



UNIVERSIDAD CARLOS III DE MADRID
DEPARTAMENTO DE MATEMÁTICAS

TESIS DOCTORAL

**Reducción de principios variacionales
con simetría y problemas de control óptimo
de Lie-Scheffers-Brockett.**

AUTORA: THALÍA RODRÍGUEZ DE LA PEÑA

DIRECTOR: ALBERTO IBORT LATRE

LEGANÉS, 2009

**Reducción de principios variacionales
con simetría y problemas de control óptimo
de Lie-Scheffers-Brockett.**

Memoria de Tesis Doctoral presentada por
Thalía Rodríguez de la Peña
Licenciada en Ciencias Matemáticas
para optar al grado de doctor en Ciencias Matemáticas.

Dirigida por
Alberto Ibort Latre
Catedrático de Matemática Aplicada.

2009

Para mi gran familia,

AGRADECIMIENTOS

Después de terminar este trabajo, que no ha sido fácil para mí, quiero agradecer a aquellas personas que me han apoyado de una manera especial durante todo este tiempo.

En primer lugar quiero agradecer especialmente a mi director Alberto Ibort por todo su apoyo y sus enseñanzas durante todos estos años. Así como las ideas aportadas y el seguimiento continuo del avance de las mismas. También quiero agradecerle el tiempo invertido y el esfuerzo dedicado, sin los cuales este trabajo nunca habría visto la luz. Gracias también por su continua atención a mis problemas familiares, que han alargado la duración de esta memoria.

También me han sido de gran ayuda las reuniones científicas mantenidas en el seminario de Geometría-Física de la UCM donde además de conocimiento he encontrado unos compañeros estupendos, que me han ayudado con su alegría y sus bromas. Quiero agradecerles a todos ellos que hayan estado ahí.

He tenido la suerte de estar en un departamento con gente maravillosa, entre ellos especial mención a Daniel y Rebecca, y sobre todo gracias a Marina Delgado-Téllez por apoyarme todos estos años.

Luego, aunque no les haya podido ver estos años tanto como haya querido, aunque su amistad ha sido fundamental, a mis amigas, Ana, Silvia y María.

A mi familia, a quién debo demasiado como para poder expresarlo con palabras, que ha estado a mi lado en todo momento, empujándome a mejorar en mi vida, con un inmenso cariño.

Índice general

1. Introducción	1
1.1. Preámbulo: Variables de Clebsch y ligaduras de Lin en fluidos incompresibles.	1
1.2. Sistemas con simetría: El sólido rígido	5
1.3. Control óptimo y sistemas lagrangianos con simetría: El sólido rígido de nuevo	6
1.4. Objetivos	7
1.5. Algunas ideas, técnicas y organización de la memoria.	8
2. El problema del cálculo de variaciones con simetría.	11
2.1. El problema del cálculo de variaciones en mecánica	11
2.1.1. El problema general del cálculo de variaciones	11
2.1.2. Métodos directos	12
2.1.3. Caracterización de puntos críticos: Ecuaciones de Euler-Lagrange	14
2.1.4. El caso del cálculo de variaciones.	15
2.1.5. El marco analítico para el cálculo de variaciones en Mecánica	15
2.1.6. Variedades de Hilbert de curvas en variedades riemannianas	16
2.1.7. Fibrados tangentes y otros fibrados sobre $\mathcal{P}_I^k(M)$	18
2.1.8. Algunas subvariedades de $\mathcal{P}_I^k(M)$	20
2.2. Principios variacionales con simetría en mecánica: reducción simpléctica	23
2.2.1. Acciones de grupos de Lie	23
2.2.2. Principios variacionales con simetría en Mecánica	25
2.2.3. Elementos geométricos en Mecánica.	28
2.2.4. Reducción de las ecuaciones de Euler-Lagrange	29
2.2.5. Reducciones y cocientes de espacios de caminos	31
2.3. El problema general del análisis de puntos críticos de funciones con simetrías y ligaduras	32
2.3.1. Puntos críticos de funcionales con simetría.	32

2.3.2.	El problema de la reducción y los principios variacionales con simetría	35
2.3.3.	Puntos críticos de funcionales con ligaduras: Teorema de los multiplicadores de Lagrange	36
3.	Variables de Clebsch y Ligaduras de Lin en problemas variacionales con simetría de Borel	41
3.1.	Conexiones y paralelismo en fibrados asociados	41
3.1.1.	Fibrados asociados a fibrados principales	42
3.2.	Reducción variacional de Lagrange-Poincaré	45
3.2.1.	Conexiones inducidas en fibrados asociados $P \times_G Q$	47
3.3.	Espacios de curvas horizontales sobre fibrados asociados	48
3.3.1.	Condiciones extremales compatibles	49
3.3.2.	Variables de Clebsch y Ligaduras de Lin para problemas variacionales con simetría	50
3.3.3.	Teorema de los multiplicadores de Lagrange sobre espacios de curvas	52
4.	Problemas de Control Óptimo de Lie-Scheffers-Brockett	55
4.1.	Elementos de Teoría de Control.	55
4.1.1.	Sistemas de Control Óptimo	55
4.1.2.	El principio variacional de Hamilton-Pontryagin	57
4.2.	La geometría del principio de Hamilton-Pontryagin: El fibrado de Pontryagin	60
4.3.	Sistemas de Control Óptimo de tipo Lie-Scheffers-Brockett.	66
4.4.	Un teorema de dualidad para problemas de Control Óptimo de Lie-Scheffers-Brockett y sistemas Lagrangianos con simetría	70
4.4.1.	Representación de Clebsch de sistemas de Lie-Scheffers-Brocket de control óptimo	70
4.4.2.	Puntos críticos del Lagrangiano de Clebsch	71
4.4.3.	Equivalencias entre las distintas representaciones del problema de Lie-Scheffers-Brockett	71
5.	Estructuras de Dirac y la Teoría de Control Óptimo	73
5.1.	Estructuras de Dirac	73
5.1.1.	La estructura de Dirac de un problema de control óptimo	75
5.1.2.	Simetrías y Sistemas de Control Óptimo	76
5.1.3.	Reducción de Lie-Dirac de sistemas presimplécticos.	78
5.1.4.	Reducción de la estructura de Dirac D_Ω de un problema de Control Óptimo	80

5.1.5. Descripción de los fibrados reducidos	81
5.1.6. El principio variacional reducido de Lagrange-Pontryagin	83
5.1.7. El principio del Máximo de Pontryagin Reducido	85
5.2. Reducción de Lie-Dirac y el Principio del Máximo de Pontryagin reducido	86
6. Aplicaciones y ejemplos.	89
6.1. Las ecuaciones de Euler para el sólido rígido	89
6.1.1. Ecuaciones para el grupo $SO(3)$	89
6.1.2. Ecuaciones para el grupo $SU(2)$	90
6.2. Ecuaciones de Ricatti	91
6.2.1. Ecuaciones para el grupo $SL(2, \mathbb{R})$	91
6.2.2. Ecuaciones para el grupo $SU(2)$	93
6.2.3. Ecuaciones para el grupo $SO(2, 1)$	93
6.3. Las ecuaciones simétricas de "Elroy's beanie"	94
6.3.1. Ecuaciones reducidas de Elroy's beanie	96
6.4. Un ejemplo simple en dimensión finita	98
6.4.1. Hamiltonianos de Control Cuántico de tipo Lie-Scheffers- Brockett.	98
6.4.2. Control Óptimo de Sistemas Cuánticos y representación de Clebsch.	100
7. Conclusiones y problemas abiertos.	105
7.1. Conclusiones y resultados	105
7.2. Algunos problemas abiertos	106
Bibliografía	109

Capítulo 1

Introducción

1.1. Preámbulo: Variables de Clebsch y ligaduras de Lin en fluidos incompresibles.

El problema que vamos a analizar en esta memoria tiene sus orígenes en el estudio de principios variacionales para fluidos.

El problema que dio lugar históricamente a la introducción de las llamadas variables de Clebsch y las ligaduras de Lin, es el análisis de la derivación a partir de un principio variacional, de las ecuaciones de Euler para un fluido ideal incompresible en una región $\Omega \in \mathbb{R}^n$ (n típicamente 2 ó 3) con frontera $\partial\Omega$ que supondremos fija y regular, descrito por un campo de velocidades $\mathbf{v}(\mathbf{x}, t)$, que satisface la ecuación de Euler:

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \cdot \mathbf{v} = -\frac{1}{\rho} \nabla p ; \mathbf{v} = \mathbf{v}(\mathbf{x}, t), \\ \mathbf{x} \in \Omega \subset \mathbb{R}^n, n = 2, 3, \mathbf{v} \parallel \partial\Omega, \operatorname{div} \mathbf{v} = 0, \\ \rho = \text{cte.} \end{array} \right. \quad (1.1)$$

En esta descripción Euleriana de la dinámica de un fluido no es trivial establecer las ecuaciones del movimiento a partir de un principio variacional. Recordemos que los principios variacionales nos ayudan a establecer teoremas que prueban la existencia de soluciones de las ecuaciones de Euler-Lagrange, lo cual resulta particularmente difícil para las ecuaciones (1.1), constituyendo incluso hoy en día un problema abierto en general.

El principio variacional natural para el problema (1.1) consistiría en considerar la densidad lagrangiana:

$$L(\mathbf{v}, \rho) = \frac{1}{2} \rho \mathbf{v}^2 - \rho (e + U)$$

que representa la energía cinética de un elemento de volumen del fluido menos su energía potencial, así e representaría la energía interna del fluido y U su energía potencial mecánica (por tanto, para fluidos libres $U = 0$).

Sin embargo, el funcional

$$S(\mathbf{v}, \rho) = \int_0^T dt \int_{\Omega} d^n x \left(\frac{1}{2} \rho \mathbf{v}^2 - e\rho \right) \quad (1.2)$$

definido en un cierto espacio de funciones $\mathbf{v}(\mathbf{x}, t)$, $\rho(\mathbf{x}, t)$ junto con las condiciones (ligaduras), $\mathbf{v} \parallel \partial\Omega$ y $\frac{\partial\rho}{\partial t} + \text{div}(\rho \mathbf{v}) = 0$, conduce tras una rápida (y heurística) aplicación del teorema de los multiplicadores de Lagrange al funcional extendido:

$$\mathbb{S}(\mathbf{v}, \rho, \phi) = \int_0^T dt \int_{\Omega} d^n x \left[\left(\frac{1}{2} \rho \mathbf{v}^2 - e\rho \right) + \phi \left(\frac{\partial\rho}{\partial t} + \text{div}(\rho \mathbf{v}) \right) \right]$$

del que se obtienen las ecuaciones:

$$\begin{cases} \mathbf{v} = \nabla\phi, \\ \frac{\partial\rho}{\partial t} + \text{div}(\rho \mathbf{v}) = 0, \\ \frac{\partial\phi}{\partial t} + (\mathbf{v} \cdot \nabla)\phi = \frac{1}{2}\mathbf{v}^2 - e. \end{cases} \quad (1.3)$$

Nótese que el campo de velocidades \mathbf{v} es potencial (1.3), por lo que sólo podemos obtener como soluciones de las ecuaciones de Euler, fluidos irrotacionales:

$$\nabla \wedge \mathbf{v} = 0.$$

En 1963 Lin [Li63] propuso un principio variacional modificando (1.2) que permite obtener directamente la representación de Clebsch del campo de velocidades de un fluido en la forma [La32]:

$$\mathbf{v} = \nabla\phi + \lambda\nabla\mu \quad (1.4)$$

con $\lambda, \mu, \phi \in \mathcal{C}^\infty(\Omega)$. Las funciones λ y μ se denominan variables de Clebsch y ϕ la función potencial de \mathbf{v} .

La idea de Lin consiste en utilizar la representación Lagrangiana del fluido. Esta representación se obtiene pensando que la posición \mathbf{x} a tiempo t de una partícula de fluido se obtiene por medio de un difeomorfismo ϕ_t que nos transforma la posición \mathbf{X} de la partícula a tiempo $t = 0$, en la posición actual. Esta transformación es invertible, y así podemos escribir:

$$\mathbf{X} = \phi_t^{-1}(\mathbf{x}(t)) \equiv \mathbf{X}(\mathbf{x}, t).$$

Pero \mathbf{X} es fijo por lo que $d\mathbf{X}/dt = 0$; y por tanto

$$0 = \frac{d\mathbf{X}}{dt} = \frac{\partial\mathbf{X}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{X}, \quad (1.5)$$

por lo que obtenemos unas ligaduras naturales a partir de la representación Lagrangiana del fluido.

De hecho, no es necesario utilizar todas las ligaduras anteriores. Si denotamos por μ a la primera componente de \mathbf{X} , podemos introducir la ligadura de Lin (1.5) a través de un multiplicador de Lagrange λ , y tendríamos entonces un nuevo funcional extendido:

$$S_L(\mathbf{v}, \rho, \phi, \lambda, \mu) \equiv S_L = \int_0^T dt \int_{\Omega} d^n x \left[\frac{1}{2} \rho \mathbf{v}^2 - e \rho + \phi \left(\frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{v}) \right) + \lambda \left(\frac{\partial \mu}{\partial t} + \mathbf{v} \cdot \nabla \mu \right) \right].$$

Cuya variación nos proporciona rápidamente las ecuaciones:

$$\left\{ \begin{array}{l} \rho \mathbf{v} - \rho \nabla \phi + \lambda \nabla \mu = 0, \\ \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi + e - \frac{1}{2} \mathbf{v}^2 = 0, \\ \frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{v}) = 0, \\ \frac{\partial \mu}{\partial t} + \mathbf{v} \cdot \nabla \mu = 0, \\ \frac{\partial \lambda}{\partial t} + \text{div}(\lambda \mathbf{v}) = 0. \end{array} \right. \quad (1.6)$$

De donde obtenemos la representación de Clebsch del campo de velocidades \mathbf{v} :

$$\mathbf{v} = \nabla \phi - \frac{\lambda}{\rho} \nabla \mu.$$

Por tanto para un fluido incompresible ($\rho = \text{cte.}$), las ecuaciones dinámicas del fluido se convierten en:

$$\left\{ \begin{array}{l} \mathbf{v} = \nabla\phi - \frac{\lambda}{\rho}\nabla\mu; \quad (\rho = cte; \operatorname{div} \mathbf{v} = 0), \\ \frac{\partial\phi}{\partial t} + \mathbf{v} \cdot \nabla\phi + e - \frac{1}{2}\mathbf{v}^2 = 0, \\ \frac{\partial\mu}{\partial t} + \mathbf{v} \cdot \nabla\mu = 0, \\ \frac{\partial\lambda}{\partial t} + \operatorname{div}(\lambda\mathbf{v}) = 0, \end{array} \right. \quad (1.7)$$

que son equivalentes a las Ecuaciones de Euler iniciales (1.1).

En los trabajos, [Ce87a] Cendra y Marsden y [Ce87b] Cendra, Ibort y Marsden, acometieron la tarea de clarificar el sentido de las manipulaciones anteriores, usando para ello una construcción geométrica que permitía interpretar las ligaduras de Lin como una condición de horizontalidad respecto de una conexión en un espacio auxiliar, el espacio de las variables de Clebsch λ, μ . Estas ideas permitieron una descripción elegante de la teoría anterior para el problema de un fluido incompresible y a su vez pueden ser aplicadas a gran variedad de situaciones con dificultades topológicas adicionales.

Es significativo reseñar que el análisis de Lin permitió una notable simplificación por ejemplo en el análisis de las propiedades del HeII superfluido. Volviendo a las cuestiones relativas a las complicaciones topológicas que pueden exhibir los sistemas, es notable señalar que la representación de Clebsch del campo de velocidades de un fluido no incluye todas las posibles configuraciones de éste. Si definimos la helicidad \mathcal{H} de un campo de velocidades \mathbf{v} como:

$$\mathcal{H} = \int_{\Omega} \mathbf{v}(\mathbf{x}, t) \wedge \omega(\mathbf{x}, t) d^n x$$

obtenemos que si $\omega = \nabla\lambda \wedge \nabla\mu$, entonces $\mathcal{H} = 0$. Por otro lado, sabemos que en un fluido con líneas de vorticidad anudadas (topología no trivial) la helicidad es no nula $\mathcal{H} \neq 0$ [Mo90].

Algunos problemas en dinámica de fluidos ó medios continuos en los que las ideas geométricas que constituyen el armazón de esta memoria se pueden aplicar (ó ya han sido aplicados) son:

- Fluidos homogéneos incompresibles.
- Fluidos ideales compresibles.
- Fluidos estratificados.
- Plasmas.

- Superconductores y superfluidos.
- Elasticidad.
- Problemas de fluidos con fronteras libres.
- Magneto hidrodinámica (MHD), etc.

1.2. Sistemas con simetría: El sólido rígido

Un ingrediente fundamental, aunque oculto, en la discusión anterior de la formulación variacional para un fluido incompresible es la existencia de un grupo G de simetrías del sistema. De hecho la respuesta correcta a alguna de las cuestiones anteriores surge cuando consideramos la descripción del sistema en su formulación lagrangiana. El espacio de configuraciones Q es entonces el grupo de difeomorfismos $G = \text{Diff}(\Omega)$ de la región Ω y el lagrangiano L es la energía cinética del fluido $L(\phi, \dot{\phi}) = \int_{\Omega} \frac{1}{2} \|\dot{\phi} \phi^{-1}\|^2 dx$.

Dicho lagrangiano es invariante con respecto a la acción del propio grupo de difeomorfismos actuando a derechas en el espacio de configuraciones, esto es $L(\phi \circ \psi, \dot{\phi} \circ \psi) = L(\phi, \dot{\phi})$.

Por otra parte, los sistemas Lagrangianos con simetría tienen propiedades estructurales que han sido ampliamente usadas en el estudio de la naturaleza cualitativa de sus soluciones. Este fue el trabajo que realizaron entre muchos otros J. E. Marsden y A. Weinstein [Ma74] sobre reducción de sistemas simplécticos con simetría donde agruparon y mejoraron las ideas clásicas sobre simetría, abriendo así una nueva puerta para la comprensión sistemática de las estructuras de algunos de los sistemas paradigmáticos en ecuaciones diferenciales y en mecánica, por ejemplo el sólido rígido y las ecuaciones de Euler. Pronto se comprobó que tales sistemas podían ser obtenidos por reducción de sistemas Lagrangianos simples con simetría, usualmente definidos sobre el fibrado tangente de un grupo de Lie (finito ó infinito dimensional) ó, más generalmente, sobre un fibrado principal sobre algún espacio de configuraciones. Esta idea ha sido ampliamente usada para describir una gran clase de sistemas, que comprenden desde la física del plasma hasta la elasticidad, y otros muchos problemas en medios continuos. Una contribución reciente importante a este campo ha sido la comprensión intrínseca de la reducción de las ecuaciones de Euler-Lagrange y de algunas de las estructuras más importantes asociadas a ella (ver [Ce01a], [Ce01b] y las referencias contenidas en ellas para una visión panorámica de estas ideas).

El sólido rígido constituye un magnífico ejemplo en dimensión finita, donde todas las ideas anteriores se encuentran condensadas. La formulación lagrangiana de un sólido rígido se realiza sobre el espacio de configuraciones dado por el grupo $G = SO(3)$. La energía cinética está definida por una métrica invariante

en el grupo caracterizada por una forma bilineal simétrica y definida positiva en $T_I SO(3)$.

Usando la métrica natural del grupo $SO(3)$, dada por $\langle A, B \rangle = -\text{Tr}(A^T B)$ el lagrangiano del sistema se escribirá como:

$$L(R, \dot{R}) = \frac{1}{2} \langle \dot{R}R^{-1}, I\dot{R}R^{-1} \rangle \quad (1.8)$$

donde $\dot{R} \in T_R SO(3)$, $\Omega = \dot{R}R^{-1} \in \mathfrak{so}(3)$ es ahora una matriz 3×3 antisimétrica. Nótese que el lagrangiano anterior es invariante bajo la acción de $SO(3)$ por la derecha.

Un sencillo cálculo muestra que las ecuaciones de Euler-Lagrange del funcional $S = \int_0^T L(R, \dot{R}) dt$ adoptan la forma de las bien conocidas ecuaciones de Euler:

$$\dot{R} = \Omega R ; I\dot{\Omega} = [\Omega, I\Omega].$$

Sin embargo, si utilizamos la invarianza del lagrangiano para escribir las ecuaciones del movimiento en $T SO(3)/SO(3) \cong \mathfrak{so}(3)$ obtenemos, al igual que en el caso de las ecuaciones de un fluido incompresible, que la variación del principio variacional reducido $\int_0^T l(\Omega) dt = \int_0^T \frac{1}{2} \langle \Omega, I\Omega \rangle dt$ conduce a las ecuaciones triviales $\Omega = 0$.

Recuperamos las ecuaciones de Euler observando que la definición de la velocidad angular $\Omega = \dot{R}R^{-1}$, establece que sus variaciones no pueden ser arbitrarias sino de la forma $\partial\Omega = \dot{\eta} + [\eta, \Omega]$, con $\eta(t) \in \mathfrak{so}(3)$ y $\eta(0) = \eta(T) = 0$.

La introducción de variables de Clebsch adecuadas y las correspondientes ligaduras de Lin permiten resolver de nuevo este problema elegantemente, como veremos más adelante.

En cualquier caso, la teoría que desarrollaremos en esta memoria será para sistemas mecánicos, de tal manera que reduciremos las dificultades analíticas al mínimo y nos concentraremos en los aspectos geométricos del problema. Aunque excluirémos de esta manera del análisis, el problema de los fluidos presentado en la sección anterior, las ideas geométricas que emergen se pueden aplicar directamente a ellos. Dicho trabajo se realizará en un futuro inmediato.

1.3. Control óptimo y sistemas lagrangianos con simetría: El sólido rígido de nuevo

El interés del estudio del sólido rígido no termina en la discusión anterior. En un breve pero interesante artículo Bloch, Crouch, Holm y Marsden [Bl00] realizaron la observación de que ciertos problemas lagrangianos con simetría pueden interpretarse como problemas de control óptimo, y al aplicar en ellos el Principio del Máximo de Pontryagin obtener nuevas interpretaciones de las ecuaciones del movimiento.

Ilustraremos esta idea de nuevo tomando el sólido rígido descrito por el lagrangiano anterior en $TSO(3)$, ecuación (1.8).

