_{C_1} = \dot{q}_{C_2} + \dot{q}_{C_3} = 0, \\ (C_1 + C_3)\dot{p}_L + C_1\dot{p}_{C_1} + (C_1 + C_3)\dot{p}_{C_2} + C_3\dot{p}_{C_3} = 0\}.$$

La ecuación $\langle d\mathcal{E}, (TM_2)^D \rangle = 0$ agrega otra restricción,

$$\begin{cases} p_L = Lv_L \\ p_{C_1} = p_{C_2} = p_{C_3} = 0 \\ v_L = v_{C_2} = v_{C_1} + v_{C_3} \\ \frac{q_{C_1}}{C_1} = \frac{q_{C_3}}{C_3} \\ \frac{v_{C_1}}{C_1} = \frac{v_{C_3}}{C_3}, \end{cases}$$

obteniéndose las ecuaciones que definen M_3 .

En el siguiente paso resulta $M_4 = M_3$, por lo que el algoritmo se detiene. La variedad M_3 puede parametrizarse por las coordenadas p_L , v_L , q_L y q_{C_1} . La variable q_L es una variable arbitraria, porque no aparece en ninguna de las ecuaciones.

Estas ecuaciones llevan al mismo resultado al que llevarían las fórmulas para combinar capacitores en serie y en paralelo que se usan habitualmente.

Cabe mencionar que, si bien en el ejemplo se utilizaron inductores y capacitores lineales, el método funciona igualmente con capacitores e inductores no lineales, utilizando el Lagrangiano propio del sistema, en el cual las inductancias y capacitancias en general dependerán de p o de q [16]. Casos más generales están contemplados por la teoría, pero tal como está no puede aplicarse a circuitos con resistencias. Nos ocuparemos de eso en la sección 4.4 (ejemplo 4.7).

Ejemplo 3.5 (Sistemas Lagrangianos con vínculos no-holónomos). En el ejemplo 3.2 (pág. 30) hemos visto que dado un Lagrangiano $L : TQ \rightarrow \mathbb{R}$, y una distribución de velocidades permitidas $\Delta \subset TQ$, el sistema mecánico con vínculos no holónomos puede describirse mediante el sistema dinámico de Dirac,

$$\dot{x} \oplus d\mathcal{E} \in D_{(\omega, \tilde{\Delta})},$$

utilizando la estructura de Dirac sobre $M = TQ \oplus T^*Q$ definida como

$$D_{(\omega, \tilde{\Delta})} = \{\dot{x} \oplus \mu \in TM \oplus T^*M \mid \dot{x} \in \tilde{\Delta}, \omega^\flat(\dot{x}) - \mu \in \tilde{\Delta}^\circ\}.$$

Si llamamos $\tau_M : M \rightarrow Q$, $\pi_{TQ} : M \rightarrow TQ$ y $\pi_{T^*Q} : M \rightarrow T^*Q$ a las proyecciones canónicas y Ω a la 2-forma canónica sobre T^*Q , entonces ω y $\tilde{\Delta}$ quedan definidas mediante

$$\omega = \pi_{T^*Q}^* \Omega \quad \tilde{\Delta} = (T\tau_M)^{-1} \Delta.$$

La función de energía $\mathcal{E} : M \rightarrow \mathbb{R}$ está definida en coordenadas como $\mathcal{E} = pv - L(q, v)$. Esta definición puede obtenerse de manera independiente de las coordenadas haciendo

$$\mathcal{E}(x) = \langle \pi_{T^*Q}(x), \pi_{TQ}(x) \rangle - L(\pi_{TQ}(x)).$$

El algoritmo CAD desarrollado a partir de la estructura de Dirac $D_{(\omega, \tilde{\Delta})}$ no es una generalización inmediata del método de Dirac, ya que este último da lugar a subvariedades en T^*Q , mientras que a partir de aplicar el algoritmo CAD se obtienen subvariedades en $TQ \oplus T^*Q$. Sin embargo cabe preguntarse si habrá alguna relación entre el algoritmo de Gotay–Nester y las proyecciones de las subvariedades obtenidas a partir del algoritmo CAD.

En [15] se muestra que el algoritmo CAD aplicado en $M = TQ \oplus T^*Q$, cuando $\Delta = TM$, es equivalente al algoritmo de Gotay–Nester en la misma variedad tomando como forma presimpléctica a la restricción a M de la 2-forma canónica de T^*TQ (considerando el embedding $(q, v, p) \mapsto (q, v, p, 0)$, esta restricción es lo mismo que el pullback de la 2-forma canónica de T^*Q). De este modo, el algoritmo CAD resulta equivalente al método de Dirac, pero aplicado sobre T^*TQ con variedad de restricciones primarias igual a $TQ \oplus T^*Q$.

Una pregunta que, hasta donde sabemos, no ha sido respondida aún, es en qué sentido y bajo qué condiciones esto es equivalente al método de Dirac aplicado sobre T^*Q . La variedad M_1 resultante del primer paso del algoritmo se proyecta sobre la variedad de restricciones primarias en T^*Q . Si esto continúa ocurriendo con las siguientes iteraciones del algoritmo es una pregunta interesante que podría ser un tema de estudio futuro.

Capítulo 4

Sistemas con fricción

En el Capítulo anterior ya se ha mencionado que la idea principal subyacente en el algoritmo de Gotay–Nester puede ser aplicada a otras clases de EDIs, en particular se desarrolló el caso de los sistemas dinámicos de Dirac, que permiten trabajar con sistemas mecánicos con vínculos no–holónomos y circuitos eléctricos LC con un formalismo bastante similar al del algoritmo de Gotay–Nester [15].

Otras formulaciones aún más generales son las que se presentan para el caso \mathcal{C}^∞ en [12] y para el caso analítico en [2], ambos enfoques serán desarrollados en el Capítulo 5.

Volviendo a los algoritmos considerados en el Capítulo anterior, hay que observar que la formulación geométrica general provino de la noción de ortogonalidad (respecto a una forma presimpléctica en el caso del algoritmo de Gotay–Nester, y respecto a una estructura de Dirac en el caso del algoritmo CAD). En este Capítulo consideraremos algunas estructuras geométricas que nos permitan definir algún concepto de ortogonalidad para poder implementar algoritmos similares al de Gotay–Nester ó al CAD, pero aplicables a sistemas donde no haya necesariamente conservación de energía.

4.1. Principio D’Alembert generalizado

La antisimetría de las formas presimplécticas está directamente vinculada con la conservación de la energía en los sistemas que pueden ser tratados mediante el algoritmo de Gotay–Nester, ya que de la ecuación de tipo (3.10)

$$\iota(\dot{x})\omega = dH,$$

surge que $\langle dH \dot{x} \rangle = \omega(\dot{x}, \dot{x}) = 0$.

Igualmente la condición de isotropía de las estructuras de Dirac (3.13) implica también la conservación de la energía, pues a partir de la ecuación de tipo (3.15)

$$\dot{x} \oplus \mu \in D,$$

reemplazando (\dot{x}, μ) en los dos pares que aparecen en (3.13) resulta que

$$\langle \mu, \dot{x} \rangle + \langle \mu, \dot{x} \rangle = 0 \implies \langle \mu, \dot{x} \rangle = 0.$$

Para poder incluir sistemas no conservativos en formalismos similares a los del Capítulo anterior, es necesario perder las hipótesis que implican antisimetría.

En el caso del principio de D'Alembert, la condición de antisimetría aparece al suponer que la fuerza que realiza el vínculo (o las variaciones) está en el anulador del subespacio Δ de las velocidades admisibles. Por lo tanto para obtener un principio de D'Alembert generalizado que incluya también sistemas disipativos es necesario incluir como datos, además de las restricciones cinemáticas (i.e. $\dot{q} \in \Delta$), a las restricciones variacionales, que tendrán la forma $\delta q \in \Delta_V$. Notar que el caso del principio de D'Alembert para sistemas conservativos aparecerá cuando $\Delta_V = \Delta$.

Diferentes trabajos han estudiado este tipo de ideas, entre ellos [1, 3, 4, 5].

Aquí consideraremos el caso en que las restricciones cinemáticas Δ son una distribución en TQ y las restricciones variacionales Δ_V son otra distribución en TQ . En esta situación el problema variacional dado por el principio de D'Alembert generalizado

$$\begin{cases} \dot{q} \in \Delta \\ \delta \int L(q, \dot{q}) \in (\Delta_V)^\circ \end{cases}$$

da lugar a las ecuaciones

$$\begin{cases} \dot{q} \in \Delta \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) - \frac{\partial L}{\partial q}(q, \dot{q}) \in (\Delta_V)^\circ. \end{cases} \quad (4.1)$$

El objetivo de este Capítulo es buscar formas adecuadas para plantear este tipo de ecuaciones sobre una estructura geométrica que nos permita aplicar un algoritmo del tipo Gotay–Nester. La clave para poder llegar a este resultado consistirá en poder definir de manera adecuada el concepto de ortogonalidad.

4.2. Algoritmo de ligaduras para formas que no son antisimétricas

Consideremos la ecuación

$$\omega(\dot{x}, \cdot) = dH \quad (4.2)$$

donde dH es una 1-forma y ω es una forma bilineal (no necesariamente antisimétrica).

Esta forma bilineal ω puede ser vista como una aplicación $\omega^\flat : TM \rightarrow T^*M$, dada por $\omega^\flat(v)(w) = \omega(v, w)$, de modo que (4.2) toma la forma

$$\omega^\flat(\dot{x}) = dH. \quad (4.3)$$

Para aplicar un algoritmo similar al de Gotay–Nester a una ecuación de la forma (4.3) es necesario ser cuidadoso al momento de definir un concepto relacionado con la ortogonalidad, ya que cuando ω no es antisimétrica, $\omega(u, v) = 0$ no implica necesariamente que $\omega(v, u) = 0$.

Definición 4.1. Dados una forma bilineal ω definida en un espacio vectorial V y un vector $v \in V$ definimos $\omega^\flat(v) \in V^*$ como la 1-forma tal que $\omega^\flat(v)(w) = \omega(v, w)$, para todo $w \in V$.

Dada una forma bilineal ω sobre una variedad M definimos $\omega^\flat : TM \rightarrow T^*M$ de modo que $\omega^\flat|_{T_x M} = (\omega|_{T_x M})^\flat$.

Definición 4.2. Dada una forma bilineal ω en V y dado un subespacio $E \subset V$ definimos el *subespacio ortogonal respecto a ω* como

$$E^\omega = \{u \in V \mid \omega(v, u) = 0, \forall v \in E\} = \bigcap_{v \in E} \ker \omega^\flat(v) = (\omega^\flat(E))^\circ,$$

donde la última igualdad se obtiene a partir de la identificación entre V y V^{**} .

Definición 4.3. Dada una forma bilineal ω sobre una variedad M y una distribución $E \subset TM$ definimos la *distribución ortogonal respecto a ω* como

$$E^\omega = \bigcup_{x \in M} (E_x)^\omega.$$

De esta última definición, y dado que estamos considerando el caso finito dimensional, resulta la fórmula $\omega^\flat(E) = (E^\omega)^\circ$.

Con estas definiciones puede establecerse un algoritmo similar al de Gotay–Nester para formas bilineales que no sean antisimétricas. Más precisamente, dada una ecuación de la forma (4.3) el conjunto

$M_1 = \{x \in M_0 \mid \text{la ecuación } \omega_x^b(v) = dH(x) \text{ tiene solución } v \in T_x M_0\}$
puede redefinirse así,

$$M_1 = \{x \in M_0 \mid \langle dH, (T_x M_0)^\omega \rangle = 0\},$$

y los siguientes pasos del algoritmo quedan determinados del siguiente modo,

$$M_{k+1} = \{x \in M_k \mid \langle dH, (T_x M_k)^\omega \rangle = 0\};$$

donde estamos considerando $M_0 := M$.

Formalmente estas definiciones parecen coincidir con las establecidas a partir del algoritmo de Gotay–Nester, pero debe tenerse en cuenta que el concepto de ortogonalidad con respecto a ω ha sido modificado convenientemente. Es necesario remarcar que $E^\omega = \{u \in V \mid \omega(v, u) = 0, \forall v \in E\} \neq \{u \in V \mid \omega(u, v) = 0, \forall v \in E\}$, es decir, es importante el orden de u y v en la definición del conjunto ortogonal.

Ejemplo 4.1. Consideremos un péndulo, con espacio de configuración $Q = S^1$, y Lagrangiano $L = \frac{1}{2}mv^2 + mg \cos q$, donde m es la masa del péndulo, g es la aceleración de la gravedad, $q \in Q$ es el ángulo que forma la varilla del péndulo con la vertical y $v = dq/dt$ (suponemos que la varilla tiene longitud 1). Tomamos $M_0 = TQ \oplus T^*Q$, con coordenadas (q, v, p) , y la función de energía $H(q, v, p) = pv - L(q, v) = pv - \frac{1}{2}mv^2 - mg \cos q$.

Consideremos la forma bilineal ω dada por $\omega^b(\dot{q}, \dot{v}, \dot{p}) = \dot{q}dp - \dot{p}dq - k\dot{q}dq$ que es la suma del pullback de la forma simpléctica canónica sobre T^*Q y de la forma lineal simétrica $-k\dot{q}dq$, donde k es el coeficiente de fricción.

Se tiene que $TM_0^\omega = \{\dot{q} = 0, \dot{p} - k\dot{q} = 0\}$, de aquí que al aplicar el primer paso del algoritmo se obtiene

$$M_1 = \{(q, v, p) \in M_0 \mid p = mv\}.$$

Además, para cada $x \in M_1$ se tiene que $T_x M_1^\omega = T_x M_0^\omega$, con lo cual el algoritmo se detiene en este paso.

Si en M_1 tomamos coordenadas (q, v) la ecuación resulta

$$\begin{cases} m\dot{q} = mv \\ -m\dot{v} - k\dot{q} = mg \sin q \end{cases}$$

o lo que es lo mismo

$$m\ddot{q} = -mg \sin q - k\dot{q},$$

que es la ecuación del péndulo con fricción. Aquí puede notarse que la parte simétrica de la forma bilineal ω representa la fuerza de roce.

4.3. Estructuras de Dirac generalizadas

Comenzamos recordando la definición de estructura de Dirac sobre un espacio vectorial. Dado un espacio vectorial V de dimensión finita, una *estructura de Dirac* sobre V es un subespacio $D \subset V \oplus V^*$, que es maximal respecto a la propiedad $\langle \alpha_1, v_2 \rangle + \langle \alpha_2, v_1 \rangle = 0$. Un ejemplo muy importante es la gráfica de una forma bilineal antisimétrica ω definida sobre V , es decir, $D = \{(v, \omega(v, \cdot)) \mid v \in V\}$.

Para dar diferentes formas equivalentes de esta definición, en $D \subset V \oplus V^*$ podemos introducir los siguientes productos bilineales:

$$\begin{aligned} \langle (v_1, \alpha_1), (v_2, \alpha_2) \rangle &= \alpha_1 v_2, \\ \langle (v_1, \alpha_1), (v_2, \alpha_2) \rangle_- &= (\alpha_1 v_2 - \alpha_2 v_1)/2, \\ \langle (v_1, \alpha_1), (v_2, \alpha_2) \rangle_+ &= (\alpha_1 v_2 + \alpha_2 v_1)/2. \end{aligned}$$

Así resulta que un subespacio $D \subset V \oplus V^*$ es una estructura de Dirac si y sólo si D es maximal isotrópico respecto a \langle, \rangle_+ .

Consideramos las proyecciones canónicas $\pi_V : V \oplus V^* \rightarrow V$ y $\pi_{V^*} : V \oplus V^* \rightarrow V^*$, a partir de ellas definimos los siguientes objetos, para cualquier subespacio $D \subset V \oplus V^*$:

- $E_D = \pi_V(D)$, que es un subespacio de V .
- Para $v \in E_D$ se define $D^b(v) = \pi_{V^*}(D \cap \pi_V^{-1}(v))$. Así $D^b(0)$ es un subespacio de V^* .
- Para $v, w \in E_D$ se define $\omega_D(v, w) = \alpha w$, $\alpha \in D^b(v)$.

Si D es una estructura de Dirac, vale que $E_D = D^b(0)^\circ$. En efecto, sea $\alpha \in D^b(0)$, entonces $(0, \alpha) \in D$. Sea $v \in E_D$, entonces existe $\beta \in V^*$ tal que $(v, \beta) \in D$. Como D es isotrópico tenemos que $0 = \langle (0, \alpha), (v, \beta) \rangle_+ = \alpha v$. Luego $v \in \alpha^\circ$. Como esto ocurre para cada $v \in E_D$ y para cada $\alpha \in D^b(0)$, resulta $E_D \subset D^b(0)^\circ$.

Por otro lado tenemos que $E_D = \text{Im}(\pi_V|_D)$ y $D^b(0) = \pi_{V^*}(\ker(\pi_V|_D)) = \pi_{V^*}(\ker(\pi_V) \cap D)$.

La restricción $\pi_{V^*}|_{\ker(\pi_V)}$ es un isomorfismo, luego podemos escribir $\dim D^b(0) = \dim \ker(\pi_V|_D) = \dim D - \dim \text{Im}(\pi_V|_D) = \dim V - \dim E_D$.

Finalmente, $\dim D^b(0)^\circ = \dim V - \dim D^b(0) = \dim E_D$. Como además vale que $E_D \subset D^b(0)^\circ$, concluimos que $E_D = D^b(0)^\circ$. □

La propiedad que acabamos de probar garantiza que ω_D esté bien definida, ya que si $\alpha, \beta \in D^b(v)$ entonces $\alpha - \beta \in D^b(0) = E_D^\circ$, luego $\alpha w = \beta w$, para todo $w \in E_D$.

4.3.1. Estructuras de Dirac generalizadas sobre un espacio vectorial

Nos ocuparemos ahora de la generalización de las estructuras de Dirac.

Dado que la propiedad de isotropía en las estructuras de Dirac es la que garantiza la conservación de la energía en los sistemas de Dirac, para generalizar el concepto de estructura de Dirac podemos simplemente eliminar esta propiedad.

Definición 4.4. Una estructura de Dirac generalizada sobre un espacio vectorial de dimensión finita V es un subespacio $D \subset V \oplus V^*$.

Esta definición es, quizá, demasiado general. Nuestro interés consiste en hallar dos modos de generalizar las estructuras de Dirac que permitan incluir sistemas no conservativos, uno incluyendo la fuerza disipativa explícitamente, como ocurrió en el caso de las formas bilineales no antisimétricas y otro tomando la forma de las ecuaciones del principio de D'Alembert generalizado.

Para el primero de estos objetivos deberían generalizarse las estructuras de Dirac de modo que la forma bilineal ω_D pudiera ser una forma bilineal cualquiera (no necesariamente antisimétrica), y para el otro objetivo deberían considerarse estructuras donde la propiedad $D^b(0)^\circ = E_D$ no se cumpla (ya veremos que E_D jugará el papel de Δ y $D^b(0)^\circ$ jugará el papel de Δ_V en el principio de D'Alembert generalizado).

La definición 4.4 engloba ambas posibilidades, y permite hacer una gran parte del trabajo de modo tal de incluir ambas situaciones. Las particularidades de cada caso las consideraremos cuando sea necesario.

Las definiciones de E_D y D^b dadas para estructuras de Dirac, se extienden sin problemas si consideramos estructuras de Dirac generalizadas (no así la de ω_D). Es conveniente, además, agregar las siguientes definiciones.

- $\tilde{E}_D = D^b(0)^\circ$, que resulta un subespacio de V .
- Para $\alpha \in \pi_{V^*}(D)$ se define $D^\sharp(\alpha) = \pi_V(D \cap \pi_{V^*}^{-1}(\alpha))$.

Notar que en general D^b y D^\sharp no son funciones inversas. En cambio se cumple que $\alpha \in D^b(v) \iff v \in D^\sharp(\alpha)$.

Para generalizar la definición de ω_D procedemos del siguiente modo.

Definición 4.5. Dado un espacio vectorial V y dado $D \subset V \oplus V^*$ un subespacio, definimos $\omega_D : E_D \times \tilde{E}_D \rightarrow \mathbb{R}$ como $\omega_D(v, \tilde{v}) = \alpha\tilde{v}$, con $\alpha \in D^b(v)$.

La función ω_D está bien definida, ya que si $\alpha, \beta \in D^b(v)$ entonces $\alpha - \beta \in D^b(0) = (\tilde{E}_D)^\circ$ y por lo tanto $\alpha\tilde{v} = \beta\tilde{v}$. Además, es inmediato verificar que se trata de una función bilineal.

Lema 4.1. *Los subespacios E_D, \tilde{E}_D junto con la función bilineal ω_D determinan unívocamente a la estructura de Dirac generalizada D . Más aún, se tiene que,*

$$D = \{(v, \alpha) \in V \oplus V^* \mid v \in E_D, \alpha|_{\tilde{E}_D} = \omega_D(v, \cdot)\}.$$

Demostración. Sea $(v, \alpha) \in D$, por definición tenemos que $v \in E_D$ y $\omega_D(v, w) = \alpha w$ para todo $w \in \tilde{E}_D$, por lo tanto $D \subset \{(v, \alpha) \in V \oplus V^* \mid v \in E_D, \alpha|_{\tilde{E}_D} = \omega_D(v, \cdot)\}$.

Por otro lado $\dim D = \dim \text{Im}(\pi_V|_D) + \dim \ker(\pi_V|_D) = \dim E_D + \dim D^b(0) = \dim V + \dim E_D - \dim \tilde{E}_D = \dim \{(v, \alpha) \in V \oplus V^* \mid v \in E_D, \alpha|_{\tilde{E}_D} = \omega_D(v, \cdot)\}$. \square

Definición 4.6. Dados dos vectores $v \in E_D$ y $w \in \tilde{E}_D$ decimos que v y w son ortogonales respecto a D si y sólo si $\omega_D(v, w) = 0$.

Dado un subespacio $W \subset E_D$, definimos el *subespacio ortogonal respecto a D* , $W^D \subset \tilde{E}_D$ como

$$W^D = \{w \in \tilde{E}_D \mid v \text{ y } w \text{ son ortogonales respecto a } D, \forall v \in W\}.$$

Cuando no se cumpla que $W \subset E_D$, por abuso de notación escribiremos W^D para referirnos a $(W \cap E_D)^D$.

Equivalentemente, dos vectores $v, w \in V$ son ortogonales respecto a D si y sólo si $w \in (D^b(0))^\circ$ y existe $\alpha \in V^*$, con $(v, \alpha) \in D$ y $\alpha w = 0$. En consecuencia, teniendo en cuenta que $0 \in W$ y entonces $(D^b(W))^\circ \subset (D^b(0))^\circ$, se tiene la siguiente fórmula,

$$W^D = (D^b(W))^\circ. \quad (4.4)$$

\square

Esta fórmula tendrá utilidad más adelante escrita de la siguiente manera: dados $\alpha \in V^*$ y $W \subset V$,

$$\alpha \in D^b(W) \iff \alpha \cdot (W^D) = \{0\}.$$

Ahora pasaremos a calcular algunas fórmulas que permitirán ver a W^D de una manera diferente.

Definición 4.7. Dada una estructura de Dirac generalizada $D \subset V \oplus V^*$ definimos

$$D^\top = \{(w, \beta) \in V \oplus V^* \mid \langle (v, \alpha), (w, \beta) \rangle_- = 0, \forall (v, \alpha) \in D\},$$

$$D^\perp = \{(w, \beta) \in V \oplus V^* \mid \langle (v, \alpha), (w, \beta) \rangle_+ = 0, \forall (v, \alpha) \in D\}.$$

Lema 4.2. *Se tienen las siguientes igualdades:*

- $E_{D^\top} = E_{D^\perp} = \tilde{E}_D.$
- $\tilde{E}_{D^\top} = \tilde{E}_{D^\perp} = E_D.$
- $\omega_{D^\top}(w, v) = \omega_D(v, w).$
- $\omega_{D^\perp}(w, v) = -\omega_D(v, w).$

Demostración. Primero notar que, dado que \langle, \rangle_+ y \langle, \rangle_- son formas bilineales no degeneradas, entonces resulta que $\dim D^\top = \dim D^\perp = 2 \dim V - \dim D.$

Notar además que si $\alpha \in D^\flat(0)$, i.e. si $(0, \alpha) \in D$, y si $w \in E_{D^\top}$ (resp. $w \in E_{D^\perp}$) entonces existe $\beta \in V^*$ tal que $\alpha w + \beta 0 = 0$ (resp. $\alpha w - \beta 0 = 0$), en consecuencia $\alpha w = 0$. Por lo tanto resulta que $E_{D^\top} \subseteq D^\flat(0)^\circ = \tilde{E}_D$ (resp. $E_{D^\perp} \subseteq \tilde{E}_D$).

Vale también que, si $(0, \beta) \in D^\top$ (resp. $(0, \beta) \in D^\perp$), resulta que $\beta v = 0$, para todo $v \in E_D$, por lo tanto $E_D \subseteq \tilde{E}_{D^\top}$ (resp. $E_D \subseteq \tilde{E}_{D^\perp}$).

Para concluir tenemos que $\dim D^\top = \dim V + \dim E_{D^\top} - \dim \tilde{E}_{D^\top} \leq \dim V + \dim \tilde{E}_D - \dim E_D = 2 \dim V - \dim D$, y la igualdad se cumple sólo si $E_{D^\top} = \tilde{E}_D$ y $E_D = \tilde{E}_{D^\top}$. Con un razonamiento semejante se prueba que $E_{D^\perp} = \tilde{E}_D$ y $E_D = \tilde{E}_{D^\perp}$.

Para terminar, si $(v, \alpha) \in D$ y $(w, \beta) \in D^\top$, entonces $\beta v = \alpha w$, de donde $\omega_{D^\top}(w, v) = \omega_D(v, w)$. Del mismo modo, si $(v, \alpha) \in D$ y $(w, \beta) \in D^\perp$, entonces $\beta v = -\alpha w$, de donde $\omega_{D^\perp}(w, v) = -\omega_D(v, w)$. □

De este resultado podemos obtener otra fórmula para calcular \tilde{E}_D que resultará útil más adelante.

Corolario 4.3.

$$\tilde{E}_D = \pi_V(D^\perp) = \pi_V(D^\top). \quad (4.5)$$

Demostración. Del lema anterior se tiene $\tilde{E}_D = E_{D^\perp} = E_{D^\top}$; además $E_{D^\perp} = \pi_V(D^\perp)$ y $E_{D^\top} = \pi_V(D^\top)$, de donde se deduce la fórmula (4.5). □

Lema 4.4. *Dados $D \subset V \oplus V^*$ una estructura de Dirac generalizada y $W \subset V$ un subespacio, consideramos $D + W \subset V \oplus V^*$ dado por $D + W = \{(v, \alpha) \in V \oplus V^* \mid \exists w \in W, (v - w, \alpha) \in D\}$. Entonces el subespacio ortogonal W^D puede calcularse como*

$$W^D = ((D + W)^\flat(0))^\circ = \tilde{E}_{(D+W)} = \pi_V((D + W)^\perp).$$

Demostración. Por definición $D^b(W) = \{\alpha \in V^* \mid \exists w \in W, (w, \alpha) \in D\}$, y $(D + W)^b(0) = \{\alpha \in V^* \mid (0, \alpha) \in D + W\} = \{\alpha \in V^* \mid \exists w \in W, (-w, \alpha) \in D\}$. De esto se deduce que $D^b(W) = (D + W)^b(0)$ y por lo tanto $W^D = D^b(W)^\circ = ((D + W)^b(0))^\circ = \tilde{E}_{(D+W)} = \pi_V((D + W)^\perp)$; la última igualdad surge del corolario 4.3. □

Para finalizar, resumimos las definiciones y fórmulas presentadas en esta sección que se utilizarán luego.