Sin embargo, ahora consideraremos el problema de control óptimo en $SO(3)$ con ecuación de control $\dot{R} = \Omega R$ y funcional objetivo $l(\Omega) = \frac{1}{2}\langle \Omega, I\Omega \rangle$ donde Ω representa ahora los controles del problema. Usando el principio del máximo de Pontryagin, obtenemos que si $R(t)$ es una trayectoria optimal, entonces existirá $P(t) \in T^*SO(3)$ tal que

$$\dot{R} = \frac{\partial H}{\partial P}, \quad \dot{P} = -\frac{\partial H}{\partial R} \quad (1.9)$$

donde $H(R, P, \Omega) = \langle P, \Omega R \rangle - l(\Omega)$ es el Hamiltoniano de Pontryagin.

Además, la síntesis optimal se obtendrá a través de la ecuación

$$\frac{\partial H}{\partial \Omega} = 0,$$

por tanto

$$I\Omega = PR^T - RP^T \quad (1.10)$$

que junto con las expresiones explícitas para las ecuaciones (1.9) nos proporcionan las ecuaciones en $T^*SO(3)$ equivalentes a las ecuaciones de Euler:

$$\dot{R} = \Omega R, \quad \dot{P} = \Omega P.$$

Estas ecuaciones constituyen la llamada representación simétrica de las ecuaciones del sólido rígido [B198].

Nótese que si calculamos, $I\dot{\Omega}$ con la ayuda de la síntesis optimal, ecuación (1.10) obtenemos de nuevo las ecuaciones de Euler.

Esta sencilla observación establece que un problema lagrangiano con simetría puede considerarse como un problema de control, pero también viceversa. Podemos estudiar una interesante clase de problemas de control óptimo como problemas lagrangianos con simetría. En el transcurso de este análisis, intentaremos dilucidar qué relación existe entre esta observación de Bloch, Crouch, Holm y Marsden y la teoría de reducción de principios variacionales con simetría. Como consecuencia de este análisis, en el capítulo 4 probaremos que existe una estrecha relación entre la construcción anterior y la teoría de variables de Clebsch y ligaduras de Lin. Esta dualidad entre sistemas lagrangianos con simetría y una clase de sistemas de control óptimo que extiende el problema del sólido rígido, constituye uno de los temas centrales de este trabajo, dualidad que denominaremos reciprocidad de Lagrange-Pontryagin asociada a un grupo G .

1.4. Objetivos

Podemos resumir la discusión de los apartados anteriores, diciendo que en esta tesis nos planteamos por un lado como objetivo fundamental el profundizar

en la relación existente entre la formulación a la Clebsch/Lin de los principios variacionales con simetría y la teoría de la reducción de sistemas Lagrangianos y Hamiltonianos con simetría. Por otro lado, siguiendo la observación en los artículos recientes [Bl98] y [Bl00] nos plantearemos dilucidar la relación existente entre la reciprocidad de Lagrange-Pontryagin, la teoría de reducción de principios variacionales con simetría y la teoría de variables de Clebsch y ligaduras de Lin.

1.5. Algunas ideas, técnicas y organización de la memoria.

En esta memoria combinaremos varias ideas para avanzar en la resolución de los problemas y objetivos propuestos.

a) Por un lado, recurriremos a una formulación analítica de los principios variacionales. Para ello utilizaremos ideas de análisis global tales como la construcción de variedades Hilbertianas de espacios de curvas, etc. De esta manera, ganaremos en precisión analítica sin descuidar la intuición geométrica, que se trasladará en muchas ocasiones de manera natural, de los espacios de dimensión finita usados para la construcción de la teoría (espacios de configuraciones, grupos de Lie de simetría, etc.), a los correspondientes espacios de curvas (ver por ejemplo [Ib96]). Esta aproximación al problema ya ha sido utilizada por diversos autores (ver por ejemplo [Kl78], [Fl89]) y de ellos tomaremos prestados resultados y notaciones. Un primer uso extensivo de estas ideas y técnicas de análisis global fue realizado en la Tesis de M. Delgado [De04]. Es posible sin embargo realizar un análisis "tradicional" de los principios variacionales introduciendo en cada caso la noción apropiada de variación y calculando la "variación" del correspondiente principio variacional, sin embargo la justificación en cada caso del uso de un teorema de los multiplicadores de Lagrange resulta tediosa. En esta memoria estableceremos una formulación de dicho teorema que puede aplicarse con seguridad en todos los casos y que resulta bastante transparente (ver sección 2.3.2.).

b) Para el estudio y discusión de los problemas de control óptimo utilizaremos la formulación presimpléctica desarrollada en varios trabajos anteriores [De03a], [De03b] y [De08]. Dicha formulación se inspira en la formulación geométrica desarrollada por E. Martínez y C. López [Lo99], [Lo00], [Mr99] y que proporciona una equivalencia entre las condiciones diferenciales de primer orden del Principio del Máximo de Pontryagin y un sistema Hamiltoniano presimpléctico. A partir de dicha representación ciertos problemas difíciles de analizar como las soluciones singulares se vuelven transparentes. En este trabajo profundizaremos dicho análisis al considerar no solamente la formulación presimpléctica, sino que ampliaremos el foco y consideraremos la formulación del Principio del Máximo de Pontryagin a través de estructuras de Dirac siguiendo las ideas precursoras de Marsden y Yoshimura [Yo07a], [Yo07b], [Yo09].

c) Por otro lado, ya hemos resaltado que una característica fundamental de los problemas que estamos considerando es que se trata siempre de sistemas con simetría. Una idea subyacente a parte de este trabajo, se obtiene de la construcción de Borel, que permite analizar propiedades de un espacio cociente a partir del uso de espacios auxiliares. Así, si X es una variedad diferenciable y G un grupo de Lie que actúa en ella por la derecha quisiéramos estudiar las propiedades del espacio de órbitas X/G ó clases de equivalencia en X introducidas por la acción del grupo G .

Ahora, debido a la sucesión exacta larga de homotopía:

$$\begin{aligned} 1 \rightarrow \pi_0(F) \rightarrow \pi_0(X) \rightarrow \pi_0(X/G) \rightarrow \\ \rightarrow \pi_1(F) \rightarrow \pi_1(X) \rightarrow \pi_1(X/G) \rightarrow \dots \rightarrow \\ \rightarrow \pi_k(F) \rightarrow \pi_k(X) \rightarrow \pi_k(X/G) \rightarrow \pi_{k+1}(F) \rightarrow \dots \end{aligned}$$

donde F denota la fibra standard de la aplicación de proyección $\pi: X \longrightarrow X/G$, cuando X es un espacio contractible, tenemos que:

$$\pi_k(X/G) \cong \pi_{k+1}(F), \quad k = 0, 1, \dots$$

Si además, G actúa libremente sobre X , entonces la fibra de π es difeomorfa a G , y la topología de X/G es computable en términos de la topología de G . Consideraremos entonces la construcción de Borel: Sea M un G -espacio por la derecha. Un modo de entender la estructura del espacio cociente M/G es considerar un G -espacio auxiliar X donde G actúa libremente, por ejemplo un espacio universal clasificante para G . Entonces, consideraremos el espacio producto $M \times X$ con la acción natural:

$$\begin{aligned} \Phi: (M \times X) \times G &\longrightarrow M \times X \\ ((m, x), g) &\rightsquigarrow (mg, g^{-1}x) \end{aligned}$$

(Notar que la acción de G sobre $M \times X$ es libre). Denotaremos al espacio cociente $M_G := M \times_G X = (M \times X)/G$. A los elementos del espacio cociente $[m, x] = \{(mg, g^{-1}x) | g \in G\}$ los denotaremos en lo que sigue simplemente por: $mx := [m, x]$.

$$\begin{array}{ccccc} M & \xleftarrow{p_M} & M \times X & \xrightarrow{p_X} & X \\ \downarrow \pi & & \downarrow \Pi & & \downarrow p \\ M/G & \xleftarrow{\pi_M} & M \times_G X & \xrightarrow{\pi_X} & G \backslash X \end{array}$$

Que el espacio X sea contractible implica que M es un retracto de deformación de $M \times X$. Lo mismo ocurre para M/G con respecto al cociente equivariante M_G . Por tanto, la cohomología equivariante de M con respecto a G es la cohomología del espacio M_G , luego $H_G^*(M) = H^*(M_G)$.

En nuestro trabajo nos inspiraremos en esta aproximación al problema de estudiar espacios cocientes ya que en lugar de trabajar directamente con ellos, usaremos G -espacios auxiliares con buenas propiedades, y los utilizaremos para construir nuevos espacios, que llamaremos espacios de Clebsch, donde podamos trasladar las cuestiones que estamos analizando de manera que ganen en claridad ó nos ofrezcan nuevas perspectivas. A las coordenadas de los espacios de Clebsch las llamaremos variables de Clebsch.

d) Finalmente utilizaremos la teoría de conexiones y transporte paralelo en fibrados principales y asociados para la realización de diversas construcciones, cálculos y ejemplos. Dichas ideas serán revisadas de manera resumida cuando sea necesario.

Capítulo 2

El problema del cálculo de variaciones con simetría.

En este capítulo se realizará una discusión general sobre el problema de la determinación de los máximos y mínimos de funciones invariantes restringidas a subvariedades, que constituyen el marco natural para el estudio de los problemas en optimización y mecánica con simetría.

2.1. El problema del cálculo de variaciones en mecánica

2.1.1. El problema general del cálculo de variaciones

El problema del cálculo de variaciones consiste en la determinación de los mínimos (ó máximos) de un funcional S de la forma:

$$S = \int_{\Omega} L(x, \varphi, \varphi_x, \dots) d^n x \quad (2.1)$$

donde $x \in \Omega \in \mathbb{R}^n$, en un espacio de funciones $\varphi: \Omega \longrightarrow \mathbb{R}^m$ llamados los campos de la teoría (suficientemente regulares para que las derivadas $\varphi_x, \varphi_{xx}, \dots$ tengan sentido) y $L(x, \varphi, \varphi_x, \dots)$ es una densidad, llamada lagrangiana ó lagrangiano de la teoría.

Determinar el espacio de funciones φ en Ω donde buscar los mínimos de S es en general parte del problema en el sentido que, en principio, podemos tratar de usar un espacio de funciones X suficientemente grande donde garantizar la existencia de mínimos de S , y posteriormente intentar probar que dichos mínimos tienen determinadas propiedades de regularidad, esto es, pertenecen efectivamente a un espacio mucho más pequeño de funciones. En muchas ocasiones el espacio X escogido para resolver el problema tendrá buenas propiedades topológicas, e

incluso muchas veces podremos suponer que se trata de un espacio de Banach y poder aplicar entonces resultados potentes del análisis funcional.

Así, de manera muy general, supondremos que el funcional S es una función $S: X \longrightarrow \mathbb{R} : \varphi \rightsquigarrow S(\varphi)$ definida en un espacio de Banach X , de la que queremos determinar un elemento $\varphi^* \in X$ tal que

$$S(\varphi^*) = \inf_{\varphi \in X} S(\varphi). \quad (2.2)$$

2.1.2. Métodos directos

En algunas ocasiones es posible probar directamente la existencia de máximos ó mínimos de S , utilizando exclusivamente propiedades de X y del funcional S (a continuación daremos un teorema típico de este tipo), aunque en general, tal procedimiento es inaplicable y nos hemos de limitar a determinar indirectamente los puntos críticos del funcional S , para posteriormente hallar los máximos ó mínimos (por ejemplo, entre las soluciones de las ecuaciones de Euler-Lagrange del problema en caso de que seamos capaces de resolverlas).

Un argumento típico que conduce a un teorema de existencia de mínimos de S (y que contiene muchos de los resultados conocidos en la literatura) es el siguiente: supondremos que S está acotado inferiormente y por tanto existe el $\inf S$. Si existe el $\inf S$, existen sucesiones minimizantes, esto es, familias de funciones φ_k tales que

$$\lim_{k \rightarrow \infty} S(\varphi_k) = \inf S.$$

En general dichas sucesiones φ_k no serán convergentes. Debemos por tanto añadir condiciones a X y S que garanticen la existencia de subsucesiones convergentes de tal sucesión minimizante.

Una de estas condiciones es pedir que el funcional S sea coercivo, esto es, S ha de ser tal que si φ_k es una sucesión tal que $\|\varphi_k\| \rightarrow \infty$, entonces $S(\varphi_k) \rightarrow \infty$, cuando $k \rightarrow \infty$.

Nótese que entonces si φ_k es una sucesión minimizante, necesariamente $\|\varphi_k\| \leq M$ para cierta constante M . Los elementos φ_k de la sucesión estarán por tanto todos ellos en una bola cerrada de radio M del espacio de Banach X . Si X es un espacio de Banach reflexivo, tendremos que $X = X'$ y por tanto por el teorema de Alaglou-Banach la bola cerrada es compacta en la topología ω^* , ó dicho en otras palabras, existe una subsucesión de φ_k (que denotaremos igual), que converge débilmente a un cierto elemento $\varphi^* \in X$, esto es, $\lim \varphi_k(x) = \varphi^*(x), \forall x \in X$.

Si pedimos que el funcional S sea inferiormente débilmente semicontinuo, se verifica que:

$$\underline{\lim} S(\varphi_k) \geq S(\varphi^*)$$

para toda sucesión φ_k que converge débilmente a φ^* (a veces se denota por $\varphi_k \rightharpoonup \varphi^*$) y si φ_k es minimizante se tendrá que:

$$\inf(S) = \lim S(\varphi_k) \geq S(\varphi^*)$$

por tanto $S(\varphi^*) = \inf S$ y $\varphi^* = \min S$, como estábamos buscando. Las propiedades de convexidad de S ayudan a caracterizar funciones que son débilmente semicontinuas. Por ejemplo, del teorema de Mazur se sigue que toda función convexa e inferiormente semicontinua es débilmente inferiormente semicontinua. Por tanto, en estas condiciones hemos probado el siguiente teorema de existencia de mínimos:

Teorema 2.1.1 *Sea S una función acotada inferiormente, coerciva, inferiormente semicontinua y convexa en un espacio de Banach X reflexivo, entonces S tiene mínimo que se alcanza como límite de sucesiones minimizantes.*

Nótese que el resultado anterior no utiliza apenas ninguna propiedad estructural de la función S y por supuesto tampoco su estructura local (2.1). Se puede refinar notablemente el resultado anterior utilizando la estructura local del funcional aunque se pueden obtener resultados generales utilizando otras teorías como la teoría de Morse, técnicas de minimax, teorías de categorías de Liusternik-Schnierelmann, teorías del índice de Benci ó de Zenhder, teorías de Floer, etc.

No proseguiremos por esta dirección en este trabajo, sino que nos concentraremos en funcionales locales del tipo de los que aparecen en los principios variacionales en mecánica y teoría de control y con una estructura adicional, el poseer un grupo de simetría.

Sin embargo, antes de continuar en esta dirección, podemos establecer una versión global del teorema anterior que es válido en problemas no lineales en mecánica y teoría de control.

El conjunto X denotará ahora una variedad de Banach, es decir una variedad diferenciable modelada sobre un espacio de Banach E . El ejemplo típico de tales variedades se obtiene a partir de espacios de curvas (ver sección §2.1.6 más adelante) en una variedad diferenciable M de dimensión finita (ó un grupo de Lie G) en la que se define una métrica dada por:

$$d(m, m') = \inf_{\gamma \in \Omega_{m, m'}(M)} \int_0^1 \|\dot{\gamma}(t)\|^2 dt,$$

donde $\Omega_{m, m'}(M)$ es el conjunto de curvas diferenciables a trozos $\gamma: [0, 1] \longrightarrow M$ tales que $\gamma(0) = m$, $\gamma(1) = m'$ y $\|\cdot\|$ denota la norma inducida por una métrica riemanniana dada en M . Diremos que una función $S: X \longrightarrow \mathbb{R}$ es coerciva si para toda familia de puntos $\{x_k\} \subset X$, tal que $\text{diam}\{x_k\} = +\infty$, entonces $\{S(x_k)\}$ no

está acotado superiormente (ó $\overline{\lim} S(x_k) = +\infty$). Se sigue inmediatamente de la definición el lema siguiente:

Lema 2.1.1 *Si S es una función coerciva acotada inferiormente, entonces posee sucesiones minimizantes acotadas.*

Demostración. Si S está acotada inferiormente $\exists \inf S < +\infty$. Sea $\{x_k\}$ una sucesión minimizante, es decir, que $\lim_{k \rightarrow \infty} S(x_k) = \inf S$. Si x_k no está acotada, entonces $\text{diam}\{x_k\} = \infty$. Por tanto $S(x_k)$ no está acotada y existe una subsucesión x_{k_l} tal que $\lim_{l \rightarrow \infty} S(x_{k_l}) = \inf S$, lo que entra en contradicción con la existencia de $\inf(S)$. \square

Podemos probar entonces fácilmente un análogo global del teorema 2.1.1:

Teorema 2.1.2 *Sea X una variedad de Banach reflexiva paracompacta y S una función acotada inferiormente, coerciva y débilmente inferiormente semicontinua, entonces S tiene mínimo.*

Demostración. La variedad X se dice reflexiva si el espacio de Banach modelo E es reflexivo. Una pequeña extensión del teorema de Alaglou-Banach permite probar la existencia de una subsucesión que converge débilmente a un cierto elemento $x^* \in X$. Los mismos razonamientos anteriores nos muestran entonces que $x^* = \inf S$. \square

2.1.3. Caracterización de puntos críticos: Ecuaciones de Euler-Lagrange

Tal y como se indicó anteriormente en muchas ocasiones no es posible usar métodos directos. Sin embargo, es posible caracterizar una familia de puntos a la que necesariamente pertenecen los máximos ó mínimos que estamos buscando. Esto ocurre en particular si el funcional S es diferenciable. Si X es un espacio de Banach (ó una variedad de Banach) y $S: X \rightarrow \mathbb{R}$ es una función diferenciable, es fácil demostrar que si el punto $x_0 \in X$ es un extremal de S , entonces $dS(x_0) = 0$.

El conjunto $C(S)$ de puntos $x_0 \in X$ tal que $dS(x_0) = 0$, se llama conjunto crítico de S , y dichos puntos son llamados puntos críticos de S . Los valores $S(x_0)$, $x_0 \in C(S)$, son llamados valores críticos de S .

El análisis de los conjuntos críticos y de los valores críticos para una función bien escogida proporciona una maravillosa herramienta para estudiar la topología de la variedad X llamada teoría de Morse.

2.1.4. El caso del cálculo de variaciones.

En muchas aplicaciones y problemas de interés, la variedad de Banach X es un espacio funcional, cuyos puntos son funciones ó secciones de fibrados apropiados. La función $S: X \longrightarrow \mathbb{R}$ muy a menudo es un funcional local, es decir, de la forma descrita en §2.1.1:

$$S(\varphi) = \int_{\Omega} L(\varphi, \varphi_x, \dots) d^n x \quad (2.3)$$

donde L es una función definida en alguna variedad auxiliar M (ó en algún espacio de jets sobre M). En estas circunstancias algunos de los resultados discutidos antes, tanto métodos directos como indirectos, pueden ser refinados considerablemente.

Para poder fijar el planteamiento de lo que sigue, nos concentraremos en el problema del cálculo de variaciones que también denominaremos Principios Variacionales, en Mecánica. Esto constituye ciertamente una restricción dentro de la generalidad de los problemas de importancia en Física ó Matemáticas. Por ejemplo, no discutiremos el problema de las ecuaciones de Euler para un fluido incompresible mencionadas en la introducción y que es una de las motivaciones históricas para este trabajo. Sin embargo, la clase de problemas que trataremos es suficientemente grande, contiene ejemplos muy importantes de la Mecánica y de la Teoría de Control, y proporciona el marco adecuado para el desarrollo de las principales ideas geométricas que ilustran los problemas que queremos resolver.

Por otra parte, tal y como ya hemos indicado con anterioridad, el desarrollo técnico en estas situaciones es más simple y en algunos casos incluso más rico (en un sentido que clarificaremos más tarde) que la situación general.

2.1.5. El marco analítico para el cálculo de variaciones en Mecánica

Como ya discutimos en la introducción (§1.5.a), hay varias posibilidades para establecer el marco analítico del cálculo de variaciones en Mecánica. Un modo, que corresponde a la aproximación clásica, consiste en escoger una clase de curvas que se adapten bien a las peculiaridades de los funcionales locales que estamos tratando, por ejemplo, la clase de las curvas diferenciables a trozos. Esta es la aproximación tomada por diversos autores clásicos y la usada de manera habitual (ver por ejemplo los recientes trabajos de Marsden et al [Ce01a],[Yo09]).

Para el propósito de la formulación geométrica de la teoría, la clase C^∞ es la más conveniente. En esta memoria tomaremos sin embargo, la aproximación de Klingerberg que consiste en seleccionar curvas de una clase de Sobolev dada. De este modo nuestros espacios de caminos serán variedades de Hilbert en sentido estricto y algunas manipulaciones analíticas se pueden realizar de manera rigurosa.

Se usarán por tanto para construir la teoría, variedades de Hilbert y algunas ideas geométricas, establecidas a nivel de espacios finito-dimensionales, y también a nivel de espacios de caminos, proporcionando en varias ocasiones nuevos e interesantes puntos de vista.

2.1.6. Variedades de Hilbert de curvas en variedades riemannianas

Comenzaremos esta discusión introduciendo las definiciones y notaciones básicas de la teoría de espacios de Sobolev que vamos a necesitar a continuación.

Definición 2.1.1 *Se define el espacio de curvas de Sobolev L^p de clase k en \mathbb{R}^r como:*

$$L_k^p(I; \mathbb{R}^r) = \left\{ \gamma: I \longrightarrow \mathbb{R}^r \mid \gamma \in L^p(I; \mathbb{R}) \text{ y } \exists \frac{d\gamma}{dt}, \dots, \frac{d^k \gamma}{dt^k} \in L^p(I; \mathbb{R}^r) \right\}.$$

La notación $d\gamma/dt$ en la definición anterior representa la derivada débil de γ , es decir, tal que $\forall \varphi \in C_0^\infty([0, T], \mathbb{R}^r)$, se tiene

$$\int_0^T \left\langle \frac{d\gamma}{dt}, \varphi \right\rangle dt = - \int_0^T \left\langle \gamma, \frac{d\varphi}{dt} \right\rangle dt,$$

donde $C_0^\infty([0, T], \mathbb{R}^r)$ denota el conjunto de funciones C^∞ de soporte compacto contenido en $(0, T)$.

El espacio de Sobolev $L_k^p(I; \mathbb{R}^r)$ es un espacio de Banach con la norma $\|\cdot\|_{k,p}$ definida de la forma siguiente:

$$\|\gamma\|_{k,p}^p = \sum_{l=0}^k \int_0^T \left\| \frac{d^l \gamma}{dt^l}(t) \right\|^p dt.$$

En el caso particular $p = 2$, tenemos el espacio $L_k^2(I; \mathbb{R}^r)$ que también se denota como $H^k(I; \mathbb{R}^r)$, ó $W^{2,k}(I, \mathbb{R}^r)$, es decir,

$$H^k(I; \mathbb{R}^r) = \left\{ \gamma: I \longrightarrow \mathbb{R}^r \mid \gamma \in L^2(I; \mathbb{R}) \text{ y } \exists \frac{d\gamma}{dt}, \dots, \frac{d^k \gamma}{dt^k} \in L^2(I; \mathbb{R}^r) \right\}.$$

Este espacio, $H^k(I; \mathbb{R}^r)$, es un espacio de Hilbert real, con respecto al producto interno

$$\langle \gamma, \tilde{\gamma} \rangle_k = \sum_{l=0}^k \int_0^T \left\langle \frac{d^l \gamma}{dt^l}(t), \frac{d^l \tilde{\gamma}}{dt^l}(t) \right\rangle dt.$$

En particular los casos $k = 0, 1$ van a ser los de mayor interés a lo largo de todo este trabajo. El espacio correspondiente a $k = 1$ es $H^1(I; \mathbb{R}^r)$, que puede identificarse con el espacio de funciones $\gamma: I \longrightarrow \mathbb{R}$ absolutamente continuas, esto

es de la forma $\gamma(t) = \int_0^t g(s) ds$, $g(s)$ integrable en I , tales que $d\gamma/dt \in L^2(I; \mathbb{R}^r)$, con el producto interno:

$$\langle \gamma, \tilde{\gamma} \rangle_1 = \int_0^T \langle \gamma(t), \tilde{\gamma}(t) \rangle dt + \int_0^T \left\langle \frac{d\gamma}{dt}(t), \frac{d\tilde{\gamma}}{dt}(t) \right\rangle dt.$$

La aplicación $H^1(I) \hookrightarrow C^0(I)$, que envía cada representante absolutamente continuo en una clase de funciones en H^1 en ella misma, no sólo es continua sino compacta. El espacio correspondiente a $k = 0$ es $H^0(I; \mathbb{R}^r) = L^2(I; \mathbb{R}^r)$.

Otra forma de definir conveniente los espacios de curvas de Sobolev es a través de la completación del espacio $C^\infty(I; \mathbb{R}^r)$ con respecto a la norma $\|\cdot\|_{k,p}$, esto es:

$$\overline{C^\infty(I; \mathbb{R}^r)}^{\|\cdot\|_{k,p}} = L_k^p(I; \mathbb{R}^r).$$

Si consideramos ahora el espacio anterior, pero con soporte compacto, $C_0^\infty(I; \mathbb{R}^r)$, es decir, formado por funciones diferenciables con soporte compacto, para $p = 2$ tenemos que

$$H_0^k(I; \mathbb{R}^r) = \overline{C_0^\infty(I; \mathbb{R}^r)}^{\|\cdot\|_{k,2}},$$

que es el espacio de Hilbert de curvas de Sobolev de clase k que se anulan en los extremos de $I = [0, T]$.

Espacios de Sobolev de curvas en variedades riemannianas

Definamos ahora los espacios de Sobolev de curvas en variedades riemannianas. Sea (M, g) una variedad diferenciable riemanniana sin borde. Definimos el espacio de curvas de Sobolev de clase L_k^p en M como el espacio de aplicaciones $\gamma: I \longrightarrow M$ tales que

$$\begin{aligned} \psi \circ \gamma \in L_k^p(\gamma^{-1}(\gamma(I) \cap U), \mathbb{R}^r) \text{ para toda carta diferenciable } (U, \psi) \text{ en } M, \\ \psi: U \subset M \longrightarrow \mathbb{R}^r. \end{aligned}$$

Este espacio se denotaría por $L_k^p(I; M)$. Estamos interesados en particular en el espacio de curvas de clase de Sobolev H^1 , es decir:

$$H^1(I; M) \equiv L_1^2(I; M),$$

que ha sido estudiado, por ejemplo, por W. Klingenberg [Kl78]. Si $\gamma \in H^1(I; M)$, entonces existe $d(\psi \circ \gamma(t))/dt$ en $[0, T]$, por lo que podemos definir $d\gamma/dt|_t$ a.e. en $[0, T]$ como el vector tangente a $\gamma(t)$ en t :

$$\left. \frac{d\gamma}{dt} \right|_t = D\psi^{-1}(\psi \circ \gamma(t)) \left. \frac{d(\psi \circ \gamma)}{dt} \right|_t \in T_{\gamma(t)}M.$$

Por lo tanto, dado $\gamma \in H^1(I; M)$, tenemos la aplicación siguiente:

$$\frac{d\gamma}{dt}: I \longrightarrow TM \quad \text{a.e. en } [0, T],$$

con la propiedad de que $(\tau_M: TM \longrightarrow M)$ denota la proyección canónica):

$$\tau_M \circ \frac{d\gamma}{dt} \Big|_t = \gamma(t).$$

Esta aplicación $d\gamma/dt$ pertenece claramente a $L^2(I; TM) \equiv H^0(I; TM)$.

El siguiente teorema-definición resume la discusión anterior:

Teorema 2.1.3 *Al espacio $L_k^p(I; M)$, $k \geq 0$, $p > 1$, se le puede dotar con la estructura de una variedad diferenciable de Banach separable paracompacta modelada en el espacio de Banach $L_k^p(I; \mathbb{R}^n)$. Al espacio $L_1^2(I; M) = H^1(I; M)$ se le puede dotar con la estructura de una variedad diferenciable de Hilbert separable paracompacta modelada en el espacio de Hilbert $H^1(I; \mathbb{R}^n)$ con espacio tangente en γ dado por $H^1(I; \gamma^*(TM))$. La variedad de Hilbert de curvas en la variedad M , $H^1(I; M)$ se va a denotar en lo que sigue como $\mathcal{P}_I(M)$, $\mathcal{P}(M)$ ó simplemente como \mathcal{P} .*

El atlas diferenciable de \mathcal{P} viene dado por la siguiente familia de cartas. Como $C^\infty(I; M) \hookrightarrow H^1(I; M) \hookrightarrow C^0(I; M)$ y como $C^\infty(I; M)$ es denso en $C^0(I; M)$, entonces $C^\infty(I; M)$ va a ser denso en $H^1(I; M)$. Con esto se definen cartas como entornos de curvas en $C^\infty(I; M)$.

Sea $\gamma \in C^\infty(I; M)$ y $\varepsilon > 0$, definimos $(\mathcal{O}_{\gamma, \varepsilon}, \Phi_{\gamma, \varepsilon})$ de la forma siguiente: sea $H_\gamma^1 = H^1(\gamma^*(TM))$ y el entorno tubular de la sección cero, es decir, $\gamma_\varepsilon^*(TM) = \{v(t) \in T_{\gamma(t)}M \mid \|v(t)\| < \varepsilon\}$. Entonces

$$\mathcal{O}_{\gamma, \varepsilon} = \{v(t) = (\exp^X)(t) \mid \|X(t)\| < \varepsilon \quad \forall t \in [0, T]\},$$

es decir,

$$\begin{aligned} \mathcal{O}_{\gamma, \varepsilon} &= \exp^{(H^1(\gamma_\varepsilon^*(TM)))} \\ \Phi_{\gamma, \varepsilon} &: \mathcal{O}_{\gamma, \varepsilon} \longrightarrow H^1(I; \mathbb{R}^n) \\ &\quad v(t) \longrightarrow X(t) \end{aligned}$$

donde $v(t) = \exp^X(t)$.

2.1.7. Fibrados tangentes y otros fibrados sobre $\mathcal{P}_I^k(M)$

Sea $E \longrightarrow M$ un fibrado vectorial sobre M , donde M está dotada de una métrica, entonces podemos definir el espacio de secciones de Sobolev de E de clase L_k^p como la completación de las secciones C^∞ de E , es decir, si denotamos por:

$$\Gamma^\infty(E) = \{\sigma: M \longrightarrow E \mid \pi \circ \sigma = id_M, \sigma \in C^\infty\}$$

al espacio de las secciones C^∞ de E , tendremos:

$$L_k^p(E) = \overline{\Gamma^\infty(E)}^{\|\cdot\|_{k,p}},$$

donde definimos la norma usual como:

$$\|\sigma\|_{k,p}^p = \sum_{l=0}^k \int_M \|\nabla^l \sigma\|^p d^n x,$$

donde $d^n x = \sqrt{g} dx^1 \wedge \dots \wedge dx^n$ es el volumen Riemanniano y ∇ denota la derivada covariante definida por la conexión de Levi-Civita en M . Como estamos interesados en la clase H^1 , en este caso tenemos

$$\|\sigma\|_{1,2}^2 = \int_M \|\sigma\|^2 d^n x + \int_M \|\nabla \sigma\|^2 d^n x.$$

Estudiemos ahora los candidatos a fibrado tangente de $H^1(I; M)$. Sea una curva $\gamma \in H^1(I; M)$, como posee derivada clásica a.e. podemos considerar el pull-back $\gamma^*(TM)$ del fibrado tangente TM , es decir,

$$\gamma^*(TM) = \{(t, v) \in I \times TM \mid v \in T_{\gamma(t)}M\}.$$

Tenemos una aplicación natural $p_1: \gamma^*(TM) \rightarrow I$ que es C^∞ a.e. En cualquier caso, la aplicación p_1 es absolutamente continua y el fibrado es trivial.

Dado por tanto un camino $\gamma \in \mathcal{P}_I^k(M)$, podemos definir una familia de espacios tangentes en el punto γ . Consideremos el espacio de secciones de la clase de Sobolev l de $\gamma^*(TM) \cong \mathbb{R}^n \times I$. Este espacio puede ser obviamente identificado con el espacio de funciones $H^l(I, \mathbb{R}^n)$. Denotaremos por $T_\gamma^l \mathcal{P}$ tal espacio tangente. La unión de todos estos espacios definen un fibrado vectorial sobre \mathcal{P} con fibra estándar $H^l(I, \mathbb{R}^n)$. Denotaremos a este fibrado vectorial por $T^l \mathcal{P}^k$:

$$T^l \mathcal{P}^k = \bigcup_{\gamma \in \mathcal{P}} T_\gamma^l \mathcal{P}^k$$

Denotaremos por letras mayúsculas $U, V, \dots \in T_\gamma^l \mathcal{P}^k$ a los vectores tangentes a \mathcal{P}^k en el camino γ ó, simplemente, siguiendo la tradición del cálculo de variaciones, como $\delta\gamma \in T_\gamma^l \mathcal{P}^k$.

Recordemos por otro lado que un vector tangente es una clase de equivalencia de curvas sobre la variedad. Por tanto, podemos considerar una curva $\gamma: (-\epsilon, \epsilon) \rightarrow \mathcal{P}^k : s \rightsquigarrow \gamma_s; \gamma_s(t) \in M; t \in I; \gamma_0(t) = \gamma(t)$. Podemos también escribir la curva de caminos como una función de dos variables $\sigma: (-\epsilon, \epsilon) \times I \rightarrow M : \sigma(s, t) = \gamma_s(t)$. Es notable señalar que la dependencia en s de la curva γ_s es C^{l+1} mientras que la dependencia en t es H^k y depende del fibrado tangente $T^l \mathcal{P}^k$ en el que estamos trabajando. Notar que:

$$\delta\gamma(t) = \frac{d}{ds} \gamma_s(t)|_{s=0} = \frac{\partial}{\partial s} \sigma(s, t)|_{s=0} \in T_{\gamma(t)}M, \quad \forall t \in I.$$

Debido a que $\sigma(0, t) = \gamma(t) \in H^k(I; M)$ los elementos $\delta\gamma(t)$ pueden ser como mucho de clase de Sobolev k . Por tanto, los fibrados tangentes $T^l \mathcal{P}^k$ están definidos para $0 \leq l \leq k$.

Dado un espacio de caminos $\mathcal{P}_I^k(M)$, $k \geq 1$, existe siempre un campo de vectores canónico Γ sobre \mathcal{P}^k definido como:

$$\Gamma_\gamma \in T_\gamma^{k-1}\mathcal{P}^k, \quad \Gamma_\gamma(t) = \dot{\gamma}(t) \in T_{\gamma(t)}M$$

Es importante señalar que Γ es una sección diferenciable del fibrado $T^{k-1}\mathcal{P}^k$.

2.1.8. Algunas subvariedades de $\mathcal{P}_I^k(M)$

Hay dos aplicaciones canónicas $S: \mathcal{P}_I^k(M) \longrightarrow M$ y $T: \mathcal{P}_I^k(M) \longrightarrow M$, la fuente y el sumidero de un camino, es decir, si $I = [a, b]$, entonces:

$$S(\gamma) = \gamma(a) \quad y \quad T(\gamma) = \gamma(b).$$

Notar que si fijamos los extremos de dos caminos, digamos $m_0, m_1 \in M$, entonces el espacio de caminos con punto inicial m_0 es $S^{-1}(m_0) \in \mathcal{P}^k$ y el espacio de caminos con punto final m_1 es $T^{-1}(m_1)$. Debido a que ambas aplicaciones S y T son continuas, ambos subespacios son variedades cerradas de \mathcal{P}^k .

Sea ahora $x_0 \in M$, entonces definimos el conjunto de curvas que tienen como punto inicial x_0 de la forma siguiente:

$$\mathcal{P}_{x_0}(M) = \{\gamma: I \rightarrow M \mid \gamma(0) = x_0, \gamma \in H^1(I; M)\}. \quad (2.4)$$

Se ve que $\mathcal{P}_{x_0}(M)$ es una subvariedad de $\mathcal{P}(M)$ cuyo espacio tangente es el siguiente:

$$\begin{aligned} T_\gamma \mathcal{P}_{x_0}(M) = \{ \delta\gamma: I \rightarrow TM \mid \tau_M(\delta\gamma(0)) &= x_0, \delta\gamma \in H^1(\gamma^*(TM)), \\ \tau_M(\delta\gamma(t)) &= \gamma(t), \delta\gamma(0) = 0 \}. \end{aligned}$$

Por lo tanto, la inmersión $i: \mathcal{P}_{x_0}(M) \subset \mathcal{P}(M)$, que envía a la curva γ en ella misma, es una aplicación C^∞ cuya derivada es $i_{*\gamma}: T_\gamma \mathcal{P}_{x_0} \longrightarrow T_\gamma \mathcal{P}$, que no es otra que la identidad y siendo la codimensión de $T_\gamma \mathcal{P}_{x_0}$ igual a la dimensión de M . De la misma forma, definimos la variedad $\mathcal{P}_{x_0, x_T}(M)$ como la variedad de curvas cuyos extremos son x_0 y x_T respectivamente, es decir:

$$\mathcal{P}_{x_0, x_T}(M) = \{\gamma: I \rightarrow M \mid \gamma(0) = x_0, \gamma(T) = x_T, \gamma \in H^1(I; M)\}, \quad (2.5)$$

cuyo espacio tangente, se construye de manera similar al caso anterior:

$$\begin{aligned} T_\gamma \mathcal{P}_{x_0, x_T}(M) = \\ = \{ \delta\gamma: I \rightarrow TM \mid \tau_M \circ \delta\gamma = \gamma, \delta\gamma(0) = \delta\gamma(T) = 0, \delta\gamma \in H^1(\gamma^*(TM)) \}. \end{aligned}$$

Se ve, claramente, que \mathcal{P}_{x_0, x_T} tiene codimensión $2n$ en $\mathcal{P}(M)$.

Denotaremos en general por $\Omega_{x_0, x_T}(M)$ el subconjunto abierto en $\mathcal{P}_{x_0, x_T}(M)$ de las curvas de clase C^∞ . Al ser un abierto es una subvariedad abierta de $\mathcal{P}_{x_0, x_T}(M)$ y las nociones de espacio tangente, etc., son las mismas que las ya estudiadas.

Finalmente, definimos la subvariedad de curvas cerradas o periódicas como sigue:

$$\begin{aligned}\mathcal{L}(M) &= \{\gamma: I \rightarrow M \mid \gamma(0) = \gamma(T), \gamma \in H^1(I; M)\} \\ &= \{\gamma: S^1 \rightarrow M \mid \gamma \in H^1(S^1; M)\},\end{aligned}\tag{2.6}$$

cuya codimensión es n en $\mathcal{P}(M)$, y como espacio tangente la siguiente variedad:

$$T_\gamma \mathcal{L}(M) = \{\delta\gamma: I \rightarrow TM \mid \tau_Q M \circ \delta\gamma = \gamma, \delta\gamma(0) = \delta\gamma(T), \delta\gamma \in H^1(\gamma^*(TM))\}.$$

La variedad de curvas solución de una ecuación diferencial

Sea M una variedad diferenciable y sea $\Gamma_t \in \mathfrak{X}(M)$ un campo vectorial en M dependiente del tiempo, que en coordenadas locales x^i en M se escribe como

$$\Gamma_t = f^i(x, t)\partial/\partial x^i,$$

y cuyas curvas integrales satisfacen las ecuaciones diferenciales dependientes del tiempo:

$$\frac{dx^i}{dt} = f^i(x, t), \quad i = 1, \dots, n.$$

Vamos a considerar ahora la subvariedad de \mathcal{P}_Γ formada por las curvas que verifican las ecuaciones anteriores, es decir, consideramos aquellas curvas $\gamma: I \rightarrow M$ de clase H^1 tal que

$$\frac{d\gamma}{dt} = \Gamma(\gamma(t), t) \quad a.e. [0, T].$$

Nótese que si M es compacta, estas curvas existen porque el flujo de Γ_t es completo. Si no es compacta para cada x_0 existe un $T = T(x_0)$ y $\gamma: [0, T] \rightarrow M$ tal que $\gamma(0) = x_0$ y que satisface la ecuación anterior.

Tenemos la siguiente aplicación $A: \mathcal{P} \rightarrow H^0(I; TM)$ definida como:

$$A(\gamma) = \frac{d\gamma}{dt} - \Gamma(\gamma(t), t).\tag{2.7}$$

Recordemos que sobre $\mathcal{P}(I; M)$ hay dos fibrados vectoriales tangentes, $H^1(I; TM)$ y $H^0(I; TM)$. El primero modelado sobre $H^1(I; \mathbb{R}^n)$ y el segundo sobre $H^0(I; \mathbb{R}^n) = L^2(I; \mathbb{R}^n)$. El espacio $H^0(I; TM)$ es un fibrado de Hilbert sobre \mathcal{P} con una aplicación de proyección $\tau_0: H^0(I; TM) \rightarrow \mathcal{P}$ dada por $\tau_0(\delta\gamma) = \tau_M \circ \delta\gamma$. La aplicación A es una sección del fibrado $H^0(I; TM)$, $A \in \Gamma(H^0(I; TM))$, porque

$(\tau_o(A))\gamma = \tau_M \circ (d\gamma/dt - \Gamma(\gamma(t), t)) = \gamma(t)$, por lo tanto, el conjunto de soluciones de la ecuación diferencial antes mencionada va a estar formada por los ceros de la aplicación anterior, esto es, $A^{-1}(\mathbf{Z}_0)$, donde \mathbf{Z}_0 representa la sección cero de $H^0(I; TM) \longrightarrow \mathcal{P}(M)$. Tal conjunto es por tanto una subvariedad.

Definición 2.1.2 *La subvariedad dinámica o de las soluciones de Γ se define como:*

$$\mathcal{W} = A^{-1}(\mathbf{Z}_0) = \left\{ \gamma \in \mathcal{P} \left| \frac{d\gamma}{dt} - \Gamma(\gamma(t), t) = 0 \right. \right\}.$$

Tenemos que ver que es una subvariedad, para ello hay que demostrar que A es transversa a la sección cero de τ_0 , es decir, que $A_*\gamma + T_\gamma Z_0 = T_\gamma H^0(I; TM)$, o lo que es lo mismo, que la aplicación tangente a A que denotamos de la siguiente forma: $\mathfrak{L} := A_*$ es sobreyectiva, así, tendremos que \mathcal{W} es una subvariedad embebida. Si la aplicación es sobreyectiva, ha de cumplirse que

$$T_\gamma \mathcal{W} = \ker \mathfrak{L}_\gamma.$$

Para ello estudiamos la aplicación $\mathfrak{L}_\gamma: T_\gamma \mathcal{P} \longrightarrow T_{A(\gamma)} H^0(TM)$, donde hemos denotado como \mathfrak{L}_γ a la aplicación tangente de A en γ . Por definición tenemos que:

$$\begin{aligned} (\mathfrak{L}_\gamma)(\delta\gamma) &= \left. \frac{\partial}{\partial s} A(\sigma(s, t)) \right|_{s=0} = \\ &= \left. \frac{\partial}{\partial s} \left(\frac{d}{dt} \sigma(s, t) - \Gamma(\sigma(s, t), t) \right) \right|_{s=0} = \\ &= \left. \frac{\partial}{\partial s} \left(\frac{d}{dt} \sigma(s, t) \right) \right|_{s=0} - \left. \frac{\partial}{\partial s} (\Gamma(\sigma(s, t), t)) \right|_{s=0}. \end{aligned}$$

Para un t fijo $d\sigma(s, t)/dt$ es una curva $U(s)$ de vectores tangentes cuya derivada en s define un vector vertical en $T_{\delta\gamma(t)} TM$.

De la misma manera $\partial\Gamma(\sigma(s, t), t)/\partial s|_{s=0} = T\Gamma_{\gamma(t)}(\delta\gamma(t))$ y donde $\tau_Q \circ \Gamma(\gamma) = \gamma$ implica que $T\tau_M \circ T\Gamma = 0 =$ es decir $\partial\Gamma(\gamma(s, t), t)/\partial s|_{s=0}$ es vertical respecto a $T\tau_M$, por lo que tenemos que el rango $\mathfrak{L}_\gamma(\delta\gamma)$ está en el espacio $H^0(I; V(TTM))_\gamma \cong H^0(I; TM)_\gamma$.