Dados $D \subset V \oplus V^*$, $W \subset V (\subset V \oplus V^*)$:

$$E_D = \pi_V(D),$$

$$D^b(W) = \{\alpha \in V \mid \exists w \in W, (w, \alpha) \in D\},$$

$$\tilde{E}_D = D^b(0)^\circ = \pi_V(D^\perp) \quad (\perp = \text{ortogonal respecto a } \langle, \rangle_+),$$

$$W^D = D^b(W)^\circ = \pi_V((D + W)^\perp),$$

$$\alpha \in D^b(W) \iff \alpha \cdot (W^D) = \{0\}.$$

4.3.2. Estructuras de Dirac generalizadas sobre una variedad

En esta sección desarrollaremos una teoría de estructuras de Dirac generalizadas sobre variedades diferenciales. Todas las definiciones y resultados serán válidas ya sea considerando que tanto la variedad M como todos los objetos utilizados en las definiciones (como campos vectoriales, secciones de fibrados, etc.) son de clase C^∞ , o también considerando que son C^ω . En trabajos previos es habitual trabajar en la clase C^∞ , pero en nuestro caso es importante que los resultados valgan también para el caso C^ω para que se puedan coordinar con los capítulos siguientes de esta tesis. Desarrollar una teoría similar a esta para el caso C^k requeriría un trabajo un poco más cuidadoso porque no todas las operaciones que utilizamos conservarían el grado de regularidad.

En general no aclararemos cuándo los objetos considerados son C^∞ o C^ω , y la palabra “suave” se utilizará como sinónimo en cualquiera de los dos casos, salvo cuando resulte estrictamente necesario ya sea porque se trabaje con objetos no habituales o porque haya alguna diferencia entre un caso u el otro; ver definiciones en el Capítulo 2.

Definiremos a las *estructuras de Dirac generalizadas sobre una variedad* M como subfibrados de la suma de Whitney $TM \oplus T^*M$. En esta sección llamaremos π_{TM} y π_{T^*M} a las proyecciones canónicas de $TM \oplus T^*M$ sobre TM y T^*M , respectivamente; y llamaremos $\tau_{TM \oplus T^*M}$, τ_{TM} y τ_{T^*M} a las

proyecciones sobre M de $TM \oplus T^*M$, TM y T^*M , respectivamente. Se tiene el siguiente diagrama,

$$\begin{array}{ccccc}
 & & TM \oplus T^*M & & \\
 & \swarrow & \downarrow & \searrow & \\
 & \pi_{TM} & \tau_{TM \oplus T^*M} & \pi_{T^*M} & \\
 TM & & & & T^*M \\
 & \searrow & \downarrow & \swarrow & \\
 & \tau_{TM} & M & \tau_{T^*M} &
 \end{array}$$

Con estas proyecciones damos la definición de estructura de Dirac generalizada.

Definición 4.8. Dada una variedad M , una *estructura de Dirac generalizada sobre M* es un subfibrado D del fibrado $TM \oplus T^*M$.

$$\begin{array}{ccc}
 D & \xleftarrow{i} & TM \oplus T^*M \\
 & \searrow & \downarrow \tau_{TM \oplus T^*M} \\
 & & M \\
 \tau_{TM \oplus T^*M}|_D & \searrow &
 \end{array}$$

Es decir, se cumplen las siguientes dos propiedades:

- para cada $x \in M$, se tiene $D_x = (\tau_{TM \oplus T^*M}|_D)^{-1}(x)$ que es un subespacio de $T_xM \oplus T_x^*M$, o equivalentemente D_x es una estructura de Dirac generalizada sobre el espacio vectorial T_xM , y la dimensión $\text{rg}D = k$ es constante;
- D_x depende de manera suave de x . Esto es, existen trivializaciones locales, o sea, para cada $x_0 \in M$ existen un entorno U , campos vectoriales $v_i : U \rightarrow TM$ y 1-formas $\alpha_i : U \rightarrow T^*M$, $i = 1, \dots, k$ tales que para cada $x \in U$ el conjunto $\{(v_i(x), \alpha_i(x)) | i = 1, \dots, k\}$ es una base de D_x .

A continuación definiremos, para una estructura de Dirac generalizada sobre una variedad M , los objetos correspondientes a aquellos que definimos en la sección anterior para una estructura de Dirac generalizada sobre un espacio vectorial.

Definición 4.9. Dada una estructura de Dirac generalizada sobre M , definimos $E_D = \pi_{TM}(D) \subset TM$.

Dada una trivialización local de D , los campos vectoriales $v_i, i = 1, \dots, k$ generan localmente a E_D , pero en general no son linealmente independientes. Por esto E_D es una distribución, pero en general no tiene rango constante.

La definición de \tilde{E}_D no sigue de manera tan directa de la definición correspondiente para espacios vectoriales. Teniendo en cuenta que para el caso lineal se definió $\tilde{E}_D = D^b(0)^\circ$, sería natural llamar $\tilde{E}_D = (\pi_{T^*M}(\ker(\pi_{TM}|_D)))^\circ$, pero a partir de esta definición se presentaría el siguiente inconveniente: dado que la dimensión de $\text{Im}(\pi_{TM}|_D) = E_D$ es semicontinua inferiormente, entonces la dimensión de $\ker(\pi_{TM}|_D)$ es semicontinua superiormente, y por lo tanto no hay sistemas de generadores de $\ker(\pi_{TM}|_D)$.

Debido a esto resulta más conveniente definir \tilde{E}_D a través de D^\perp , guiándose por las fórmulas del lema 4.2.

Para poder definir D^\perp notar que el producto $\langle \cdot, \cdot \rangle_+$, definido en la sección anterior sobre $V \oplus V^*$, se extiende a un producto de secciones del fibrado $TM \oplus T^*M$ definido punto a punto.

Entonces, dada una estructura de Dirac generalizada D , consideramos $(v_i, \alpha_i), i = 1, \dots, k$, una trivialización local de D . Esta trivialización puede extenderse a una trivialización local de $TM \oplus T^*M$, dada por $(v_i, \alpha_i), i = 1, \dots, 2n$. Las ecuaciones $\langle (v_i, \alpha_i), (v_j^\perp, \alpha_j^\perp) \rangle_+ = \delta_i^j$ definen unívocamente los campos vectoriales $v_j^\perp : U \rightarrow TM$ y las 1-formas $\alpha_j^\perp \rightarrow T^*M$. Así, definimos D^\perp a la estructura de Dirac generalizada generada por $(v_j^\perp, \alpha_j^\perp), j = k + 1, \dots, 2n$. Además se define $\tilde{E}_D = E_{(D^\perp)}$.

Lema 4.5. *Dados D una estructura de Dirac generalizada sobre M , y un punto $x \in M$, se cumple que $(0, \alpha) \in D_x$ si y sólo si $(\alpha \cdot (\tilde{E}_D)_x) = \{0\}$.*

Demostración. si $(\alpha \cdot (\tilde{E}_D)_x) = \{0\}$, entonces para todo $(v, \beta) \in D_x^\perp$ se tiene $\langle (0, \alpha), (v, \beta) \rangle_+ = \alpha v + \beta 0 = 0$ y por lo tanto $(0, \alpha) \in (D_x^\perp)^\perp = D_x$.

Recíprocamente, si $(0, \alpha) \in D_x$ y $v \in (\tilde{E}_D)_x$, entonces existe $\beta \in T_x^*M$ tal que $(v, \beta) \in (D^\perp)_x$ y por lo tanto $\alpha \cdot v = -\beta 0 = 0$. \square

Nota. Este resultado, visto en cada fibra, se lee como $\tilde{E}_D = D^b(0)^\circ$.

Corolario 4.6. *Dada D una estructura de Dirac generalizada sobre una variedad M de dimensión n resulta que $\text{rg}E_D - \text{rg}\tilde{E}_D$ es una cantidad constante y $\text{rg}D = n + \text{rg}E_D - \text{rg}\tilde{E}_D$*

Demostración. Consideremos la fórmula $\text{rg}D = \text{rg}\text{Im}(\pi_{TM}) + \text{rg}\ker(\pi_{TM}) = \text{rg}E_D + \text{rg}D^b(0)$.

Del lema anterior resulta $\text{rg}D^b(0) = \text{rg}(\tilde{E}_D)^\circ = n - \text{rg}\tilde{E}_D$, luego reemplazando en la fórmula anterior se tiene $\text{rg}D = n + \text{rg}E_D - \text{dim}(\tilde{E}_D)_x$.

Finalmente, como $\text{rg}D$ es constante, resulta que $\text{rg}E_D - \text{rg}\tilde{E}_D$ es una cantidad constante.

□

Hasta ahora, para D una estructura de Dirac generalizada sobre una variedad, hemos definido E_D y \tilde{E}_D . El siguiente paso será definir ω_D y ver bajo qué condiciones la terna $(E_D, \tilde{E}_D, \omega_D)$ define a la estructura D .

Definición 4.10. Dada D una estructura de Dirac generalizada sobre M , sea $\omega_D : E_D \times_M \tilde{E}_D \rightarrow \mathbb{R}$ la función definida por $((x, v), (x, w)) \mapsto \alpha w$, donde $\alpha \in T_x^*M$ es tal que $(x, v, \alpha) \in D$.

El resultado no depende de la elección de α , ya que si $(x, v, \alpha') \in D$ entonces $(x, 0, \alpha - \alpha') \in D$ y por el lema 4.5 resulta que $\alpha w = \alpha' w$.

La función ω_D es suave (como función de x) en el sentido de la definición 2.13 del capítulo 2, considerando a la imagen como el fibrado trivial $M \times \mathbb{R}$.

En el caso en que E_D sea una distribución de rango constante la definición anterior es equivalente a pedir que ω_D sea suave al evaluarla en campos vectoriales que formen una base local.

Establecer si la terna $(E_D, \tilde{E}_D, \omega_D)$ define o no una estructura de Dirac generalizada no es trivial, como puede verse en el siguiente ejemplo.

Ejemplo 4.2. Sea $M = \mathbb{R}$. Sean $E = \tilde{E} \subset TM \simeq \mathbb{R}^2$ distribuciones generadas por el campo vectorial $x \mapsto (x, x)$. Sea ω definida por $\omega((x, v), (x, w)) = k(x)vw$ donde $k(x)$ es una función dada.

Con estas definiciones, tenemos entonces que E y \tilde{E} son distribuciones y ω es una función bilineal suave. Si existiera una estructura de Dirac generalizada D tal que $E = E_D$, $\tilde{E} = \tilde{E}_D$ y $\omega = \omega_D$, entonces D tendría rango 1 y estaría localmente generada por una sección de $TM \oplus T^*M$.

Supongamos que tal D existe y sea $(x, v(x), \alpha(x))$ la sección que genera a D en un entorno de $x = 0$. Como $v(x)$ genera a E_D resulta que $v(0) = 0$, y $\alpha(0) \neq 0$ puesto que D_0 tiene dimensión 1. Por otro lado, para $x \neq 0$ se tiene $\alpha(x)w = k(x)v(x)w$, para todo $w \in T_xM = \mathbb{R}$, de donde $k(x) = \alpha(x)/v(x)$. Como $\alpha(0) \neq 0$ y $v(0) = 0$ resulta que $k(x)$ diverge cuando $x \rightarrow 0$. Esto muestra que no todas las posibles elecciones de ω provienen de una estructura de Dirac generalizada.

Para analizar mejor qué es lo que puede fallar, consideremos el caso $k(x) = 1$. Para $x \neq 0$ debería ser $(x, v, \alpha) \in D$ si y sólo si $\alpha w = vw$, para todo w , esto es si y sólo si $\alpha = v$; por otro lado, para $x = 0$ debería ser $(0, v, \alpha) \in D$ si y sólo si $v = 0$ y $v0 = \alpha 0$ (la segunda ecuación es así porque $\tilde{E}_0 = \{0\}$). Por lo tanto debería ser $D = \{(x, v, \alpha) | v = \alpha, x \neq 0\} \cup \{(x, v, \alpha) | x = 0, v = 0\}$, pero esto no es un subfibrado de $TM \oplus T^*M$, y por lo tanto D no es una estructura de Dirac generalizada.

Respecto a bajo qué condiciones las ternas $(E_D, \tilde{E}_D, \omega_D)$ definen una estructura de Dirac generalizada se tienen los siguientes resultados.

Lema 4.7. *Dada una variedad M , dos distribuciones E y \tilde{E} tales que $\dim E_x - \dim \tilde{E}_x$ es constante, y una función bilineal $\omega : E \times_M \tilde{E} \rightarrow \mathbb{R}$, existe a lo sumo una estructura de Dirac generalizada D tal que $E = E_D$, $\tilde{E} = \tilde{E}_D$ y $\omega = \omega_D$.*

Demostración. En la sección anterior (lema 4.1) vimos que una estructura de Dirac generalizada D sobre un espacio vectorial está unívocamente determinada por E_D , \tilde{E}_D y ω_D . En este contexto, ese resultado implica que D_x está unívocamente determinada por $(E_D)_x$, $(\tilde{E}_D)_x$ y $\omega_D|_{(E_D)_x \times (\tilde{E}_D)_x}$.

Entonces la terna (E, \tilde{E}, ω) determina unívocamente a las fibras D_x . Por esto, si existe una estructura de Dirac generalizada tal que tal que $E = E_D$, $\tilde{E} = \tilde{E}_D$ y $\omega = \omega_D$, ésta es única. □

Nota. Esta demostración prueba la unicidad de D , pero no la existencia, porque la unión de las fibras D_x podría no ser un subfibrado de $TM \oplus T^*M$ como se vio en el ejemplo 4.2.

El siguiente es un resultado de existencia.

Lema 4.8. *Dadas una variedad M , dos distribuciones E y \tilde{E} de rango constante, y una función bilineal $\omega : E \times_M \tilde{E} \rightarrow \mathbb{R}$, existe una única estructura de Dirac generalizada D tal que $E = E_D$, $\tilde{E} = \tilde{E}_D$ y $\omega = \omega_D$.*

Demostración. Del lema 4.7 sabemos que, si existiera una estructura de Dirac generalizada como en el enunciado, debería ser

$$D = \{(x, v, \alpha) \in TM \oplus T^*M \mid (x, v) \in E \text{ y } \alpha w = \omega_x(v, w), \forall w \in (\tilde{E})_x\}.$$

Supongamos que D así definido es un subfibrado, vamos a demostrar que $E = E_D$, $\tilde{E} = \tilde{E}_D$ y $\omega = \omega_D$. Primero observar que $E_D = \pi_{TM}(D) = E$ por construcción. Como segundo paso, dado $w \in (\tilde{E})_x$, resulta que para todo $\alpha \in D_x^\flat(0)$ (es decir $(0, \alpha) \in D_x$), resulta $\alpha w = \omega_x(0, w) = 0$, y por el lema 4.5 resulta $w \in (\tilde{E}_D)_x = (D_x^\flat(0))^\circ$. Luego $\tilde{E} \subset \tilde{E}_D$. Por otro lado, el rango $\text{rg} D$ es igual a $n + \text{rg} E - \text{rg} \tilde{E}$ y $\dim(\tilde{E}_D)_x = n - \text{rg} D + \dim E_x$. De aquí resulta que \tilde{E}_D tiene rango constante igual al rango de \tilde{E} , y en consecuencia $\tilde{E} = \tilde{E}_D$. Finalmente $\omega_D = \omega$ por construcción.

Falta demostrar que D es efectivamente un subfibrado de $TM \oplus T^*M$. Para esto vamos a construir trivializaciones locales de D . Como E_D y \tilde{E}_D son distribuciones regulares tienen trivializaciones locales. Esto quiere decir que para todo $x_0 \in M$ existen un entorno U y campos vectoriales $v_i : U \rightarrow TM$, $i = 1, \dots, m$ y $\tilde{v}_j : U \rightarrow M$, $j = 1, \dots, \tilde{m}$ tales que $\{v_i(x), i = 1, \dots, m\}$ es una base de E_x y $\{\tilde{v}_j(x), j = 1, \dots, \tilde{m}\}$ es una base de \tilde{E}_x , para cada $x \in U$. Cada una de estas trivializaciones puede extenderse a trivializaciones de TM

obteniéndose $v_i : U \rightarrow TM$, $i = 1, \dots, n$ y $\tilde{v}_j : U \rightarrow M$, $j = 1, \dots, n$, ambas trivializaciones locales de TM .

Utilizando estas trivializaciones definiremos 1-formas α_i , $i = 1, \dots, n + m - \tilde{m}$, donde n es la dimensión de M , m es el rango de E y \tilde{m} es el rango de \tilde{E} . Primero definimos α_i , $i = 1, \dots, m$ mediante las ecuaciones

$$\alpha_i \cdot \tilde{v}_j = \begin{cases} \omega_D(v_i, \tilde{v}_j), & j = 1, \dots, \tilde{m} \\ 0, & j = \tilde{m} + 1, \dots, n. \end{cases}$$

Luego, si $\tilde{m} < n$, definimos α_{m+k} , $k = 1, \dots, n - \tilde{m}$ mediante las ecuaciones

$$\alpha_{m+k} \cdot \tilde{v}_j = \begin{cases} 1, & \tilde{m} + k = j \\ 0, & \tilde{m} + k \neq j \end{cases}$$

Con estas definiciones resulta que:

1. $(x, v_i(x), \alpha_i(x)) \in D_x$, para $i = 1, \dots, m$.
2. Las secciones $x \mapsto (x, v_i(x), \alpha_i(x))$, con $i = 1, \dots, m$, son linealmente independientes porque los v_i son linealmente independientes.
3. $(x, 0, \alpha_{m+k}(x)) \in D_x$, para $k = 1, \dots, n - \tilde{m}$, ya que $\alpha_{m+k} \cdot \tilde{v}_j = 0$, para $j \leq \tilde{m}$, luego $\alpha_{m+k} \cdot w = 0 = \omega_x(0, w)$, para todo $w \in \tilde{E}$.
4. Las secciones $x \mapsto (x, 0, \alpha_{m+k}(x)) \in D_x$, para $k = 1, \dots, n - \tilde{m}$, son linealmente independientes entre sí porque los α_{m+k} son linealmente independientes, y además son linealmente independientes de las $x \mapsto (x, v_i(x), \alpha_i(x))$, para $i = 1, \dots, m$, porque los vectores $v_i(x)$ son no nulos.

Así hemos construido $n + m - \tilde{m}$ secciones linealmente independientes, a saber: $x \mapsto (x, v_i(x), \alpha_i(x))$, para $i = 1, \dots, m$, y $x \mapsto (x, 0, \alpha_{m+k}(x)) \in D_x$, para $k = 1, \dots, n - \tilde{m}$. La imagen de estas secciones está contenida en D , y $\text{rg}D = n + m - \tilde{m}$, para todo x . Por lo tanto estas secciones son una trivialización local de D .

Como esto vale para todo $x_0 \in M$ resulta que D es un subfibrado de $TM \oplus T^*M$ con lo que se completa la demostración. □

Estructuras de Dirac generalizadas con núcleo de rango constante. Para llevar adelante el algoritmo gCAD, necesitaremos una condición de regularidad para las estructuras de Dirac generalizadas con las que trabajaremos.

Definición 4.11. Definimos el núcleo de D , $\ker D \subset TM$, como

$$\ker D = \pi_{TM}((TM \oplus \{0\}) \cap D).$$

El conjunto $\ker D$ es el conjunto de los vectores $(x, v) \in TM$ tales que $(x, v, 0)$ está en D . En general no es una distribución, ya que el rango de $\ker D$ es semicontinuo superiormente (en lugar de serlo inferiormente como ocurre con las distribuciones).

Definición 4.12. Una *estructura de Dirac generalizada con núcleo de rango constante* es una estructura de Dirac generalizada D tal que $\ker D$ es una distribución de rango constante.

Lema 4.9. *Dada una estructura de Dirac generalizada D tal que $\ker D$ tiene rango constante entonces:*

1. $F = \pi_{T^*M}(D)$ es un subfibrado de T^*M .
2. Existe una aplicación $\nu : F \rightarrow TM$ tal que se cumple $(v, \mu) \in D \iff v - \nu(\mu) \in \ker D$.
3. Dada una distribución E^C tal que $TM = E^C \oplus \ker D$ la aplicación ν del inciso anterior puede elegirse de modo que $\text{Im}\nu \subset E^C$.

Demostración. 1. Observar que $\ker D$ tiene rango constante si y sólo si $\ker(\pi_{T^*M}|_D)$ tiene rango constante, y luego $F = \text{Im}(\pi_{T^*M}|_D)$ tiene también rango constante y resulta, por lo tanto, un subfibrado.

2. Dado que $\ker(\pi_{T^*M}|_D)$ es un subfibrado de D , entonces existe un complemento D' , $D = D' \oplus \ker(\pi_{T^*M}|_D)$. Por construcción se cumple que $\ker(\pi_{T^*M}|_{D'}) = 0$, y por lo tanto $\pi_{T^*M}|_{D'}$ es invertible. Luego llamamos $\nu = \pi_{TM} \circ (\pi_{T^*M}|_{D'})^{-1}$. De este modo, para todo $\mu \in F$ se cumple que $(\nu(\mu), \mu) = (\pi_{T^*M}|_{D'})^{-1}(\mu) \in D'$, y todo elemento de D' se escribe de esta manera. Por otro lado, como $D = D' \oplus \ker(\pi_{T^*M}|_D)$, todo elemento de D se escribe de manera única como la suma de un elemento $(\nu(\mu), \mu) \in D'$ más un elemento $(v_0, 0) \in \ker(\pi_{T^*M}|_D)$. Dicho de otro modo, todo elemento de D se escribe de manera única como $(\nu(\mu) + v_0, \mu)$, de donde se deduce que $(v, \mu) \in D \iff v - \nu(\mu) \in \ker D$.

3. Para demostrar el último punto, alcanza con tomar $D' = D \cap E^C$ en la construcción del párrafo anterior.

□

En el siguiente teorema vamos a mostrar que la existencia de F y ν es una condición necesaria y suficiente para que la terna (E, \tilde{E}, ω) defina una estructura de Dirac generalizada con núcleo de rango constante.

Teorema 4.10. *Dadas dos distribuciones $E, \tilde{E} \subset TM$ y una aplicación bilineal $\omega : E \times \tilde{E} \rightarrow \mathbb{R}$, entonces existe una única estructura de Dirac generalizada D con núcleo de rango constante tal que $E_D = E$, $\tilde{E}_D = \tilde{E}$ y $\omega_D = \omega$ si y solo si se cumplen las siguientes condiciones:*

1. $\dim E_x - \dim \tilde{E}_x$ es una cantidad constante.
2. El conjunto $F = \{\mu \in T^*M \mid \exists v \in E, \omega(v, \cdot) = \mu|_{\tilde{E}}\}$ es un subfibrado de T^*M .
3. El conjunto $\ker \omega = \{v \in E \mid \forall w \in \tilde{E}, \omega(v, w) = 0\}$ es una distribución de rango constante
4. Dada una distribución E^C tal que $TM = E^C \oplus \ker \omega$, la función $\nu : F \rightarrow E^C$ dada por $\nu(\mu) = v \iff \omega(v, \cdot) = \mu|_{\tilde{E}}$ es una aplicación suave.

Demostración. Primero veamos que son condiciones necesarias.

El punto 1 debe satisfacerse por el corolario 4.6.

Por el lema 4.7, si existe una estructura de Dirac generalizada como en el enunciado, resulta $D = \{(x, v, \mu) \in TM \oplus T^*M \mid v \in E_x \text{ y } \omega(v, \cdot) = \mu|_{\tilde{E}}\}$. Suponiendo que se trata de una estructura de Dirac generalizada con núcleo de rango constante, el conjunto mencionado en el inciso 2 es $F = \pi_{T^*M}(D)$, y de acuerdo al lema 4.9 debe ser un subfibrado de T^*M .

Además $\ker D = \ker \omega$, por lo tanto si D tiene núcleo de rango constante esto implica que $\ker \omega$ es una distribución de rango constante.

Finalmente, la función ν coincide con la construida en el lema 4.9, ya que si $v \in E^C$ y $(x, v, \mu) \in D \implies \nu(\mu) = v$, y por lo tanto debe ser una aplicación suave.

Para la recíproca, definimos $D = \{(x, v, \mu) \in TM \oplus T^*M \mid \mu \in F, v - \nu(\mu) \in \ker(\omega)\}$.

Si llamamos $\{\mu_i\}$ a una trivialización local de F , y $\{v_j\}$ a una trivialización local de $\ker D$, entonces el conjunto $\{(\nu(\mu_i), \mu_i)\} \cup \{(v_j, 0)\}$ es una trivialización local de D , y por lo tanto D es una estructura de Dirac generalizada. Y tiene núcleo de rango constante ya que, por construcción, $\ker D = \ker \omega$.

Por construcción $E_D = \nu(F) + \ker \omega = E$.

Por otro lado, $(0, \mu) \in D \iff \nu(\mu) = 0 \iff \mu w = 0 \forall w \in \tilde{E}$. Entonces por el lema 4.5 resulta que $\tilde{E} = \tilde{E}_D$.

Finalmente, todo $v \in E_D$ se puede escribir de la forma $v = \nu(\mu) + v_0$ con $\mu \in F$ y $v_0 \in \ker \omega$, por lo tanto $\omega_D(\nu(\mu) + v_0, w) = \mu w = \omega(\nu(\mu) + v_0, w)$, de donde $\omega = \omega_D$.

La unicidad sigue del lema 4.7.

□

Nota. . En este punto es conveniente repetir lo que dijimos al principio de la sección: las estructuras de Dirac generalizadas que estamos estudiando pueden ser de clase C^∞ o C^ω según el grado de regularidad que se pida a los campos que las generan. En el teorema anterior, el mismo grado de regularidad debe pedirse a todos los objetos involucrados: E_D , \tilde{E}_D , F y ν .

Ejemplo 4.3 (continuación del ejemplo 4.2). En el ejemplo 4.2 hemos considerado $M = \mathbb{R}$, E y \tilde{E} generadas por el campo $x \mapsto (x, x)$ y $\omega((x, u), (x, v)) = k(x)uv$.

Las distribuciones E y \tilde{E} son suaves por construcción.