Lema 2.1.2 *Sea $\tau_0: H^0(I; TM) \longrightarrow \mathcal{P}$ el fibrado H^0 -tangente de \mathcal{P} , entonces $TH^0(I; TM) \cong V(\tau_0) \oplus \tau_0^* H^0(I; TM)$, donde $V(\tau_0)$ es el subfibrado vertical respecto a la proyección τ_0 , es decir, $V(\tau_0) = \ker \tau_{0*}$. Más aún, $V(\tau_0) = H^0(I; V(TTM))$ y se puede identificar $V(TTM)$ con TM ; por lo que existe una identificación natural entre $V(\tau_0)$ y $H^0(I; TM)$.*

Utilizando la identificación anterior, podemos considerar la aplicación \mathfrak{L}_γ como una aplicación de $T_\gamma \mathcal{P}$ a $H^0(I; TM)$ y, por tanto, tenemos el siguiente lema.

Lema 2.1.3 *La aplicación $\mathfrak{L}_\gamma: T_\gamma\mathcal{P} \longrightarrow T_{A(\gamma)}H^0(I; TM)$ se define como la aplicación*

$$\begin{aligned} \mathfrak{L}_\gamma: H^1(\gamma^*(TM)) &\longrightarrow H^0(\gamma^*(TM)) = L^2(\gamma^*(TM)) \\ \delta\gamma &\rightsquigarrow \mathfrak{L}_\gamma(\delta\gamma) = \frac{d}{dt}\delta\gamma - D\Gamma(\gamma(t), t)\delta\gamma \end{aligned}$$

donde $D\Gamma$ es la diferencial de Γ en $(\gamma(t), t)$.

Lema 2.1.4 *El operador \mathfrak{L}_γ es un operador elíptico Fredholm cuyo rango es denso en $H^0(\gamma^*TM)$ y, por lo tanto, su núcleo tiene dimensión finita y $T_\gamma\mathcal{W} = \ker \mathfrak{L}_\gamma$.*

Los resultados previos nos muestran que \mathcal{W} es una variedad finita dimensional, lo que es obvio, debido a que para una condición inicial dada, hay una única curva $\gamma(t)$ que es solución de la ecuación diferencial dada, por lo tanto, el espacio \mathcal{W} tiene la misma dimensión que M , es decir, n , y la dimensión de $\ker \mathfrak{L}_\gamma$ es de nuevo n porque la solución de la ecuación lineal

$$\frac{d}{dt}\delta\gamma = D\Gamma(\gamma(t), t)\delta\gamma$$

es un espacio vectorial n dimensional.

2.2. Principios variacionales con simetría en mecánica: reducción simpléctica

Discutiremos en esta sección brevemente algunos aspectos del papel de las simetrías en mecánica.

2.2.1. Acciones de grupos de Lie

Revisaremos en primer lugar algunos conceptos y notaciones básicas de la teoría de grupos de Lie y de sus acciones sobre variedades diferenciables.

Consideremos un grupo de Lie G . Definimos las aplicaciones diferenciables $L_g: G \longrightarrow G, L_g h = gh$ y $R_g: G \longrightarrow G, R_g h = hg, \forall g, h \in G$; llamadas traslaciones por la izquierda y la derecha por el elemento $g \in G$ respectivamente. Es claro que $L_g L_h = L_{gh}$; $R_g R_h = R_{hg}$; $L_g R_h = R_h L_g$; $(L_g)^{-1} = L_{g^{-1}}$ y $(R_g)^{-1} = R_{g^{-1}} \forall g \in G$.

Un campo vectorial X sobre G se dice invariante a izquierda si se verifica que $(L_g)_* X = X \circ L_g, \forall g \in G$. Una forma ω sobre G es invariante a izquierda si $L_g^* \omega = \omega, \forall g \in G$ (de manera análoga definimos campos y formas invariantes a derecha). Llamamos $\mathfrak{X}_L(G)$ al subespacio vectorial de los campos vectoriales invariantes a

izquierda sobre G . Se puede ver que dicho subespacio es una subálgebra de Lie de $\mathfrak{X}(G)$ ya que $(L_g)_*([X, Y]) = [(L_g)_*X, (L_g)_*Y] = [X, Y] \circ L_g$. Identificamos los espacios $\mathfrak{X}_L(G)$ y T_eG a través de la aplicación $\xi \in T_eG \rightarrow X_\xi(g) = (L_g)_*(e)\xi$.

El álgebra de Lie \mathfrak{g} de G es T_eG con la estructura de álgebra de Lie dada por:

$$[\xi, \mu] = [X_\xi, X_\mu](e)$$

Dado $\xi \in \mathfrak{g}$, consideremos el campo X_ξ asociado. Se sabe que X_ξ genera, por ser invariante a izquierda, un subgrupo uniparamétrico completo $\phi_\xi(t)$ de G . Es decir, $\phi_\xi(t)$ es la curva integral de X_ξ . Se comprueba fácilmente que $\phi_\xi(t+s) = \phi_\xi(s) \cdot \phi_\xi(t)$. La función $\exp: \mathfrak{g} \rightarrow G$ dada por $\exp(\xi) = \phi_\xi(1)$ es la exponencial de \mathfrak{g} en G . La aplicación adjunta de g es la diferencial de $I_g = L_g R_{g^{-1}} = R_{g^{-1}} L_g$, en la identidad, esto es $Ad_g = (I_g)_*(e)$. Se tiene que $\exp(Ad_g \xi) = g(\exp \xi)g^{-1}$.

Acción de un grupo de Lie en una variedad diferenciable

Sea M una variedad diferenciable. Una acción a izquierda de un grupo de Lie G sobre M es una aplicación diferenciable $\Phi: G \times M \rightarrow M$ tal que

1. $\Phi(e, m) = m, \forall m \in M$,
2. $\Phi(g, \Phi(h, m)) = \Phi(gh, m), \forall g, h \in G$ y $\forall m \in M$.

Una acción Φ de G sobre M , define una familia de difeomorfismos Φ_g de M de modo que $\Phi_g(m) = \Phi(g, m)$ y se cumple que $\Phi_e = id|_M$, $\Phi_{gh} = \Phi_g \circ \Phi_h$ y por último $(\Phi_g)^{-1} = \Phi_{g^{-1}}$. En otros términos, una acción Φ de G sobre M puede pensarse como una aplicación $G \rightarrow \text{Diff}(M), g \rightarrow \Phi_g$ que es un difeomorfismo de grupos. Denotaremos habitualmente si no hay riesgo de confusión, $g.m \equiv gm \equiv \Phi_g(m) = \Phi(g, m)$. Si M es un espacio vectorial y cada Φ_g es una transformación lineal, la acción G sobre M es una representación de G sobre M . La acción coadjunta de G sobre \mathfrak{g}^* se define como $\langle Ad_g^* \alpha, \xi \rangle = \alpha(Ad_g \xi) = \langle \alpha, Ad_g \xi \rangle$. Dada una acción Φ de G sobre M , para cada $m \in M$, la órbita de m bajo la acción de Φ es $\mathcal{O}_m = \{gm \mid g \in G\}$. Para $m \in M$, el grupo de isotropía de m es el subgrupo de G , $G_m = \{g \in G \mid gm = m\}$. Algunas propiedades importantes de las acciones de grupos son las siguientes

- Una acción Φ es libre si $\forall m \in M, \Phi_m: G \rightarrow M, g \mapsto \Phi(g, m)$, es inyectiva.
- Una acción Φ es transitiva si existe una sola órbita.
- Una acción es efectiva si $\Phi_g = id \Rightarrow g = e$.
- Una acción Φ es propia si $\bar{\Phi}: G \times M \rightarrow M \times M$ dada por $\bar{\Phi}(g, m) = (m, \Phi_g(m))$ es propia, es decir, $\forall K \subset M$ compacto se tiene $\bar{\Phi}^{-1}(K \times K)$ es compacto.

Una acción de G sobre M define en M una relación de equivalencia:
 $m \sim m'$ iff $\exists g \in G, m' = gm$.

Dada Φ una acción de G sobre M , sea el conjunto

$$R = \{ (m, \Phi_g(m)) \mid g \in G \} \subset M \times M.$$

Si R es cerrado, la topología cociente en M/G es Hausdorff.

Si $\Phi_m|_{G/G_m}: G/G_m \longrightarrow \mathcal{O}_m$ es inyectiva y Φ es propia entonces \mathcal{O}_m es una subvariedad cerrada de M y $\Phi_m|_{G/G_m}$ es un difeomorfismo. Es decir, $G/G_m \cong \mathcal{O}_m$.

Si $\xi \in \mathfrak{g}$, llamaremos campo de Killing inducido por la acción Φ de G en M al campo ξ_M dado por

$$\xi_M(m) = \left. \frac{d}{dt} \Phi(\exp(t\xi), m) \right|_{t=0}, \quad m \in M.$$

Se verifica inmediatamente que dada una acción Φ de G sobre M , se tiene

1. $(Ad_g \xi)_M = (\Phi_{g^{-1}})_*(\xi_M) \quad \forall \xi \in \mathfrak{g} \text{ y } \forall g \in G$.
2. $[\xi_M, \eta_M] = -[\xi, \eta]_M \quad \forall \xi, \eta \in \mathfrak{g}$.

Se tiene también que:

$$T_m \mathcal{O}_m = \{ \xi_M(m) \mid \xi \in \mathfrak{g} \}.$$

Sea la acción dada por la multiplicación a izquierda de G sobre si mismo $L: G \times G \longrightarrow G$. Si $\xi \in \mathfrak{g}$, $\Phi^\xi: \mathbb{R} \times G \longrightarrow G$ está dada por $\Phi^\xi(t, g) = \exp(t\xi)g$. Entonces,

$$\xi_G(g) = \left. \frac{d}{dt} R_g(\exp(t\xi)) \right|_{t=0} = (R_g)_*(e)\xi.$$

Sea la acción adjunta de G sobre \mathfrak{g} . Si $\xi \in \mathfrak{g}$, se tiene que $\xi_{\mathfrak{g}} = ad_\xi$ donde

$$ad_\xi(\eta) = ad(\xi, \eta) = [\xi, \eta] = [X_\xi, X_\eta](e)$$

Sea la acción coadjunta de G sobre \mathfrak{g}^* . Entonces, $\xi_{\mathfrak{g}^*} = -ad_\xi^*$, $\forall \xi \in \mathfrak{g}$ donde ad_ξ^* denota la aplicación dual de ad_ξ , esto es, $\langle ad_\xi^* \mu, \zeta \rangle = \langle \mu, ad_\xi \zeta \rangle$, $\forall \zeta \in \mathfrak{g}$, $\mu \in \mathfrak{g}^*$.

2.2.2. Principios variacionales con simetría en Mecánica

A partir de ahora nos centraremos en el estudio de principios variacionales del tipo que son comunes en Mecánica, esto es, partiremos de un espacio de configuraciones Q , que supondremos una variedad diferenciable con coordenadas locales q^i , y una función diferenciable $L: TQ \longrightarrow \mathbb{R}$ definida en el fibrado tangente $\tau_Q: TQ \longrightarrow Q$ de Q . Denotaremos las coordenadas locales naturales en TQ como (q^i, v^i) donde v^i también se denotará ocasionalmente como \dot{q}^i . Consideremos finalmente el funcional S definido a través de la función L como:

$$S(\gamma) = \int_0^T L(\gamma(t), \dot{\gamma}(t)) dt$$

donde $(\gamma(t), \dot{\gamma}(t))$ es la curva en TQ definida por la curva γ en Q . La función L se llama lagrangiano del sistema y la función S , se denomina acción. Un espacio natural de curvas donde definir S será por ejemplo $\Omega([0, T], Q)$, esto es, el conjunto de curvas de clase C^∞ del intervalo $[0, T]$ en Q . Sin embargo recordemos que dicho espacio es una subvariedad abierta en el espacio de curvas de clase de Sobolev 1 en $[0, T]$, $H^1([0, T], Q) =: \mathcal{P}_{[0, T]}(Q)$ (ver sección §2.1.6).

El funcional S resulta ser diferenciable en $\mathcal{P}(Q)$ bajo condiciones suficientemente generales de la función L . Supondremos en el contexto de este trabajo que la función L es suficientemente regular de tal manera que S sea diferenciable (ver por ejemplo [De04]) aunque tenemos una condición suficiente de diferenciabilidad como veremos en la siguiente proposición.

Diferenciabilidad de S .

Consideremos el espacio de caminos en Q de clase de Sobolev (k, p) , esto es, el espacio que hemos denotado por $\mathcal{P}^{k,p}(Q)$, $k \geq 1$, $p \geq 1$. Definamos en este espacio como anteriormente el funcional $S(\gamma) = \int_0^T L(\gamma(t), \dot{\gamma}(t)) dt$, $\gamma \in \mathcal{P}^{k,p}(Q)$ y determinemos bajo qué condiciones es diferenciable.

Calculemos en primer lugar la derivada direccional de S en la dirección del vector $\xi \in T_\gamma \mathcal{P}^{k,p}(Q)$. Nótese que por ser $\xi: [0, T] \rightarrow \gamma^*(TQ)$ una sección de clase de Sobolev (k, p) , se tiene que las derivadas de ξ con respecto a t existen hasta orden k y pertenecen a L^p . Por tanto la familia de curvas $\sigma(s, t)$ tal que $\frac{\partial \sigma}{\partial s}(s, t) |_{s=0} = \xi(t)$, es de clase C^∞ en s y tiene derivadas débiles hasta orden k en L^p con respecto a t . Por otro lado tenemos que

$$dS_\gamma(\xi) = \frac{\partial}{\partial s} \left(\int_0^T L(\sigma(s, t), \dot{\sigma}(s, t)) dt \right) \Big|_{s=0}.$$

Para poder intercambiar la integral con la derivada tenemos que imponer condiciones de regularidad en L . Si por ejemplo $k \geq 2$ y $p = 2$, entonces la derivada $\dot{\sigma}(s, t)$ está en H^1 y por tanto es una función absolutamente continua en t . Si asumimos que L es diferenciable (por ejemplo de clase C^∞) entonces la función $L(\sigma(s, t), \dot{\sigma}(s, t))$ es integrable, diferenciable respecto a s y continua respecto a t , por lo que podemos derivar dentro del signo integral. Si por el contrario $k = 1$, entonces $\dot{\sigma}(s, t)$ está en L^p y tenemos que imponer una condición de crecimiento en L con respecto a v , por ejemplo de $L(q, v) \leq c(|q|)|v|^p$ para alguna función continua c en q . Tenemos por tanto el siguiente resultado sobre diferenciabilidad de S que va a ser suficiente para los objetivos de esta tesis.

Proposición 2.2.1 Sea $L: TQ \longrightarrow \mathbb{R}$ una función diferenciable de clase C^∞ y consideremos el funcional acción $S: \mathcal{P}^k(Q) \longrightarrow \mathbb{R}$ definido en el espacio de caminos de clase de Sobolev $(k, 2)$, $k \geq 2$ en la variedad Q . Entonces S es diferenciable de clase C^1 , y además:

$$dS_\gamma(\xi) = \int_0^T \left\langle \frac{\partial L}{\partial q}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)), \xi(t) \right\rangle dt, \quad \forall \xi \in T_\gamma \mathcal{P}^k(Q).$$

Demostración. Como ya comentamos con anterioridad, si $k \geq p = 2$, entonces podemos intercambiar la derivación con respecto a la integral. Tras una integración por partes obtenemos:

$$dS_\gamma(\xi) = \int_0^T \left\langle \frac{\partial L}{\partial q}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)), \xi(t) \right\rangle dt$$

y por tanto $|dS_\gamma(\xi)| \leq C \|\xi\|_{k,2}$ ya que las funciones $\frac{\partial L}{\partial q}$, $\frac{\partial^2 L}{\partial q \partial v}$ y $\frac{\partial^2 L}{\partial^2 v}$ son acotadas en $[0, T]$ al evaluarlas sobre $\gamma(t), \dot{\gamma}(t)$.

Por otro lado notemos que:

$$\int_0^T \frac{d}{dt} \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)) \xi(t) dt = \int_0^T \frac{\partial^2 L}{\partial q \partial v} \dot{\gamma}(t) \xi(t) dt + \int_0^T \frac{\partial^2 L}{\partial^2 v} \ddot{\gamma}(t) \xi(t) dt.$$

Por tanto,

$$\left| \int_0^T \frac{d}{dt} \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)) \xi(t) dt \right| \leq \int_0^T \left| \frac{\partial^2 L}{\partial q \partial v} \right| |\dot{\gamma}(t) \xi(t)| dt + \int_0^T \left| \frac{\partial^2 L}{\partial^2 v} \right| |\ddot{\gamma}(t) \xi(t)| dt,$$

pero

$$\begin{aligned} & \int_0^T \left| \frac{\partial^2 L}{\partial q \partial v}(\gamma(t), \dot{\gamma}(t)) \right| |\dot{\gamma}(t) \xi(t)| dt \leq \\ & \leq C_1 \int_0^T |\dot{\gamma}(t) \xi(t)| dt \leq C_1 \|\dot{\gamma}\|_2 \|\xi\|_2 \leq C_1 \|\gamma\|_{k,2} \|\xi\|_{k,2} \end{aligned}$$

y por otro lado

$$\int_0^T \left| \frac{\partial^2 L}{\partial^2 v}(\gamma(t), \dot{\gamma}(t)) \right| |\ddot{\gamma}(t) \xi(t)| dt \leq C_2 \int_0^T |\ddot{\gamma}(t) \xi(t)| dt.$$

Como $\ddot{\gamma} \in H^{k-2}$ con $k \geq 2$, entonces $\ddot{\gamma} \in L^2$ y por tanto se tiene que:

$$\int_0^T |\ddot{\gamma}(t) \xi(t)| dt \leq \|\ddot{\gamma}\|_2 \|\xi\|_2 \leq C_3 \|\gamma\|_{k,2} \|\xi\|_{k,2}.$$

Por tanto tenemos que las derivadas parciales existen y son continuas por tanto S es diferenciable con continuidad. \square

Las ecuaciones de Euler-Lagrange determinando los puntos críticos del funcional S se obtendrán computando $dS(\gamma) = 0$ (ver por ejemplo la Tesis de Carlos López [Lo89] para una derivación intrínseca de dichas fórmulas).

Para el propósito general de este trabajo, es conveniente derivar las ecuaciones de Euler-Lagrange de manera más intrínseca en un contexto más general.

2.2.3. Elementos geométricos en Mecánica.

El objeto básico de la formulación geométrica de los sistemas hamiltonianos es una variedad simpléctica. Esto es, una variedad diferenciable M con una 2-forma diferencial cerrada y no degenerada. Una 2-forma ω sobre M es no degenerada si $\omega(X, Y) = 0, \forall X \in \chi(M) \Rightarrow Y = 0$. El ejemplo por excelencia de las variedades simplécticas es el fibrado cotangente de una variedad diferenciable Q que representa el espacio de configuración del sistema. Sea Q una variedad de dimensión n y (q_i) un sistema local de coordenadas en Q . Consideremos su fibrado cotangente $M = T^*Q$ y $\pi_Q^*: M \rightarrow Q$ la proyección canónica. Se define sobre M la 1-forma θ como

$$\theta(\alpha_q)(v_{\alpha_q}) = \alpha_q(d\pi_Q^*v_{\alpha_q})$$

donde $\alpha_q \in T_q^*Q$ y $v_{\alpha_q} \in T_{\alpha_q}(T_q^*Q)$. La forma simpléctica canónica sobre M está dada por $\omega = -d\theta$. Si (q_i, p_i) son coordenadas en T^*Q ,

$$\theta = p^i dq_i, \omega = dq_i \wedge dp^i$$

La energía del sistema se representa con una función llamada hamiltoniano del sistema $H: T^*Q \rightarrow \mathbb{R}$ que permite escribir las ecuaciones de movimiento. Por otro lado, el fibrado tangente TQ es el espacio de velocidades del sistema y es uno de los elementos fundamentales en la descripción Lagrangiana de la mecánica. Dada la aplicación lagrangiana $L: TQ \rightarrow \mathbb{R}$ se define la transformada de Legendre $FL: TQ \rightarrow T^*Q$ como

$$FL(v) \cdot w = \left. \frac{d}{dt} (L(v + tw)) \right|_{t=0}$$

Es decir, $FL(v) \cdot w$ es la derivada de L en v a lo largo de la fibra T_qQ en la dirección de w y también es llamada derivada a lo largo de la fibra. La energía del sistema está dada por la aplicación $E: TQ \rightarrow \mathbb{R}$ donde $E(v) = FL(v) \cdot v - L(v)$. En coordenadas (q_i, \dot{q}_i) de TQ , FL se escribe como

$$FL(q_i, \dot{q}_i) = \left(q_i, \frac{\partial L}{\partial \dot{q}_i} \right)$$

Es decir, $p^i = \frac{\partial L}{\partial \dot{q}_i}$. La transformada de Legendre permite pasar de la formulación lagrangiana a la formulación hamiltoniana de la mecánica. Con FL se definen en TQ las 1-forma y 2-forma

$$\theta_L = FL^*\theta, \omega_L = FL^*\omega$$

En coordenadas,

$$\theta_L = \frac{\partial L}{\partial \dot{q}_i}$$

y

$$\omega_L = \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} dq_i \wedge dq_j + \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j} dq_i \wedge d\dot{q}_j$$

Un lagrangiano se dice regular si FL es un difeomorfismo. En este caso, ω_L define una estructura simpléctica canónica sobre TQ .

Definimos un funcional S_H como:

$$S_H(\sigma) = \int_0^T (\sigma^*(\theta - H)) dt = \int_0^T (\theta_{\sigma(t)}(\dot{\sigma}(t)) - H(\sigma(t))) dt$$

El funcional S_H es a menudo llamado el funcional de Hamilton. De nuevo S_H está definido sobre el espacio de curvas diferenciables σ sobre M , pero es conveniente considerar el marco analítico dado por $H^1([0, T], M) =: \mathcal{P}(M)$. Ahora tenemos:

$$\begin{aligned} dS_H(\sigma)(\delta\sigma) &= \left. \frac{d}{ds} S_H(\sigma_s(t)) \right|_{s=0} = \frac{d}{ds} \int_0^T (\sigma_s^* \theta - \sigma_s^* H) dt = \\ &= \int_0^T (\omega(X_t, \dot{\sigma}) - i_{X_t} dH)(\sigma(t)) dt = \int_0^T i_{X_t} \lrcorner (\omega_{\sigma(t)}(\dot{\sigma}, \cdot) - dH(\sigma(t))) dt \end{aligned}$$

donde X_t es un campo vectorial cualquiera a lo largo de $\sigma(t)$, es decir $X_t(\sigma(t)) = \delta\sigma(t)$. Por tanto la ecuación de Euler-Lagrange para el funcional S_H resultan ser:

$$i_{\Gamma}\omega = dH$$

esto es la ecuación de Hamilton en forma simpléctica.

2.2.4. Reducción de las ecuaciones de Euler-Lagrange

Hemos visto por tanto que si tenemos un funcional mecánico $S = \int L dt$, donde L es una función definida en el fibrado tangente TQ de un espacio de configuraciones y si además hay un grupo de Lie G actuando en Q y cuya acción natural en TQ deja invariante L , esto es:

$$L(gq, g\dot{q}) = L(q, \dot{q}), \quad \forall (q, \dot{q}) \in TQ, \quad g \in G,$$

entonces G actúa como un grupo de simetría del funcional S . El levantamiento natural de la acción de G a TQ es Por tanto, si la curva γ es un punto crítico de dicho funcional, la curva $g.\gamma$ definida como $(g.\gamma)(t) = g(\gamma(t))$, $\forall t$ también será un punto crítico de S .

Dado que el Lagrangiano L es invariante podemos plantearnos si podemos simplificar de alguna manera el cálculo de las curvas críticas de S . En efecto, como ya vimos en la sección §2.2.2 las curvas críticas de S son curvas integrales del campo vectorial Γ que satisface la ecuación:

$$i_{\Gamma}\omega_L = -dE_L$$

donde $\omega_L = d\theta_L$ es la 2-forma de Cartan, θ_L es la 1-forma de Poincaré-Cartan y E_L es la energía del Lagrangiano L .

Si el grupo G por ejemplo es compacto, es fácil comprobar que la cantidad

$$J_{\xi}(q, \dot{q}) = \xi_Q^i(q) \frac{\partial L}{\partial v^i}(q, \dot{q}), \quad \xi \in \mathfrak{g}$$

donde ξ_Q es el campo vectorial en Q correspondiente al elemento $\xi \in \mathfrak{g}$, es una constante del movimiento (teorema de Noether). Podemos definir una aplicación $J: TQ \longrightarrow \mathfrak{g}^*$ a través de:

$$J_{\xi}(q, \dot{q}) := \langle J(q, \dot{q}), \xi \rangle.$$

que verifica la propiedad de ser invariante con respecto a la dinámica Γ , esto es, las curvas integrales de Γ están contenidas en conjuntos de nivel de J . Dicha aplicación es comúnmente conocida como aplicación momento. Por tanto, a la hora de determinar puntos críticos de S podemos restringirnos a considerar los puntos críticos de S en una subvariedad de $\mathcal{P}(Q)$ definida por la condición $J = \text{cte}$. Consideremos en concreto la subvariedad $J = 0$. En dicho caso nuestro espacio sería ahora $\mathcal{P}(J^{-1}(0))$ y todavía en dicho espacio tenemos una acción del grupo G . En efecto, si $\gamma \in \mathcal{P}(J^{-1}(0))$, i.e. $J(\gamma(t)) = 0$, $\forall t$, entonces $g.\gamma \in \mathcal{P}(J^{-1}(0))$ ya que $J(g.\gamma(t)) = Ad_{g^{-1}}^* J(\gamma(t)) = 0$ ya que es fácil comprobar que la aplicación momento es equivariante, i.e.:

$$J(g.\gamma) = Ad_{g^{-1}}^* J(\gamma)$$

Ahora el grupo G actúa en el conjunto de nivel $J^{-1}(0)$, así que podemos definir el espacio cociente $J^{-1}(0)/G$. Por tanto, si tenemos las curvas γ y γ' que son puntos críticos de S y están relacionadas por un elemento de G , proyectaran en la misma curva en $J^{-1}(0)/G$. Por tanto podríamos intentar resolver el problema de la determinación de los puntos críticos de S mirando solamente en el espacio

reducido $J^{-1}(0)/G$. Ocurre que este espacio $J^{-1}(0)/G$ tiene una estructura simpléctica natural ω_G y el campo vectorial Γ proyecta a un campo vectorial Γ_G tal que:

$$i_{\Gamma_G}\omega_G = -dE_G$$

donde otra vez E_G es la proyección de la función de energía E_L . Este procedimiento de pasar de (Γ, ω_L, E_L) a $(\Gamma_G, \omega_G, E_G)$ es llamado reducción simpléctica y fue formalizado por Marsden y Weinstein [Ma74] y parcialmente por Arnold [Ar76].

2.2.5. Reducciones y cocientes de espacios de caminos

Como hemos visto, una situación habitual en el cálculo de variaciones consiste en la presencia de un grupo de simetría. Si tenemos un funcional $S(\gamma) = \int_0^T F(\gamma(t)) dt$ definido en el espacio de caminos $\mathcal{P}^k(M)$ a través de una función $F: M \rightarrow \mathbb{R}$, las ecuaciones de Euler-Lagrange serían simplemente $dS_\gamma = 0$, y los puntos críticos γ de S se caracterizarán como aquellas curvas $\gamma: [0, T] \rightarrow M$ tales que $dF(\gamma(t)) = 0$, $t \in [0, T]$. Nótese que si $k \geq 1$ y F es diferenciable, entonces S es diferenciable y además $dS_\gamma = \int_0^T dF(\gamma(t)) dt$, esto es la derivada direccional de S ó dF . Por tanto si F tiene puntos críticos aislados (por ejemplo F es una función de Morse), entonces $\gamma(t) = cte \in C(F)$. Los puntos críticos de S son por tanto curvas constantes con valores en puntos críticos de S .

Supongamos que el grupo G actúa en M y F es G -invariante. Está claro entonces que si M/G es una variedad, F desciende a una función f en M/G ; $f([x]) = F(x)$, $\forall x \in M$, y el principio variacional definido por F puede reducirse al principio variacional $S_G([\gamma]) = \int_0^T f([\gamma(t)]) dt$ donde $[\gamma]: [0, T] \rightarrow M/G$ denota una curva en M/G . El funcional S_G está definido en $\mathcal{P}^k(M/G)$ y sus puntos críticos son curvas $[\gamma]$ en M/G , tales que $[\gamma(t)] = cte \in C(f)$. Dado que los puntos críticos de f están en correspondencia uno a uno con G -órbitas críticas de F en M , tenemos que $C(F)/G = C(f)$.

Por otro lado observamos que el funcional $S: \mathcal{P}^k(M) \rightarrow \mathbb{R}$ es invariante con respecto al grupo $\mathcal{G} = \mathcal{P}^k(G)$ de caminos de clase k en G . En efecto si $g(t) \in \mathcal{P}^k(G)$, entonces $S(g(t)\gamma(t)) = \int_0^T F(g(t)\gamma(t)) dt = \int_0^T F(\gamma(t)) dt = S(\gamma)$. Por tanto vemos que S_G es realmente el funcional inducido por S en $\mathcal{P}(M)/\mathcal{P}(G) \cong \mathcal{P}(M/G)$.

Las cosas se complican si la densidad que empleamos para definir S es de orden superior, como por ejemplo en el caso de la mecánica. En efecto, si $S(\gamma) = \int_0^T L(\gamma(t), \dot{\gamma}(t)) dt$, con $L: TQ \rightarrow \mathbb{R}$, entonces aunque $S: \mathcal{P}^k(Q) \rightarrow \mathbb{R}$ sea diferenciable como antes y sea G -invariante con respecto a la acción de un grupo G , $L(gq, gv) = L(q, v)$, $\forall q, v \in TQ$, $g \in G$, no se verifica que S desciende al espacio $\mathcal{P}^k(Q/G)$. En efecto el principio variacional reducido está definido por la densidad lagrangiana $l: TQ/G \rightarrow \mathbb{R}$; definido por $l([q, v]) = L(q, v)$.

Podemos identificar $TQ/G \cong T(Q/G) \oplus \tilde{\mathfrak{g}}$ donde $\tilde{\mathfrak{g}}$ denota el fibrado de Atiyah (ver sección 3.1.1) y los elementos $[q, v_q] \in TQ/G$ los podemos denotar por $[q, v_q] = (x, v_x; \xi)$, $x = \pi(Q)$, $v_x = \pi_*(v_q)$, $\pi: Q \longrightarrow Q/G$, y $\xi \in \mathfrak{g}$.

Por tanto el principio variacional reducido está definido en el espacio de curvas $\mathcal{P}^k(\tilde{\mathfrak{g}})$. Nótese que $\tilde{\mathfrak{g}} \longrightarrow Q/G$ es el fibrado Ad-asociado a $Q \longrightarrow Q/G$ y por tanto con base Q/G . Así las curvas σ en $\mathcal{P}^k(\tilde{\mathfrak{g}})$ pueden escribirse como $\sigma(t) = (x(t); \xi(t))$, $x(t) \in Q/G$, $\xi(t) \in \mathfrak{g}$. Podemos escribir por tanto $S_G: \mathcal{P}^k(\tilde{\mathfrak{g}}) \longrightarrow \mathbb{R}$; $S_G(\sigma) = \int_0^T l(x(t), \dot{x}(t); \xi(t)) dt$. Falta por precisar en qué sentido S_G es el cociente de S por la acción del grupo G . Nótese en primer lugar que $\mathcal{P}^k(Q)/G \cong \mathcal{P}^k(\tilde{\mathfrak{g}})$. En efecto, si $\gamma(t) \in \mathcal{P}^k(Q)$, la curva $\sigma(t) = (\pi\gamma(t), \xi(t) = A_{\gamma(t)}(\dot{\gamma}(t)) \in \tilde{\mathfrak{g}})$. En efecto, nótese que si reemplazamos $\gamma(t) \rightsquigarrow \gamma'(t) = g\gamma(t)$ entonces $\xi(t) \rightsquigarrow \xi'(t) = A_{g\gamma(t)}(g\dot{\gamma}(t)) = (L_g^*A)_{\gamma(t)}(\dot{\gamma}(t)) = Ad_g^{-1}A_{\gamma(t)}(\dot{\gamma}(t)) = Ad_g^{-1}\xi(t)$; donde hemos usado una conexión A en $Q \longrightarrow Q/G$.

2.3. El problema general del análisis de puntos críticos de funciones con simetrías y ligaduras

2.3.1. Puntos críticos de funcionales con simetría.

El planteamiento general para los problemas de optimización consiste en determinar los máximos ó mínimos de una función S en la que estamos interesados, sobre un cierto espacio X , típicamente un espacio funcional con alguna estructura topológica ó métrica adicional (preferiblemente una estructura de espacio ó variedad de Banach). Podemos aproximarnos a la determinación de los extremos de S desde varias perspectivas. Los llamados métodos directos pretenden probar la existencia (y unicidad) de ellos poniendo directamente condiciones adicionales sobre el espacio X y la función S (ver capítulo I, §2.1.2). Un acercamiento diferente consiste en caracterizar los puntos críticos de S (ver capítulo §2.1.3) y determinar entre ellos, los correspondientes máximos ó mínimos. Si la función S es suficientemente regular, los puntos críticos son simplemente las soluciones de la condición de Euler-Lagrange:

$$dS(x) = 0 \tag{2.8}$$

donde dS denota la diferencial de S . Para funciones S definidas sobre espacios de funciones con propiedades locales apropiadas, la condición de Euler-Lagrange (2.8) se convierte en un sistema de ecuaciones diferenciables (ordinarias ó parciales) cuyas soluciones podemos probar con métodos alternativos. Históricamente ambos caminos han sido seguidos. Lagrange resolvió varios problemas mecánicos que fueron formulados como problemas de optimización, resolviendo las correspondientes ecuaciones de Euler-Lagrange, y alternativamente, Riemann resolvió

la ecuación de Laplace considerándola como una ecuación de Euler-Lagrange de un problema de optimización.

Los problemas de Control Óptimo son problemas de optimización sujetos a ligaduras adicionales (la ecuación de control y posiblemente junto con otras ligaduras). Los métodos directos son difíciles de implementar debido a la delicada naturaleza de la ecuación de control (típicamente una ecuación diferencial), sin embargo el principio del máximo de Pontryagin proporciona una maravillosa herramienta para buscar las soluciones al problema combinando ambos, el desarrollo de Euler-Lagrange y la caracterización directa de los extremales.

Desafortunadamente (aunque a veces también afortunadamente) las cosas no siempre son tan simples. En muchas ocasiones los problemas, tanto de la Física como de la Ingeniería, exhiben simetrías, es decir, hay un grupo de transformaciones G actuando sobre el espacio X que dejan a S invariante. Como indicamos anteriormente, si denotamos por $g \in G$ un elemento típico del grupo, $g \cdot x \in X$ al elemento transformado de $x \in X$, entonces:

$$S(g \cdot x) = S(x) \quad (2.9)$$

La primera observación que emerge de la propiedad de invarianza anterior (2.9) es que los extremales aparecen en familias.

Si x es un extremal, $g \cdot x$ es también un extremal de S . Más aún, si G es un grupo continuo (por ejemplo un grupo de Lie), entonces los extremales no pueden ser aislados y deben ser degenerados. Un modo de deshacer las dificultades planteadas por estas consideraciones es eliminar la ambigüedad causada por las simetrías de G , pasando al espacio cociente $X_G = X/G$. Efectivamente, la función S pasa al cociente definiendo una función $S_G: X_G \rightarrow \mathbb{R}$, tal que:

$$S = S_G \circ \pi \quad (2.10)$$

con $\pi: X \rightarrow X/G$ la proyección canónica.

Discutiremos ahora las propiedades de los puntos críticos de S y S_G . Para ello asumiremos por simplicidad que el espacio X es una variedad diferenciable y S es una función diferenciable. También asumiremos que X/G es una variedad diferenciable y la proyección canónica $\pi: X \rightarrow X/G$ es una submersión. Denotaremos el punto $\pi(x)$, es decir, la órbita $G \cdot x$ de G a través de x , como x_G ó \bar{x} .

Ahora supongamos que $x_0 \in X$ es un máximo local (ó mínimo) de S , es decir, $\exists U$ un entorno abierto de x_0 tal que $S(x) \geq S(x_0), \forall x \in U$. Por tanto, $\bar{U} = \pi(U)$ es un entorno abierto de $\pi(x_0) \in X_G$, and $S_G(\bar{x}) \geq S_G(\bar{x}_0), \forall \bar{x} \in \bar{U}$, luego \bar{x}_0 es un máximo local de S_G . Inversamente, si \bar{x}_0 es un máximo local de S_G , entonces $\exists \bar{U}$ un entorno abierto de \bar{x}_0 tal que $S_G(\bar{x}) \geq S_G(\bar{x}_0), \forall \bar{x} \in \bar{U}$. Consideremos ahora el entorno abierto $U = \pi^{-1}(\bar{U})$ de x_0 . Ahora, por definición,

$S(x_0) = S_G(\bar{x}_0)$, $\forall x_0 \in \pi^{-1}(\bar{x}_0)$. Entonces, $S(x) = S_G(\bar{x}) \geq S_G(\bar{x}_0) = S(x_0)$ y x_0 es un máximo local de S .

Estudiemos ahora los puntos críticos de S y S_G . Denotaremos por $E = T_{x_0}X$ y $\bar{E} = T_{\bar{x}_0}\bar{X} = T_{\bar{x}_0}X_G$, el espacio tangente de X y X_G en x_0 y \bar{x}_0 respectivamente. Si la acción de G es regular ó el grupo G es compacto por ejemplo, existen secciones locales σ para la función de proyección π . Entonces podemos usar una de ellas para descomponer el espacio E como suma directa $E = H \oplus V$ donde $V = \text{Ker}T\pi(x_0)$ es el espacio vertical y H es un subespacio suplementario tal que $T\pi(x_0)$ restringido a H es un isomorfismo sobre \bar{E} , $H \cong \bar{E}$. Por ejemplo H puede definirse como $T\sigma_{\bar{x}_0}(\bar{E}) \subset E$ si $\sigma(\bar{x}_0) = x_0$.

Por ejemplo, si X es una variedad de Hilbert, siempre podemos construir tal subespacio horizontal H , simplemente tomando el complemento ortogonal a V , es decir, $H = V^\perp$. El subespacio H , sin embargo, no será en general isomorfo (isométrico) a \bar{E} . La diferencial de S en x_0 será una aplicación lineal continua $s = dS(x_0): E \rightarrow \mathbb{R}$ y la diferencial de S_G en \bar{x}_0 será también una aplicación lineal continua $\bar{s} = dS_G(\bar{x}_0): \bar{E} \rightarrow \mathbb{R}$.

Como $E = H \oplus V$, entonces $s = s^H + s^V$ donde $s^H(u) = s(h)$ para $u = h + v$, $h \in H$, $v \in V$ y $s^V(u) = s(v)$. Ahora, es claro que $s^V = 0$ porque si $x(t) \in X$ es una curva en X con $x(0) = x$ tal que $\dot{x}(0) = v \in V$, entonces $\pi(x(t)) = \bar{x}_0$ para $t \in (-\varepsilon, \varepsilon)$, y

$$s(v) = \left. \frac{d}{dt} S(x(t)) \right|_{t=0} = \left. \frac{d}{dt} S_G(\bar{x}_0) \right|_{t=0} = 0 \quad (2.11)$$

Por tanto, concluimos que

$$s = \bar{s} \circ T\pi(x_0) \quad (2.12)$$

porque si $x(t) \in X$ es cualquier curva con $x(0) = x_0$, entonces $\pi x(t) = \bar{x}(t)$ y si $u = \bar{x}(0)$, $\bar{u} = \dot{\bar{x}}(0)$ y $T\pi(x_0)u = \bar{u}$, tenemos:

$$s(u) = \left. \frac{d}{dt} S(x(t)) \right|_{t=0} = \left. \frac{d}{dt} S_G(\bar{x}(t)) \right|_{t=0} = \bar{s}(\bar{u}) = \bar{s}T\pi(x_0)(u) \quad (2.13)$$

que también puede ser obtenido aplicando la regla de la cadena a (2.10). Ahora bien, la ecuación (2.12) nos indica que los puntos críticos de S proyectan sobre los puntos críticos de S_G y viceversa. Si denotamos por $C(S)$ al conjunto de puntos críticos de S , entonces tenemos:

$$C(S)/G = C(S_G).$$

2.3.2. El problema de la reducción y los principios variacionales con simetría

En el formalismo Lagrangiano que estamos usando para esta discusión, podemos decir incluso más, ya que se puede ver fácilmente que $J^{-1}(0)/G$ es $T(Q/G)$. Por tanto podemos concluir que G -órbitas de puntos críticos del funcional S que yacen en el conjunto de nivel $J = 0$ de la aplicación momento, están en correspondencia con las curvas integrales del sistema Hamiltoniano definido en el espacio reducido $(T(Q/G), \omega_G)$.

Usando las ideas anteriores habríamos conseguido una importante simplificación del problema, ya que hemos reducido de un modo significativo el número de grados de libertad (la dimensión de Q) del problema al tomar el cociente por G y eliminar la simetría. El mismo argumento puede ser repetido cambiando el conjunto de nivel de J por un $\mu \neq 0$. Ahora, sin embargo, aparecen términos extra en el formalismo que hace la discusión menos clara. Ver [Ce01a], [Ma90].

Desafortunadamente, aquí tenemos una dificultad: el funcional S es G -invariante, sin embargo el cociente del espacio $\mathcal{P}(Q)$ (donde S está definido) por G no es el espacio donde está definido el funcional reducido S_G . De hecho S_G está definido en $\mathcal{P}(Q/G)$, mientras que $\mathcal{P}(Q)/G$ es un espacio mucho mayor. Si consideramos ahora el grupo infinito dimensional $\mathcal{P}(G)$, entonces es fácil demostrar que:

$$\mathcal{P}(Q)/\mathcal{P}(G) = \mathcal{P}(Q/G).$$

Además el funcional S no es invariante por la acción natural de $\mathcal{P}(G)$ en $\mathcal{P}(Q)$. Esto puede ser visualizado inmediatamente ya que si $\gamma(t) \in \mathcal{P}(Q)$ y $g(t) \in \mathcal{P}(G)$, entonces

$$\widehat{g\dot{\gamma}}(t) = [\dot{g}(t)g^{-1}(t)]g(t)\gamma(t) + g(t)\dot{\gamma}(t)$$

Por tanto el Lagrangiano correspondiente a la curva $g(t)\gamma(t)$, será de la forma:

$$L(g(t)\gamma(t), \widehat{g(t)\dot{\gamma}}(t)) = L(\gamma(t), \hat{\xi}(t)_Q + \dot{\gamma}(t))$$

donde $\xi(t) = \dot{g}(t)g^{-1}(t) \in \mathfrak{g}$ y $\hat{\xi}(t)_Q$ es el campo vectorial sobre Q definido por el vector dependiente del tiempo $\xi(t)$ sobre \mathfrak{g} . Ahora es obvio que si:

$$L(\gamma(t), \dot{\gamma}(t)) \neq L(\gamma(t), \hat{\xi}(t)_Q + \dot{\gamma}(t))$$

entonces el Lagrangiano no será invariante por $\mathcal{P}(G)$ y el funcional S no será $\mathcal{P}(G)$ invariante.

2.3.3. Puntos críticos de funcionales con ligaduras: Teorema de los multiplicadores de Lagrange

En esta sección describiremos una forma del teorema de los multiplicadores de Lagrange que es apropiada para los propósitos de esta tesis.