La condición que ω sea suave puede verificarse del siguiente modo: cualquier campo que genere a E tendrá la forma $(x, u(x))$ con $u(x)$ suave, $u(0) = 0$ y $u(x) \neq 0$ para $x \neq 0$. Algo similar puede decirse para \tilde{E} y un campo $(x, v(x))$. Entonces ω es suave si y sólo si existen u y v en estas condiciones que verifiquen que el producto $k(x)u(x)v(x)$ sea suave, esto es equivalente a que $k(x)$ sea suave para $x \neq 0$ (en el caso C^ω hay que pedir además que k no tenga una singularidad esencial en $x = 0$, en el caso C^∞ no hace falta pedir ninguna condición en $x = 0$).

Se tiene además que $\dim E_x = \dim \tilde{E}_x$, luego se verifica la condición 1.

Por otro lado, $(\ker \omega)_0 = \{0\}$ y para $x \neq 0$ $(\ker \omega)_x = \{0\}$ si y sólo si $k(x) \neq 0$. Luego, la condición 3 se cumple si y sólo si $k(x) \neq 0$ para $x \neq 0$. Suponiendo que esto se cumple, tenemos $F = T^*M$, con lo que la condición 2 se cumple automáticamente.

Finalmente, para la condición 4, ν queda definida, para $x \neq 0$, a partir de $\nu(x, k(x)u) = (x, u)$, es decir $\nu(x, \mu) = (x, \mu/k(x))$, y para $x = 0$ por $\nu(x, \mu) = (x, 0)$. La terna (E, \tilde{E}, ω) definirá una estructura de Dirac generalizada (con núcleo de rango constante) si y sólo si la función ν así construida es suave (i.e. C^∞ o C^ω según corresponda).

Distribución ortogonal. Para completar el proceso de extender las definiciones que dimos en el caso lineal al caso de una estructura de Dirac generalizada sobre una variedad M , sólo nos falta definir el ortogonal respecto a D .

Definición 4.13. Dada una distribución de rango constante $W \subset TM$ tal que se cumple que $W \cap \ker D$ también tiene rango constante, definimos

$$W^D = \pi_{TM} (((W \oplus \{0\}) + D)^\perp).$$

Nota. Dado que $(W \oplus \{0\}) \cap D = (W \cap \ker D) \oplus \{0\}$, entonces la condición que $W \cap \ker D$ tiene rango constante garantiza que $(W \oplus \{0\}) + D$ también tiene rango constante y es por lo tanto una estructura de Dirac generalizada.

Esto asegura que se puede calcular $(\cdot)^\perp$ en la definición anterior, y que W^D será una distribución.

Lema 4.11. *Dadas una estructura de Dirac generalizada D sobre M , una 1-forma μ y una distribución de rango constante W , tal que $W \cap \ker D$ tiene rango constante, resulta que, para cada $x \in M$, existe $v \in W_x$ tal que $(v, \mu_x) \in D_x$ si y sólo si*

$$\langle \mu_x, W_x^D \rangle = 0. \quad (4.6)$$

Demostración. Por el lema 4.4 se tiene que el W^\perp de la definición 4.13 es equivalente a aplicar $(\cdot)^\perp$ en cada fibra, y por lo tanto vale que $W^D = (D^b(W))^\circ$.

En consecuencia, $\mu_x \in D^b(W)$ si y solo si

$$\langle \mu_x, W_x^D \rangle = 0.$$

□

El hecho que W^D sea una distribución implica que el conjunto

$$S = \{x \in M \mid \langle \mu_x, W_x^D \rangle = 0\}$$

es un conjunto cerrado definido por ecuaciones. En efecto, si $\{v_i\}$ es un conjunto de generadores de W^D , el conjunto S queda definido por las ecuaciones $\langle \mu_x, v_i(x) \rangle = 0$.

Este hecho va a ser importante para poder generalizar el algoritmo CAD.

4.3.3. Sistemas de Dirac generalizados y gCAD

Definición 4.14. Un sistema de Dirac generalizado sobre una variedad M es una ecuación de la forma

$$\dot{x} \oplus \mu \in D \quad (4.7)$$

donde μ es una 1-forma y D es una estructura de Dirac generalizada sobre M .

Ejemplo 4.4. Los sistemas dinámicos de Dirac son casos particulares de sistemas de Dirac generalizados, ya que las estructuras de Dirac son estructuras de Dirac generalizadas D que además verifican $D^\perp = D$.

Ejemplo 4.5. Dadas una variedad Q , dos distribuciones de rango constante Δ y Δ_V y un Lagrangiano $L : TQ \rightarrow \mathbb{R}$, las ecuaciones del principio de D'Alembert generalizado (4.1) son equivalentes a un sistema de Dirac generalizado sobre $M = TQ \oplus T^*Q$. Para verlo llamamos $\tau_M : M \rightarrow Q$, $\pi_{TQ} : M \rightarrow TQ$ y $\pi_{T^*Q} : M \rightarrow T^*Q$ a las proyecciones canónicas, y sean

$E = (T\tau_M)^{-1}(\Delta)$, $\tilde{E} = (T\tau_M)^{-1}(\Delta_V)$ y $\omega = (\pi_{T^*Q})^*\Omega$, donde Ω es la 2-forma canónica en T^*Q . Entonces definimos $D = D(E, \tilde{E}, \omega)$ y $\mathcal{E} : M \rightarrow \mathbb{R}$, $\mathcal{E}(x) = \langle \pi_{T^*Q}(x), \pi_{TQ}(x) \rangle - L(\pi_{TQ}(x))$.

Con estas definiciones el sistema de Dirac generalizado

$$\dot{x} \oplus d\mathcal{E} \in D \quad (4.8)$$

es equivalente a las ecuaciones del principio de D'Alembert generalizado.

En efecto, si en M tomamos coordenadas (q, v, p) la ecuación (4.8) toma la forma

$$\begin{cases} v \in \Delta \\ -\dot{p} + \frac{\partial L}{\partial q}(q, v) \in \Delta_V^\circ \\ -p + \frac{\partial L}{\partial v}(q, v) = 0 \end{cases}$$

Ejemplo 4.6. Dada una variedad M y una forma bilineal (no necesariamente antisimétrica) ω , entonces $D = \{(x, \dot{x}, \mu) \in TM \oplus T^*M \mid \mu = \omega(\dot{x}, \cdot)\}$ es una estructura de Dirac generalizada.

Los últimos dos ejemplos muestran cómo las estructuras de Dirac generalizadas permiten desarrollar una teoría común para los problemas presentados en las secciones 4.1 y 4.2.

Algoritmo de Ligaduras de Dirac generalizado (gCAD). Analizaremos ahora cómo resulta el algoritmo de ligaduras para una ecuación de la forma (4.7).

Cada paso del algoritmo debe tomar la siguiente forma: si M_k es la subvariedad obtenida en el paso k , entonces M_{k+1} deberá estar formado por aquellos puntos $x \in M_k$ tales que exista un $v_x \in T_x M_k$ con $(v_x, \mu_x) \in D_x$.

A partir del lema 4.11 se tiene una ecuación para resolver el problema planteado en el párrafo anterior, pero reemplazando TM_k por una distribución de rango constante W que además verifique que $W \cap \ker D$ tenga rango constante.

Ahora bien, TM_k siempre puede extenderse a una distribución de rango constante definida en un entorno de M_k (ver corolario 2.6). Por lo tanto tiene sentido hablar de $(TM_k)^D$ siempre que se cumpla que $TM_k \cap \ker D$ tenga rango constante.

Entonces el lema 4.11 indica que

$$M_{k+1} = \{x \in M_k \mid \langle \mu_x, (T_x M_k)^D \rangle = 0\}. \quad (4.9)$$

Notar que esta definición coincide con (3.19), con la salvedad que ahora D es una estructura de Dirac generalizada.

Entonces el algoritmo gCAD queda especificado así:

1. $M_0 = M$.
2. Si $TM_k \cap \ker D$ no tiene rango constante el algoritmo no se puede continuar.
3. Si $TM_k \cap \ker D$ tiene rango constante se calcula

$$M_{k+1} = \{x \in M_k \mid \langle \mu_x, (T_x M_k)^D \rangle = 0\}.$$

4. Si $M_{k+1} = M_k$, el algoritmo termina.
5. Si M_{k+1} es una subvariedad propia de M_k , se vuelve al punto 2 para el siguiente valor de k .

Si el algoritmo termina, o sea si en cada paso $TM_k \cap \ker D$ tiene rango constante y M_k es una subvariedad, llamamos M_c a la subvariedad que se obtiene en la última iteración.

Luego, la subvariedad M_c verifica:

- Para cada $x \in M_c$ existe $v \in T_x M_c$ tal que $(v_x, \mu_x) \in D_x$.
- $TM_k \cap \ker D$ tiene rango constante.

Vamos a demostrar el siguiente teorema:

Teorema 4.12. *Dado un sistema de la forma (4.7) y un punto $x_0 \in M$, suponemos que el algoritmo gCAD termina, entonces existe una curva solución que pasa por x_0 si y sólo si $x_0 \in M_c$*

Para demostrar este teorema utilizaremos los siguientes lemas:

Lema 4.13. *Dado un sistema de la forma (4.7), suponemos que el algoritmo gCAD termina. Luego, la imagen de toda curva solución $x(t)$ está contenida en M_c .*

Demostración. Demostraremos por inducción sobre k , que la imagen de $x(t)$ está contenida en M_k , $\forall k$. Supongamos que la imagen de $x(t)$ está contenida en M_k , para algún k . Para cada t , se cumple por un lado que $(x(t), \dot{x}(t)) \in TM_k$ y, por el otro, que $(x(t), \dot{x}(t), \mu(x(t))) \in D$. Teniendo en cuenta ambas condiciones concluimos que $x(t) \in M_{k+1}$. \square

Lema 4.14. *Dado un sistema de la forma (4.7), suponemos que el algoritmo gCAD termina. Luego, la aplicación ν del lema 4.9 puede elegirse de modo que $\nu(\mu_x) \in T_x M_c$ para todo $x \in M_c$.*

Demostración. Sea ν_1 una aplicación como la del lema 4.9.

Sabemos que TM_c puede extenderse a una distribución de rango constante en un entorno de M_c , en los párrafos siguientes nos referiremos a dicha extensión usando el símbolo \tilde{TM}_c (ver comentario previo a la ecuación (4.9))

Como $\ker D$ y $\ker D \cap \tilde{TM}_c$ tienen rango constante, entonces puede elegirse K' un subfibrado de $\ker D$ complementario de $\ker D \cap \tilde{TM}_c$ (es decir, $\ker D = (\ker D \cap \tilde{TM}_c) \oplus K'$).

Como $\tilde{TM}_c \cap K'$ es la sección nula, entonces puedo extender \tilde{TM}_c a un fibrado complementario de K' , al que llamaremos T' . Sea $P_{K'}$ la proyección sobre T' , paralela a K' .

Definimos $\nu = P_{K'} \circ \nu_1$. Para cada $\mu' \in \pi_2(D)$ se cumple que $\nu(\mu') - \nu_1(\mu) \in K' \subset \ker D$, por lo tanto ν es una aplicación que está en los términos del lema 4.9.

Ahora demostraremos que $\nu(\mu_x) \in T_x M_c$ para todo $x \in M_c$. Como M_c es la variedad final del algoritmo gCAD, sabemos que para cada $x \in M_c$ existe $v_x \in T_x M_c$ tal que $(x, v_x, \mu) \in D_x$, entonces $v_x - \nu(\mu_x) \in \ker D_x$ por el lema 4.9. Además $\nu(\mu_x) \in T'_x$ por construcción, y $v_x \in T_x M_c \subset T'_x$, de donde $v_x - \nu(\mu_x) \in \ker D_x \cap T'_x = \ker D_x \cap T_x M_c$. En particular esto implica que $\nu(\mu_x) \in T_x M_c$

□

Demostración del teorema 4.12. Si hay una curva solución que pase por x_0 , la imagen de esa curva está contenida en $M_0 = M$, y por lema 4.13 resulta que su imagen (y en particular x_0) está contenida en M_c .

Para la recíproca, por los lemas 4.9 y 4.14 sabemos que existe una aplicación $\nu : \pi_2(D) \rightarrow TM$ tal que

$$(\nu(\mu_x), \mu_x) \in D \text{ para todo } x \quad (4.10)$$

y

$$\nu(\mu_x) \in T_x M_c \text{ para todo } x \in M_c. \quad (4.11)$$

Por (4.11) resulta que

$$\dot{x} = \nu(\mu_x) \quad (4.12)$$

es una ecuación diferencial ordinaria en M_c , y por lo tanto existen curvas solución de (4.12) por cada punto de M_c .

Finalmente, por (4.10) resulta que las curvas solución de (4.12) son curvas solución del sistema (4.7).

□

4.4. Ejemplos

Los siguientes ejemplos corresponden a sistemas no conservativos y queremos mostrar que pueden ser tratados como sistemas dinámicos de Dirac generalizados, para lo cual hallaremos en cada caso la estructura de Dirac generalizada correspondiente.

Ejemplo 4.7 (Circuitos RLC). Un circuito RLC es un circuito eléctrico formado por capacitores, inductores y resistencias.

La ecuación que rige la dinámica de este tipo de circuitos es similar a la de los circuitos LC (ver ejemplo 3.4), agregando un término adicional correspondiente a la resistencia. Las ecuaciones correspondientes a un circuito RLC se obtienen a partir de las ecuaciones de Lagrange–D’Alembert para un circuito LC (3.21) agregando el término correspondiente a la resistencia:

$$\begin{cases} \dot{q} \in \Delta \\ \sum_i \frac{d}{dt}(L_i \dot{q}_{L_i}) + R_i \dot{q}_{R_i} + q_{C_i}/C_i \in \Delta^\circ. \end{cases} \quad (4.13)$$

Por tratarse de un sistema muy similar al del ejemplo 3.4 pero con un término que implica disipación de energía, podemos usar las ideas de la sección 4.2 y construir una estructura similar a la del ejemplo 3.4. Para esto, utilizaremos el lema 4.8, definiendo dos distribuciones de rango constante E y \tilde{E} y una forma bilineal ω ; la parte simétrica de la forma bilineal ω corresponderá a la resistencia.

Dados Q el espacio de cargas y Δ el subespacio de TQ correspondiente a la ley de Kirchoff para las corrientes, tomamos $M = TQ \oplus T^*Q$ y consideramos en M coordenadas (q, v, p) . Sean $E = \tilde{E} = (T\tau_M)^{-1}(\Delta)$, donde $\tau_M : M \rightarrow Q$ es la proyección, y $\omega(u_1, u_2) = \Omega(T\pi_{T^*Q}(u_1), T\pi_{T^*Q}(u_2)) - (T\tau_M(u_1))^\top R T\tau_M(u_2)$, donde Ω es la 2-forma canónica en T^*Q y R es la matriz diagonal con los valores de las resistencias de cada rama (o cero si no hubiera resistencia). Así, el primer término de ω es el pull-back de la 2-forma canónica de T^*Q y el segundo término es el pullback de un producto simétrico en Q pesado por los valores de las resistencias. En coordenadas, si tomamos $u = (u_q, u_v, u_p) \in TM$ y $w = (w_q, w_v, w_p) \in T^*M$, resulta $\omega(u, w) = \sum_i (u_q)_i (w_p)_i - (u_p)_i (w_q)_i - R_i (u_q)_i (w_q)_i$.

La terna (E, \tilde{E}, ω) define una estructura de Dirac generalizada D porque E y \tilde{E} tienen rango constante (lema 4.8).

Al igual que en el ejemplo 3.4 la función de energía será,

$$\mathcal{E} = p \cdot v - \sum_i \frac{L_i}{2} (v_i)^2 - \frac{1}{2C_i} (q_i)^2.$$

conviniendo que $1/C_i = 0$ en las ramas donde no haya capacitores.

En estas condiciones, el sistema de Dirac generalizado

$$\dot{x} \oplus d\mathcal{E} \in D \quad (4.14)$$

es equivalente al sistema (4.13). En efecto, $x(t) = (q(t), i(t), p(t))$ es solución de (4.14) si y sólo si $\dot{x}(t) \in E_D = E$ y $\omega_D(\dot{x}, \cdot) = \omega(\dot{x}, \cdot) = d\mathcal{E}|_{\tilde{E}}$. En primer lugar tenemos que $\dot{x}(t) \in E_D$ si y sólo si $\dot{q}(t) \in \Delta$, que es la primera ecuación en (4.13). Además,

$$\begin{aligned} \omega_D(\dot{x}, \cdot) &= \sum_j \dot{q}_j dp_j - \dot{p}_j dq_j - R_j \dot{q}_j dq_j, \\ d\mathcal{E} &= \sum_j \frac{1}{C_j} q_j dq_j + (p_j - L_j v_j) dv_j + v_j dp_j \end{aligned}$$

y como $\tilde{E}^\circ = \Delta^\circ \oplus \{0\} \oplus \{0\}$, se tiene que $\omega_D(\dot{x}, \cdot) = d\mathcal{E}|_{\tilde{E}}$ si y sólo si

$$\begin{cases} \dot{p}_j(t) + R_j \dot{q}_j(t) + \frac{1}{C_j} q_j(t) \in \Delta^\circ \\ p_j(t) - L_j v_j(t) = 0 \\ v_j(t) - \dot{q}_j(t) = 0; \end{cases}$$

es sencillo verificar que estas ecuaciones son equivalentes a la segunda ecuación en (4.13).

En este caso, M es un espacio vectorial y D es constante (no depende de $x \in M$), luego $\ker D$ es también constante y entonces también lo es su dimensión. Además, como todas las ecuaciones son lineales, los conjuntos M_k que se obtienen a partir del algoritmo gCAD resultan subespacios de M , y por lo tanto también serán constantes los espacios $TM_k \cap \ker D$, para cada k . Esto implica que el algoritmo gCAD siempre puede aplicarse.

Ejemplo 4.8 (Sistemas mecánicos con fricción). Las ecuaciones del principio de D'Alembert generalizado (4.1) también pueden escribirse como un sistema de Dirac generalizado.

Supongamos dados Q la variedad de configuraciones, $\Delta_K \subset TQ$ la distribución de las velocidades permitidas, $\Delta_V \subset TQ$ la distribución de las variaciones y $L : TQ \rightarrow \mathbb{R}$ el Lagrangiano. Consideramos $M = TQ \oplus T^*Q$, y sean τ_M , π_{TQ} y π_{T^*Q} las proyecciones canónicas de M en Q , TQ y T^*Q , respectivamente.

Para definir la estructura de Dirac generalizada D sobre M tomamos $E = (T\tau_M)^{-1}(\Delta_K)$, $\tilde{E} = (T\tau_M)^{-1}(\Delta_V)$ y $\omega = (\pi_{T^*Q})^*\Omega$, con Ω la 2-forma canónica en T^*Q . Si Δ_K y Δ_V tienen rango constante entonces E y \tilde{E} también

tienen rango constante, y por lo tanto la terna (E, \tilde{E}, ω) define la estructura de Dirac generalizada D .

La función de energía queda definida como $\mathcal{E} : M \rightarrow \mathbb{R}$, siendo $\mathcal{E}(x) = \pi_{T^*Q}(x) \cdot \pi_{TQ}(x) - L(\pi_{TQ}(x))$.

Si en M establecemos coordenadas $x = (q, v, p)$, entonces se tiene $D = \{(q, v, p, \dot{q}, \dot{v}, \dot{p}, \alpha, \gamma, \beta) \in TM \oplus T^*M \mid \dot{q} \in \Delta_K, \alpha + \dot{p} \in (\Delta_V)^\circ, \beta = \dot{q}, \gamma = 0\}$, $\mathcal{E} = p \cdot v - L(q, v)$ y $d\mathcal{E} = -\frac{\partial}{\partial q}L(q, v)dq + (p - \frac{\partial}{\partial q}L(q, v))dv + vdp$.

Luego, el sistema de Dirac generalizado

$$\dot{x} \oplus d\mathcal{E} \in D$$

es equivalente al sistema de ecuaciones

$$\begin{cases} \dot{q} \in \Delta_K \\ \dot{p} - \frac{\partial L}{\partial q}(q, v) \in (\Delta_V)^\circ \\ v = \dot{q} \\ p - \frac{\partial L}{\partial v}(q, v) = 0 \end{cases}$$

que luego de realizar las sustituciones correspondientes resulta equivalente al sistema (4.1)

$$\begin{cases} \dot{q} \in \Delta_K \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) - \frac{\partial L}{\partial q}(q, \dot{q}) \in (\Delta_V)^\circ. \end{cases}$$

En cuanto a la aplicación del algoritmo gCAD, el núcleo $\ker D$ está formado por los elementos $(q, v, p, \dot{q}, \dot{v}, \dot{p}) \in E$ tales que $\dot{p} \in \Delta_V^\circ$. Las distribuciones Δ_K y Δ_V tienen ambas rango constante, entonces $\ker D$ tiene rango constante. Debido a esto, al menos la primera iteración del algoritmo puede llevarse a cabo. La continuidad del algoritmo hasta la variedad final dependerá de la geometría de M y del Lagrangiano L .

Ejemplo 4.9 (Dinámica de una rueda neumática). En este ejemplo estudiaremos la dinámica de una rueda que se mueve sin deslizar sobre un plano y que conserva la posición vertical. Supondremos que el material que forma la rueda se puede deformar levemente en el punto de apoyo debido a las fuerzas que actúan sobre la rueda, de manera que la dirección de avance no coincida con la orientación de la rueda.

Las variables que caracterizan el estado del sistema serán: $x = (x_1, x_2) \in \mathbb{R}^2$ punto del plano donde está apoyada la rueda, $\psi \in S^1$ ángulo de rotación de la rueda sobre su propio eje, $\theta \in S^1$ ángulo hacia donde está orientada la rueda, $\varepsilon \in (-h, h)$ deformación de la rueda en su punto de apoyo, de manera que la dirección real de avance de la rueda es la correspondiente al ángulo $\theta + \varepsilon$.

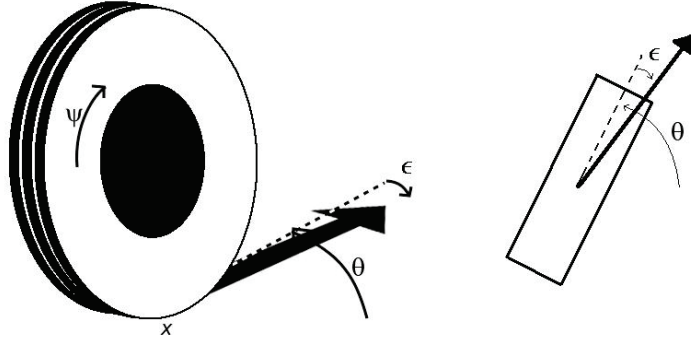


Figura 4.1: Variables de estado del sistema.

En [31] se presenta una generalización del principio de D'Alembert para vínculos cinemáticos de mayor orden, trabajando en fibrados de orden superior de tipo $T^{(k)}Q$. En ese trabajo se estudia el ejemplo de la rueda en base a la teoría de [39], proponiendo los siguientes vínculos cinemáticos

$$\begin{cases} \dot{x}_1 = \dot{\psi} \cos(\theta - \varepsilon) \\ \dot{x}_2 = \dot{\psi} \sin(\theta - \varepsilon) \\ -\ddot{\psi} \tan \varepsilon + \dot{\psi}(\dot{\theta} - \dot{\varepsilon}) = \frac{\dot{\psi}}{|\dot{\psi}|} \frac{a}{M} \tan \varepsilon \end{cases}$$

y los vínculos variacionales

$$\begin{cases} \delta\psi \cos \theta - \delta x_1 = 0 \\ \delta\psi \sin \theta - \delta x_2 = 0 \\ \delta\theta - \delta\varepsilon = 0, \end{cases}$$

con el Lagrangiano

$$L(\psi, x, \theta, \varepsilon) = \frac{1}{2}I\dot{\psi}^2 + \frac{1}{2}J\dot{\theta}^2 + \frac{1}{2}M\|\dot{x}\|^2 - \frac{1}{2}K\varepsilon^2,$$

donde I es el momento angular de la rueda sobre su eje, J es el momento angular de la rueda sobre el eje vertical que pasa por el punto de apoyo, M es la masa de la rueda, K es la constante de elasticidad del material que forma la rueda y a es una constante física.

Considerando $M = TQ \oplus T^*Q$ como en el ejemplo anterior, podemos incluir los vínculos de segundo orden (como los que aparecen en este ejemplo) en los sistemas de Dirac. Para ello definiremos una estructura de Dirac $D \subset TM \oplus T^*M$. En M tomamos coordenadas $(q, v, p) = (\psi, x, \theta, \varepsilon, v_\psi, v_x, v_\theta, v_\varepsilon, p_\psi, p_x, p_\theta, p_\varepsilon)$, y a las coordenadas correspondientes en $TM \oplus T^*M$ las

llamaremos $(\psi, x, \theta, \varepsilon, v_\psi, v_x, v_\theta, v_\varepsilon, p_\psi, p_x, p_\theta, p_\varepsilon, \dot{\psi}, \dot{x}, \dot{\theta}, \dot{\varepsilon}, \dot{v}_\psi, \dot{v}_x, \dot{v}_\theta, \dot{v}_\varepsilon, \dot{p}_\psi, \dot{p}_x, \dot{p}_\theta, \dot{p}_\varepsilon, \alpha_\psi, \alpha_x, \alpha_\theta, \alpha_\varepsilon, \gamma_\psi, \gamma_x, \gamma_\theta, \gamma_\varepsilon, \beta_\psi, \beta_x, \beta_\theta, \beta_\varepsilon)$, y entonces definimos a la estructura de Dirac D mediante las ecuaciones

$$\begin{cases} \dot{x}_1 = \dot{\psi} \cos(\theta - \varepsilon) \\ \dot{x}_2 = \dot{\psi} \sin(\theta - \varepsilon) \\ -\dot{v}_\psi \tan \varepsilon + \dot{\psi}(v_\theta - v_\varepsilon) = \frac{\dot{\psi}}{|v_\psi|} \frac{a}{M} \tan \varepsilon \\ \gamma = 0 \\ \beta = \dot{q} \\ \alpha_\theta + \dot{p}_\theta + \alpha_\varepsilon + \dot{p}_\varepsilon = 0 \\ \alpha_\psi + \dot{p}_\psi + (\alpha_{x_1} + \dot{p}_{x_1}) \cos \theta + (\alpha_{x_2} + \dot{p}_{x_2}) \sin \theta = 0 \end{cases}$$

Estas ecuaciones definen una estructura de Dirac generalizada para $\varepsilon \neq 0$ y $v_\psi \neq 0$.

La función de energía será

$$\mathcal{E} = pv - \frac{1}{2}Iv_\psi^2 - \frac{1}{2}Jv_\theta^2 - \frac{1}{2}M\|v_x\|^2 + \frac{1}{2}K\varepsilon^2.$$

Entonces el sistema de Dirac generalizado

$$\dot{x} \oplus d\mathcal{E} \in D,$$

describe la dinámica de la rueda siempre que esta se mantenga en movimiento ($v_\psi \neq 0$) y el movimiento no sea en línea recta ($\varepsilon \neq 0$).