Teorema 2.3.1 (Teorema de los multiplicadores de Lagrange) *Sea*

$E \xrightarrow{\pi} M$ un fibrado vectorial sobre una variedad de Hilbert diferenciable para-compacta M , modelado sobre el espacio de Hilbert real F con conexión simétrica ∇ . Sea $s \in \Gamma(E)$ una sección diferenciable de E transversa a la sección cero de E , y W un subconjunto abierto de $s^{-1}(0)$. Sea $f: M \rightarrow \mathbb{R}$ una función C^1 y $f_W: W \rightarrow \mathbb{R}$ la restricción de f a W . Entonces son equivalentes:

i.- El punto $x_0 \in W$ es un punto crítico de f_W .

ii.- Sea $E^* \xrightarrow{\tilde{\pi}^*} M$ el fibrado dual de E . Existe $\alpha_0 = (\tilde{\pi}^*)^{-1}(x_0) \in E_{x_0}^*$ tal que el punto $(x_0, \alpha_0) \in E^*$ es un punto crítico de la función $F: E^* \rightarrow \mathbb{R}$ dada por:

$$F(x, \alpha) = f(x) + \langle \alpha, s(x) \rangle.$$

El fibrado E^* es el fibrado dual de E . La fibra E_x^* en cada punto x de M es el dual topológico del espacio de Hilbert $E_x \cong F$ el cual, debido al teorema de Riesz, puede ser naturalmente identificado con E_x . La función F establecida en el teorema anterior puede escribirse también como sigue:

$$F = f\tilde{\pi}^* + P_s,$$

donde $\tilde{\pi}^*: E^* \rightarrow M$ es la proyección y P_s denota la función lineal a lo largo de las fibras de E inducida por la sección s , $P_s(x, \alpha) = \langle \alpha, s(x) \rangle$

Demostración. Debido a que s es transversa a la sección cero de E , $s^{-1}(0)$ es una subvariedad diferenciable de M . Además, si W es un subconjunto abierto de $s^{-1}(0)$, también es una subvariedad diferenciable de M . Además, $T_x(s^{-1}(0)) = \ker \nabla s(x)$, $x \in s^{-1}(0)$. Si $x \in W$, como W es abierto en $s^{-1}(0)$, entonces $T_x W = T_x s^{-1}(0) = \ker \nabla s(x)$.

Consideremos ahora un punto $x_0 \in W$ el cual es un punto crítico de f_W . Entonces, $df_W(x_0) = 0$, es decir, $df(x_0)(U) = 0$ para todo $U \in T_x W$, luego $df(x_0) \in (\ker \nabla s(x_0))^0$. Si calculamos ahora la diferencial de la función F obtenemos,

$$dF(x, \alpha)(X) = df(x)(\pi_* X) + d\langle \alpha, s(x) \rangle(X) = df(x)(U) + \langle X^V, s(x) \rangle + \langle \alpha, \nabla_U s(x) \rangle,$$

donde el vector tangente $X \in T_{(x_0, \alpha)}E$ está descompuesto en sus componentes horizontal y vertical $X = X^H + X^V$ con respecto a la conexión ∇ y $\pi_*(X) = U$. Si $x_0 \in W$, entonces $s(x_0) = 0$, y obtenemos:

$$dF(x_0, \alpha)(X) = df(x_0)(U) + \langle \alpha, \nabla_U s(x) \rangle.$$

Pero la ecuación $\langle \alpha, \nabla s(x) \rangle = -df(x_0)$ tiene una solución si y sólo si $df(x_0) \in (\ker \nabla s(x_0))^0$.

Inversamente, si $(x_0, \alpha_0) \in E^*$ es un punto crítico de la función F entonces, debido a que $dF(x_0, \alpha_0) = 0$ y $x_0 \in W$, obtenemos:

$$df(x_0) + \langle \alpha_0, \nabla s(x_0) \rangle = 0.$$

Pero, como antes, la ecuación $\langle \alpha, \nabla s(x) \rangle = -df(x_0)$ tiene una solución si y sólo si $df(x_0) \in (\ker \nabla s(x_0))^0$, y esto implica que $df_W(x_0) = 0$. \square

Sea M una variedad de Hilbert modelada en un espacio de Hilbert \mathcal{H} y sea $S: M \rightarrow \mathbb{R}$ una función C^1 cuya diferencial es dS , es decir:

$$dS: M \rightarrow T^*M, \quad dS(x) \in T_x^*M$$

y $\forall V \in T_xM$

$$dS(x)(V) = \left. \frac{d}{dt} S(\gamma_t(x)) \right|_{t=0},$$

donde $d\gamma_t(x)/dt|_{t=0} = V$ y $\gamma_0 = x$.

El espacio tangente $T_xM \cong \mathcal{H}$ está dotado con un producto interno $\langle \cdot, \cdot \rangle_x$ y, por lo tanto, debido al teorema de Riesz ($\mathcal{H} \cong \mathcal{H}'$), tenemos que $\exists \nabla S(x)$ tal que

$$\langle \nabla S(x), V \rangle_x = dS(x)(V) \quad \forall V \in T_xM.$$

El vector $\nabla S(x) \equiv \nabla_x S$ se denomina el gradiente de S en x .

Una subvariedad de M es una variedad de Hilbert N y una inmersión $i: N \rightarrow M$ tal que la aplicación $T_x i: T_x N \rightarrow T_{i(x)}M$ es una aplicación lineal inyectiva y continua. Por tanto, $i(T_x N)$ es un subespacio cerrado de $T_x M$. Identificamos $i(x) = x$, y denotaremos como $T_x N^\perp$ al subespacio ortogonal de $T_x N$ dentro de $T_x M$, es decir, al conjunto:

$$T_x N^\perp = \{V \in T_x M \mid \langle U, V \rangle_x = 0 \quad \forall U \in T_x N\}.$$

Corolario 2.3.1 (Teorema de los multiplicadores de Lagrange II) *Sea S una función C^1 sobre M y $N \hookrightarrow M$ una subvariedad de M . Sea S_N la restricción de S a N . Entonces $x \in N$ es un punto crítico de $S_N \iff \nabla S(x) \in T_x N^\perp$.*

Demostración. x es un punto crítico de $S_N \iff dS_N(x) = 0 \iff dS_N(x)(U) = 0 \quad \forall U \in T_x N \iff \langle \nabla S(x), U \rangle = 0 \quad \forall U \in T_x N \iff \nabla S(x) \in T_x N^\perp.$ \square

Debemos aplicar ahora el teorema de multiplicadores de Lagrange al problema que tenemos aquí.

Sea $S: \mathcal{P} \rightarrow \mathbb{R}$ una función C^1 y consideremos la subvariedad \mathcal{W} antes mencionada, sea $S_{\mathcal{W}} = S|_{\mathcal{W}}$ la restricción de S a \mathcal{W} . Queremos caracterizar los puntos críticos de $S_{\mathcal{W}}$; debido al teorema de los multiplicadores de Lagrange, Teorema 2.3.1, sabemos que γ es un punto crítico si y sólo si

$$\nabla_{\gamma} S \in T_{\gamma} \mathcal{W}^{\perp},$$

como se tiene que $T_{\gamma} \mathcal{W} = \ker \mathfrak{L}_{\gamma}$, entonces debemos calcular $(\ker \mathfrak{L}_{\gamma})^{\perp} = \overline{\text{Ran } \mathfrak{L}_{\gamma}^+}$. Por lo tanto, tenemos que $\gamma \in \mathcal{W}$ va a ser un punto crítico si y sólo si

$$\nabla_{\gamma} S \in \overline{\text{Ran } (\mathfrak{L}_{\gamma}^+)}.$$

Diremos, como en la sección anterior, que γ es un extremal ordinario si

$$\nabla_{\gamma} S \in \text{Ran } (\mathfrak{L}_{\gamma}^+),$$

lo que significa que $\exists \lambda \in H^0(\gamma^*(TQ))$ tal que

$$\langle \nabla_{\gamma} S, \delta\gamma \rangle_1 = \langle \mathfrak{L}_{\gamma}^+(\lambda), \delta\gamma \rangle_1, \quad \forall \delta\gamma \in H^1(\gamma^*(TQ)).$$

Como se tiene que

$$\langle \mathfrak{L}_{\gamma}^+(\lambda), \delta\gamma \rangle_1 = \langle \lambda, \mathfrak{L}_{\gamma}(\delta\gamma) \rangle_0 = \int_0^T \left\langle \lambda(t), \frac{d}{dt} \delta\gamma - D\Gamma(\gamma(t), t) \delta\gamma \right\rangle dt,$$

identificando TQ con T^*Q mediante la métrica riemanniana g escribimos el resultado previo como sigue.

Corolario 2.3.2 (Teorema de los multiplicadores de Lagrange III) *Con la notación previa, $\gamma \in \mathcal{W}$ es un punto crítico ordinario de $S_{\mathcal{W}} \iff \exists p \in H^0(\gamma^*(T^*Q))$ tal que*

$$\langle \nabla_{\gamma} S, \delta\gamma \rangle_1 - \int_0^T p(t) \left(\frac{d}{dt} \delta\gamma(t) - D\Gamma(\gamma(t), t) \delta\gamma(t) \right) dt = 0, \quad \forall \delta\gamma \in H^1(\gamma^*(TQ)).$$

Existe otra formulación de este teorema más conocida. Sea $\pi_Q: T^*Q \rightarrow Q$ el fibrado cotangente a Q . Entonces el fibrado cotangente a \mathcal{P} se construye de forma similar al fibrado tangente de \mathcal{P} y en γ vendrá dado por $T_{\gamma}^* \mathcal{P} = H^1(\gamma^*(T^*Q))$. También tenemos el fibrado H^0 -cotangente a \mathcal{P} definido en γ como $H^0(\gamma^*(T^*Q))$, que denotaremos como $H^0(I; T^*Q)$, dual del fibrado tangente H^0 . La proyección natural la denotamos como $\pi_0: H^0(I; T^*Q) \rightarrow \mathcal{P}$. Una curva en $H^0(I; T^*Q)$ se va a denotar por el par $(\gamma(t), p(t))$, donde $p(t) \in T_{\gamma(t)}^* Q$ es un campo covectorial a lo largo de $\gamma(t)$. Entonces definimos la función

$$\begin{aligned} \mathbb{S}(\gamma, p) &= S(\gamma) + \int_0^T p(t) \left(\frac{d\gamma}{dt} - \Gamma(\gamma(t), t) \right) dt \\ &= S(\gamma) + \langle p, A(\gamma) \rangle_0 \end{aligned}$$

y, por lo tanto, para cualquier par $(\delta\gamma, \delta p) \in T_{(\gamma,p)}H^0(I; T^*Q)$ tenemos que:

$$\langle \nabla_{(\gamma,p)}\mathbb{S}, (\delta\gamma, \delta p) \rangle = \langle \nabla_\gamma S, \delta\gamma \rangle + \langle \delta p, A(\gamma) \rangle_0 + \langle p, \mathfrak{L}_\gamma(\delta\gamma) \rangle_0.$$

Corolario 2.3.3 *Sea $(\gamma, p) \in H^0(I; T^*Q)$ una curva cotangente a \mathcal{P} . Entonces (γ, p) es un punto crítico de $\mathbb{S} \iff \gamma$ es un punto crítico ordinario de $S_{\mathcal{W}}$.*

Demostración. Si $(\gamma, p) \in H^0(I; T^*Q)$ es un punto crítico de \mathbb{S} , entonces

$$\langle \nabla_{(\gamma,p)}\mathbb{S}, (\delta\gamma, \delta p) \rangle = 0 \quad \forall \delta\gamma, \delta p \rightarrow \langle \delta p, A(\gamma) \rangle_0 = 0 \quad \forall \delta p \rightarrow A(\gamma) = 0 \rightarrow \gamma \in \mathcal{W}.$$

Y, por otra parte, se ha de cumplir que

$$\langle \nabla_\gamma S, \delta\gamma \rangle + \langle p, \mathfrak{L}_\gamma(\delta\gamma) \rangle_0 = 0 \quad \forall \delta\gamma \rightarrow \nabla_\gamma S \in \text{Ran } \mathfrak{L}_\gamma^+ = T_\gamma \mathcal{W}^\perp$$

$\rightarrow \gamma$ es un punto crítico de $S_{\mathcal{W}}$.

La demostración recíproca es trivial. □

Capítulo 3

Variabes de Clebsch y Ligaduras de Lin en problemas variacionales con simetría de Borel

En este capítulo vamos a desarrollar y resolver la primera pregunta que nos planteábamos en la introducción a esta memoria y que consiste en entender la estructura de las variables de Clebsch y las ligaduras de Lin como una manera de tratar problemas variacionales con simetría. Desarrollaremos en profundidad las ideas comenzadas a tratar en Cendra y Marsden [Ce87a] y Cendra, Ibort y Marsden [Ce87b], y las aplicaremos a pares de Borel $P \times_G Q$. Finalizaremos el capítulo escribiendo las ecuaciones reducidas de un problema variacional sobre pares de Borel con Lagrangiano $L: TQ \longrightarrow \mathbb{R}$ y conexión principal B .

3.1. Conexiones y paralelismo en fibrados asociados

Consideremos un fibrado principal diferenciable $Q(N, G)$ a izquierda sobre la variedad base N con un grupo de estructura G , esto es, $Q(N, G)$ está dado por una variedad diferenciable Q tal que; G es un grupo de Lie que actúa libremente a izquierda sobre Q , $\Psi: G \times Q \longrightarrow Q$, $(g, q) \mapsto \Psi(g, q) = gq = L_g q$; $N \cong Q/G$ y la proyección $\pi: Q \longrightarrow N$ es una aplicación diferenciable, además se verifica que Q es localmente trivial, es decir, $\forall n \in N \exists U \subseteq N$ entorno de n tal que $\pi^{-1}(U)$ es isomorfo a $G \times U$, esto es existe un difeomorfismo $\psi: \pi^{-1}(U) \longrightarrow G \times U$ tal que $\psi(u) = (\varphi(u), \pi(u))$ donde φ es una aplicación de $\pi^{-1}(U)$ en G tal que $\varphi(gu) = g\varphi(u)$, $\forall g \in G$.

La fibra de n en Q es el subconjunto de Q definido por $\pi^{-1}(n) = \{q \in Q \mid \pi(q) = n\}$. Si $q \in \pi^{-1}(n)$, podemos identificar $\pi^{-1}(n)$ con G a través de la aplicación $i_q: G \longrightarrow \pi^{-1}(n)$ dada por $i_q(g) = gq$. Así $\pi^{-1}(n)$ es una copia del grupo de estructura G .

Si tenemos un grupo G que actúa en una variedad Q y la acción de G sobre Q es libre y propia se sabe que Q/G es una variedad diferenciable, $\pi: Q \rightarrow Q/G$ es una submersión y la fibra $\pi^{-1}(n)$ es una subvariedad de Q , entonces $Q(Q/G, G)$ es un fibrado principal. Dado el fibrado principal $Q(N, G)$, la acción de G sobre Q induce un homomorfismo $\mathfrak{g} \rightarrow \mathfrak{X}(Q)$, $\xi \rightarrow \xi_Q$ de \mathfrak{g} en $\mathfrak{X}(Q)$, el conjunto de campos vectoriales verticales sobre Q , esto es el conjunto de campos X tal que $\pi_*(X) = 0$. Nótese que si X e Y son verticales, $[X, Y]$ también es vertical. Finalmente, observemos que el espacio tangente a una fibra $\pi^{-1}(n)$ en el punto q es el espacio tangente a la órbita \mathcal{O}_q con q tal que $\pi(q) = n$, $T_q\pi^{-1}(n) = \{ \xi_Q(q) \mid \xi \in \mathfrak{g} \}$.

Conexiones en fibrados principales

Una conexión principal en $Q(N, G)$ es elegir un complemento de V_q en T_qQ , $\forall q \in Q$; es decir, determinar un subespacio H_q de T_qQ tal que $T_qQ = V_q \oplus H_q$ que llamaremos subespacio horizontal de T_qQ , y tal que $H_{gq} = (L_g)_*H_q, \forall g \in G$, y H_q depende diferenciablemente de q . Todo vector tangente $v_q \in T_qQ$ puede escribirse de manera única como $v_q = v_q^V + v_q^H$ donde $v_q^V \in V_q$ y $v_q^H \in H_q$. Dada una conexión H en $Q(N, G)$, definimos una 1-forma B sobre Q a valores en \mathfrak{g} del siguiente modo: dado $q \in Q$, $B_q(\xi_Q(q)) = \xi$ y $B_q(v_q^H) = 0$. Alternativamente llamaremos conexión principal sobre Q a una 1-forma B con valores en \mathfrak{g} tal que:

1. $B(\xi_Q) = \xi, \forall \xi \in \mathfrak{g}$.
2. $(L_g)^*B = Ad_g B, \forall g \in G$. Esto es, $B((L_g)_*X) = Ad_{g^{-1}}(B(X)), \forall X \in \mathfrak{X}(Q)$.

y entonces la distribución horizontal H se define como $H_q = \text{Ker}B_q$. La 1-forma de Maurer-Cartan a izquierdas sobre G es la 1-forma θ^L invariante a izquierda a valores en \mathfrak{g} dada por

$$\theta_g^L(X) = L_{g^{-1}*}(X).$$

3.1.1. Fibrados asociados a fibrados principales

Recordemos ahora la construcción de fibrados asociados a un fibrado principal y varios ejemplos que aparecerán una y otra vez en lo sucesivo. Sea $Q(N, G)$ un fibrado principal. Supongamos que G actúa a la derecha sobre una variedad P , $\Phi: P \times G \rightarrow P, (x, g) \mapsto \Phi(x, g) = xg$. Vamos a construir sobre N un fibrado cuya fibra sea una copia de P y que preserve la posible no trivialidad de Q . Consideremos la acción a la izquierda de G sobre la variedad producto $P \times Q$, dada por:

$$\begin{aligned} \Xi: \quad G \times (P \times Q) &\longrightarrow P \times Q \\ (g, (x, q)) &\longrightarrow \Xi(g, (x, q)) = (xg^{-1}, gq). \end{aligned}$$

Notemos en primer lugar que la acción de G en $P \times Q$ también es libre. Denotemos por $P \times_G Q$ al espacio cociente $(P \times Q)/G$ y denotaremos por $[x, q]$ a la clase de equivalencia de $(x, q) \in P \times Q$ definida por la órbita de G que pasa a través del par (x, q) , es decir,

$$[x, q] = \{ (xg^{-1}, gq) \in P \times Q \mid g \in G \},$$

esto es, $(x, q) \sim (x', q')$ si y solo si $\exists g \in G$, $x' = xg^{-1}$ y $q' = gq$. La clase de equivalencia $[x, q]$ también se denotará simplemente por xq , es decir $[x, q] = xq$. Nótese que con esta notación $(xg)q = x(gq)$ y además $xg^{-1}gq = xq$, $\forall g \in G$. La proyección principal $\pi_2: Q \longrightarrow N$ induce una proyección $\pi_Q: P \times_G Q \longrightarrow N$, $\pi_Q(xq) = \pi_2(q)$. Nótese que la fibra $\pi_Q^{-1}(n)$ es difeomorfa a P lo que puede verse simplemente usando la aplicación natural $i_q: P \longrightarrow P \times_G Q$, definida por $i_q(x) = xq$, con cualquier $q \in \pi_2^{-1}(n)$. Entonces es fácil observar que $P \times_G Q$ es un fibrado sobre N con fibra P y proyección π_Q . Llamaremos al fibrado $E_P = P \times_G Q$ con base N y proyección π_Q el fibrado asociado al fibrado principal $Q(N, G)$ con respecto a la acción de G en P .

Notar que junto a π_Q hay otra proyección natural sobre $P \times_G Q$ inducida por la proyección $\pi_1: P \longrightarrow P/G$ definida por $\pi_P: P \times_G Q \longrightarrow M$, $\pi_P(xq) = \pi_1(x)$. De nuevo, para cualquier $x \in P$, hay una aplicación natural $i_x: Q \longrightarrow P \times_G Q$ definida por $i_x(q) = xq$, que aplica Q en la fibra de π_P sobre el punto $\pi_1(x)$.

El fibrado de Atiyah y su dual

Consideremos dos ejemplos importantes de fibrados asociados a un fibrado principal $Q(N = G \backslash Q, G)$, el fibrado de Atiyah y su dual. Consideremos el álgebra de Lie \mathfrak{g} del grupo de Lie G y la acción de G en \mathfrak{g} dada por la representación adjunta, $Ad_g: \mathfrak{g} \longrightarrow \mathfrak{g}$, $\xi \rightsquigarrow Ad_g \xi$, $g \in G$, $\xi \in \mathfrak{g}$. Podemos considerar una acción a derechas de G en \mathfrak{g} a través de $\Phi(\xi, g) = Ad_{g^{-1}} \xi$. Llamaremos fibrado de Atiyah (a derechas) de $Q(N, G)$ al fibrado asociado a $Q(N, G)$ por la acción a derechas Φ de G en \mathfrak{g} definida arriba, es decir, $\mathfrak{g} \times_G Q \longrightarrow N$, denotaremos al espacio total de dicho fibrado por $\tilde{\mathfrak{g}}_R \longrightarrow N$, esto es $\mathfrak{g} \times_G Q \equiv \tilde{\mathfrak{g}}_R$. Omitiremos el subíndice R si no hay riesgo de confusión. Escribiremos los puntos en $\tilde{\mathfrak{g}}_R$ como $[\xi, g]$ ó ξg .

Nótese que de manera análoga podemos definir un fibrado de Atiyah para fibrados principales por la derecha $Q(Q/G, G)$ y la acción a izquierda de G en \mathfrak{g} dada por la representación adjunta, esto es $Q \times_G \mathfrak{g} \equiv \tilde{\mathfrak{g}}_L$.

Si consideramos el levantamiento natural de la acción a izquierdas de G a TQ , tendremos el fibrado principal $TQ \longrightarrow G \backslash TQ$ sobre $G \backslash TQ$ con fibra G . Sin embargo hay una aplicación natural $G \backslash TQ \longrightarrow T(G \backslash Q)$ dada por $[v_q] \longrightarrow v_{\pi_2(q)} = (\pi_2)_*(v_q)$. Entonces se ve fácilmente que el fibrado anterior es el pull-back del fibrado $\tilde{\mathfrak{g}}$ a $T(G \backslash Q)$ a lo largo de la aplicación $\tau_{G \backslash Q}$, es decir $G \backslash TQ \cong \tau_{G \backslash Q}^* \tilde{\mathfrak{g}}$. Para ello basta escoger una conexión principal B en Q y definimos la siguiente aplicación:

$$\Lambda: G \backslash TQ \longrightarrow \tau_{G \backslash Q}^* \tilde{\mathfrak{g}}_R: \Lambda([q, v_q]) = ((\pi_2)_*(v_q), [B_q(v_q), q]).$$

Notemos en primer lugar que Λ está bien definido:

si $[q, v] = [q', v'] \in G \backslash TQ$, entonces $\exists g \in G$, $q' = gq$, $v' = (L_g)_*v$. Por tanto, $(\pi_2)_*(v) = (\pi_2)_*(v')$ y dado que $B_{q'}(v') = B_{gq}((L_g)_*v) = ((L_g)^*B)_q(v) = Ad_g B_q(v)$, entonces $[B_{q'}(v'), q'] = [Ad_g B_q(v), gq] = [B_q(v), q]$. Es también fácil ver que la aplicación $\tau_{G \backslash Q}^* \tilde{\mathfrak{g}}_R \rightarrow G \backslash TQ$ definida por $(v_{\pi_2(q)}, [\xi, q]) \rightsquigarrow [(q, v_q^H + \xi_Q(q))]$ es la inversa de Λ .

Levantamiento horizontal y transporte paralelo

Veamos cómo una conexión B en $Q(N, G)$ proporciona una noción de levantamiento horizontal de campos sobre N . Consideremos la diferencial de la proyección π , $(T\pi)_q: T_q Q \longrightarrow T_n N$ con $\pi(q) = n$. Dado $X \in T_n N$, un levantado de X a $q \in Q$ es un vector $\tilde{X}_q \in T_q Q$ tal que $(T\pi)_q(\tilde{X}_q) = X$. Como es claro, Y es vertical si y sólo si $(T\pi)_q(Y) = 0$. Entonces, de todos los levantados de X a q sólo uno es horizontal. El levantado horizontal de un campo $X \in \mathfrak{X}(N)$, X^h es el único campo horizontal sobre Q tal que $(T\pi)_q(X_q^h) = X_{\pi(q)}$. Está claro que X_q^h es invariante por L_g , $\forall g \in G$, esto es $L_{g*}(X_q^h) = X_{gq}^h$. Inversamente, todo campo horizontal X^h sobre Q invariante por G es el levantado horizontal de algún campo X sobre N . Se verifica fácilmente que si X^h e Y^h son los levantados horizontales de X e $Y \in \mathfrak{X}(N)$ respectivamente. Entonces, $X^h + Y^h$ es el levantado horizontal de $X + Y$, y si f es una función diferenciable sobre N , $f^h X^h$ es el levantado horizontal de fX siendo $f^h: Q \longrightarrow \mathbb{R}$ la función dada por $f^h = f \circ \pi$. Nótese que la componente horizontal de $[X^h, Y^h]$ es el levantado horizontal de $[X, Y]$.

Veamos ahora cómo se levantan curvas de N . Sea $n(t)$ una curva diferenciable sobre N con $0 \leq t \leq T$. Un levantamiento horizontal de $n(t)$ es una curva horizontal $q(t) \subset Q$ tal que $\pi(q(t)) = n(t) \forall t$; es decir, es una curva tal que sus vectores tangentes son horizontales: $\dot{q}(t) \in H_q(t)$. Es claro que las nociones de levantamiento de curvas y levantamiento de campos están relacionadas. Se puede ver que si X^h es el levantado horizontal de X , la curva integral de X^h que comienza en el punto q_0 , $\pi(q_0) = n(0) = n_0$ es el levantamiento de la curva integral de X comenzando en n_0 .

Se tiene por tanto que si $\gamma(t)$ es una curva de clase $C^1(N)$ con $0 \leq t \leq T$. Para todo punto $q_0 \in Q$ tal que $\pi(q_0) = n_0 = n(0)$ existe un único levantamiento horizontal $\gamma^h(t) = q(t)$ tal que $\gamma^h(0) = q_0 = q(0)$.

Con este resultado se puede definir el concepto de transporte paralelo de fibras a lo largo de una curva en la base. Sea $\gamma(t)$ una curva en N con $0 \leq t \leq T$, $q_0 \in \pi^{-1}(n_0)$. Sea $\gamma^h(t)$ el levantamiento horizontal de $\gamma(t)$ al punto q_0 cuyo punto final es $\gamma^h(T) = q_T \in \pi^{-1}(n_T)$. Variando q_0 en la fibra $\pi^{-1}(n_0)$ se obtiene una aplicación de la fibra $\pi^{-1}(n_0)$ en $\pi^{-1}(n_T)$ que es un isomorfismo.

El transporte paralelo a lo largo de la curva $n(t)$ es la aplicación

$$\mathbb{P}_T: \pi^{-1}(n_0) \longrightarrow \pi^{-1}(n_T)$$

que le asigna a q_0 el valor en $t = T$ del único levantamiento horizontal de $n(t)$ cuyo valor en $t = 0$ es q_0 . El transporte paralelo a lo largo de la curva $n(t)$ conmuta con la acción de G sobre Q . Es decir, $n^h \circ L_g = L_g \circ n^h, \forall g \in G$.

Es sencillo ver también que

1. El transporte paralelo es independiente de la parametrización de la curva en la base.
2. El transporte paralelo a lo largo de $-\gamma$ es el inverso del transporte paralelo a lo largo de γ .
3. El transporte paralelo a lo largo de $\alpha \cdot \beta$ es la composición de los transportes paralelos a lo largo de α y β .

Para cada $n \in N$ denotamos por $C(n)$ el espacio de los loops en n , es decir, el espacio de caminos cerrados en n . Dado $\gamma \in C(n)$, el transporte paralelo a lo largo de γ es un isomorfismo de la fibra $\pi^{-1}(n)$. Por la observación anterior resulta claro que $\gamma^h: \pi^{-1}(n) \longrightarrow \pi^{-1}(n)$ es un isomorfismo de grupos y que $\gamma^h(gq) = g\gamma^h(q)$. El grupo de los isomorfismos γ^h de la fibra $\pi^{-1}(n)$ es el grupo de holonomía de la conexión B en n .

3.2. Reducción variacional de Lagrange-Poincaré

Consideraremos por tanto tal y como se indicó en la introducción y en el capítulo §2 un lagrangiano L definido en el fibrado tangente TQ de un fibrado principal a izquierdas $Q(N, G)$ con $\pi: Q \longrightarrow G \backslash Q = N$ la fibración principal. Supondremos que L es invariante con respecto al levantamiento natural de la acción de G a TQ , esto es, $L(gq, gv_q) = L(q, v_q), \forall g \in G, (q, v_q) \in TQ$, y queremos resolver el problema de determinar los puntos críticos del funcional $S = \int_0^T L(q, v_q) dt$ sobre el espacio de curvas $\Omega_{q_0, q_T}(Q)$ con extremos fijos q_0, q_T . Dado que el lagrangiano L es G -invariante desciende a una función $l: G \backslash TQ \longrightarrow \mathbb{R}$ definida como $l([q, v_q]) = L(q, v_q)$, donde $[q, v_q]$ denota la clase de equivalencia definida por la acción de G en TQ , es decir $[q, v_q] = \{(gq, gv_q) \mid g \in G\}$.

El fibrado cociente $G \backslash TQ$ sobre $T(G \backslash Q)$ se puede identificar como se describió con anterioridad con el fibrado obtenido por pull-back del fibrado de Atiyah a $T(G \backslash Q)$ a lo largo de la proyección canónica $\tau_{G \backslash Q}$.

Podemos por tanto identificar los puntos de $G \backslash TQ$ como $(n, v_n; \xi)$ con $(n, v_n) \in TN = T(G \backslash Q)$ y $\xi \in \mathfrak{g}$.

Tal y como se indicó en la introducción el análisis del principio variacional reducido, esto es el definido por l en el espacio de curvas

$$\Omega_{n_0, n_T}(\tilde{\mathfrak{g}}) = \{ \gamma: [0, T] \rightarrow \tilde{\mathfrak{g}} = \mathfrak{g} \times_G Q \mid \pi_Q(\gamma(0)) = \pi(q_0) = n_0, \\ , \pi_Q(\gamma(T)) = \pi(q_T) = n_T \},$$

ha sido realizado por Cendra, Marsden y Ratiu [Ce01a] con el nombre de reducción de Lagrange-Poincaré. Aunque en el capítulo 5 volveremos sobre este tema, vamos a abordar el estudio del principio variacional reducido $S_l = \int_0^T l(n, v_n, \xi) dt$ utilizando variables auxiliares, esto es desarrollando el punto de vista de Clebsch y Lin.

Consideraremos por tanto en primer lugar un espacio auxiliar P con una acción que supondremos a derechas del grupo G . Denotamos, como en párrafos anteriores dicha acción por xg , $x \in P$, $g \in G$. El espacio T^*P , el fibrado cotangente de P se denominará el espacio de Clebsch del problema.

El espacio auxiliar de Clebsch T^*P tiene una estructura simpléctica canónica ω_P y la acción a derecha de G en P se levanta de manera natural a una acción en T^*P . El levantamiento de dicha acción es simpléctica y deja invariante, no sólo la 2-forma canónica ω_P , sino también la 1-forma de Liouville θ_P en T^*P . Existirá por tanto una aplicación momento para dicha acción definida por $J_p: T^*P \rightarrow \mathfrak{g}^*$ y $\langle J_P(x, p_x), \xi \rangle = \langle \xi_P(x), p_x \rangle$ con $(x, p_x) \in T^*P$, $\xi \in \mathfrak{g}$ y $\xi_P(x)$ denota el campo vectorial asociado a ξ en P .

El espacio de Clebsch auxiliar podría sustituirse por una variedad simpléctica (M, ω) donde actúa el grupo G y que tenga aplicación momento $J: M \rightarrow \mathfrak{g}^*$, aunque no insistiremos en dicho punto de vista.

La extensión trivial del principio variacional $\int L(q, v_q) dt$ a $P \times Q$, esto es, consideramos el mismo funcional $S_L = \int L(q, v_q) dt$ pero ahora definido en el espacio de curvas $\Omega_{q_0, q_T}(P \times Q)$ no nos resuelve de momento nada, ya que el funcional S_L es degenerado en él. Lo que haremos es definir un subespacio de curvas $\tilde{\Omega}$ en $\Omega_{q_0, q_T}(P \times Q)$ difeomorfo a $\Omega_{q_0, q_T}(Q)$ y podremos aplicar el teorema de los multiplicadores de Lagrange (ver capítulo §3.3.3) a la subvariedad $\tilde{\Omega}$ de $\Omega_{q_0, q_T}(P \times Q)$. De esta manera escribiremos los puntos críticos de S_L como puntos críticos de un funcional \mathbb{S} definido en $\Omega_{q_0, q_T}(P \times Q)$ donde habremos incorporado las ecuaciones que definen $\tilde{\Omega}$ como subvariedad de $\Omega_{q_0, q_T}(P \times Q)$. Dichas ecuaciones podrán interpretarse como ligaduras que relacionan curvas en P y curvas en Q y constituyen la geometrización de las ligaduras de Lin comentadas en la introducción. Finalmente nuestro funcional \mathbb{S} en $\Omega_{q_0, q_T}(P \times Q)$ continúa siendo G -invariante por lo que podemos proceder a su reducción, bien de Lagrange-Poincaré, ó bien de Hamilton-Poisson. Discutiremos estos aspectos más adelante.

El subespacio $\tilde{\Omega} \subset \Omega(P \times Q)$ estará dado en general por una condición de horizontalidad. Para describir esta idea vamos a revisar algunos conceptos y notaciones sobre horizontalidad y conexiones en fibrados asociados.

3.2.1. Conexiones inducidas en fibrados asociados $P \times_G Q$

Recordemos que una conexión (principal) B sobre Q puede ser caracterizada por sus espacios verticales y horizontales en $q \in Q$. Los denotamos respectivamente por

$$V_q = \ker T_q\pi, \quad H_q = \ker B_q.$$

donde π es la proyección principal $\pi: Q \rightarrow N$. Estos espacios nos proporcionan la siguiente descomposición $TQ = H(TQ) \oplus V(TQ)$ donde $H(TQ) = \bigcup_{q \in Q} H_q$ y $V(TQ) = \bigcup_{q \in Q} V_q$, los cuales son invariantes bajo la acción de G .

Las componentes verticales y horizontales de un vector $v \in T_qQ$ las denotaremos como anteriormente por $V(v)$ y $H(v)$ ó v^V y v^H respectivamente. Por definición, $V(v) = B_q(v)_Q$ y $H(v) = v - B_q(v)_Q$. Un vector tangente v es llamado horizontal si su componente vertical es cero; es decir, si $B_q(v) = 0$ y es llamado vertical si su componente horizontal es cero; es decir, $T_q\pi(v) = 0$. Una curva $q(t)$ diremos que es horizontal si $\dot{q}(t)$ es horizontal para todo t . Notar que $T_q\pi: H_q \rightarrow T_{\pi(q)}N$ es un isomorfismo. Por tanto, una curva $q(t)$ sobre Q es horizontal si $B_{q(t)}(\dot{q}(t)) = 0$, $\forall t$.

Dado un vector $X \in T_{\pi(q)}N$ el levantamiento horizontal X_q^h de X en q es el único vector horizontal en T_qQ tal que $T_q\pi(X_q^h) = X$.

Para cualquier curva $n(t) \in N$, $t \in [0, T]$ podemos definir su levantamiento horizontal $n_{q_0}^h$ como la única curva horizontal $n_{q_0}^h(t)$ que se proyecta sobre $n(t)$ y tal que $n_{q_0}^h(0) = q_0$, $\pi(q_0) = n_0 = n(0)$.

Consideremos una curva $q(t) \in Q$, $t \in [0, T]$. Entonces existe una única curva horizontal $q_h(t)$ tal que $q_h(t_0) = q(t_0)$ y $\pi(q_h(t)) = \pi(q(t))$ para todo $t \in [0, T]$. Por tanto, podemos definir una curva $g(t)$, $t \in [0, T]$ en G usando la descomposición $q(t) = g(t)q_h(t)$. También, notar que si $n(t) = \pi(q(t))$ entonces $q_h(t) = n_{q_0}^h(t)$. Además, $\dot{q}(t) = \dot{g}(t)q_h(t) + g(t)\dot{q}_h(t)$ y $\dot{q}(t) = H(\dot{q}(t)) + V(\dot{q}(t)) = H(\dot{q}(t)) + B_{q(t)}(\dot{q}(t))_Q$. Por definición de vector horizontal, $B_{q(t)}(g(t)\dot{q}_h(t)) = 0$, luego $B_{q(t)}(\dot{q}(t)) = B_{q(t)}(\dot{g}(t)q_h(t)) = B_{q(t)}(\dot{g}(t)g^{-1}(t)g(t)q_h(t)) = \dot{g}(t)g^{-1}(t)$

La conexión B induce una conexión sobre cualquier fibrado asociado declarando horizontales las curvas de la forma $x(t)q(t)$ donde $q(t)$ es horizontal. Esto define una distribución H^P sobre $T(P \times_G Q)$ donde $v \in H^P(xq)$ si existe una curva horizontal $\gamma(t) = x(t)q(t)$ sobre $P \times_G Q$ tal que $v = \dot{\gamma}(0)$.

Los vectores tangentes $v \in T_{xq}(P \times_G Q)$ pueden ser fácilmente descritos como sigue. Sea $x(t)q(t)$ una curva tal que $x(0)q(0) = x_0q_0$ y $d/dt(x(t)q(t))|_{t=0} = v$. Entonces no es difícil demostrar que $v = T_qi_x(\dot{q}) + T_xi_q(\dot{x})$. Simplemente denotaremos por $x\dot{q} := T_qi_x(\dot{q})$ y $\dot{x}q := T_xi_q(\dot{x})$. Con esta notación tenemos

$$\left. \frac{d}{dt} (x(t)q(t)) \right|_{t=0} = x\dot{q} + \dot{x}q.$$

Por tanto, debido a la definición de conexión asociada, los vectores horizontales en H^P serán de la forma $x\dot{q}_h$.

Notemos que si $\xi \in \mathfrak{g}$ es un elemento del álgebra de Lie de G , entonces

$$(x \cdot \exp(t\xi))q = x(\exp(t\xi) \cdot q), \quad \forall t,$$

y tomando derivadas con respecto a t obtenemos,

$$x\xi_Q(q) = -\xi_P(x)q, \quad (3.1)$$

donde $\xi_Q(q) = d/dt(\exp(t\xi) \cdot q) |_{t=0}$ y $\xi_P(x) = d/dt(x \cdot \exp(-t\xi)) |_{t=0}$.

Dados $x(t)$ y $q(t)$ curvas en P y Q respectivamente, tendremos:

$$\frac{d}{dt}(x(t)q(t)) = \dot{x}(t)q(t) + x(t)\dot{q}(t) = \dot{x}(t)q(t) + x(t)g(t)\dot{q}_h(t) + x(t)B_{q(t)}(\dot{q}(t))_Q(q(t))$$

y el vector tangente $d/dt(x(t)q(t))$ será horizontal si y sólo si

$$x(t)B_{q(t)}(\dot{q}(t))_Q(q(t)) + \dot{x}(t)q(t) = 0.$$

Por consiguiente definimos una 1-forma de conexión B^P sobre el fibrado asociado $P \times_G Q \rightarrow N$ cuyo kernel es la distribución horizontal H^P definida antes y dada por:

$$B_{xq}^P(\dot{x}\dot{q}) = xB_q(\dot{q})_Q(q) + \dot{x}q = -B_q(\dot{q})_P(x)q + \dot{x}q,$$

donde hemos usado la eq. (3.1).

3.3. Espacios de curvas horizontales sobre fibrados asociados

Denotaremos por $\Omega_{x_0}(P)$ al espacio de curvas diferenciables en P con origen fijo x_0 y por $\Omega_{x_0, x_T}(P)$ al que tiene los extremos fijos x_0, x_T . También denotaremos por $\Omega_{q_0}(Q)$ al espacio de curvas diferenciables en Q con origen q_0 y por $\Omega_{q_0, q_T}(Q)$ al que tiene los extremos fijos q_0, q_T . Como establecimos anteriormente es claro que dada una curva $q(t)$ en Q hay una única descomposición $q(t) = g(t)q_h(t)$, donde $g(0) = e$ y $q_h(t)$ es horizontal con respecto a la conexión B , es decir, $B_{q(t)}(\dot{q}_h(t)) = 0$. Dada una curva $q(t) \in \Omega_{q_0, q_T}(Q)$ y $x_0 \in P$, existe una única curva denotada por $x_h(t)$ en $\Omega_{x_0}(P)$ tal que $x_h(t)q(t)$ es horizontal con respecto a la conexión B^P sobre $P \times_G Q$ y $x_h(0)q(0) = x_0q_0$. Se observa con facilidad que esta curva está definida por $x_h(t) = x_0g^{-1}(t)$, ya que

$$x_h(t)q(t) = x_0g^{-1}(t)q(t) = x_0g^{-1}(t)g(t)q_h(t) = x_0q_h(t),$$

que es horizontal.

Fijemos un punto q_0 en Q . Denotaremos por $\Omega_{x_0, x_T; q_0}^H(P \times Q)$ al conjunto de curvas $(x(t), q(t))$ con dominio $[0, T]$, $q(t) \in \Omega_{q_0}(Q)$, $x(t)$ con extremos fijos x_0, x_T y tal que $x(t)q(t)$ es horizontal para todo t . Notar que este último requisito implica que $x(t)g(t) = x_0$, donde $q(t) = g(t)q_h(t)$, y que $x_T = x_0g^{-1}(T) = x_0g_T^{-1}$, $g(T) = g_T$.

3.3.1. Condiciones extremales compatibles

Sea $x_0 \in P$ y x_0G denotará la órbita del grupo G sobre P pasando a través de x_0 , es decir, $x_0G = \{x_0g \in P \mid g \in G\}$. Ahora, dada una curva diferenciable $g(t)$ en G y denotando por $x(t) = x_0g(t)$ a la correspondiente curva que comienza en x_0 , el extremo x_T estará en la órbita x_0G . Si denotamos por $\xi(t)$ la curva en \mathfrak{g} tal que se verifica

$$\dot{g}(t) = \xi(t)g(t)$$

sobre el grupo G con condición inicial $g(0) = e$, un simple cálculo demuestra que la curva $x(t) = x_0g(t; \xi)$ es una curva integral del sistema $\dot{x} = \xi_P(t)(x)$ con condición inicial x_0 , esto es:

$$\frac{d}{dt}x(t) = \frac{d}{dt}(x_0g(t; \xi)) = x_0\dot{g}(t; \xi) = x_0\xi_P(t)g(t; \xi) = \xi_P(t)(x(t)),$$

además $x_0g(0; \xi) = x_0e = x_0$.

Diremos que dos puntos $x_0, x_T \in P$ en el espacio auxiliar P son compatibles si existe $\xi(t) \in \mathfrak{g}$, tal que $x_T = x(T; \xi)$ donde $x(t; \xi)$ denota la curva integral de $\dot{x} = \xi_P(t)(x)$ con condición inicial x_0 .

Dadas dos curvas $\xi(t), \xi'(t)$ en \mathfrak{g} , puede ocurrir que $\xi_P(t) = \xi'_P(t)$. La no-unicidad de la elección de la curva $\xi(t)$ correspondiente a $\xi(t)_P$ depende del álgebra de isotropía $\mathfrak{g}_x(t)$ a lo largo de las curvas integrales del campo vectorial $\xi_P(t)$.

Por tanto para que el extremo x_T sea accesible desde x_0 es necesario que $x_T \in x_0G$. Si $x_T \in x_0G$, diremos que los extremos x_0 y x_T son compatibles.

Sin embargo, cualquier punto $x_T \in x_0G$ no es compatible con x_0 . Sin embargo, tenemos la siguiente proposición:

Proposición 3.3.1 *Si G es un grupo exponencial, cualquier punto $x_T \in x_0G$ es compatible con x_0 .*

Demostración. Si el grupo es exponencial, es conexo. Tomemos cualquier curva diferenciable $g(t)$ tal que $g^{-1}(0) = e$ y $g^{-1}(T) = g_T^{-1}$ donde $x_T = x_0g_T^{-1}$. Ahora la curva $x(t) = x_0g^{-1}(t)$ satisface que $x(0) = x_0$ y $x(T) = x_T$. Además $\dot{x}(t)q(t) = -x(t)\dot{g}(t)g^{-1}(t)q(t) = -x(t)B_{q(t)}(\dot{g}(t))_Q(q(t)) = B_{q(t)}(\dot{g}(t))_P(x(t))q(t)$ donde $q(t) = g(t)q_0$. \square

Fijemos $x_T = x_0g_T^{-1}$. Notar que para cualquier $g' \in x_0G$, $g'_T = g_Tg'$ define el mismo extremo x_T .

Para cualquier punto inicial dado q_0 , la curva horizontal que pasa por él está únicamente determinada. Por tanto, si $q(t)$ es una curva arbitraria y $x(t)q(t)$ es horizontal, tenemos $x(t)q(t) = x_0q_h(t)$. Luego $x_Tq(T) = x_0q_h(T)$, y $x_0g_T^{-1}q(T) = x_0q_h(T)$. Si x_0G denota el grupo de isotropía de x_0 , tenemos que $g_T^{-1}q(T) =$

$hq_h(T)$, $h \in x_0G$. Además, $q(T) = g(T)q_h(T)$, y como G actúa libremente sobre Q , tendremos que $g(T) \in x_0Gg_T$.