Al aplicar el algoritmo gCAD a este sistema resulta que después de la primera iteración obtenemos M_1 definido por las ecuaciones

$$\begin{cases} p_\psi = Iv_\psi \\ p_x = Mv_x \\ p_\theta = Jv_\theta \\ p_\varepsilon = 0 \\ v_{x_1} = v_\psi \cos(\theta - \varepsilon) \\ v_{x_2} = v_\psi \sin(\theta - \varepsilon) \end{cases}$$

Para realizar el paso siguiente del algoritmo calculamos $TM_1 \cap E_D$, que

está definido por las ecuaciones

$$\begin{cases} \dot{x}_1 = \dot{\psi} \cos(\theta - \varepsilon) \\ \dot{x}_2 = \dot{\psi} \sin(\theta - \varepsilon) \\ -\dot{v}_\psi \tan \varepsilon + \dot{\psi}(v_\theta - v_\varepsilon) = \frac{\dot{\psi}}{|v_\psi|} \frac{a}{M} \tan \varepsilon \\ \dot{p}_\psi = I \dot{v}_\psi \\ \dot{p}_x = M \dot{v}_x \\ \dot{p}_\theta = J \dot{v}_\theta \\ \dot{p}_\varepsilon = 0 \\ \dot{v}_{x_1} = \dot{v}_\psi \cos(\theta - \varepsilon) - v_\psi(\dot{\theta} - \dot{\varepsilon}) \sin(\theta - \varepsilon) \\ \dot{v}_{x_2} = \dot{v}_\psi \sin(\theta - \varepsilon) + v_\psi(\dot{\theta} - \dot{\varepsilon}) \cos(\theta - \varepsilon) \end{cases}$$

Las tres ecuaciones que involucran a \dot{v} (y no a \dot{p}) pueden reescribirse como

$$\begin{aligned} \dot{v}_\psi &= \dot{\psi} \left(\frac{v_\theta - v_\varepsilon}{\tan \varepsilon} - \frac{a}{M} \frac{1}{|v_\psi|} \right) \\ \dot{v}_{x_1} &= \dot{\psi} \left(\frac{v_\theta - v_\varepsilon}{\tan \varepsilon} - \frac{a}{M} \frac{1}{|v_\psi|} \right) \cos(\theta - \varepsilon) - v_\psi(\dot{\theta} - \dot{\varepsilon}) \sin(\theta - \varepsilon) \\ \dot{v}_{x_2} &= \dot{\psi} \left(\frac{v_\theta - v_\varepsilon}{\tan \varepsilon} - \frac{a}{M} \frac{1}{|v_\psi|} \right) \sin(\theta - \varepsilon) + v_\psi(\dot{\theta} - \dot{\varepsilon}) \cos(\theta - \varepsilon) \end{aligned}$$

y $D^b(TM_1)$ queda definido por las ecuaciones

$$\begin{cases} \gamma = 0 \\ \beta_{x_1} = \beta_\psi \cos(\theta - \varepsilon) \\ \beta_{x_2} = \beta_\psi \sin(\theta - \varepsilon) \\ \alpha_\psi + \alpha_{x_1} \cos \theta + \alpha_{x_2} \sin \theta = -\beta_\psi \left(\frac{v_\theta - v_\varepsilon}{\tan \varepsilon} - \frac{a}{M} \frac{1}{|v_\psi|} \right) (I + M \cos(\varepsilon)) \\ \quad \quad \quad - M v_\psi (\beta_\theta - \beta_\varepsilon) \sin(\varepsilon) \end{cases}$$

(en la última ecuación aplicamos las equivalencias $\cos(\theta) \cos(\theta - \varepsilon) + \sin(\theta) \sin(\theta - \varepsilon) = \cos(\varepsilon)$ y $-\cos(\theta) \sin(\theta - \varepsilon) + \sin(\theta) \cos(\theta - \varepsilon) = \sin(\varepsilon)$).

Entonces M_2 está formado por aquellos puntos de M_1 que, además, satisfacen la ecuación

$$v_\psi(v_\theta - v_\varepsilon) \left(\frac{I + M \cos(\varepsilon)}{\tan(\varepsilon)} + M \sin(\varepsilon) \right) = \frac{a}{M} \frac{v_\psi}{|v_\psi|} (1 + \cos(\varepsilon)).$$

A esta ecuación, por comodidad, la escribiremos como

$$v_\theta - v_\varepsilon = g(v_\psi, \varepsilon).$$

Entonces TM_2 queda definido por las mismas ecuaciones que TM_1 , más la ecuación adicional

$$\dot{v}_\theta - \dot{v}_\varepsilon = \dot{v}_\psi g_1(v_\psi, \varepsilon) + \dot{v}_\varepsilon g_2(v_\psi, \varepsilon),$$

donde g_1 y g_2 son las derivadas parciales de g respecto a v_ψ y ε , respectivamente.

La variable \dot{v}_ε puede tomar cualquier valor, por lo tanto esta última ecuación no implica ninguna restricción nueva sobre $D^b(TM_2)$, y entonces $D^b(TM_2) = D^b(TM_1)$ y $M_3 = M_2$. Luego, el algoritmo termina en el paso 2.

Capítulo 5

Algoritmos de restricciones

En el Capítulo 3 hemos introducido el algoritmo de Gotay-Nester para resolver una EDI del tipo (3.10). La forma presimpléctica ω juega un rol fundamental en la descripción de este algoritmo y en hacerlo de fácil aplicación en distintos ejemplos, especialmente en casos infinito-dimensionales. Sin embargo, algunos aspectos básicos no están relacionados necesariamente a la estructura presimpléctica y entonces el algoritmo puede aplicarse a EDIs más generales, por ejemplo a las cuasilineales, ya definidas en el Capítulo 2, que son el tipo de ecuaciones con que el trabajaremos de aquí en más.

En adelante cuando digamos solución de una EDICL nos estaremos refiriendo a curva solución como fue definida en el Capítulo 2.

Usando la notación del Capítulo 2, consideramos una EDICL (2.1) y una condición inicial $x_0 \in M$. Con el objetivo de establecer existencia de solución presentamos dos *algoritmos de restricciones* ya conocidos, que llamaremos, por simplicidad, *regular* y *singular*, aunque esta no sea una nomenclatura estándar en la literatura. En el caso regular [12], se trabaja en el contexto \mathcal{C}^∞ , mientras que en el caso singular, desarrollado en [2], se trabaja en el contexto analítico. Daremos una breve descripción de ambos algoritmos, buscando enfatizar bajo qué hipótesis pueden aplicarse y en qué casos se resuelve el problema de existencia de solución.

Por último, analizaremos para qué condiciones iniciales el problema de existencia de solución no puede resolverse aplicando estos algoritmos. Para tratar estos casos serán necesarios nuevos resultados que se establecerán en el Capítulo 6, que contiene los principales resultados originales de esta tesis.

5.1. Algoritmo de restricciones regular

Para resolver una EDICL del tipo (2.1) se introduce un algoritmo que llamaremos *algoritmo de restricciones regular* [12]. Sea $x(t)$ una curva solución dada de (2.1), según se definió en el Capítulo 2, entonces, para cada t , el sistema lineal algebraico asociado tiene al menos una solución, $v(t) = \dot{x}(t)$. Esto implica que, para cada t , $x(t)$ debe pertenecer al subconjunto

$$M_1 = \{x \in M \mid a(x)v = f(x) \text{ tiene al menos una solución } v \in T_x M\}.$$

Supongamos que M_1 es una subvariedad de M . Como $x(t) \in M_1$, para todo t , debe cumplirse que $\dot{x}(t) \in T_{x(t)}M_1$, para todo t . Esto implica que, para cada t , $x(t)$ debe pertenecer al subconjunto

$$M_2 = \{x \in M_1 \mid a(x)v = f(x) \text{ tiene al menos una solución } v \in T_x M_1\}.$$

Se puede continuar de modo similar y definir M_{k+1} recursivamente así,

$$M_{k+1} = \{x \in M_k \mid a(x)v = f(x) \text{ tiene al menos una solución } v \in T_x M_k\},$$

bajo la hipótesis que M_k es una subvariedad para $k = 1, 2, \dots$. Como M es de dimensión finita la secuencia $M \supseteq M_1 \supseteq \dots \supseteq M_k \supseteq \dots$ se estabiliza, es decir, existe $q \in \mathbb{N}$ tal que $M_q = M_{q+p}$, para $p \in \mathbb{N}$. El menor entero q tal que $M_q = \emptyset$ o, $M_q \neq \emptyset$ y $M_q = M_{q+1}$ se llama *índice geométrico de la ecuación* (2.1). Si $M_q = \emptyset$, la ecuación no admite soluciones. Si en cambio $M_q \neq \emptyset$, suponiendo que la matriz $a(x)$ tiene rango localmente constante en M_q , podemos garantizar la existencia local de soluciones para (2.1). En efecto, supongamos que se toman coordenadas locales (x_1, \dots, x_n) en M centradas en un punto dado $x_0 \in M_q$, y que M_q está definida por las ecuaciones $x_{c+1} = 0, \dots, x_n = 0$, siendo c la dimensión de M_q . Se puede probar que si $r = \dim(\ker a(x)|_{T_x M_q})$, las soluciones del sistema pueden describirse así: luego de una permutación de índices si fuera necesario se eligen funciones arbitrarias $x_1(t), \dots, x_r(t)$, y luego se resuelve (2.1) unívocamente para $x_{r+1}(t), \dots, x_c(t)$. La ecuación (2.1) restringida a M_q , es un ejemplo de una EDI de rango localmente constante.

Ejemplo 5.1. Consideremos la ecuación presentada en el ejemplo 1.1, que puede reescribirse como una ecuación cuasilínea

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ 1 \\ x_1 - \cos x_2 \end{bmatrix}$$

En este caso $M = \mathbb{R}^3$ y $M_1 = \{x = (x_1, x_2, x_3) \in M \mid x_1 - \cos x_2 = 0\}$. Para calcular M_2 consideramos que $\dot{x} \in TM_1$ implica que $\dot{x}_1 = -\sin(x_2)\dot{x}_2$,

y por lo tanto $a(x)\dot{x}$ tendrá la forma $(-\sin(x_2)\dot{x}_2, \dot{x}_2, 0)^\top$, para que esto sea igual a $(x_3, 1, x_1 - \cos x_2)$ es necesario que $x_3 = -\sin(x_2)$, y entonces

$$M_2 = \{x = (x_1, x_2, x_3) \in M \mid x_1 - \cos x_2 = 0, x_3 + \sin(x_2) = 0\}.$$

En el paso siguiente resulta $M_3 = M_2$, por lo que el algoritmo se detiene en el segundo paso.

En M_2 podemos tomar como coordenada a x_2 , y la ecuación restringida a M_2 adquiere entonces la forma $\dot{x}_2 = 1$.

La hipótesis que en cada paso k cada M_k resulta una subvariedad es demasiado restrictiva y no se satisface en muchos ejemplos de interés. Un ejemplo sencillo en el que esta condición no se satisface es el ejemplo 1.2, donde si intentamos aplicar el algoritmo regular obtenemos

$$M_1 = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 x_2 = 0\},$$

que no es una subvariedad.

Mostraremos que para evitar en parte estas limitaciones se pueden utilizar resultados de la *teoría de desingularización analítica real*. Esto es posible teniendo en cuenta la teoría de conjuntos semianalíticos y subanalíticos desarrollada originalmente por Lojasiewicz [22], y también a partir de trabajos posteriores de Hironaka [23], Bierstone y Milman [24], Sussmann [25], entre otros. A continuación introduciremos el *algoritmo de restricciones singular*, que permite, dada una EDICL del tipo (2.1), hallar una EDICL equivalente restringida a una subvariedad de dimensión menor, que resulta ser de rango localmente constante, donde se puede garantizar existencia de solución.

5.2. Algoritmo de restricciones singular

Se presenta en esta subsección una breve descripción del algoritmo de restricciones singular desarrollado detalladamente en [2]. Se trabajará con variedades y aplicaciones analíticas reales. Luego, si (2.1) es una EDICL dada con dominio M , entonces M será una variedad analítica real y a, f serán aplicaciones analíticas reales. A continuación enunciaremos algunas definiciones y algunos resultados ya conocidos que nos permitirán describir el algoritmo, los resultados están demostrados en el trabajo de Bierstone y Milman [24].

Conjuntos semianalíticos y subanalíticos. Sea \mathcal{A} un anillo de todas las funciones a valores reales definidas sobre un conjunto E . $S(\mathcal{A})$ está formado

por todos los subconjuntos de E de la forma $X = \bigcup_{i=1}^p \bigcap_{j=1}^q X_{ij}$, donde cada X_{ij} es, o bien $\{x \in E : f_{ij}(x) = 0\}$, o bien $\{x \in E : f_{ij}(x) > 0\}$, con $f_{ij} \in \mathcal{A}$.

Sea M una variedad analítica real. Si U es un subconjunto abierto de M , llamamos $\mathcal{O}(U)$ al anillo de las funciones analíticas reales definidas sobre U .

Definición 5.1. Un subconjunto X de M es *semianalítico* si para cada $a \in M$ existe un entorno U de a tal que $X \cap U \in S(\mathcal{O}(U))$.

Definición 5.2. Un subconjunto X de M es *subanalítico* si cada punto de M admite un entorno U tal que $X \cap U$ es una proyección de un conjunto semianalítico relativamente compacto (es decir, existe una variedad analítica real N y un conjunto semianalítico relativamente compacto A de $M \times N$ tal que $X \cap U = \pi(A)$, con $\pi : M \times N \rightarrow M$ la proyección usual).

Definición 5.3. Sea M una variedad analítica real y sea X un subconjunto subanalítico de M . Una *desingularización* de X es una aplicación analítica real propia $f : N \rightarrow M$ tal que $f(N) = X$, donde N es una variedad analítica real de la misma dimensión de X .

El siguiente teorema garantiza la existencia de desingularizaciones.

Teorema 5.1 (Hironaka 1973, [23]). *Sea M una variedad analítica real y sea X un subconjunto subanalítico cerrado. Entonces existe una desingularización $f : N \rightarrow M$ de X .*

En el algoritmo que describiremos a continuación sólo necesitaremos el siguiente resultado de desingularización, más débil que el anterior [24].

Teorema 5.2. *Sea M una variedad analítica real y sea X un subconjunto analítico cerrado. Entonces existe una desingularización $f : N \rightarrow M$ de X .*

Descripción del algoritmo. Sea M una variedad de dimensión d y sea (2.1) una EDICL dada con dominio M y rango F , que simbolizaremos $(a, f)|M$. Se probará que esta ecuación puede transformarse en una EDICL equivalente,

$$\tilde{a}_2(y)\dot{y} = \tilde{f}_2(y), \quad (5.1)$$

sobre una variedad analítica \tilde{M}_2 , que es una EDICL de rango localmente constante. Las dos ecuaciones resultan equivalentes en el siguiente sentido: existe una proyección de \tilde{M}_2 en M tal que las soluciones de (5.1) se proyectan como soluciones de (2.1), y recíprocamente, todas las soluciones de (2.1) se obtienen concatenando proyecciones de soluciones de (5.1). Para construir la

variedad \tilde{M}_2 se utilizará un algoritmo que involucra un proceso de desingularización.

Dado que M es una variedad de dimensión d , las componentes conexas de M serán de dimensión menor o igual que d . Sea $M_{\text{máx}} = \bigcup_j W_j$ la unión de todas las componentes conexas de M con dimensión d . Se puede probar que M es la unión disjunta

$$M = M_{(0)} \cup M_{(1)} \cup M_{(2)},$$

donde $M_{(1)}$ y $M_{(2)}$ son subvariedades abiertas de M de igual dimensión que M , mientras que $M_{(0)}$ es una unión finita de subconjuntos definidos por ecuaciones analíticas sobre cada W_j , unión $M - M_{(\text{máx})}$, siendo $\dim M_{(0)} < d$. Se tiene que el sistema lineal algebraico asociado a (2.1) no tiene solución para $x \in M_{(1)}$ y tiene solución para todo $x \in M_{(2)}$, más aún, el sistema (2.1) restringido a $M_{(2)}$, $(a, f)|_{M_{(2)}}$, es una EDICL de rango localmente constante. Por otro lado, $M_{(0)}$ es unión de conjuntos analíticos cerrados, que se podrán desingularizar, si fuera necesario. Los detalles de como se obtiene la descomposición $M = M_{(0)} \cup M_{(1)} \cup M_{(2)}$ se pueden ver en [2].

Nota. Si M es una variedad conexa se tiene que, o bien $M_{(1)} = \emptyset$, o bien $M_{(2)} = \emptyset$; además, en cualquiera de las dos situaciones $M_{(0)}$ resulta un subconjunto analítico cerrado de la variedad M y de dimensión menor.

Sabemos que dada una condición inicial $x_0 \in M$ se tiene que si $x_0 \in M_{(2)}$, podemos garantizar existencia de solución, mientras que si $x_0 \in M_{(1)}$ no existe solución. Luego resta analizar que sucede si $x_0 \in M_{(0)}$. Aplicando el teorema (5.2) se obtendrá una desingularización de $M_{(0)}$,

$$\pi_0 : M^1 \rightarrow M, \quad \text{donde } \pi_0(M^1) = M_{(0)}.$$

El sistema (2.1) restringido a $M_{(0)}$ puede ser “*levantado*” de manera natural, usando la operación pull-back. a un sistema cuasilineal $(a_1, f_1) = \pi_0^*((a, f)|_{M_{(0)}})$ sobre M^1 así:

$$a_1(y)\dot{y} = a(\pi_0(y))T_y\pi_0(y, \dot{y}), \quad f_1(y) = f(\pi_0(y)),$$

donde si $y(t)$ es una curva analítica en M^1 entonces

$$T_y\pi_0(y, \dot{y}) = \left. \frac{d\pi_0(y(t))}{dt} \right|_{t=0}$$

es un vector tangente a $M_{(0)}$ en $\pi_0(y(0))$.

Nota. Observar que M^1 es una variedad de dimensión $\dim M^1 = \dim M_{(0)} = d_1 < d$. Además, si $a(x)$ es una matriz cuadrada, entonces $a_1(x)$ es una matriz cuyo número de filas coincide con el número de filas de $a(x)$, es decir es d , mientras que su número de columnas es igual a $\dim M^1 = \dim M_{(0)} = d_1 < d$. El análisis de este sistema sobredeterminado en principio puede reducirse a un número finito de sistemas cuadrados dados por submatrices de orden d_1 de $a_1(x)$.

Este proceso se repite para la EDICL (a_1, f_1) con dominio M^1 y rango F . Obtenemos una descomposición

$$M^1 = M^1_{(0)} \cup M^1_{(1)} \cup M^1_{(2)}, \quad (5.2)$$

donde no existe solución para el sistema lineal asociado si $y \in M^1_{(1)}$, existe solución para el mismo sistema lineal asociado si $y \in M^1_{(2)}$, además, $(a_1, f_1)|_{M^1_{(2)}}$ es una EDICL de rango localmente constante. Como antes se desingulariza $M^1_{(0)}$ y se repite el proceso. Se obtiene así una secuencia de variedades y aplicaciones

$$M^q \xrightarrow{\pi_{q-1}} M^{q-1} \xrightarrow{\pi_{q-2}} \dots \xrightarrow{\pi_1} M^1 \xrightarrow{\pi_0} M,$$

donde $\pi_i(M^{i+1}) = M^i_{(0)}$, para $i = 0, \dots, q-1$, con $M^0 \equiv M$.

La colección de EDICLs $(a_k, f_k)|_{M^k_{(2)}}$, $k = 0, \dots, q$, define una EDICL $(\tilde{a}_2, \tilde{f}_2)$ de rango localmente constante en la unión disjunta $\tilde{M}_{(2)} = \bigsqcup_{k=0}^q M^k_{(2)}$. Se tiene la proyección natural $\tilde{\pi}_2 : \tilde{M}_{(2)} \rightarrow M$. La EDICL $(\tilde{a}_2, \tilde{f}_2)$ con dominio $\tilde{M}_{(2)}$ y rango F se llama *EDICL desingularizada*.

Resolución de una EDICL vía desingularización. Se resume en lo que sigue cual es la estrategia a seguir para hallar las soluciones analíticas de una EDICL dada vía desingularización.

1. Dada una EDICL del tipo (2.1) encontrar el sistema desingularizado $(\tilde{a}_2, \tilde{f}_2)$, como se explicó más arriba.
2. Encontrar todas las soluciones analíticas $y(t)$, $t \in (t_0, t_1)$ de $(\tilde{a}_2, \tilde{f}_2)$. Observar que dada una condición inicial $y_0 \in M^k_{(2)}$, $k = 0, 1, \dots, q$, todas las soluciones $y(t) \subset M^k_{(2)}$, $t \in (t_0, t_1)$, pueden encontrarse aplicando el algoritmo regular descripto antes.
3. Para cada solución proyectada $x(t) = \tilde{\pi}_2(y(t))$, $t \in (t_0, t_1)$, determinar si existen extensiones continuas a los intervalos $[t_0, t_1)$, $(t_0, t_1]$ y si estas extensiones son derivables a izquierda en t_1 y a derecha en t_0 , respectivamente. Finalmente, determinar si estas extensiones satisfacen la EDICL dada en los puntos t_0 y t_1 .

Observamos que dada una solución analítica $x(t)$, $t \in (t_0, t_1)$, existe una descomposición de $x(t)$ a trozos del tipo $x|_{(\tau_i, \tau_{i+1})}$, $i = 1, 2, \dots$, y cada trozo es una proyección vía $\tilde{\pi}_2$ de una solución $y(t)$, $t \in (\tau_i, \tau_{i+1})$ de $(\tilde{a}_2, \tilde{f}_2)$. Concluimos entonces que cualquier solución analítica de (a, f) puede recuperarse pegando extensiones por los extremos derechos de los intervalos donde están definidas, si tales extensiones existen, de proyecciones vía $\tilde{\pi}_2$ de soluciones de $(\tilde{a}_2, \tilde{f}_2)$.

Ejemplo 5.2. Revisitamos el ejemplo 1.2 aplicando el algoritmo singular.

El sistema

$$\begin{cases} \dot{x}_1 + \dot{x}_2 &= 1 \\ x_1 x_2 &= 0 \end{cases}$$

no puede ser analizado con el algoritmo regular debido a que el conjunto $\{(x_1, x_2) \mid x_1 x_2 = 0\}$ no es una subvariedad.

Aplicando el algoritmo singular y considerando que $M = \mathbb{R}^2$, en el primer paso obtenemos

$$\begin{aligned} M_{(0)} &= \{(x_1, x_2) \mid x_1 x_2 = 0\} \\ M_{(1)} &= \mathbb{R}^2 - M_{(0)} \\ M_{(2)} &= \emptyset \end{aligned}$$

Una desingularización de $M_{(0)}$ se obtiene tomando M^1 igual a la unión disjunta de dos rectas, proyectándose la primera sobre $x_1 = 0$ y la segunda sobre $x_2 = 0$. Por ejemplo, podemos considerar $M^1 = \{(y_1, y_2) \in \mathbb{R}^2 \mid y_2 = 1 \text{ o } y_2 = 2\}$ y $\pi_0 : M^1 \rightarrow M_{(0)}$, dada por $\pi_0(y_1, 1) = (y_1, 0)$ y $\pi_0(y_1, 2) = (0, y_1)$. El sistema levantado tiene la forma

$$\begin{cases} \dot{y}_1 &= 1 \\ 0 &= 0 \end{cases}$$

Entonces tenemos

$$\begin{aligned} M^1_{(0)} &= \emptyset \\ M^1_{(1)} &= \emptyset \\ M^1_{(2)} &= M^1. \end{aligned}$$

El algoritmo concluye en este punto, siendo $\tilde{M}_{(2)} = M^1_{(2)} = M^1$. Las soluciones del sistema desingularizado son de la forma: $y_1(t) = C + t$ e $y_2(t)$ igual a 1 o 2, siendo C una constante arbitraria. Las proyecciones de estas soluciones resultan $x_1(t) = C + t$ con $x_2 \equiv 0$ y $x_2(t) = C + t$ con $x_1 \equiv 0$, ambas soluciones de la ecuación original. Notar que para el problema de valores iniciales con $x(0) = (0, 0)$ hay dos soluciones diferentes.

Ejemplo 5.3 (Bola rodando sobre un plano horizontal sin deslizamiento y sin girar sobre su eje vertical). En este ejemplo trabajamos con las ecuaciones que modelan la dinámica de una bola rígida rodando sobre un plano horizontal. Asumimos que dos de los momentos de inercia principales de la bola son iguales entre sí, mientras que el tercero podría ser diferente. También supondremos que la bola rueda sin deslizamiento y que no gira alrededor de su eje vertical. Este modelo se representa mediante un sistema no-holónimo sobre la variedad $SO(3) \times \mathbb{R}^2$, donde para un elemento $(A, x) \in SO(3) \times \mathbb{R}^2$, A representa una rotación rígida y x la posición del punto de contacto de la bola sobre el plano. Más precisamente, consideramos una base ortonormal fija (e_1, e_2, e_3) en el espacio y otra base ortonormal que se mueve con la bola (Ae_1, Ae_2, Ae_3) donde $A = A(t) \in SO(3)$ depende del tiempo. Definimos la variable $z \in S^2$, $z = Ae_3$, entonces la velocidad angular espacial ω puede escribirse como $\omega = v_0 z + z \times \dot{z}$, donde $v_0 = \langle \omega, z \rangle$ es la componente de ω en la dirección de z .

Las variables z , v_0 y x son suficientes para describir el sistema ya que a partir de ellas se puede reconstruir ω , y a partir de ω se pueden definir ecuaciones diferenciales ordinarias para recuperar A .

El vínculo no holónimo correspondiente al no deslizamiento es $\dot{x} = \omega \times re_3$, donde r es el radio de la bola, y el vínculo correspondiente al no giro alrededor del eje vertical es $\omega_3 = \langle \omega, e_3 \rangle = 0$.

El Lagrangiano del sistema está dado por la energía cinética

$$\frac{1}{2}I_1\dot{z}^2 + \frac{1}{2}I_3v_0^2 + \frac{1}{2}M\dot{x}^2$$

donde $I_1 = I_2$ son los dos momentos de inercia iguales, I_3 es el momento de inercia que podría ser diferente y M es la masa de la bola.

En [32] se deducen las ecuaciones de este sistema, recurriendo a un cambio de variables para el caso $z_3 \neq 0$. En [2] se analiza el mismo problema recurriendo al algoritmo singular que presentamos en este capítulo. Las ecuaciones del sistema resultan:

$$\begin{aligned} \dot{z} &= z \times u \\ (1 + \beta)z_3(z_1\dot{u}_2 - z_2\dot{u}_1) &= (\alpha + \beta)u_3^2 \\ 0 &= (1 + \beta)u^2 + (\alpha + \beta)v_0^2 - 2E \\ 0 &= (1 + \beta)z_3^2u^2 + (\alpha + \beta)u_3^2 - 2Ez_3^2 \\ 0 &= z^2 - 1 \\ 0 &= \langle z, u \rangle \end{aligned}$$

donde E representa la energía normalizada, $\alpha = I_3/I_1$ y $\beta = Mr^2/I_1$.

Al aplicar el algoritmo de restricciones singular a estas ecuaciones, se tiene

$$M_{(0)} = \{z_3 = 0\} \cup \{z_1 = z_2 = 0\} \cup \{(1 + \beta)u^2 + (\alpha + \beta)v_0^2 = 2E, z_3 \neq 0, u_3 = z_3v_0, z^2 = 1, \langle z, u \rangle = 0, z_1z_2 \neq 0\};$$

$$M_{(1)} = \mathbb{R}^7 - M_{(0)}; \quad M_{(2)} = \emptyset.$$

La desingularización M^1 de $M_{(0)}$ será la union disjunta de las tres subvariedades de \mathbb{R}^7 , es decir, $M^1 = M^{1a} \sqcup M^{1b} \sqcup M^{1c}$, siendo

$$M^{1a} = \{z_3 = 0\},$$

$$M^{1b} = \{z_1 = z_2 = 0\},$$

$$M^{1c} = \{(1 + \beta)u^2 + (\alpha + \beta)v_0^2 = 2E, z_3 \neq 0, u_3 = z_3v_0, z^2 = 1, \langle z, u \rangle = 0, z_1z_2 \neq 0\}.$$

Al levantar el sistema a la componente M^{1a} las ecuaciones resultantes son:

$$\begin{aligned} \dot{z}_1 &= z_2u_3 \\ \dot{z}_2 &= -z_1u_3 \\ 0 &= z_1u_2 - z_2u_1 \\ 0 &= u_3 \\ 0 &= (1 + \beta)u^2 + (\alpha + \beta)v_0^2 - 2E \\ 0 &= z^2 - 1 \\ 0 &= \langle z, u \rangle \end{aligned}$$

Se ve rápidamente que las soluciones contenidas en esta componente de M^1 son las soluciones en las que z es constante y horizontal, es decir, corresponden a la esfera rodando alrededor del eje que tiene momento de inercia distinto.

Al levantar el sistema a la componente M^{1b} las ecuaciones que se obtienen son:

$$\begin{aligned} 0 &= u_2 \\ 0 &= u_1 \\ \dot{z}_3 &= 0 \\ 0 &= u_3 \\ 0 &= +(\alpha + \beta)v_0^2 - 2E \\ 0 &= -2Ez_3^2 \\ 0 &= z_3 \pm 1 \end{aligned}$$

En este caso no existen soluciones si $E \neq 0$, si en cambio $E = 0$ la bola no se mueve.

Finalmente, al levantar el sistema a la tercera componente M^{1c} las ecuaciones resultan ser un ODE, por lo cual, en esa componente hay existencia y unicidad de soluciones.

Problema abierto. El algoritmo de restricciones singular descrito en esta sección, si bien permitiría encontrar todas las soluciones de una EDICL del tipo (2.1), no da respuesta al siguiente problema de valores iniciales: dada una condición inicial $x(t_0) = x_0 \in M_0$, determinar si existe una solución $x(t)$ con $t \in [t_0, t_0 + \epsilon)$ (resp. $t \in (t_0 - \epsilon, t_0]$) tal que $x(t) \in M_2$, para $t \in (t_0, t_0 + \epsilon)$ (resp. $t \in (t_0 - \epsilon, t_0)$) y $x(t_0) = x_0$. Más generalmente, dada una condición inicial $y(t_0) = y_0$, donde $y(t)$ con $t \in [t_0, t_0 + \epsilon)$ (resp. $t \in (t_0 - \epsilon, t_0]$) tal que $y(t) \in M_2^k$, para $t \in (t_0, t_0 + \epsilon)$ (resp. $t \in (t_0 - \epsilon, t_0)$) y $y(t_0) = x_0$.

5.3. Comparación entre ambos algoritmos

El algoritmo de restricciones regular descrito en la sección 5.1 es un algoritmo muy general, del cual se han estudiado casos particulares. El motivo por el cual es útil especializar este algoritmo en ciertas clases de ecuaciones es que no siempre resulta sencillo calcular la subvariedad M_1 de los puntos donde el sistema lineal asociado a la EDICL tiene solución. El algoritmo de Gotay–Nester descrito en la sección 3.2 es un ejemplo de estos casos particulares. El algoritmo gCAD que introdujimos en la subsección 4.3.3 es otro. En ambos casos las estructuras geométricas utilizadas permiten que las subvariedades que aparecen en cada algoritmo queden definidas explícitamente por ecuaciones. En particular, recordemos que en el caso del algoritmo gCAD se pedía, con el objeto que los conjuntos quedaran definidos por ecuaciones, que el núcleo de la estructura de Dirac generalizada tuviera rango constante.

El algoritmo de restricciones singular descrito en la sección 5.2 requiere la hipótesis adicional que los datos sean analíticos, pero tiene la ventaja que no se necesitan estructuras geométricas adicionales para que los subconjuntos que aparecen en el algoritmo resulten ser conjuntos definidos por ecuaciones e inecuaciones.

Para colocar ambos algoritmos de restricciones, regular y singular, en un contexto común que permita compararlos debemos suponer que los datos $a(x)$ y $f(x)$ sean funciones analíticas. Por otro lado, por simplicidad, supondremos que el rango de la matriz $a(x)$ es constante en M .

Con estas hipótesis, y suponiendo que M es una variedad conexa, resulta que, para el caso del algoritmo singular, o bien $M_{(2)} = M$ y $M_{(0)} = M_{(1)} = \emptyset$,

o bien $M_{(2)} = \emptyset$, $M_{(0)} = \{f(x) \in \text{Im}(a(x))\}$, $M_{(1)} = M - M_{(0)}$.

En el primer caso, sabemos que en todo $M_{(2)}$ existen soluciones de la ecuación (2.1), y por lo tanto el conjunto M_1 que se obtiene en el primer paso del algoritmo regular también coincide con M . Es decir, ambos algoritmos concluyen en el primer paso y $M_{(2)} = M_1 = M$.

En el otro caso, el conjunto M_1 que se obtiene en el primer paso del algoritmo regular es el conjunto de los puntos x donde el sistema lineal $a(x)v = f(x)$ tiene solución, es decir que $M_1 = M_{(0)}$. Una de las hipótesis necesarias para poder aplicar el algoritmo regular es que M_1 sea una subvariedad, luego la desingularización de $M_{(0)}$ es el mismo conjunto $M_{(0)}$. Se concluye entonces en este caso que $M_{(0)} = M^1 = M_1$.

Si suponemos que el pullback de $a(x)$ a M^1 tiene rango localmente constante, entonces las mismas consideraciones que hicimos en los párrafos anteriores pueden hacerse para cada componente conexa de M^1 , es decir que si iteramos sobre cada componente conexa ambos algoritmos continuarán dando las mismas subvariedades.

En resumen, bajo la hipótesis de que $a(x)$ y de que el pullback de $a(x)$ a cada uno de los M^k tengan rango localmente constante, resulta que el algoritmo singular es una generalización del algoritmo regular.

Capítulo 6

Singularidades de Ecuaciones Diferenciales Implícitas Cuasilineales

En este capítulo consideraremos EDICLs de la forma

$$a(x)\dot{x} = f(x), \tag{6.1}$$

donde $a(x) : T_x M \rightarrow \mathbb{R}^m$ es una transformación lineal para cada $x \in M$.

Se estudiará la existencia de soluciones para este tipo de ecuaciones, considerándose especialmente soluciones por cierta clase de puntos singulares que definiremos en forma precisa más adelante.

Supondremos a lo largo de este Capítulo que M es una variedad conexa. Si bien en la sección 6.4 supondremos que $a(x)$ y $f(x)$ son aplicaciones analíticas, varios de los resultados previos valdrán con menores requisitos de regularidad, que serán mencionados explícitamente en cada caso. Cuando no se indique, se asumirá que $a(x)$ y $f(x)$ son analíticas.

6.1. Definición de puntos de cruce y puntos de impasse

En lo que sigue se recuerda el concepto general de *curva solución* de una EDICL como fue establecido en el Capítulo 2, y se da una clasificación de curvas solución con distintas propiedades de regularidad. Recordemos que una solución de (6.1) en un punto $x \in M$ es un vector $(x, v) \in T_x M$ que satisface (6.1).

Definición 6.1. Una curva $x(t)$ es solución de (6.1) si es una función continua $x: I \rightarrow M$ con $I = (t_0, t_1)$ un intervalo real (resp. $I = (t_0, t_1]$, $I = [t_0, t_1)$ o $I = [t_0, t_1]$), x diferenciable en el interior de I y $(x(t), \dot{x}(t))$ es solución de la ecuación en el punto $x(t)$, para todo t en el interior de I . Además, si $t_0 \in I$ entonces $(x(t_0), \dot{x}(t_0^+))$ es solución y si $t_1 \in I$ entonces $(x(t_1), \dot{x}(t_1^-))$ es solución, siendo $\dot{x}(t_0^+) = \lim_{t \rightarrow t_0^+} \frac{x(t) - x(t_0)}{t - t_0}$ y $\dot{x}(t_1^-) = \lim_{t \rightarrow t_1^-} \frac{x(t) - x(t_1)}{t - t_1}$.

Definición 6.2. Una curva solución $x(t)$ de (6.1) es C^k ($1 \leq k \leq \infty$) en I si se cumplen las siguientes condiciones:

- Las derivadas de x de orden menor o igual a k existen y son continuas en el interior de I .
- Si $I = (t_0, t_1]$ o $I = [t_0, t_1]$ entonces las derivadas de orden menor o igual a k tienen límite cuando $t \rightarrow t_1^-$; notar que en este caso $\dot{x}(t_1^-)$, $\ddot{x}(t_1^-)$, $x^{(3)}(t_1^-)$, etc. coinciden con estos límites.
- Si $I = [t_0, t_1)$ o $I = [t_0, t_1]$ entonces las derivadas de orden menor o igual a k tienen límite cuando $t \rightarrow t_0^+$; notar que en este caso $\dot{x}(t_0^+)$, $\ddot{x}(t_0^+)$, $x^{(3)}(t_0^+)$, etc. coinciden con estos límites.

Definición 6.3. Decimos que una curva solución $x(t)$ de (6.1) es analítica en I si es C^∞ en I y además x es una función analítica en el interior de I . Decimos que x es una *solución-as* en I si es una solución analítica en I y además la imagen $x(I)$ es un subconjunto semianalítico de M (en este caso se trata de un subconjunto semianalítico de dimensión 1, ver definición 5.1).

Hipótesis: En adelante supondremos que $\dim M = m$, es decir, que $a(x)$ es una matriz cuadrada. Supondremos también que $\det(a(x))$ no es idénticamente nulo sobre M .

Nota. Todo sistema del tipo (6.1) puede ser localmente estudiado como un sistema cuadrado (si fuera $\dim M < m$ tomamos $\dim M$ filas que sean linealmente independientes y si fuera $\dim M > m$ agregamos filas con ceros). Nuestro objetivo es resolver el problema que queda abierto del algoritmo singular del capítulo anterior, a saber, resolver el problema de valores iniciales con condiciones iniciales en $M_{(0)}$ y con la curva solución en $M_{(2)}$. Podemos suponer que el sistema ha sido levantado a una variedad donde hay existencia de soluciones. En los casos en que en $M_{(2)}$ haya unicidad de soluciones se cumplirá que $\det a(x)$ es no nulo en $M_{(2)}$. Los casos en que no haya unicidad de soluciones en $M_{(2)}$ pueden ser estudiados como un sistema con parámetros.

A partir de las hipótesis anteriores y teniendo en cuenta la descomposición $M = M_{(0)} \cup M_{(1)} \cup M_{(2)}$ introducida en (5.2), se tiene que $M_{(1)} = \emptyset$, luego M será la unión disjunta $M = M_{(0)} \cup M_{(2)}$. Más precisamente,

$$M_{(0)} = \{x \in M \mid \det(a(x)) = 0\},$$

que lo llamaremos conjunto de *puntos singulares* o *conjunto singular* y

$$M_{(2)} = M - M_{(0)},$$

que lo llamaremos *conjunto de puntos regulares* o *conjunto regular*. Observemos que $M_{(2)}$ es una subvariedad abierta de M , y luego de igual dimensión, donde hay existencia y unicidad de solución, mientras que $M_{(0)}$ es un subconjunto analítico cerrado de M de dimensión menor.

Las soluciones de (6.1) con condiciones iniciales en $M_{(0)}$ y totalmente contenidas en $M_{(0)}$ pueden ser estudiadas a partir de un levantamiento del sistema restringido a $M_{(0)}$, como fue planteado en el Capítulo anterior. Sin embargo, el algoritmo singular planteado en el Capítulo 5 no da respuesta al problema abierto planteado al final de dicho Capítulo.

Más precisamente, queremos estudiar el siguiente problema. Dado $x_0 \in M_{(0)}$ queremos determinar condiciones para la existencia de soluciones tales que $x(t) \in M_{(2)}$ para $t \in (t_0, t_0 + \epsilon)$ y $x(t) \rightarrow x_0$ cuando $t \rightarrow t_0^+$ o soluciones tales que $x(t) \in M_{(2)}$ para $t \in (t_0 - \epsilon, t_0)$ y $x(t) \rightarrow x_0$ cuando $t \rightarrow t_0^-$. En el caso de la existencia de tal solución queremos determinar si puede extenderse más allá de x_0 . En particular, nos interesa estudiar la existencia de soluciones con cierto tipo de singularidades, que definiremos a continuación.

Definición 6.4. Dado un punto $x_0 \in M_{(0)}$ decimos que una curva $x(t)$ solución de (6.1) en $I = (t_0, t_1)$, con $x(t) \in M_{(2)}$ para $t \in (t_0, t_1)$, tiene un *punto de impasse de entrada* (resp. *punto de impasse de salida*) en x_0 si $x(t) \rightarrow x_0$ cuando $x \rightarrow t_1^-$ (resp. $x \rightarrow t_0^+$) y además no existe $\dot{x}(t_1^-)$ (respectivamente $\dot{x}(t_0^+)$).

Si existe una curva solución $x(t)$ que tiene un punto de impasse (de entrada o de salida) en x_0 , decimos que $x(t)$ es una *solución de impasse* de (6.1) por x_0 .

Nota. Observar que si a una solución de impasse $x(t)$, con $t \in I$, la extendemos por continuidad hasta el borde de I , la curva resultante *no* es una solución.

Definición 6.5. Una curva $x(t)$ solución de (6.1) es una *solución de cruce* si está definida como solución en $(t_0 - \epsilon, t_0 + \epsilon)$, para algún $\epsilon > 0$, $x(t_0) = x_0 \in M_{(0)}$ y $x(t) \in M_{(2)}$ para $t \neq t_0$.

En este caso decimos que $x_0 = x(t_0)$ es un *punto de cruce* de (6.1).

Nota. Observar que si suponemos que los datos son analíticos, las soluciones $x(t)$ contenidas en M_2 serán soluciones analíticas. Dada una solución de cruce $x(t)$, las restricciones de x a los intervalos $(t_0 - \epsilon, t_0)$ y $(t_0, t_0 + \epsilon)$ serán soluciones analíticas, luego el grado de regularidad de una solución de cruce dependerá del grado de regularidad de dicha solución en el punto x_0 .

Presentamos ahora una serie de ejemplos simples, donde, a partir del estudio de la dinámica en cada uno de ellos, podemos detectar soluciones de cruce y/o soluciones de impasse, según sea el caso.

Ejemplo 6.1. Consideremos el sistema

$$\begin{cases} (x_1 + x_2)\dot{x}_1 = x_1 \\ \dot{x}_2 = -1. \end{cases}$$

En este caso $M_{(0)} = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$. Las soluciones en $\mathbb{R}^2 - M_{(0)}$ son de la forma

$$\begin{cases} x_1(t) = t \pm \sqrt{t^2 - C} \\ x_2(t) = -t. \end{cases}$$

Para $C < 0$ resultan soluciones que nunca se aproximan a $M_{(0)}$.

Para $C > 0$, $t > \sqrt{C}$ resultan soluciones que tienen un punto de impasse de salida en $(\sqrt{C}, -\sqrt{C})$.

Para $C > 0$, $t < -\sqrt{C}$ resultan soluciones que tienen un punto de impasse de entrada en $(-\sqrt{C}, \sqrt{C})$.

Para $C = 0$ se obtienen cuatro soluciones, que combinándolas de manera adecuada dan lugar a dos soluciones de cruce: $t \mapsto (0, -t)$ y $t \mapsto (2t, -t)$.

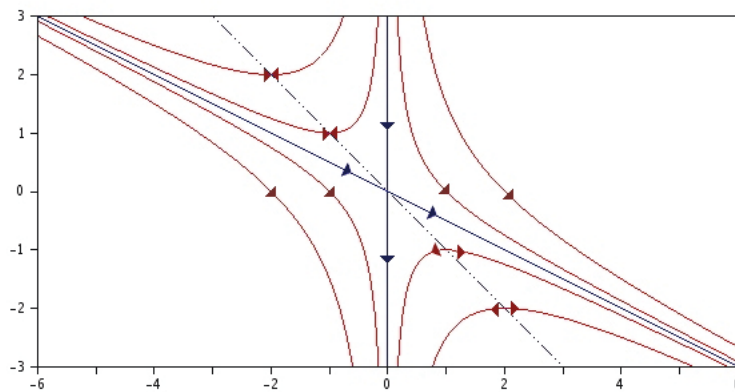


Figura 6.1: Soluciones del ejemplo 6.1

Ejemplo 6.2.

$$\begin{cases} \dot{x}_1 = 1 \\ -2x_2\dot{x}_1 + x_1\dot{x}_2 = x_1^2 \end{cases}$$

En este caso el conjunto singular es $M_{(0)} = \{x_1 = 0\}$.

Las soluciones para $x_1 \neq 0$ son de la forma $x(t) = (t, t^2(K + \ln |t|))$. Todas las soluciones de este sistema llegan al punto $(0,0)$ o salen de ese

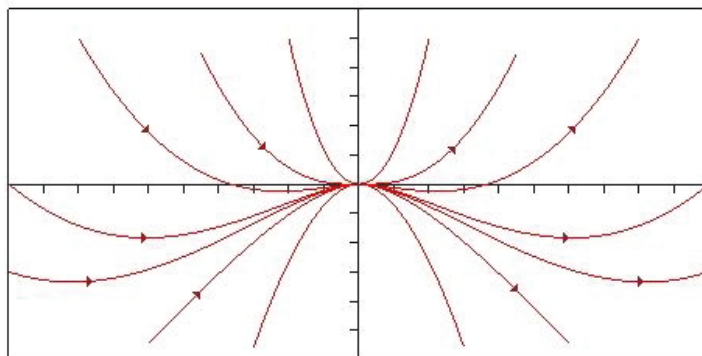


Figura 6.2: Soluciones del ejemplo 6.2. Las curvas se recorren de izquierda a derecha.

punto. Combinando cualquiera de las que llegan al punto con cualquiera de las que salen resulta una solución de cruce que es C^1 (pero no C^2).

Ejemplo 6.3.

$$\begin{cases} x_1\dot{x}_1 + x_2\dot{x}_2 = 2x_3^3 \\ -x_2\dot{x}_1 + x_1\dot{x}_2 = x_3^2 \\ \dot{x}_3 = 1 \end{cases}$$

En este caso se tiene $M_{(0)} = \{(x_1, x_2, x_3) \mid x_1 = x_2 = 0\}$. Una solución particular para $t \neq 0$ es

$$\begin{cases} x_1 = t^2 \sin(1/t) \\ x_2 = t^2 \cos(1/t) \\ x_3 = t, \end{cases}$$

que puede completarse en $t = 0$ de manera que resulte una curva continua que intersecta a $M_{(0)}$ en $(0, 0, 0)$. Además la curva es diferenciable en $t = 0$, con vector tangente $(0, 0, 1)$, y por lo tanto es una solución del sistema para todo

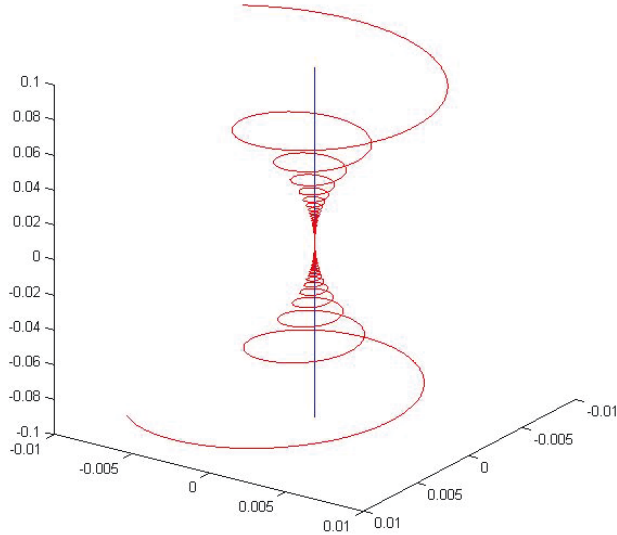


Figura 6.3: La curva (recorrida de abajo hacia arriba) es una solución de cruce del sistema. En el punto $(0, 0, 0)$ es diferenciable, pero no es C^1 .

t , resultando entonces una solución de cruce por $(0, 0, 0) \in M_{(0)}$. Observar sin embargo que no es una solución C^1 , ya que \dot{x} no es continua en $t = 0$.

6.2. Reparametrización

Queremos analizar la existencia de puntos de impasse y puntos de cruce de (6.1), o equivalentemente estudiar la existencia de soluciones de cruce y soluciones de impasse. Para ello introducimos una nueva ecuación asociada a la EDICL (6.1), que se obtiene premultiplicando (6.1) por la matriz $\text{adj}(a(x))$,

$$h(x)\dot{x} = g(x), \quad (6.2)$$

donde $h(x) = \det(a(x))$ y $g(x) = \text{adj}(a(x))f(x)$.

Los siguientes lemas preliminares serán importantes para estudiar la existencia de puntos de impasse y puntos de cruce.

Lema 6.1. *Sea $x: I \rightarrow M$ una curva al menos C^1 tal que $x(t_0) = x_0 \in M_{(0)}$ y $x(t) \in M_{(2)}$ para $t \neq t_0$, entonces $x(t)$ es solución de (6.1) si y sólo si $x(t)$ es solución de (6.2).*

Demostración. La ecuación (6.2) se obtiene a partir de premultiplicar (6.1) por $\text{adj}(a(x))$, por lo tanto toda solución de (6.1) es también solución de (6.2).

Por otro lado, la matriz $\text{adj}(a(x))$ es invertible si $x \in M_{(2)}$, luego si $x(t)$ es solución de (6.2) resulta que $(x(t), \dot{x}(t))$ es solución de (6.1) para $t \neq t_0$. Como además $x(t)$, $\dot{x}(t)$, $a(x)$ y $f(x)$ son continuas resulta que $(x(t_0), \dot{x}(t_0))$ también es solución de (6.1).

La deducción del párrafo precedente sigue siendo válida si las hipótesis de regularidad para las soluciones son reemplazadas por cualesquiera otras que garanticen la continuidad de \dot{x} . Notar en particular que para el caso en que t_0 sea un borde del intervalo I , es necesario que la derivada de x sea continua hasta ese borde (ver definición 6.2). □

Lema 6.2. *Si $x_0 \in M_{(0)}$ es un punto de cruce de (6.1), entonces $g(x_0) = 0$.*

Demostración. Sea $x(t)$ una solución de cruce de (6.1) con $x(t_0) = x_0$. Luego $x(t)$ también es solución de (6.2), entonces $g(x_0) = g(x(t_0)) = h(x_0)\dot{x}(t_0) = 0$. □

Lema 6.3. *1. Las soluciones de (6.1) y (6.2) contenidas en $M_{(2)}$ son reparametrizaciones de las soluciones de la ecuación reparametrizada*

$$y' = g(y), \tag{6.3}$$

siendo $y'(s) = \frac{dy(s)}{ds}$.

2. Sea $y(s)$ una solución de (6.3) con $y(s_0) = x_0 \in M_{(0)}$ e $y(s) \in M_{(2)}$ para $s \neq s_0$, y consideramos las dos curvas $y_1(s) = y(s)$, para $s < s_0$ e $y_2(s) = y(s)$, para $s > s_0$. Mediante la reparametrización de cada una de estas curvas se obtienen dos nuevas curvas $x_1(t)$, $x_2(t)$, ambas soluciones de (6.1); además, $x_1(t)$ tiende a $x_0 = y(s_0)$ o bien cuando $t \rightarrow 0^+$ o bien cuando $t \rightarrow 0^-$; lo mismo ocurre con $x_2(t)$.

Demostración. 1. Sea $y(s)$ una curva solución de (6.3).

Supongamos que $y(s)$ está contenida en $M_{(2)}$ y consideremos la función $t(s) = \int_{s_0}^s h(y(u))du$ y su función inversa $s(t)$ (la inversa siempre existe puesto que, al ser $h(y)$ de signo constante resulta que $t(s)$ es monótona), entonces $x(t) = y(s(t))$ es una solución de (6.2) y también de (6.1). Como en $M_{(2)}$ hay existencia y unicidad de soluciones, las soluciones obtenidas mediante reparametrización son las únicas soluciones de (6.2) y (6.1) en $M_{(2)}$.

2. Por otro lado, si $y(s_0) = x_0 \in M_0$ e $y(s) \in M_{(2)}$ para $s \neq s_0$, entonces $h(y(s_0)) = 0$ y existe un $\epsilon > 0$ tal que $h(y(s))$ no se anula en $(s_0 - \epsilon, s_0)$ ni en $(s_0, s_0 + \epsilon)$. Luego, cada uno de estos intervalos da un dominio donde la función $t(s)$ es invertible, por lo tanto hay dos inversas $s_1(t)$ y $s_2(t)$ con imagen en $(s_0 - \epsilon, s_0)$ y en $(s_0, s_0 + \epsilon)$ respectivamente. Para cada una de estas inversas se obtiene una reparametrización diferente $x_i(t) = y_i(s_i(t))$, y para cada una se tiene $x_i(t) \rightarrow y(s_0) = x_0$ cuando $t \rightarrow 0^+$ o $t \rightarrow 0^-$, dependiendo del signo de $t(s)$. \square

Nota. En [19] se prueba lo anterior en el contexto C^∞ , agregándose la hipótesis adicional que $\nabla h(x_0)\dot{y}(s_0) \neq 0$. En el caso analítico la función $s \mapsto h(y(s))$ debe tener ceros aislados porque no es idénticamente nula, sin embargo en el caso C^∞ es necesaria una hipótesis adicional que garantice que su derivada no se anule, al menos para $s \neq s_0$, con s pequeño.

Proposición 6.4. *Dado $x_0 \in M_{(0)}$, si $g(x_0)$ no es tangente a $M_{(0)}$ entonces x_0 es un punto de impasse de (6.1).*

Demostración. Sea $y(s)$ la curva solución de (6.3) con $y(0) = x_0$. Como $g(x_0)$ no es tangente a $M_{(0)}$, entonces existe $\epsilon > 0$ tal que $y(s) \in M_{(2)}$ para $0 < |s| < \epsilon$.

Por el lema anterior existen dos curvas $x_1(t)$ y $x_2(t)$ soluciones de (6.1) y (6.2) tales que para cada una de ellas $x_i(t) \rightarrow x_0$ cuando $t \rightarrow 0^+$ o cuando $t \rightarrow 0^-$. Consideremos el caso que $t \rightarrow 0^+$, el otro se prueba de manera similar.

Vamos a demostrar que $\dot{x}(0^+)$ no existe.

Como $h(x(t))\dot{x}(t) = g(x(t))$ resulta, usando teorema del valor medio, en una carta local de $M_{(2)}$, que

$$\dot{x}(0^+) = \lim_{t \rightarrow 0^+} \frac{x(t) - x_0}{t} = \lim_{t \rightarrow 0^+} \dot{x}(\xi) = \lim_{t \rightarrow 0^+} g(x(\xi))/h(x(\xi))$$

donde ξ es un valor intermedio entre 0 y t .

Como además $h(x(\xi)) \rightarrow 0$ y $g(x(\xi)) \rightarrow g(x_0) \neq 0$, entonces $\|\dot{x}(\xi)\| \rightarrow \infty$ y por lo tanto no existe $\dot{x}(0^+)$, es decir, x_0 es un punto de impasse de (6.1). \square

Lema 6.5. *Sea $x(t)$ una solución de cruce C^k , $k = 1, \dots, \infty, \omega$ de (6.1) con $x(t_0) = x_0 \in M_{(0)}$ y $\dot{x}(t_0) \neq 0$, entonces existe una subvariedad W de dimensión 1, con el mismo grado de regularidad que x , que contiene a x_0 y que es invariante por el flujo de la ecuación (6.3).*

Demostración. Podemos suponer que \dot{x} no se anula nunca (achicando el dominio de x si fuera necesario), entonces la imagen de x es una variedad W de dimensión 1 con el grado de regularidad necesario, que contiene a x_0 . Además $W - \{x_0\}$ es invariante por el flujo de (6.1) (recordar que (6.1) es una EDO en $M_{(2)}$).

Como las soluciones de (6.3) contenidas en $M_{(2)}$ son reparametrizaciones de las soluciones de (6.1), entonces $W - \{x_0\}$ es invariante por el flujo de (6.3). Además, por lema 6.2, x_0 es un punto de equilibrio de (6.3), entonces W es invariante por el flujo de (6.3). □

En el siguiente teorema [28], del cual daremos una idea de su demostración, se establecen condiciones suficientes para la existencia de puntos de cruce.

Teorema 6.6. *Consideramos la ecuación (6.1) y supongamos que $a(x)$ y $f(x)$ son aplicaciones C^1 . Sea $x_0 \in M_{(0)}$, si se cumple que*

1. $g(x_0) = 0$,
2. $\nabla h(x_0) \neq 0$,
3. *existe una subvariedad W de dimensión 1, invariante por el flujo de (6.3), transversal a $M_{(0)}$ y tal que el sistema (6.3) restringido a W tiene un punto de equilibrio exponencialmente atractivo (o repulsivo) en x_0 ,*

entonces x_0 es un punto de cruce de (6.1).

Demostración. Daremos un esquema de la demostración, los detalles se pueden ver en [28]. Primero, se restringe el sistema a la subvariedad W .

Del hecho de que W sea transversal a $M_{(0)}$ y que $\nabla h(x_0) \neq 0$ se deduce que la derivada de $h|_W$ es no nula.

Del hecho de que el equilibrio del sistema reparametrizado restringido a W sea exponencialmente atractivo, se deduce que la derivada de $g|_W$ es estrictamente negativa.

Como $h|_W$ y $g|_W$ ambas se anulan en x_0 , y ninguna de sus derivadas se anula, entonces el cociente g/h puede extenderse de manera continua a todo W , por lo tanto el sistema restringido a W es equivalente a una ecuación diferencial ordinaria en una sola dimensión, que tendrá en consecuencia solución única. □

Para la aplicación práctica de este teorema, la hipótesis 3 sobre la existencia de una subvariedad de dimensión 1 que cumpla con determinadas condiciones resulta algo difícil de verificar. En [19] se menciona que en ciertas circunstancias el teorema de la variedad estable puede resolver el problema, y es común que eso suceda en sistemas de dimensión 2.

En la sección siguiente, luego de pasar por varios resultados intermedios, algunos de ellos ya conocidos, enunciaremos y probaremos un resultado nuevo, que puede verse, bajo ciertas condiciones, como un recíproco del teorema 6.6.

6.3. Algunos resultados sobre existencia de curvas invariantes

Comenzaremos esta sección con algunos resultados intermedios vinculados con la existencia de curvas invariantes.

El Teorema de la Variedad Estable es un resultado clásico que puede encontrarse demostrado en diferentes textos, como por ejemplo en [34] (Teorema 4.1, pag. 330). Presentamos aquí un resultado levemente más general, cuya demostración (que se agrega como apéndice) no varía demasiado de la demostración original del teorema.

Teorema 6.7. *Consideramos la ecuación (6.3), y asumimos que g es C^k , $k = 1, \dots, \infty, \omega$. Sea $x_0 \in M_{(0)}$ y supongamos que:*

1. $g(x_0) = 0$,
2. *existen $\alpha > 0$ y dos subespacios E y F , invariantes por $Dg(x_0)$, con $\mathbb{R}^m = E \oplus F$, tales que $Dg(x_0)|_E$ tiene todos sus autovalores con parte real menor que $-\alpha$ y $Dg(x_0)|_F$ tiene todos sus autovalores con parte real mayor que $-\alpha$,*

entonces existe una subvariedad W , con el mismo grado de regularidad que g , invariante por el flujo de (6.3), tangente a E en x_0 y tal que todas las soluciones con valores iniciales en W suficientemente cercanos a x_0 satisfacen

$$\|x(t) - x_0\| \leq \delta e^{-\alpha t}.$$

Demostración. Ver Apéndice A. □

El siguiente Corolario garantiza la hipótesis 3 del teorema 6.6.

Corolario 6.8. *Si $g(x_0) = 0$, $Dg(x_0)$ tiene un autovalor real simple $\lambda < 0$ y todos los otros autovalores de $Dg(x_0)$ tienen parte real mayor que λ ,*

entonces existe una variedad W de dimensión 1, tangente en x_0 al autovector correspondiente al autovalor λ , invariante por el flujo de (6.3) y tal que el sistema (6.3) restringido a W tiene un equilibrio exponencialmente atractivo en x_0 .

Demostración. El resultado vale considerando en el teorema anterior $\alpha = -\lambda + \epsilon$, con $\epsilon > 0$ suficientemente pequeño.

arriba es $\alpha = -\lambda + \epsilon$, o $\alpha = -\lambda - \epsilon$?

□

En el siguiente corolario se determinan condiciones suficientes para la existencia de puntos de cruce.

Corolario 6.9. *Dado $x_0 \in M_{(0)}$, si se cumple que*

1. $g(x_0) = 0$,
2. $\nabla h(x_0) \neq 0$,
3. $Dg(x_0)$ tiene un autovalor real simple $\lambda < 0$,
4. todos los demás autovalores de $Dg(x_0)$ tienen parte real mayor que λ ,
5. el autovector correspondiente al autovalor λ no es tangente a $M_{(0)}$,

entonces x_0 es un punto de cruce de (6.1).

Demostración. Las hipótesis 3 y 4 garantizan que existe una subvariedad W de dimensión 1, invariante por el flujo de (6.3) y tal que el sistema (6.3) restringido a W tiene un equilibrio exponencialmente atractivo en x_0 ; mientras que la hipótesis 5 garantiza que W es transversal a $M_{(0)}$. Con esto se cumple la hipótesis 3 del teorema 6.6. Las hipótesis 1 y 2 coinciden con las hipótesis 1 y 2 de ese teorema. Luego, x_0 resulta un punto de cruce de (6.1).

□

Nota. Un resultado similar puede obtenerse para el autovalor real máximo. Para verlo alcanza con invertir el tiempo en el sistema, cambiando entonces g por $-g$, con lo que se invierten todos los signos de los autovalores.

El siguiente teorema [38] es una versión un poco más fuerte del teorema de Hartman–Grobman [33], y se debe al mismo Hartman.

Teorema 6.10. *Consideremos el sistema (6.3), supongamos que g es de clase C^2 , que $g(x_0) = 0$ y que todos los autovalores de $Dg(x_0)$ tienen el mismo signo, entonces existe un difeomorfismo C^1 de un entorno de x_0 en un entorno de 0 que lleva las soluciones del sistema (6.3) en soluciones del sistema linealizado.*

Más precisamente, existe un difeomorfismo C^1 $R : U \rightarrow V$, con U, V abiertos, $x_0 \in U$, $R(x_0) = 0$, tal que si $x(t)$ es una curva solución de (6.3), contenida en U , entonces $R(x(t)) = \exp(t Dg(x_0))R(x(0))$.

Combinando los resultados anteriores puede probarse el siguiente corolario.

Corolario 6.11. *Consideremos el sistema (6.3) y supongamos que g es al menos C^2 . Sea x_0 tal que $g(x_0) = 0$ y sea v un autovector de $Dg(x_0)$ tal que el autovalor correspondiente es no nulo, entonces existe una subvariedad W de dimensión 1, tangente a v en x_0 , que es invariante por el flujo de (6.3).*

Demostración. Supongamos que el autovalor correspondiente a v es negativo (si el autovalor fuera positivo, cambiando g por $-g$ se obtiene el mismo resultado).

Por el teorema 6.7, existe una subvariedad invariante por el flujo de (6.3) que es tangente a todos los autovectores cuyos autovalores sean negativos (tomamos λ_1 un autovalor con parte real máxima entre los autovalores con parte real negativa, tomamos $\alpha = \text{Re}(\lambda_1)/2$ y aplicamos el teorema 6.7). Llamemos \tilde{W} a esa subvariedad.

Consideramos el sistema (6.3) restringido a \tilde{W} . Como $D(g|_{\tilde{W}})$ tiene todos sus autovalores con parte real negativa, entonces por el teorema 6.10, existe un difeomorfismo R de clase C^1 que transforma las trayectorias de (6.3) en trayectorias del sistema linealizado.

El sistema linealizado tiene curvas invariantes tangentes a v , entonces si llamamos W a la preimagen por R de cualquiera de estas curvas, resulta que W es una subvariedad de dimensión 1 tangente a v e invariante por el flujo de (6.3). \square

Veamos ahora que cualquier curva invariante será tangente a algún autovector.

Proposición 6.12. *Sea x_0 tal que $g(x_0) = 0$, sea W una subvariedad de dimensión 1 que pasa por x_0 invariante por el flujo de (6.3), entonces W es tangente a un autovector de $Dg(x_0)$.*

Demostración. Para cada punto $x \in W$ en un entorno de x_0 consideramos el subespacio $(T_x W)^\circ \subset T_x^* M$. Sean $\alpha_1, \alpha_2, \dots, \alpha_{m-1} \in T^* M$ los elementos de

una base de $(TW)^\circ$ a lo largo de W . Estos covectores pueden extenderse de manera suave a un entorno de W .

Entonces $f_i(x) := \alpha_i(x)g(x)$ son funciones definidas en un entorno de x_0 que se anulan en W .

Sea v un campo vectorial no nulo definido sobre la variedad M tal que $v(x_0) \in T_{x_0}W$,

Como f_i se anula en W resulta $\partial_v f_i|_{x=x_0} = 0$.

$$\begin{aligned}\partial_v f_i &= (\partial_v \alpha_i)g + \alpha_i(\partial_v g) \\ &= (\partial_v \alpha_i)g + \alpha_i(Dg \cdot v).\end{aligned}$$

Evaluando en $x = x_0$ resulta

$$\begin{aligned}(\partial_v f_i)|_{x=x_0} &= (\partial_v \alpha_i)|_{x=x_0}g(x_0) + \alpha_i(x_0)(Dg(x_0)v) \\ 0 &= 0 + \alpha_i(x_0)(Dg(x_0)v)\end{aligned}$$

Entonces $Dg(x_0)v$ está en el anulador del subespacio generado por los $\alpha_i(x_0)$, $i = 1, \dots, n-1$, que es precisamente la recta tangente $T_{x_0}W$, generada por v . Luego v es un autovector de $Dg(x_0)$.

□

Teorema 6.13. *Consideramos la ecuación (6.1) con datos suficientemente suaves y sea $x_0 \in M_{(0)}$. Supongamos que existe una curva $x(t)$ solución C^k , $k = 1, \dots, \infty, \omega$ de (6.1) que es solución de cruce por x_0 , esto es $x(t_0) = x_0$ y $x(t) \in M_{(2)}$ para todo $t \neq t_0$. Supongamos también que $\dot{x}(t_0) \neq 0$. Entonces $g(x_0) = 0$ y $\dot{x}(t_0)$ es un autovector de $Dg(x_0)$.*

precisar lo de "suficientemente" suaves

Si además el autovalor correspondiente a $\dot{x}(t_0)$ es no nulo, entonces $\nabla h(x_0) \neq 0$ y $\dot{x}(t_0)$ no es tangente a $M_{(0)}$.

Demostración. Observemos primero que $g(x_0) = 0$ por el lema 6.2. Además, por lema 6.5, $x(t)$ está contenida en una subvariedad W de dimensión 1, de clase C^k , e invariante por el flujo de (6.3). En particular $\dot{x}(t_0)$ es tangente a W y por la proposición 6.12 resulta que $\dot{x}(t_0)$ es un autovector de $Dg(x_0)$.

Ahora consideremos el hecho de que para $t \neq t_0$ se tiene $\dot{x} = g(x)/h(x)$ para todo $x \in W$, $x \neq x_0$. Esto implica que

$$\frac{g(x)}{h(x)} \xrightarrow[\substack{x \rightarrow x_0 \\ x \in W}]{} \dot{x}(t_0).$$

A partir de lo anterior vamos a demostrar que si λ , el autovalor correspondiente a $\dot{x}(t_0)$, también es distinto de cero entonces $\nabla h(x_0)\dot{x}(t_0) \neq 0$.

Sea $v(x)$ un campo vectorial no nulo tangente a W en un entorno de x_0 , supongamos que $v(x_0) = \dot{x}(t_0)$ (si hiciera falta multiplicamos por un escalar

no nulo). El campo vectorial v puede utilizarse para definir una parametrización de W así: $\gamma(t) = \int_0^t v(s)ds$. De este modo $\gamma'(t) = v(\gamma(t))$ y para cualquier función F definida sobre W se tiene que $\frac{d}{dt}F(\gamma(t)) = \partial_v F(\gamma(t))$. Entonces, por la regla de L'Hôpital se tiene

$$\lim_{\substack{x \rightarrow x_0 \\ x \in W}} \frac{g(x)}{h(x)} = \lim_{\substack{x \rightarrow x_0 \\ x \in W}} \frac{\partial_v g(x)}{\partial_v h(x)}. \quad (6.4)$$

Además $\partial_v g(x) \xrightarrow{x \rightarrow x_0} \partial_{v(x_0)} g(x_0) = Dg(x_0)\dot{x}(t_0) = \lambda\dot{x}(t_0)$.

Luego el numerador del límite (6.4) tiende a un vector no nulo, por lo tanto el denominador debe tender a un valor no nulo, es decir

$$\lim_{\substack{x \rightarrow x_0 \\ x \in W}} \partial_v h(x) \neq 0;$$

y como $\partial_v h(x) \rightarrow \nabla h(x_0)\dot{x}(t_0)$ resulta que

$$\nabla h(x_0)\dot{x}(t_0) \neq 0.$$

En particular esto implica que $\nabla h(x_0) \neq 0$, por lo que la ecuación $h(x) = 0$ define regularmente a $M_{(0)}$ en un entorno de x_0 . Además $\dot{x}(t_0)$ no es ortogonal a $\nabla h(x_0)$, o lo que es lo mismo $\dot{x}(t_0)$ es transversal a $M_{(0)}$. □

Nota. Como ya lo adelantamos, el teorema 6.13 puede verse, bajo ciertas consideraciones, como el recíproco del teorema (6.6). En efecto, se tiene que algunas de las hipótesis del teorema 6.13 son también condiciones necesarias para la existencia de un punto de cruce (estas son: $g(x_0) = 0$, la existencia de la curva W y consecuentemente que $\dot{x}(t_0)$ sea un autovector de $Dg(x_0)$ o $\dot{x}(t_0) = 0$). Además, si $\dot{x}(t_0) \neq 0$ y $Dg(x_0)\dot{x}(t_0) \neq 0$, entonces todas las hipótesis del teorema 6.13 son condiciones necesarias y suficientes para la existencia de un punto de cruce.

Nota. Queda como un problema a resolver qué es lo que ocurre cuando $Dg(x_0)$ tiene autovalores nulos o cuándo existen soluciones de cruce con $\dot{x}(t_0) = 0$.

6.4. Existencia de soluciones de cruce y de impasse

En esta sección nos ocuparemos de estudiar la existencia de soluciones para una ecuación del tipo (6.1) y una condición inicial $x_0 \in M$. Se enuncian

más abajo las hipótesis que suponemos válidas a lo largo de la sección.

Hipótesis: M es una variedad conexa; $a(x)$ y $f(x)$ son aplicaciones analíticas sobre M ; $\dim M = m$ (luego, $a(x)$ resulta una matriz cuadrada); $\det(a(x))$ no es idénticamente nulo sobre M .

Bajo estas hipótesis M se descompone en la unión disjunta $M = M_{(0)} \cup M_{(2)}$, con $M_{(0)} = \{x \in M : \det(a(x)) = 0\} \neq \emptyset$. En particular, estudiaremos si existen *soluciones de cruce o soluciones de impasse* por x_0 .

Sabemos que si $x_0 \in M_{(2)}$ entonces existe y es única la solución de (6.1), pues en este caso la ecuación resulta equivalente a una EDO, al menos localmente en un entorno U de x_0 , $U \subset M_{(2)}$. Luego, supondremos de ahora en más que $x_0 \in M_{(0)} = M - M_{(2)}$.

En adelante utilizaremos una clasificación de singularidades dada en [29]. Comenzaremos dando algunas definiciones y enunciando algunos resultados de ese trabajo, que aplicaremos en particular considerando $g(x) = \text{adj}(a(x))f(x)$ y $h(x) = \det(a(x))$.

La siguiente clasificación de singularidades corresponde al análisis del campo $F(x) = \frac{g(x)}{h(x)}$, y se basa en el estudio de la existencia de una extensión continua de este campo. Como primera observación se tiene que, si $x_0 \in M_{(0)}$ y $g(x_0) \neq 0$, entonces no es posible extender el campo $F(x)$ en forma continua por x_0 , puesto que

$$\lim_{x \rightarrow x_0} \|F(x)\| = \infty,$$

Teniendo en cuenta esto se tienen las siguientes definiciones.

Definición 6.6 (*Singularidades esenciales*).

$$M_{(0)}^e = \{x \in M_{(0)} : g(x) \neq 0\}.$$

Definición 6.7 (*Singularidades no esenciales*).

$$M_{(0)}^{ne} = M_{(0)} - M_{(0)}^e.$$

Además, al conjunto $M_{(0)}^{ne}$ lo podemos descomponer en la unión disjunta de *singularidades no esenciales fuertes* $M_{(0)}^s$ y *singularidades no esenciales débiles* $M_{(0)}^w$, siendo

$$M_{(0)}^s = \partial_{M_{(0)}} M_{(0)}^e, \quad M_{(0)}^w = M_{(0)} - \overline{M_{(0)}^e}.$$

A partir de las definiciones anteriores se puede concluir que el conjunto de singularidades fuertes corresponde a la frontera del conjunto de singularidades esenciales relativa a la topología de $M_{(0)}$, es decir, a los ceros de g que son puntos de acumulación de singularidades esenciales, mientras que el conjunto de singularidades débiles es el interior en $M_{(0)}$ del conjunto de ceros de g . Por último observemos que el conjunto singular $M_{(0)}$ puede escribirse como la unión disjunta:

$$M_{(0)} = M_{(0)}^e \cup M_{(0)}^s \cup M_{(0)}^w. \quad (6.5)$$

La siguiente proposición, probada en general en [29], puede aplicarse al conjunto regular $M_{(2)}$ correspondiente a la ecuación (6.1), para $h(x) = \det(a(x))$ y $g(x) = \text{adj}(a(x))f(x)$.

Proposición 6.14. *Sea $h \in \mathcal{C}^\omega(M, \mathbb{R})$. Si $h \not\equiv 0$, entonces $M_{(2)} = \{x \in M : h(x) \neq 0\}$ es denso en M .*

Los siguientes teoremas, probados en [29], establecen una caracterización del conjunto maximal $\widetilde{M}_{(2)} \supset M_{(2)}$ donde el campo vectorial F puede extenderse en forma continua.

Teorema 6.15. *Sean $g(x)$ y $h(x)$ funciones continuas sobre M . Entonces el conjunto maximal $\widetilde{M}_{(2)}$ donde F puede extenderse en forma continua verifica*

$$\widetilde{M}_{(2)} \subset M_{(2)} \cup M_{(0)}^w.$$

Teorema 6.16. *Supongamos $\nabla h(x) \neq 0, \forall x \in M_{(0)}^w$. Si $g \in \mathcal{C}^k(M, \mathbb{R}^n)$, $h \in \mathcal{C}^k(M, \mathbb{R})$, $k = 1, \dots, \infty, \omega$, entonces*

$$\widetilde{M}_{(2)} = M_{(2)} \cup M_{(0)}^w.$$

Nota. A partir de las hipótesis dadas al comienzo de esta Sección sabemos que las aplicaciones $h(x) = \det(a(x))$ y $g(x) = \text{adj}(a(x))$ son analíticas, luego vale el teorema 6.16 si suponemos $\nabla h(x) \neq 0, \forall x \in M_{(0)}^w$.

Nota. A partir de las definiciones y resultados dados arriba y de las definiciones de punto de impasse y punto de cruce dadas en la sección anterior observamos que:

- Los puntos de cruce de (6.1) son singularidades no esenciales, consecuencia inmediata del Lema 6.2.
- La observación anterior puede generalizarse aún más: si x_0 es una singularidad esencial, entonces no existen soluciones de (6.1) que pasen por x_0 . En efecto, dado que $g(x_0) \neq 0$, resulta que $f(x_0) \notin \ker \text{adj}(a(x_0))$,

y luego $f(x_0) \notin rg(a(x_0))$ (porque $adj(a(x_0))a(x_0) = 0$). Se tiene entonces que no pueden existir soluciones de cruce por x_0 ni tampoco soluciones contenidas en $M_{(0)}$ que pasen por x_0 .

Nos ocuparemos ahora de estudiar la existencia de soluciones de cruce y soluciones de impasse para la ecuación (6.1). Es decir, dada (6.1) y dado $x_0 \in M_{(0)}$, nos preguntamos si existe una solución de cruce o una solución de impasse por x_0 . Teniendo en cuenta la descomposición disjunta de $M_{(0)}$ dada en (6.5), consideraremos 3 casos.

Caso I: $x_0 \in M_{(0)}^e$.

Como dijimos más arriba, en este caso no existen soluciones de cruce de (6.1). Luego, buscamos resultados sobre la existencia de soluciones de impasse de (6.1) por x_0 . Para ello consideramos la ecuación diferencial algebraica - EDA-

$$y' = g(y) \tag{6.6}$$

$$0 = h(y), \tag{6.7}$$

con $y \in M$ e $y'(s) = \frac{dy(s)}{ds}$.

Nota. Observar que la ecuación (6.6) tiene una única solución por x_0 . Si esa solución estuviera contenida en $M_{(0)}$, entonces x_0 no sería un punto de impasse de (6.1). En efecto, si x_0 fuera un punto de impasse de (6.1), habría una solución de (6.1) (y por lo tanto de (6.6)) en $M_{(2)}$ que tendería en tiempo finito a x_0 , esta solución de (6.6) en $M_{(2)}$ podría extenderse de modo tal de obtener otra solución de (6.6) por x_0 , lo cual no puede ser por unicidad de solución de la EDO por x_0 .

Llamamos Y a la derivación correspondiente al campo vectorial $g(y)$. Se definen $\varphi_0, \varphi_1, \dots, \varphi_k, \dots$ así

$$\varphi_0(y) = h(y)$$

$$\varphi_1(y) = Y[\varphi_0](y) = \sum \frac{\partial \varphi_0}{\partial y^j} g^j(y)$$

⋮

$$\varphi_k(y) = Y[\varphi_{k-1}](y)$$

⋮

$\varphi_0, \varphi_1, \dots, \varphi_k, \dots \in \mathcal{C}^\omega(M)$ generan una cadena de ideales

$$\mathcal{J}_k = \langle \varphi_0, \dots, \varphi_k \rangle,$$

con $\mathcal{J}_0 \subseteq \mathcal{J}_1 \subseteq \dots \subseteq \mathcal{J}_k \subseteq \dots$

El anillo $\mathcal{C}^\omega(M)$ es Noetheriano, luego $\mathcal{J}_{\tilde{k}} = \mathcal{J}_{\tilde{k}+p}$ para un cierto \tilde{k} y $\forall p \in \mathbb{N}$. Se tiene entonces el siguiente conjunto correspondiente a los ceros del ideal final $\mathcal{J}_{\tilde{k}}$:

$$N = Z(\mathcal{J}_{\tilde{k}}) = \{\varphi_0(y) = 0, \dots, \varphi_{\tilde{k}}(y) = 0\}.$$

El siguiente resultado establece cuando un punto $x_0 \in M^e_{(0)}$ es un punto de impasse de (6.1).

Teorema 6.17. *Sea la ecuación (6.1) y sea $x_0 \in M^e_{(0)}$. Supongamos válidas todas las hipótesis enunciadas al comienzo de esta sección. Consideramos la EDA (6.6)-(6.7), los ideales \mathcal{J}_k definidos a partir de ella y los conjuntos de ceros $Z(\mathcal{J}_k)$, $k = 0, 1, \dots, \tilde{k}$. Entonces x_0 es un punto de impasse de (6.1) sii $x_0 \notin N = Z(\mathcal{J}_{\tilde{k}})$.*

Demostración. Supongamos que $x_0 \in N$. Sea $y(s)$ la solución al problema de valores iniciales (6.6) con $y(0) = x_0$. Se tiene que,

$$\begin{aligned} \varphi_n(x_0) &= 0, \quad \forall n; \\ \frac{d}{ds} \varphi_n(y(s)) &= \nabla \varphi_n(y(s)) \cdot \dot{y}(s) = Y[\varphi_n](y(s)) = \varphi_{n+1}(y(s)); \\ \frac{d^k}{ds^k} \varphi_n(y(s)) &= \varphi_{n+k}(y(s)); \end{aligned}$$

luego,

$$\left. \frac{d^k}{ds^k} \varphi_n(y(s)) \right|_{s=0} = 0, \quad \forall k.$$

Como $y(s)$ y la φ_n son funciones analíticas, entonces $\varphi(y(s)) \equiv 0$, y por lo tanto $y(s) \in N$, $\forall s$.

Si x_0 es punto de impasse de (6.1), la solución de impasse correspondiente resulta, vía una reparametrización, una solución de (6.6), que por tratarse de una EDO puede extenderse hasta x_0 . Luego, (6.6) tiene dos soluciones distintas por x_0 , la solución de impasse de (6.1) reparametrizada y extendida y la solución de (6.6) con condición inicial en $x_0 \in N$, que está contenida en N , y luego en $M_{(0)}$, lo cual no puede ser. Concluimos entonces que si $x_0 \in N$, entonces no puede ser punto de impasse de (6.1). O equivalentemente, si x_0 es punto de impasse de (6.1) entonces $x_0 \notin N$.

Recíprocamente, supongamos que $x_0 \notin N$, luego la única solución $y(s)$ de (6.6) por $y(0) = x_0$ no está contenida en N y luego tampoco en $M_{(0)}$.

poner referencias internas

En efecto, como $y(s)$ no es solución de (6.6)-(6.7) entonces $h(x_0) = 0$ pero $h(y(s))$ no es idénticamente nula; como además $h(y(s))$ es analítica, no puede tener acumulación de ceros, es decir, existe $\epsilon > 0$ tal que si $0 < |s| < \epsilon$ entonces $h(y(s)) \neq 0$. Luego, $y(s)$ sólo corta a $M_{(0)}$ en x_0 , a través de una reparametrización de esta solución obtenemos un punto de impasse de (6.1). Observemos finalmente que este análisis incluye también el caso en que $g(x_0)$ es transversal a $M_{(0)}$. \square

poner referencias internas

Caso II: $x_0 \in M_{(0)}^w$.

Por ser x_0 una singularidad esencial débil, sabemos que existe un entorno U de x_0 , tal que si $x \in U \cap M_0$, entonces $g(x) = 0$. Sabemos también por Teorema 6.16 que si $\nabla h(x) \neq 0$ en $U \cap M_0$ entonces el campo $F(x) = g(x)/h(x)$ puede extenderse en forma analítica a todo U . Luego, llamamos \tilde{F} a la extensión analítica del campo F , es decir, \tilde{F} está definido y es continuo en U .

Consideramos ahora la EDO

$$\dot{x} = \tilde{F}(x), \quad x \in U \tag{6.8}$$

que para $x \notin M_{(0)}$ tiene las mismas soluciones que (6.1).

Sabemos que la EDO (6.8) con condición inicial x_0 tiene una única solución; sea $x(t)$ esa solución. Para dar una respuesta al problema de existencia de solución de (6.1) con condición inicial x_0 , deberemos considerar las siguientes posibilidades.

1. La única solución de (6.8) por x_0 está contenida en $M_{(0)}$.
2. La única solución de (6.8) por x_0 intersecta a $M_{(0)}$ sólo en ese punto (es decir, es una curva contenida en $M_{(2)}$ salvo x_0).

Como las intersecciones de $x(t)$ con $M_{(0)}$ corresponden a los ceros de $h(x(t))$ que es una función analítica, estas dos son las únicas posibilidades (o $h(x(t))$ es constante igual a cero o los ceros son aislados).

Para determinar en cuál de los dos casos nos encontramos consideramos la EDA

$$\dot{x} = \tilde{F}(x) \tag{6.9}$$

$$0 = h(x) \tag{6.10}$$

Como en el Caso I se define X como la derivación correspondiente al campo

\tilde{F} y

$$\begin{aligned}\hat{\varphi}_0(x) &= h(x), \\ \hat{\varphi}_1(x) &= \frac{\partial \hat{\varphi}_0}{\partial x^j}(x) \cdot G^j(x) = X[\hat{\varphi}_0](x), \\ &\vdots \\ \hat{\varphi}_{\hat{k}}(x) &= X[\hat{\varphi}_{\hat{k}-1}](x)\end{aligned}$$

que generan una cadena de ideales

$$\hat{\mathcal{J}}_k = \langle \hat{\varphi}_0, \dots, \hat{\varphi}_k \rangle,$$

con $\hat{\mathcal{J}}_0 \subseteq \hat{\mathcal{J}}_1 \subseteq \dots \subseteq \hat{\mathcal{J}}_k \subseteq \dots$, y nuevamente se tiene una variedad final

$$\hat{N} = Z(\hat{\mathcal{J}}_k) = \{\hat{\varphi}_0(y) = 0, \dots, \hat{\varphi}_{\hat{k}}(y) = 0\}.$$

El siguiente resultado determina que un punto $x_0 \in M_{(0)}^w$, con $\nabla h(x_0) \neq 0$, no puede ser un punto de impasse, y establece una condición necesaria y suficiente para que resulte punto de cruce.

Teorema 6.18. *Sea la ecuación (6.1) y sea $x_0 \in M_{(0)}^w$. Supongamos válidas las hipótesis del comienzo de esta sección y supongamos también que $\nabla h(x_0) \neq 0$. Consideramos la EDA (6.9)-(6.10) y los conjuntos $Z(\hat{\mathcal{J}}_k)$, $k = 0, 1, \dots, \hat{k}$, definidos a partir de ella. Supongamos que $Z(\hat{\mathcal{J}}_k)$ es una subvariedad de $Z(\hat{\mathcal{J}}_0)$ para todo $k = 1, \dots, \hat{k}$. Entonces,*

1. *Si $x_0 \in \hat{N}$, entonces x_0 no es punto de cruce ni de impasse de (6.1) (la única solución de (6.8) con condición inicial en x_0 se mantiene en \hat{N} y luego en $M_{(0)}$).*
2. *Si $x_0 \notin \hat{N}$, entonces x_0 es punto de cruce de (6.1) (la única solución de (6.8) por x_0 no está contenida en \hat{N} y tampoco en $M_{(0)}$, entonces corta a $M_{(0)}$ sólo en x_0 , esta solución será una solución de cruce de (6.1) por x_0 .)*

Demostración. 1. Supongamos que $x_0 \in \hat{N}$, luego, la única solución de (6.8) por x_0 está contenida en $M_{(0)}$, veamos que x_0 no es punto ni de cruce ni de impasse de (6.1). En efecto, si x_0 fuera punto de cruce o de impasse, habría una solución de (6.1) (y por lo tanto de (6.8)) en $M_{(2)}$ que tendería en tiempo finito a x_0 , esta solución de (6.8) en $M_{(2)}$ podría extenderse de modo de tal obtener otra solución de (6.8) por x_0 , lo cual no puede ser por unicidad de solución de la EDO.

2. Supongamos que $x_0 \notin \hat{N}$. Luego, la única solución de (6.8) por x_0 intersecta a $M_{(0)}$ sólo en ese punto (es decir, es una curva contenida en $M_{(2)}$ salvo x_0). Veamos que x_0 resulta un punto de cruce de (6.1). En efecto, la misma curva es solución de (6.1) y (6.8) en $M_{(2)}$, y es solución de (6.1) en x_0 por continuidad, luego resulta una solución de cruce de (6.1). □

Caso III: $x_0 \in M_{(0)}^s$.

La técnica que describiremos para este caso puede aplicarse también a aquellas singularidades débiles a las que no pueda aplicarse el teorema 6.18.

El primer paso consiste en identificar las curvas invariantes. Una vez identificadas las curvas invariantes, el siguiente paso será analizar el sistema restringido a cada uno de esas curvas, para determinar si las soluciones contenidas en ellas tienen en x_0 un punto de cruce o un punto de impasse.

Para identificar curvas invariantes consideramos el Jacobiano $Dg(x_0)$. Sabemos por el corolario 6.11 que, para cada autovalor real no nulo λ , y para cada autovector v_λ , existen curvas C^1 tangentes a v_λ invariantes por el campo $g(x)$. En general estas curvas no serán únicas, salvo en el caso del autovalor con máxima parte real positiva y el autovalor con mínima parte real negativa.

Además de estas curvas invariantes, podría haber otras tangentes a autovectores de autovalor cero, pero para analizar la dinámica en estas direcciones sería necesario considerar información de mayor orden respecto del campo $g(x)$.

Existen resultados que en base a hipótesis adicionales respecto a ciertas ecuaciones diofánticas entre los autovectores garantizan una regularidad mayor que C^1 , pero para los objetivos de este trabajo nos alcanza con este nivel de regularidad.

Si v_λ no es tangente a $M_{(0)}$, entonces sabemos que las curvas invariantes tangentes a v_λ cortan a $M_{(0)}$ sólo en x_0 (localmente). Pero si v_λ es tangente a $M_{(0)}$ no tenemos forma de determinar si x_0 es la única intersección entre las curvas invariantes tangentes a v_λ o si, por el contrario, existen infinitas intersecciones, siendo x_0 un punto de acumulación de tales intersecciones. La técnica utilizada en los dos casos anteriores no puede ser utilizada aquí porque lo único que sabemos de las curvas en cuestión es que son C^1 , y para aplicar aquella técnica necesitaríamos que fueran C^ω .

Debido a esto nos concentraremos principalmente en los casos de autovalores no nulos y de autovectores que no sean tangentes a $M_{(0)}$.

Teorema 6.19. *Sea $x_0 \in M_{(0)}^s \cup M_{(0)}^w$. Consideremos el Jacobiano $Dg(x_0)$, un autovalor λ y un autovector correspondiente v_λ .*

1. Si $\lambda \neq 0$ y v_λ no es tangente a $M_{(0)}$ entonces las soluciones a lo largo de las curvas invariantes tangentes a v_λ alcanzan x_0 en tiempo finito y además
 - a) si $\nabla h(x_0) \cdot v_\lambda \neq 0$ entonces, estas soluciones son soluciones de cruce por x_0 ,
 - b) si $\nabla h(x_0) \cdot v_\lambda = 0$ entonces, estas soluciones tienen un punto de impasse en x_0 .
2. Si $\lambda \neq 0$ y v_λ es tangente a $M_{(0)}$ entonces
 - a) no hay soluciones de cruce tangentes a v_λ ,
 - b) si existen curvas invariantes tangentes a v_λ que corten a $M_{(0)}$ sólo en x_0 , entonces las soluciones a lo largo de estas curvas tienen un punto de impasse en x_0 .

Este resultado deja pendiente, por un lado, el caso de $\lambda = 0$, que requiere información de mayor orden, y por otro lado no da ninguna técnica que permita determinar la existencia de curvas invariantes como las requeridas para el punto 2b). A pesar de eso, este resultado es más abarcativo que resultados semejantes anteriores.

Observar además que este teorema no es redundante con el teorema 6.18, debido a que, bajo las hipótesis del teorema 6.18 resulta que $Dg(x_0)$ tiene a lo sumo un autovalor no nulo, ya que para todas las direcciones v tangentes a $M_{(0)}$ se cumple $Dg(x_0)v = 0$, y el teorema 6.19 no dice nada respecto a estas direcciones.

Previamente a la demostración del teorema 6.19 demostraremos el siguiente lema.

Lema 6.20. *Dado el sistema*

$$\tilde{h}(u)\dot{u} = \tilde{g}(u) \tag{6.11}$$

donde $\tilde{h}, \tilde{g} : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ son funciones C^1 , suponiendo que $\tilde{g}(0) = 0$, $\tilde{h}(0) = 0$, y \tilde{h} no se anula en ningún punto de $(-\epsilon, \epsilon) - \{0\}$ entonces el sistema (6.11) tiene existencia y unicidad de soluciones para $u \neq 0$ y además

1. Si $\tilde{g}'(0) \neq 0$ y $\tilde{h}'(0) \neq 0$ entonces $u = 0$ es un punto de cruce.
2. Si $\tilde{g}'(0) \neq 0$ y $\tilde{h}'(0) = 0$ entonces $u = 0$ es un punto de impasse de las soluciones de (6.11)

Demostración. Si $u \neq 0$ entonces el sistema (6.11) es equivalente a $\dot{u} = \tilde{g}(u)/\tilde{h}(u)$ que es C^1 ya que $\tilde{h}(u) \neq 0$. Por lo tanto hay existencia y unicidad de soluciones para $u \neq 0$.

Además,

1. Si $\tilde{g}'(0) \neq 0$ podemos suponer que $\tilde{g}(u) \neq 0$ en $(-\epsilon, \epsilon) - \{0\}$.

Si además $\tilde{h}'(0) \neq 0$ entonces $\tilde{g}/\tilde{h} \rightarrow \tilde{g}'(0)/\tilde{h}'(0)$ cuando $u \rightarrow 0$, supongamos que $\tilde{g}'(0)/\tilde{h}'(0)$ es positiva (el otro caso es análogo). Entonces tomamos dos soluciones de (6.11), u_1 y u_2 con valores iniciales $u_-(0) = -\epsilon/2$ y $u_+(0) = \epsilon/2$. Como $\tilde{g}(u)/\tilde{h}(u)$ está acotada fuera del cero, entonces existen $t_- > 0$ y $t_+ < 0$ tales que $u_-(t) \rightarrow 0$ cuando $t \rightarrow t_-^-$ y $u_+(t) \rightarrow 0$ cuando $t \rightarrow t_+^+$.

Construimos entonces la curva

$$\begin{cases} u_-(t + t_-) & t \leq 0 \\ u_+(t + t_+) & t \geq 0 \end{cases}$$

que es, por construcción, una curva continua, solución de (6.11) para $t \neq 0$, y $u(0) = 0$. Como para $t \neq 0$ se tiene $\dot{u}(t) = \tilde{g}(u(t))/\tilde{h}(u(t))$ resulta que $\lim_{t \rightarrow 0} \dot{u}(t) = \tilde{g}'(0)/\tilde{h}'(0)$, y por tratarse u de una función continua de una variable tal que existe el límite de su derivada, resulta que u es una curva C^1 . Además, u también es solución de (6.11) en $u = 0$ ya que en dicho punto la ecuación toma la forma $0\dot{u} = 0$. Por lo tanto $u = 0$ es un punto de cruce.

2. Supongamos en cambio que $\tilde{g}'(0) \neq 0$ pero $\tilde{h}'(0) = 0$. Entonces $\tilde{g}(u)/\tilde{h}(u) \rightarrow \pm\infty$. Por lo tanto ocurrirá, al igual que en el punto anterior, que $\tilde{g}(u)/\tilde{h}(u)$ está acotado fuera de cero, y por lo tanto las soluciones con valor inicial $u(0) \neq 0$ tenderán a cero en tiempo finito (positivo o negativo dependiendo del signo de $u(0)$ y de $\tilde{g}(u)/\tilde{h}(u)$).

Pero en este caso $\dot{u}(t)$ no está acotada y por lo tanto u no es derivable en $u = 0$, entonces las soluciones tienen un punto de impasse en $u = 0$ (ver definición 6.4).

□

Demostración del teorema 6.19. Consideremos una curva invariante $\gamma(u)$ tangente a v_λ en $u = 0$, esto es $\gamma(0) = x_0$ y $g(\gamma(u)) = \tilde{g}(u)\frac{d}{du}\gamma(u)$. Entonces la dinámica del sistema (6.2) restringido a la curva γ es equivalente a la dinámica del sistema

$$\tilde{h}(u)\dot{u} = \tilde{g}(u) \tag{6.12}$$

donde $\tilde{h}(u) = h(\gamma(u))$.

Dado que x_0 es una singularidad, resulta que $\tilde{h}(0) = 0$. Además $\tilde{h}'(0) = \nabla h(x_0) \frac{d}{du} \gamma(0) = \nabla h(x_0) v_\lambda$.

Por otro lado, como x_0 es una singularidad no esencial, resulta que $\tilde{g}(0) = 0$. Además $Dg(x_0) v_\lambda = \tilde{g}'(0) \gamma'(0) + \tilde{g}(0) \gamma''(0) = \frac{d}{du} \tilde{g}(0) v_\lambda$, de donde resulta que $\tilde{g}'(0) = \lambda$.

1. La existencia de curvas invariantes tangentes a v_λ está probada por el corolario 6.11. Dada una de esas curvas, consideramos el sistema (6.12). Como $\tilde{h}(0) = \tilde{g}(0) = 0$ y $\tilde{g}'(0) = \lambda \neq 0$, entonces por el lema 6.20
 - a) si $\tilde{h}'(0) = \nabla h(x_0) v_\lambda \neq 0$ entonces $x_0 = \gamma(0)$ es un punto de cruce de las soluciones contenidas en la curva γ ;
 - b) si $\tilde{h}'(0) = \nabla h(x_0) v_\lambda = 0$ entonces $x_0 = \gamma(0)$ es un punto de impasse de las soluciones contenidas en la curva γ ;
2. Si v_λ es tangente a $M_{(0)}$, resulta que $\tilde{h}'(0) = 0$, por lo tanto, en caso de existir curvas invariantes que corten a $M_{(0)}$ sólo en x_0 resultará que las soluciones a lo largo de esas curvas tendrán un punto de impasse en x_0 , por los mismos argumentos utilizados en el punto 1b).

□

Revisitamos ahora los ejemplos presentados en la sección 6.1, para aplicarles los resultados teóricos obtenidos en esta sección.

Ejemplo 6.4. En el ejemplo 6.1 se obtuvo que $M_{(0)} = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$. Se tiene además que $g(x_1, x_2) = [x_1 \quad -x_1 - x_2]^T$. Luego, $M_{(0)}^e = M_{(0)} - \{(0, 0)\}$. A partir de la secuencia:

$$\begin{aligned} \varphi_0 & : \quad x_1 + x_2 = 0, \\ \varphi_1 & : \quad x_2 = 0, \\ \varphi_2 & \equiv \varphi_1, \end{aligned}$$

obtenemos $\mathcal{J}_1 = \langle \varphi_0, \varphi_1 \rangle$ y $N = Z(\mathcal{J}_1) = \{(0, 0)\}$. Se tiene entonces por teorema 6.17 que $(x_1, x_2) \in M_{(0)}$ es punto de impasse si $(x_1, x_2) \neq (0, 0)$.

Por otro lado es fácil ver que $(0, 0)$ es una singularidad no esencial fuerte. En este caso $Dg(0, 0)$ tiene autovalores $\lambda_{1,2} = \pm 1$, y autovectores $v_1 = [1 \quad -2]^T$ y $v_2 = [0 \quad 1]^T$, respectivamente. Además $\nabla h(0, 0) = [1 \quad 1]$; $\nabla h(0, 0) \cdot v_1 = -1$ y v_1 no es tangente a $M_{(0)}$; y $\nabla h(0, 0) \cdot v_2 = 1$ y v_2 no es tangente a $M_{(0)}$. Luego, aplicando el teorema 6.19 (parte 1.a)), sabemos que las soluciones tangentes a las rectas con vectores directores v_1 y v_2 alcanzan el $(0, 0)$ en tiempo finito y son soluciones de cruce, en particular, en este ejemplo, estas soluciones coinciden con las rectas cuyos vectores directores son v_1 y v_2 .

Ejemplo 6.5. En el ejemplo 6.2 se obtuvo que $M_{(0)} = \{(x_1, x_2) \mid x_1 = 0\}$. Además se tiene que $g(x_1, x_2) = [x_1 \ 2x_2 + x_1^2]^T$. Luego, $M_{(0)}^e = M_{(0)} - \{(0, 0)\}$. En este caso se obtiene la secuencia:

$$\begin{aligned}\varphi_0 & : \ x_1, \\ \varphi_1 & \equiv \ \varphi_0,\end{aligned}$$

luego, $\mathcal{J}_0 = \langle \varphi_0 \rangle$ y $N = Z(\mathcal{J}_0) = \{x_1 = 0\}$. Por teorema 6.17 (x_1, x_2) es punto de impasse sii $x_1 \neq 0$. En este caso N coincide con $M_{(0)}$, luego no existen puntos de impasse en el conjunto de las singularidades esenciales.

Sólo falta analizar que ocurre en $(0, 0)$, que es una singularidad no esencial fuerte. En este caso $Dg(0, 0)$ tiene como autovalores $\lambda_1 = 1$ con autovector correspondiente $v_1 = [1 \ 0]^T$, y $\lambda_2 = 2$ con autovector correspondiente $[0 \ 1]^T$. A partir del teorema 6.19 podemos obtener las siguientes conclusiones:

- Se verifica que: $\lambda_1 \neq 0$, v_1 no es tangente a $M_{(0)}$ y $\nabla h(0, 0) \cdot v_1 = 1$, luego existe una solución de cruce tangente a v_1 por $(0, 0)$ (parte 1.a) del teorema 6.19). Como además λ_1 no es el máximo autovalor, existen infinitas soluciones de cruce tangentes a v_1 por $(0, 0)$, esto es así porque en el sistema linealizado hay infinitas soluciones que tienden al equilibrio de manera tangente al autovector cuyo autovalor no es el de parte real más alejada de cero, y las soluciones de cruce que estamos considerando se obtienen mediante un cambio de coordenadas C^1 y una reparametrización de estas soluciones.
- Como $\lambda_2 = 2$ es el máximo autovalor positivo hay una única curva invariante tangente a v_2 (es única porque en la linealización de g hay una única curva invariante tangente al autovector de autovalor máximo) y v_2 es tangente a $M_{(0)}$, entonces que no es solución de cruce (parte 2.a) del teorema 6.19). Pero dado que en este ejemplo $M_{(0)}$ es invariante por el flujo del sistema, resulta que la única curva invariante está contenida en $M_{(0)}$, y luego tampoco es una solución de impasse (parte 2.b) del teorema 6.19).

Ejemplo 6.6. En el ejemplo 6.3 se obtuvo $M_{(0)} = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 = 0\}$. Mediante cálculos simples se tiene que

$$g(x_1, x_2, x_3) = [2x_1x_3^2 - x_2x_3^2 \quad 2x_2x_3^2 + x_1x_3^2 \quad x_1^2 + x_2^2]^T.$$

Entonces $M_{(0)}^e = \emptyset$. Además todas las singularidades resultan ser no esenciales débiles. Pero como $\nabla h|_{M_{(0)}} \equiv 0$ no puede aplicarse el teorema 6.18.

Si calculamos el Jacobiano de g para intentar utilizar el teorema 6.19 obtenemos

$$Dg(x) = \begin{pmatrix} 2x_3^2 & -x_3^2 & (4x_1 - 2x_2)x_3 \\ x_3^2 & 2x_3^2 & (2x_1 + 4x_2)x_3 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}$$

que para $x_3 \neq 0$ tiene un autovalor nulo y dos autovalores complejos conjugados $-2 \pm i$. En este caso, esos autovalores complejos conjugados corresponden a soluciones que tienden en espiral al conjunto singular. Si bien no se aplican los teoremas que hemos probado antes, la misma técnica permite determinar que las reparametrizaciones de estas curvas dan lugar a soluciones que tienden en tiempo finito al conjunto singular. En este caso particular resultan ser soluciones con puntos de impasse, pero las técnicas desarrolladas no nos permitirían determinarlo.

Para el caso $x_3 = 0$ la matriz Jacobiana se anula, de modo que las técnicas que desarrollamos en este capítulo no resultan útiles.

Para terminar esta sección consideraremos nuevamente el ejemplo presentado en la sección anterior sobre la dinámica de una bola que rueda sobre un plano vertical. En este caso supondremos que uno de sus momentos de inercia es nulo y que la bola puede girar alrededor de su eje vertical.

Ejemplo 6.7. Consideramos un ejemplo similar al ejemplo 5.3: una bola rodando sobre un plano. Supondremos en este caso que uno de los momentos de inercia es nulo y que es posible que la bola gire alrededor del eje vertical (a diferencia de lo que ocurría en el ejemplo 5.3). Con cálculos muy similares a los hechos en aquel ejemplo se obtienen las ecuaciones siguientes (las variables y constantes que aparecen son como en el ejemplo 5.3) .

$$\begin{cases} (1 + M)\dot{u} \times e_3 + M\dot{v}_0 z \times e_3 = -M(v_0 u \times z) \times e_3 \\ \langle \dot{u}, e_3 \rangle = 0 \\ \dot{z} = u \times z \\ 0 = \langle u, z \rangle \\ 0 = \|z\|^2 - 1 \end{cases}$$

Las últimas tres ecuaciones son las mismas que se obtuvieron en el ejemplo 5.3, la segunda ecuación corresponde a que no haya torque en el eje vertical (de modo que la bola pueda rotar libremente en torno a ese eje), y la primera ecuación corresponde a la componente horizontal de la ecuación de conservación de momentos.

Al aplicar el algoritmo singular (visto en el capítulo 5) obtenemos $M_{(0)} = \{\|z\| = 1, \langle u, z \rangle = 0\}$, $M_{(1)} = \mathbb{R}^7 - M_{(0)}$. $M_{(0)}$ es la imagen encajada de $TS^2 \times \mathbb{R}$ en \mathbb{R}^7 , de modo que M^1 será $TS^2 \times \mathbb{R}$.

Al restringirse a M^1 resulta que $\langle \dot{u}, z \rangle = 0$, con esto, cuando $z \times e_3 \neq 0$ puede despejarse \dot{v}_0 , con lo cual resulta que el sistema tendrá solución única. Por lo tanto tenemos que $M^1_{(0)} = \{z_1 = z_2 = 0\}$, $M^1_{(2)} = M^1 - M^1_{(0)}$.

En $M^1_{(0)}$ puede verse que v_0 es una variable libre, y que una vez elegido el valor de v_0 las otras variables quedan unívocamente determinadas.

Intuitivamente, lo que hemos determinado hasta ahora es que, cuando el eje con momento de inercia nulo no está en posición vertical hay existencia y unicidad de soluciones, y que mientras dicho eje permanece en posición vertical la bola puede girar de cualquier manera (no hay unicidad de soluciones). Falta determinar de qué maneras puede la bola atravesar la posición con el eje vertical. Para eso necesitamos una parametrización adecuada de un entorno de $M^1_{(0)}$. Una parametrización útil en este caso consiste en tomar coordenadas polares

$$\begin{cases} z = (\rho \cos \theta, \rho \sin \theta, \sqrt{1 - \rho^2}) \\ v_0 = v_0 \\ u = \tau(-\sin \theta, \cos \theta, 0) + \mu(\cos \theta, \sin \theta, \rho/\sqrt{1 - \rho^2}) \end{cases}$$

Aquí debemos admitir que ρ tome también valores negativos, porque de otro modo las soluciones de cruce deberían presentar un salto en el valor de θ .

Pasado a estas coordenadas el sistema de ecuaciones toma la forma $a(x)\dot{x} = f(x)$ con

$$a(x) = \begin{pmatrix} 0 & -\tau \sin \theta + \mu \cos \theta & \cos \theta & \sin \theta & \lambda \rho \sin \theta \\ 0 & -\tau \cos \theta - \mu \sin \theta & -\sin \theta & \cos \theta & \lambda \rho \cos \theta \\ 0 & 0 & 0 & 0 & \rho \\ \cos \theta & -\rho \sin \theta & 0 & 0 & 0 \\ \sin \theta & \rho \cos \theta & 0 & 0 & 0 \end{pmatrix}$$

$$f(x) = \begin{pmatrix} -\lambda v_0(-\mu \cos \theta + \tau(1 - \rho^2) \sin \theta)/\sqrt{1 - \rho^2} \\ -\lambda v_0(\mu \sin \theta + \tau(1 - \rho^2) \cos \theta)/\sqrt{1 - \rho^2} \\ -\tau \rho \sqrt{1 - \rho^2} v_0 \\ (\mu \sin \theta + \tau(1 - \rho^2) \cos \theta)/\sqrt{1 - \rho^2} \\ (-\mu \cos \theta + \tau(1 - \rho^2) \sin \theta)/\sqrt{1 - \rho^2} \end{pmatrix}.$$

Notar que la tercera ecuación puede simplificarse dividiendo por ρ (este es el motivo por el cual resulta importante pasar a coordenadas polares, de otro modo más adelante quedaría $\nabla h = 0$).

Luego de simplificar resulta,

$$a(x) = \begin{pmatrix} 0 & -\tau \sin \theta + \mu \cos \theta & \cos \theta & \sin \theta & \lambda \rho \sin \theta \\ 0 & -\tau \cos \theta - \mu \sin \theta & -\sin \theta & \cos \theta & \lambda \rho \cos \theta \\ 0 & 0 & 0 & 0 & 1 \\ \cos \theta & -\rho \sin \theta & 0 & 0 & 0 \\ \sin \theta & \rho \cos \theta & 0 & 0 & 0 \end{pmatrix}$$

$$f(x) = \begin{pmatrix} -\lambda v_0(-\mu \cos \theta + \tau(1 - \rho^2) \sin \theta) / \sqrt{1 - \rho^2} \\ -\lambda v_0(\mu \sin \theta + \tau(1 - \rho^2) \cos \theta) / \sqrt{1 - \rho^2} \\ -\tau \sqrt{1 - \rho^2} v_0 \\ (\mu \sin \theta + \tau(1 - \rho^2) \cos \theta) / \sqrt{1 - \rho^2} \\ (-\mu \cos \theta + \tau(1 - \rho^2) \sin \theta) / \sqrt{1 - \rho^2} \end{pmatrix}$$

y premultiplicando por la adjunta de a se obtiene,

$$\rho \dot{x} = \begin{pmatrix} \tau \rho \sqrt{1 - \rho^2} \\ -\mu / \sqrt{1 - \rho^2} \\ (\lambda \rho v_0 \mu + \mu^2) / \sqrt{1 - \rho^2} \\ (-\lambda \tau \rho v_0 + \lambda \tau \rho^2 v_0) \sqrt{1 - \rho^2} - \tau \mu / \sqrt{1 - \rho^2} \\ -\tau \rho v_0 \sqrt{1 - \rho^2} \end{pmatrix} =: g(x)$$

Evaluando en $\rho = 0$ obtenemos

$$g|_{\rho=0} = \begin{pmatrix} 0 \\ -\mu \\ \mu^2 \\ -\tau \mu \\ 0 \end{pmatrix}.$$

Entonces para $\mu \neq 0$ obtenemos singularidades esenciales y para $\mu = 0$ obtenemos singularidades no esenciales fuertes.

Como $M_{(0)}^1$ es invariante por el flujo de g , entonces no hay puntos de impasse en las singularidades esenciales.

Falta analizar qué pasa en las singularidades fuertes. Para ello calculamos el Jacobiano de g

$$Dg|_{\rho=\mu=0} = \begin{pmatrix} \tau & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\lambda \tau v_0 & 0 & 0 & -\tau & 0 \\ -\tau v_0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Los autovalores de $Dg|_{\rho=\mu=0}$ son τ , $-\tau$ y 0 . (El caso $\tau = 0$ podemos obviarlo, porque en este caso la velocidad angular de la bola sería vertical y estaríamos en el caso que la bola permanece con el eje de inercia nula vertical).

Los autovectores de autovalor 0 son tangentes a $M^1_{(0)}$.

Los autovectores de autovalor $-\tau$ están generados por

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

Los autovectores de autovalor τ están generados por

$$\begin{pmatrix} 2\tau \\ \lambda v_0 \\ 0 \\ -\lambda\tau v_0 \\ -2\tau v_0 \end{pmatrix}$$

Analizamos primero los autovectores de autovalor $-\tau$. Como es el autovalor negativo más alejado del cero, resulta que hay una única solución tangente a este autovector, además el autovector es tangente a $M^1_{(0)}$ y $M^1_{(0)}$ es invariante por el flujo de g , por lo tanto la única curva invariante por el flujo de g tangente al autovector de autovalor $-\tau$ permanecerá dentro de $M^1_{(0)}$, y por lo tanto no puede contener soluciones de cruce ni de impasse.

Analizamos ahora los autovectores de autovalor τ . Como es el autovalor positivo más alejado del cero, hay una única curva invariante tangente a cada autovector. Además el autovector no es tangente a $M^1_{(0)}$ y el autovalor es no nulo, de modo que esa curva corresponde a una solución de cruce. Es decir, para cada combinación de valores de θ , τ , v_0 con $\tau \neq 0$ hay una única solución de cruce.

Para analizar que sucede con los autovectores de autovalor nulo los resultados obtenidos no son suficientes. Una manera de mirar este caso sería estudiar la dinámica en la variedad centro [40].

Capítulo 7

Conclusiones

Para resolver ecuaciones diferenciales implícitas $\varphi(x, \dot{x}) = 0$, una estrategia bastante común consiste en construir una secuencia de conjuntos (M_k) , de modo que M_{k+1} sea el conjunto de los puntos x tales que existe un vector v tangente en x a M_k que resuelva la ecuación $\varphi(x, v) = 0$ en ese punto. En esta tesis hemos visto este tipo de algoritmos desde diferentes puntos de vista en los distintos capítulos.

Quizá uno de los algoritmos de este tipo más conocidos sea el algoritmo de Gotay–Nester (que describimos en el capítulo 3). La particularidad de la formulación de este algoritmo es que se basa completamente en una estructura geométrica subyacente en el problema. Dos ventajas que se obtienen por formular el algoritmo de esta manera es que la descripción del algoritmo obtenida es global y que el algoritmo resultará invariante por transformaciones que no modifiquen la estructura geométrica. En particular, en el caso del algoritmo de Gotay–Nester, esto implica que puede reducirse un problema en base a simetrías que conserven la estructura presimpléctica y el algoritmo puede pasarse al cociente sin inconvenientes.

Uno de los aportes de esta tesis consiste en la definición de las estructuras de Dirac generalizadas y la definición de un algoritmo similar al de Gotay–Nester para estas estructuras (capítulo 4). De este modo se obtiene un algoritmo basado en una estructura geométrica para una clase más amplia de problemas, entre ellos los sistemas mecánicos con vínculos que no conservan la energía y circuitos eléctricos RLC. Esta estructura permite una descripción global del algoritmo.

Lo que no permite el algoritmo basado en estructuras de Dirac generalizadas es el paso al cociente por un grupo de simetrías, debido a que la estructura geométrica depende fuertemente del fibrado $TM \oplus T^*M$, pero al reducir por un grupo de simetrías que actúen en M , obteniendo una varie-

dad cociente N , el fibrado que resulta de reducir $TM \oplus T^*M$ no resulta ser $TN \oplus T^*N$, sino un fibrado más grande. Cabe aclarar que este mismo tipo de dificultades aparecen en las estructuras de Dirac (no generalizadas). Una posible estrategia para avanzar en trabajos futuros sería definir estructuras geométricas similares a estas en fibrados de la forma $V \oplus V^*$ donde V sea cualquier algebroid de Lie sobre M . Los algebroides de Lie son generalizaciones del fibrado tangente, suficientemente generales como para permitir la reducción al cociente por un grupo de Lie, pero suficientemente particulares como para permitir definir ecuaciones diferenciales en términos de ellas.

La limitación en todos estos algoritmos es que requieren que los conjuntos M_k obtenidos sean subvariedades. Cuando aparecen singularidades el algoritmo no puede continuarse preservando la estructura geométrica. Una solución a esta cuestión es el algoritmo de desingularización, que de manera local encuentra todas las soluciones cuando los datos del problema son analíticos (presentado en el capítulo 5). Otro de los aportes de esta tesis consistió en avanzar en la solución del problema de valores iniciales, cuando los valores iniciales están en el conjunto singular (capítulo 6).

La solución obtenida para este problema utiliza sólo información de primer orden. Un camino a explorar consiste en intentar utilizar información de mayor orden. La principal limitación para esto es la necesidad de garantizar la existencia de curvas invariantes con mayor grado de regularidad. En algunos casos tales resultados existen (por ejemplo, cuando los autovectores no satisfacen ciertas ecuaciones diofánticas), pero no cuando los autovalores son cero.

Otro problema que no fue estudiado en este trabajo es el de soluciones de cruce tales que su derivada se anule en el punto de cruce, es decir, soluciones que tienden en tiempo finito a un equilibrio singular.

Finalmente, otro problema interesante y en el que no hay avances es el siguiente: cuando aparecen singularidades en un algoritmo de restricciones que utiliza alguna estructura geométrica, como por ejemplo el de Gotay–Nester o el de estructuras de Dirac generalizadas, ¿puede utilizarse esa estructura geométrica para calcular más fácilmente la desingularización? ¿puede pasarse la estructura geométrica a la desingularización de modo de poder continuar con el algoritmo?

Resumiendo, los aportes de esta tesis consisten en: la definición de una estructura geométrica que permite aplicar un algoritmo similar al de Gotay–Nester a una familia más amplia de problemas; para los casos en que este

tipo de algoritmos no funcionan porque aparecen singularidades, avanzamos en la solución del problema de valores iniciales cuando los valores iniciales están en el conjunto singular.

Apéndice A

Demostración del teorema de la variedad estable

La siguiente es una demostración del Teorema (6.7).

Demostración. Consideraremos $x_0 = 0$, cualquier otro caso puede ser reducido a este mediante un cambio de coordenadas.

Si llamamos $A = Dg(0)$ entonces la ecuación (6.3) puede escribirse como

$$\dot{x} = Ax + G(x) \tag{A.1}$$

donde $G(0) = 0$ y $DG(0) = 0$. Debido a esto, dado $L > 0$ arbitrario existe $\delta > 0$ tal que

$$\|x\|, \|\bar{x}\| < \delta \implies \|G(x) - G(\bar{x})\| < L\|x - \bar{x}\|. \tag{A.2}$$

Como los autovalores de $A|_E$ tienen parte real menor que $-\alpha$ y los de $A|_F$ tienen parte real mayor que $-\alpha$, entonces existe un $\sigma > 0$ tal que los autovalores de $A|_E$ tienen parte real menor que $-\alpha - \sigma$ y los de $A|_F$ tienen parte real mayor que $-\alpha + \sigma$.

Como E y F son subespacios invariantes para A , entonces también lo serán para $\exp(tA)$, de modo que podemos escribir

$$\exp(tA) = U_E(t) + U_F(t)$$

de manera tal que se cumplan $\text{Im}U_E(t) = \ker U_F(t) = E$ y también $\text{Im}U_F(t) = \ker U_E(t) = F$.

Ahora, los autovalores de $A|_E$ tiene parte real menor que $-\alpha - \sigma$, esto implica que existe un $K > 0$ tal que para $t \geq 0$

$$\|U_E(t)\| = \|U_E(t)|_E\| = \|\exp(tA|_E)\| \leq Ke^{-t(\alpha+\sigma)}. \tag{A.3}$$

De modo similar, tenemos que los autovalores de $-A|_F$ tienen parte real menor que $\alpha - \sigma$, y obtenemos

$$\|U_F(-t)\| = \|U_F(-t)|_F\| = \|\exp(-tA|_E)\| \leq Ke^{t(\alpha-\sigma)} \quad (\text{A.4})$$

para todo $t \geq 0$.

Consideramos ahora la ecuación integral

$$\theta(t, a) = U_E(t)a + \int_0^t U_E(t-s)G(\theta(s, a))ds - \int_t^\infty U_F(t-s)G(\theta(s, a))ds. \quad (\text{A.5})$$

Un cálculo directo usando que $\frac{d}{dt}U_E = AU_E$, $\frac{d}{dt}U_F = AU_F$ y que $U_E(0) + U_F(0) = I$ muestra que toda solución de (A.5) es también solución de (A.1).

Ahora construiremos soluciones de (A.5) por el método de aproximaciones sucesivas

$$\begin{aligned} \theta_0(t, a) &= 0 \\ \theta_{n+1}(t, a) &= U_E(t)a + \int_0^t U_E(t-s)G(\theta_n(s, a))ds \\ &\quad - \int_t^\infty U_F(t-s)G(\theta_n(s, a))ds. \end{aligned} \quad (\text{A.6})$$

Notar que en esta construcción sólo importa la componente de a en E , ya que F es el núcleo de $U_E(t)$.

Para demostrar la convergencia de esta sucesión, demostraremos por inducción que para a suficientemente chico y $t \geq 0$ se cumple

$$\|\theta_{n+1}(t, a) - \theta_n(t, a)\| \leq \frac{K\|a\|e^{-\alpha t}}{2^n} \quad (\text{A.7})$$

donde K es la constante que aparece en las inecuaciones (A.3) y (A.4). Elegimos la constante L en (A.2) suficientemente pequeña para que se cumpla

$$4KL/\sigma < 1$$

y elegimos a suficientemente chico para que $\|a\| \leq \frac{1}{2}$ y

$$2K\|a\| < \delta.$$

Notar que (A.7) implica que para $t \geq 0$

$$\|\theta_n(t, a)\| \leq 2K\|a\|e^{-\alpha t} < \delta e^{-\alpha t} \leq \delta,$$

por lo tanto las curvas $\theta_n(\cdot, a)$ permanecen dentro de la región donde vale (A.2).

El caso base de la inducción, para $n = 0$, la ecuación (A.7) sigue inmediatamente de (A.3).

Para el paso inductivo, el término $U_E(t)a$ se cancela, y debemos acotar las dos integrales. Para la primera usamos (A.3) y (A.2) junto con la hipótesis inductiva (A.7) para obtener

$$\begin{aligned} & \left\| \int_0^t U_E(t-s)(G(\theta_n(s,a)) - G(\theta_{n-1}(s,a)))ds \right\| \\ & \leq \int_0^t \|U_E(t-s)\| \cdot L \|\theta_n(s,a) - \theta_{n-1}(s,a)\| ds \\ & \leq \int_0^t K e^{-(t-s)(\alpha+\sigma)} \cdot L \frac{K\|a\|e^{-\alpha s}}{2^{n-1}} ds \\ & = \frac{K\|a\|e^{-\alpha t}}{2^n} 2KL \int_0^t e^{-(t-s)\sigma} ds < \frac{K\|a\|e^{-\alpha t}}{2^n} \frac{2KL}{\sigma}. \end{aligned}$$

Una cuenta similar usando (A.4) en lugar de (A.3) nos lleva a

$$\begin{aligned} & \left\| \int_t^\infty U_F(t-s)(G(\theta_n(s,a)) - G(\theta_{n-1}(s,a)))ds \right\| \\ & \leq \int_t^\infty K e^{-(t-s)(\alpha-\sigma)} \cdot L \frac{K\|a\|e^{-\alpha s}}{2^{n-1}} ds \\ & = \frac{K\|a\|e^{-\alpha t}}{2^n} 2KL \int_t^\infty e^{(t-s)\sigma} ds = \frac{K\|a\|e^{-\alpha t}}{2^n} \frac{2KL}{\sigma}. \end{aligned}$$

Usando estas dos cotas se obtiene

$$\|\theta_{n+1}(t,a) - \theta_n(t,a)\| < \frac{4KL}{\sigma} \frac{K\|a\|e^{-t\alpha}}{2^n} < \frac{K\|a\|e^{-t\alpha}}{2^n}$$

con lo que se completa la inducción.

Por lo tanto, la sucesión θ_n converge uniformemente a θ y se cumple la desigualdad

$$\|\theta(t,a)\| \leq 2K\|a\|e^{-t\alpha} < \delta e^{-t\alpha}.$$

Ahora consideremos una solución $x(t)$ que cumple $\|x(t)\| < \delta e^{-t\alpha}$; vamos a demostrar que $x(t) = \theta(t, x(0))$.

$$x(t) = U_E(t)x(0) + U_F(t)b + \int_0^t U_E(t-s)G(x(s))ds - \int_t^\infty U_F(t-s)G(x(s))ds \quad (\text{A.8})$$

donde

$$b = x(0) + \int_0^\infty U_F(-s)G(x(s))ds.$$

Cálculos similares a los hechos más arriba muestran que tres de los términos de (A.8) están acotados por múltiplos de $e^{-t\alpha}$, salvo el término $U_F(t)b$. Para evaluar este término, sean $b_E \in E$ y $b_F \in F$ la descomposición de $b = b_E + b_F$. Entonces $U_F(t)b = U_F(t)b_F$. Además

$$\|b_F\| = \|U_F(-t)U_F(t)b_F\| \leq \|U_F(-t)\| \|U_F(t)b_F\| \leq Ke^{t(\alpha-\sigma)}\|U_F(t)b_F\|$$

de donde

$$\|U_F(t)b\| = \|U_F(t)b_F\| \geq \frac{\|b_F\|}{K}e^{-t(\alpha-\sigma)}.$$

Si $\|b_F\| \neq 0$, para t suficientemente grande el término $U_F(t)b$ es más grande que todos los otros términos de (A.8), haciendo imposible que se cumpla la cota $x(t) \leq \delta e^{-t\alpha}$. En consecuencia tenemos que

$$x(t) \leq \delta e^{-t\alpha} \quad \forall t > 0 \implies x(t) = \theta(t, x(0)).$$

Además tenemos

$$x(0) = U_E(0)x(0) - \int_0^\infty U_F(-s)G(\theta(s, x(0)))ds.$$

El término $a = U_E(0)x(0)$ es la componente de $x(0)$ en E . Por otro lado $\theta(s, x(0))$ sólo depende de esta componente (como ya habíamos notado en la construcción de θ). Si definimos $\Psi: E \rightarrow F$,

$$\Psi(a) = - \int_0^\infty U_F(-s)G(\theta(s, a))ds$$

resulta que

$$x(0) = a + \Psi(a)$$

es decir, $x(0)$ está en la gráfica de la función Ψ . Como Ψ es una función con el mismo grado de regularidad que g , su imagen es una variedad a la que llamamos W . Además, por unicidad de soluciones, cualquier punto de la curva $x(t)$ con $t \geq 0$ sirve como valor inicial de una solución que cumple la misma cota $\|x(t)\| < \delta e^{-t\alpha}$, por lo tanto toda la curva $x(t)$ $t > 0$ está contenida en W .

Luego W es una variedad invariante tal que las soluciones $x(t)$ con valores iniciales en W cumplen $\|x(t)\| < \delta e^{-t\alpha}$ siempre que $a = U_E(0)x(0)$ sea suficientemente pequeño. Para ver que es tangente a E acotamos

$$\|\Psi(a)\| = \left\| - \int_0^\infty U_F(-s)G(\theta(s, a))ds \right\| \leq \int_0^\infty Ke^{s(\alpha-\sigma)}L\delta e^{-s\alpha}ds \leq \frac{KL\delta}{\sigma}.$$

Esta cota es válida para $\|a\| < \delta/2K$. Si recordamos que L podía ser elegido arbitrariamente pequeño (y que K y σ no dependen de δ) esto demuestra que W , la gráfica de Ψ , es tangente al plano E . \square

Bibliografía

- [1] **Balseiro P. and Solomín J.** *On generalized non-holonomic systems* Lett. Math. Phys., vol. 84 n. 1 pp. 15–30 (2008)
- [2] **Cendra H. and Etchechoury M.** *Desingularization of implicit analytic differential equations* J. Phys. A:Mat. Gen., vol. 39, pp. 10975–11001 (2006)
- [3] **Cendra H. and Grillo S.** *Generalized nonholonomic mechanics, servomechanisms and related brackets* J. Math. Phys., vol. 47, n. 2 (2006)
- [4] **Cendra H. and Grillo S.** *Lagrangian Systems with higher order constraints* J. Math. Phys., vol. 48, n. 5 (2007)
- [5] **Cendra H., Ibort A., de León M. and Martín De Diego D.** *A generalization of Chetaev's for a class of higher order nonholonomic constraints* J. Math. Phys., vol. 45, pp. 2785–2801 (2004)
- [6] **Courant T. J.** *Dirac manifolds* Transactions of the AMS, vol. 319, n. 2 (1990)
- [7] **Dirac P.** *Generalized Hamiltonian dynamics* Canadian Journal of Mathematics, vol. II, n. 2, pp. 129–148 (1950)
- [8] **Dirac P.** *Generalized Hamiltonian dynamics* Proc. R. Soc. London, vol. 246, n. 1246, pp. 326–332 (1958)
- [9] **Dirac P.** *Lectures on Quantum Mechanics* Belfer Graduate School of Science, Yeshiva University (1964)
- [10] **Gotay M. Nester J. and Hinds G.** *Presymplectic manifolds and the Dirac–Bergmann theory of constraints* J. Math. Phys., vol. 19, n. 11, pp. 2388–2399 (1978)
- [11] **Pritchard F. L.** *On implicit systems of differential equations* J. Diff. Equations, vol. 194, pp. 328–363 (2003)

- [12] **Rabier P. and Rheinboldt W.** *A geometric treatment of implicit differential–algebraic equations* J. Diff. Equations, vol. 109, pp. 110–146 (1994)
- [13] **Rheinboldt W.** *Differential–algebraic systems as differential equations on manifolds* Mathematics of Computation, vol. 43, n. 186, pp. 73–482 (1984)
- [14] **Yoshimura H. and Marsden J.** *Dirac Structures in Lagrangian Mechanics. Part I: Implicit Lagrangian Systems* J. Geom. Phys., vol. 57, pp. 133–156 (2006)
- [15] **Cendra H., Etchehoury M. and Ferraro S.** *An extension of the Dirac and Gotay-Nester theories of constraints for Dirac dynamical systems*, aceptado para su publicación en J. Geom. Mechanics., arXiv:1208.1953 (2012)
- [16] **Maschke B.M., van der Schaft A.J. and Breedveld P.C.** *An Intrinsic Hamiltonian Formulation of the Dynamics of LC-Circuits* IEEE Trans. Cir. and Syst., vol. 42, No. 2, pp. 73–82 (1995)
- [17] **Marsden J.E. and Ratiu T.S.** *Introduction to Mechanics and Symmetry* (Berlin: Springer) (1994)
- [18] **Cendra H., Marsden J.E. and Ratiu T.S.** *Geometric mechanics, Lagrangian reduction and nonholonomic systems* Mathematics Unlimited-2001 and Beyond (Berlin: Springer), pp. 221–273 (2001)
- [19] **Riaza R.** *Differential-Algebraic Systems. Analytical Aspects and Circuit Applications* (World Scientific) (2008)
- [20] **Delgado Tellez M. and Iborn A.** *A Panorama of Geometrical Optimal Control Theory*. Extracta Mathematicae, vol. 18 No. 2, pp. 129-151 (2003).
- [21] **Gear C.W.** *The simultaneous numerical solution of differential-algebraic equations* IEEE Trans. Circ. Theory pp. 89-95 (1971)
- [22] **Lojasiewicz S.** *Sur les ensembles semi-analytiques*. Actes du Congress International des Mathématiciens (Nice) vol. 2 pp. 237-241 (1970)
- [23] **Hironaka H.** *Subanalytic sets*. Number Theory, Algebraic Geometry and Commutative Algebra (Tokyo: Kinokuniya) pp. 453-493 (1973)

- [24] **Bierstone E. and Milman P.** *Semianalytic and subanalytic sets.* Publ. Math. l'I.H.E.S. 67 pp. 5-42 (1988)
- [25] **Sussmann H.J.** *Real analytic desingularization and subanalytic sets: an elementary approach.* Trans. Am. Math. Soc. 317 pp. 417-461 (1990)
- [26] **Rabier P.J. and Rheinboldt W.C.** *On impasse points of quasilinear differential-algebraic equations.* J. of Math Anal. and Appl. 181 pp. 429-454 (1994)
- [27] **Chua L.O. and Deng A.C.** *Impasse Points I, Numerical Aspects.* Int. J. Circ. Theory Appl. 17 pp. 213-235 (1989)
- [28] **Marszalek, W., Amdeberhan, T. and Riaza, R.** *Singularity Crossing Phenomena in DAEs: A Two-Phase Fluid Flow Application Case Study* Comp. and Math. with Appl. J. Math. Anal. and Appl. 49 pp. 303-319 (2005)
- [29] **Riaza R. and Zufiria P.** *Weak Singularities and the Continuous Newton Method.* J. Math. Anal. and Appl. 236 pp. 438-462 (1999).
- [30] **Reich, S.** *On an existence and uniqueness theory for nonlinear differential-algebraic equations.* Circuits Systems Signal Process. 281 pp. 343-359 (1991).
- [31] **Cendra H., Ibort A., de León M. and Martin de Diego D.** *A generalization of Chetaev's principle for a class of higher order nonholonomic constraints* J. Math. Phys., v. 45, n. 7, pp. 2785–2801 (2004).
- [32] **Cendra H. and Etchehoury M.** *Rolling of a symmetric sphere on a horizontal plane without sliding or spinning* Rep. Math. Phys., v. 57, n. 3, pp. 367–374 (2006)
- [33] **Perko L.** *Differential Equations and Dynamical Systems* Texts in Applied Mathematics 7. Springer. (1991)
- [34] **Coddington E. and Levinson N.** *Theory of Ordinary Differential Equations* (Bombay-Nueva Delhi: Tata McGraw-Hill) (1972)
- [35] **Hirsch M. W.** *Differential Topology.* Graduate Texts in Mathematics 33. Springer-Verlag. (1976)
- [36] **Abraham R., Marsden J. E. and Ratiu T.** *Manifolds, Tensor Analysis, and Applications* Applied Mathematical Sciences 75. Springer-Verlag. (segunda edición, 1988)

- [37] **Lang S.** *Differential and Riemannian manifolds* Graduate Texts y Mathematics 160. Springer–Verlag. (tercera edición, 1995)
- [38] **Hartman P.** *On local homeomorphisms of Euclidean spaces.* Bol. Soc. Mat. Mexicana 5 pp. 220-241 (1960).
- [39] **Rocard Y.** *L instabilite en mecanique; automobiles, avions, ponts suspendus.* Masson, Paris (1954).
- [40] **Guckenheimer J. and Holmes P.** *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* Applied Mathematical Sciences 42. Springer. (1983).