Por tanto, denotando $g_Tq_h(T)$ por q_T , tenemos que para cada curva $q(t) \in \Omega_{q_0, x_0Gq_T}(Q)$ podemos asociar una curva en $\Omega_{x_0, x_T; q_0}^H(P \times Q)$ por medio de la correspondencia natural $q(t) \mapsto (x(t), q(t))$, donde $x(t) = x_0g^{-1}(t)$ y $q(t) = g(t)q_h(t)$ como antes. Entonces $x(0) = x_0g^{-1}(0) = x_0$ y $x(T) = x_0g^{-1}(T) = x_0hg_T^{-1} = x_0g_T^{-1} = x_T$, $h \in x_0G$.

Los comentarios anteriores pueden ser resumidos en la siguiente proposición:

Proposición 3.3.2 *Sea Q un G -fibrado principal con conexión B y P un G -espacio por la derecha. Sea $g \in G$, dos puntos compatibles $x_0, x_T = x_0g_T^{-1}$ en P y $q_0 \in Q$, entonces hay una correspondencia uno-uno entre el conjunto de curvas $\Omega_{x_0, x_T}(P)$, el conjunto de curvas horizontales $\Omega_{x_0, x_T; q_0}^H(P \times Q)$ y el conjunto de curvas $\Omega_{q_0, x_0Gq_T}(Q)$, donde $q_T = g_Tq_h(T)$, $q_h(t)$ la única curva horizontal que comienza en q_0 , y x_0G el grupo de isotropía de x_0 . Las correspondencias están dadas explícitamente por $q(t) \mapsto x_0q_h(t) = x(t)q(t)$ con $x(t) = x_0g^{-1}(t)$ y $q(t) = g(t)q_h(t)$, y finalmente $x(t)q(t) \mapsto x(t)$.*

Las curvas $(x(t), q(t))$ pertenecientes al espacio de curvas horizontales en $P \times Q$ con extremos compatibles x_0, x_T y tales que $q(0) = q_0$ denotadas por $\Omega_{x_0, x_T; q_0}^H(P \times Q)$ están en correspondencia uno-uno con las curvas en $\Omega_{x_0, x_T}(P)$.

3.3.2. Variabales de Clebsch y Ligaduras de Lin para problemas variacionales con simetría

Puesto que $B^P(\frac{d}{dt}(x(t)q(t))) = 0$ es equivalente a que $x(t)q(t)$ sea horizontal, se sigue que $\Omega_{x_0, x_T; q_0}^H(P \times Q)$ es el subconjunto de $\Omega_{x_0, x_T; q_0, x_0Gq_T}(P \times Q)$ definido por la ligadura $B^P(\frac{d}{dt}(x(t)q(t))) = 0$. Esta ligadura es también llamada ligadura de Lin [Ce87a]. Ahora la introduciremos en el principio variacional, usando un multiplicador de Lagrange. Utilizando el espacio de representación de Clebsch T^*P permitiremos variaciones arbitrarias de las curvas a lo largo de las direcciones verticales en T^*P y así, el nuevo término en el Lagrangiano será del siguiente modo:

$$\left\langle B^P\left(\frac{d}{dt}(x(t)q(t))\right), p(t)q(t) \right\rangle \tag{3.2}$$

siendo $p(t)$ cualquier levantamiento de la curva $x(t) \in \Omega_{x_0, x_T}(P)$. Notar que la acción de G sobre P puede ser levantada naturalmente a una acción de G sobre T^*P y podemos considerar el espacio cociente $T^*P \times_G Q$ como en §3.1.1.

Consideremos la curva $p(t)$ como una curva en T^*P con extremos en $\tau_P^{-1}(x_0)$ y $\tau_P^{-1}(x_T)$, ó como una curva en T^*P con extremos libres a lo largo de las fibras de la proyección canónica $\tau_P: T^*P \rightarrow P$. La ligadura horizontal de Lin dada

anteriormente, puede ser escrita en términos de la 1-forma canónica de Liouville θ_P sobre T^*P como

$$\begin{aligned} \langle B^P(\frac{d}{dt}(xq)), pq \rangle &= \langle (\dot{x} - B_q(\dot{q})_P(x))q, pq \rangle \\ &= -\theta_P(x, p)(\dot{x}, \dot{p}) - \langle J(x, p), B_q(\dot{q}) \rangle. \end{aligned} \quad (3.3)$$

Finalmente, consideremos L como un Lagrangiano G -invariante sobre Q , esto es, $L(q, \dot{q}) = L(gq, g\dot{q})$ con $(q, \dot{q}) \in TQ$ y $g \in G$. Como indicamos anteriormente, la acción de G sobre P también nos permite definir el fibrado asociado $T^*P \times_G Q$ sobre N con fibra T^*P . Entonces definimos el Lagrangiano \mathbb{L} sobre $T(T^*P \times_G Q)$ mediante la fórmula:

$$\mathbb{L}(x, p, \dot{x}, \dot{p}; q, \dot{q}) = L(q, \dot{q}) - \langle B^P(\frac{d}{dt}(xq)), pq \rangle, \quad (3.4)$$

ó usando las fórmulas anteriores, obtenemos para B^P lo siguiente:

$$\begin{aligned} \mathbb{L}(x, p, \dot{x}, \dot{p}; q, \dot{q}) &= L(q, \dot{q}) - \langle B^P(\frac{d}{dt}(xq)), pq \rangle \\ &= L(q, \dot{q}) + \langle J(x, p), B_q(\dot{q}) \rangle + \theta_P(x, p)(\dot{x}, \dot{p}) \end{aligned} \quad (3.5)$$

Ahora podemos hacer precisa la correspondencia entre los puntos críticos de L y los de \mathbb{L} .

Teorema 3.3.1 *Fijando $g_T \in G$, $q_0 \in Q$ y $x_0 \in P$, sea $x_T = x_0 g_T^{-1}$, y $q_T = g_T q_0$. Entonces las siguientes afirmaciones son equivalentes:*

i.- La curva diferenciable $q(t) \in \Omega_{q_0, x_0 G q_T}(Q)$ es un punto crítico del funcional $S_L : \Omega_{q_0, x_0 G q_T}(Q) \rightarrow \mathbb{R}$ definido por

$$S_L[q] = \int_0^T L(q, \dot{q}) dt \quad (3.6)$$

*ii.- Hay una curva diferenciable $(x(t), p(t)) \in \Omega_{x_0, x_T}(T^*P)$ tal que la curva $(x(t), p(t); q(t))$ es un punto crítico del funcional $S_{\mathbb{L}} : \Omega_{x_0, x_T; q_0}(T^*P \times_G Q) \rightarrow \mathbb{R}$ definido por:*

$$\begin{aligned} S_{\mathbb{L}}[x, p; q] &= \int_0^T \mathbb{L}(x(t), p(t), \dot{x}(t), \dot{p}(t); q(t), \dot{q}(t)) dt \\ &= \int_0^T (L(q, \dot{q}) + \langle J(x, p), B_q(\dot{q}) \rangle + \theta_P(x, p)(\dot{x}, \dot{p})) dt. \end{aligned} \quad (3.7)$$

Demostración. Supongamos que $q(t) \in \Omega_{q_0, x_0 Gq_T}(Q)$ es un punto crítico del funcional S_L , dado que el espacio de curvas $\Omega_{q_0, x_0 Gq_T}(Q)$ es difeomorfo al espacio de curvas horizontales $\Omega_{x_0, q_0}^H(P \times Q)$, el funcional S_L define una función que denotamos igual en $\Omega_{x_0, q_0}^H(P \times Q)$. El espacio de curvas $\Omega_{x_0, q_0}^H(P \times Q)$ es una subvariedad cerrada de $\Omega_{x_0, q_0}(P \times Q)$ definida por la aplicación $\Upsilon(x, q) = B^P(\dot{x}\dot{q}) = 0$. Esta aplicación es una sección regular del fibrado asociado dado por P y por tanto el teorema de los multiplicadores de Lagrange detallada a continuación nos permite establecer que los puntos críticos de S_L en $\Omega_{x_0, q_0}^H(P \times Q)$ están en correspondencia 1-1 con los puntos críticos del funcional extendido $\mathbb{S} = \int_0^T (L(q, v_q) + \langle p, B^P(\dot{x}\dot{q}) \rangle) dt$, pero utilizando ahora las fórmulas (3.5) obtenemos la expresión (3.7). \square

3.3.3. Teorema de los multiplicadores de Lagrange sobre espacios de curvas

El uso del Teorema de los multiplicadores de Lagrange Thm. 2.3.1 en el contexto de esta tesis requiere establecer el marco adecuado Utilizaremos para este propósito el marco analítico de Klingerberg para funcionales en espacios de curvas [Kl78].

Consideraremos los espacios de curvas $\gamma(t) = (x(t), q(t))$ de clase de Sobolev k , $k \geq 1$, sobre el espacio $P \times Q$, es decir $H^k(P \times Q)$ [Kl78]. Tal espacio de curvas es una variedad de Hilbert paracompacta modelada sobre el espacio de Hilbert $H^k([0, T], \mathbb{R}^N)$, $N = \dim P + \dim Q$, de aplicaciones $\gamma: [0, T] \longrightarrow \mathbb{R}^N$ en $L^2([0, T], \mathbb{R}^N)$ con derivadas débiles $\gamma^{(l)}(t)$, $0 \leq l \leq k$, en $L^2([0, T], \mathbb{R}^N)$. El espacio tangente a $H^k(P \times Q)$ en la curva $\gamma(t) = (x(t), q(t))$ está dado por las secciones de clase de Sobolev k del fibrado de los pull-back $\gamma^*(TP \times TQ)$. Debido a que el fibrado $\gamma^*(TP \times TQ)$ sobre $[0, T]$ es trivial, tal espacio de secciones puede ser identificado con el espacio de Hilbert de las aplicaciones de Sobolev $H^k([0, T], \mathbb{R}^N)$. Denotaremos por $T(H^k(P \times Q))$ al fibrado tangente construido a partir de este espacio de curvas. Además podemos considerar en cada aplicación γ el espacio de Hilbert de las secciones de clase de Sobolev l , $0 \leq l \leq k$, del fibrado de los pull-back $\gamma^*(TP \times TQ)$. La colección de tales espacios definen un fibrado vectorial sobre $H^k(P \times Q)$ cuya fibra está dada por $H^l([0, T], \mathbb{R}^N)$. Denotaremos a este fibrado vectorial como $T^{(l)}(H^k(P \times Q))$.

Varias condiciones extremales para las curvas que estamos considerando definen subvariedades de Hilbert de las variedades de Hilbert $H^k(P \times Q)$. Por ejemplo, la condición extremal que va a ser considerada a lo largo de la tesis, $x(0) = x_0$, $x(T) = x_T$, define una subvariedad de Hilbert, que será denotada a partir de ahora como \mathcal{P}^k . Por consiguiente

$$\mathcal{P}^k = \{ (x(t), q(t)) \in H^k(P \times Q) \mid x(0) = x_0, x(T) = x_T \}.$$

El espacio tangente a \mathcal{P}^k está dado por los subespacios de Hilbert

$$T_\gamma \mathcal{P}^k = \{ (\delta x(t), \delta q(t)) \in H^k(\gamma^*(P \times Q)) = T_\gamma H^k(P \times Q) \mid \delta x(0) = \delta x(T) = 0 \}$$

Por tanto tenemos definido el fibrado tangente $T\mathcal{P}^k$. Además, simultáneamente, podemos considerar para cada curva $\gamma \in \mathcal{P}^k$, el conjunto de las secciones de clase de Sobolev l , $l \leq k$, del fibrado $\gamma^*(TP \times TQ)$ que se anula en los extremos. El espacio total de estas secciones definen otro fibrado vectorial $T^{(l)}\mathcal{P}^k \longrightarrow \mathcal{P}^k$ con fibra standard $H_0^l([0, T], \mathbb{R}^N)$. Por tanto la aplicación $\gamma \mapsto \dot{\gamma}$ es una sección de $T^{(k-1)}\mathcal{P}^k$.

Podemos considerar ahora la variedad de Hilbert de curvas de clase de Sobolev k sobre el espacio $P \times_G Q$ considerado en esta tesis. Tendremos, como en el razonamiento anterior, el fibrado $T^{(k-1)}H^1(P \times_G Q)$ sobre él. Hay una proyección natural $\Pi: P \times Q \longrightarrow P \times_G Q$ que induce otra proyección, que denotaremos con la misma letra $\Pi: \mathcal{P}^k \longrightarrow H^k(P \times_G Q)$. Denotaremos por E^k al fibrado de los pull-back de $T^{(k-1)}H^k(P \times_G Q)$ a \mathcal{P}^k a lo largo de la aplicación Π . Notar que la fibra $E^k \longrightarrow \mathcal{P}^k$ en $\gamma = (x, q)$ es el espacio de Hilbert de las H^k secciones del fibrado $(xq) * (T(P \times_G Q))$ que se anulan en los extremos. Además, el fibrado de Hilbert E^k tiene una conexión hermitica ∇ [La71]. Tal conexión puede también, ser explícitamente construida desde una métrica canónica global definida sobre E^k , pero no insistiremos en estos aspectos aquí.

Consideraremos ahora la sección s de E^k definida por la aplicación

$$s(\gamma)(t) = B^P\left(\frac{d}{dt}(x(t)q(t))\right)$$

donde $\gamma(t) = (x(t), q(t))$. Claramente, la sección s es diferenciable y transversa a la sección zero de E^k . Notar que el espacio tangente al punto de la sección cero γ de E^k puede escribirse como $T_\gamma E^k = T_\gamma^{(k-1)}\mathcal{P}^k \oplus E^k$. Pero los vectores "verticales" $\xi(t) \in E^k$ pueden siempre ser obtenidos en el rango de ∇s .

Por tanto, aplicando el teorema de los multiplicadores de Lagrange, Thm. 2.3.1, la curva γ será un punto crítico de la función $S: \mathcal{P}^k \longrightarrow \mathbb{R}$ restringida a la subvariedad definida por el conjunto de zeros de la sección s si y sólo si existe un elemento $p \in (E^k)^*$ tal que (γ, p) es un punto crítico de la función $F: (E^k)^* \longrightarrow \mathbb{R}$ dada por

$$F(\gamma, p) = S(\gamma) + \langle p, s(\gamma) \rangle,$$

donde $\langle \cdot, \cdot \rangle$ es el producto interior en L^2 a lo largo de las fibras de E^k (notar que debido al teorema del embedding de Sobolev podemos identificar E^k con $(E^k)^*$ por el uso del producto interior en L^2). Además por el teorema del embedding de Sobolev, también obtenemos que el punto crítico γ es de clase C^r con $r = k - 1$. Por tanto si γ está en \mathcal{P}^k para todo $k \geq 0$, entonces el par crítico (γ, p) será de clase k para todo $k \geq 0$, por tanto de clase C^∞ .

Por tanto resumimos la discusión anterior en el siguiente teorema:

Teorema 3.3.2 *Con la notación anterior, una curva diferenciable $\gamma(t) = (x(t), q(t))$ es un punto crítico del funcional $S = \int_0^T L(x, q) dt$ sujeto a las condiciones de ligadura horizontal*

$$B^P\left(\frac{d}{dt}(x(t)q(t))\right) = 0$$

si y sólo si existe un levantamiento diferenciable $(\gamma(t), p(t))$ de esta curva tal que es un punto crítico del funcional extendido

$$S_{\mathbb{L}} = \int_0^T \left(L(x, q) + \langle p(t), B^P \left(\frac{d}{dt}(x(t)q(t)) \right) \rangle \right) dt$$

Capítulo 4

Problemas de Control Óptimo de Lie-Scheffers-Brockett

4.1. Elementos de Teoría de Control.

Comenzaremos este capítulo recapitulando algunos elementos de teoría de control y control óptimo que nos van a ser necesarios en lo que sigue. Así estableceremos las notaciones y terminología básica para sistemas de control y algunos resultados analíticos sobre espacios de curvas que nos serán necesarios posteriormente.

4.1.1. Sistemas de Control Óptimo

Adoptaremos para la descripción de sistemas de control óptimo la formulación geométrica iniciada por E. Martínez [Mr99] aunque no utilizaremos la formulación más estilizada que usa la teoría de algebroides. El espacio de estados de un sistema de control es el objeto geométrico que describe las variables que caracterizan los distintos estados del sistema y que supondremos en lo que sigue que es una variedad diferenciable de dimensión finita, con o sin borde. Así a la variedad diferenciable de los estados la denotamos como P y sus coordenadas locales las denotamos como x^i , $i = 1 \dots n$, donde n es la dimensión de P .

Como generalmente los controles van a ser fuerzas externas aplicadas dependiendo del estado, las variables de control las podemos modelar utilizando la idea geométrica de fibrado; generalmente, este fibrado va a ser un fibrado vectorial ó afín, es decir, la fibra sobre un punto arbitrario va a ser un espacio vectorial ó afín. Por lo tanto, tendremos un fibrado con espacio total que denotamos como C sobre el espacio base P y una aplicación de proyección $\pi: C \longrightarrow P$. Las variables de control, es decir, las coordenadas locales a lo largo de las fibras $\pi^{-1}(x)$ de π , las denotamos como u^a , $a = 1, \dots, m$. La fibra $C_x = \pi^{-1}(x)$ constituye el conjunto de controles que actúan en el punto x del espacio de estados. Por tanto, las

coordenadas locales en el fibrado de controles se denotarán por (x^i, u^a) , $i = 1, \dots, n$ y $a = 1, \dots, r$.

La dinámica del sistema es descrita a través de las relaciones físico/mecánicas entre las variables que describen el fibrado de control y sus derivadas. Dichas ecuaciones típicamente son ecuaciones diferenciables de primer o segundo orden y dependen de los controles como parámetros. Utilizaremos como ya hemos dicho la formulación geométrica de estas de E. Martínez que se apoya en la noción de cálculo diferencial a lo largo de aplicaciones. Por tanto estableceremos que la ecuación dinámica de un sistema de control, también llamada ecuación de estado o de control, está descrita por un campo vectorial en el fibrado de los estados y controles a lo largo de la proyección π , i.e., el campo vectorial es una aplicación $\Gamma: C \rightarrow TP$ tal que $\tau_P \circ \Gamma = \pi$ y que en coordenadas tiene la expresión $\Gamma = f^i(x, u)\partial/\partial x^i$.

$$\begin{array}{ccc} & & TP \\ & \nearrow \Gamma & \downarrow \tau_P \\ C & \xrightarrow{\pi} & P \end{array}$$

El funcional objetivo en un problema de control óptimo es la función a minimizar y, en general, es un funcional local definido en un cierto espacio de curvas.

Supondremos por tanto que el funcional objetivo S de un problema de control óptimo está definido mediante una densidad lagrangiana $L: C \rightarrow \mathbb{R}$, es decir,

$$S = \int_0^T L(x, u) dt,$$

y es una función diferenciable en un cierto espacio de curvas admisibles en C .

El funcional objetivo va a estar definido en un espacio de curvas, $\gamma: [0, T] \rightarrow C$ que en esta memoria consideraremos con extremos fijos $x(0) = \pi \circ \gamma(0) = x_0$, $x(T) = \pi \circ \gamma(T) = x_T$.

La solución del problema de control óptimo enunciado antes fue obtenida de manera muy general por L. Pontryagin y sus colaboradores [Po62] y está dada por el conocido como principio del máximo de Pontryagin, enunciado en el siguiente teorema.

Teorema 4.1.1 (Principio del máximo de Pontryagin) [Po62] *La curva $\gamma(t) = (x(t), u(t))$, $t \in [0, T]$, absolutamente continua, es una trayectoria optimal normal si existe un levantamiento de $x(t)$ al espacio de coestados, $(x(t), p(t))$, tal que se satisfagan las ecuaciones de Hamilton,*

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \quad (4.1)$$

y tal que se cumpla:

$$H(x(t), p(t), u(t)) = \max_v H(x(t), p(t), v), \quad a.e. t \in [0, T],$$

donde H denota la función hamiltoniana

$$H(x, p, u) = p_i f^i(x, u) - L(x, u). \quad (4.2)$$

4.1.2. El principio variacional de Hamilton-Pontryagin

En esta sección probaremos que las curvas integrales de los campos vectoriales que satisfacen la ecuación dinámica (4.13) son las trayectorias extremales del problema de control óptimo definido por las ecuaciones:

$$\dot{x}^i(t) = f^i(x(t), u(t)), \quad i = 1, \dots, n, \quad (4.3)$$

y

$$S(\gamma) = \int_0^T L(x(t), u(t)) dt. \quad (4.4)$$

Para esto debemos hallar los puntos críticos de S restringido a la subvariedad de las soluciones del campo vectorial, es decir, restringido a \mathcal{W} . Vimos que los puntos críticos de $S|_{\mathcal{W}}$ son los mismos que los de \mathbb{S} con

$$\mathbb{S}(\gamma) = \int_0^T [L(x(t), u(t)) + \langle p(t), \dot{x}(t) - \Gamma(x(t), u(t)) \rangle] dt$$

donde $\mathbb{S} : \mathcal{C}_{x_0, x_T}(M) \rightarrow \mathbb{R}$, es decir, está definida en el espacio de curvas diferenciables sobre M con puntos fijos x_0, x_T y $\gamma(t) = (x(t), p(t), u(t)) \in \mathcal{C}_{x_0, x_T}(M)$ es una curva $\gamma : [0, T] \rightarrow M$. Entonces si denotamos por ξ a un vector tangente en $T_\gamma \mathcal{C}_{x_0, x_T}(M)$ en la curva $\gamma(t)$, tenemos que si γ es un punto crítico de \mathbb{S} entonces

$$\mathbb{S} \in \mathcal{C}^1, \quad d\mathbb{S}(\gamma) = 0 \iff \frac{\partial \mathbb{S}}{\partial \xi} = 0, \quad \forall \xi = \delta\gamma = (\delta x, \delta p, \delta u).$$

Por lo tanto, tenemos que calcular los puntos que anulan la derivada direccional del funcional objetivo en la dirección $\xi(t) = \delta\gamma(t) = (\delta x(t), \delta p(t), \delta u(t))$.

La definición de derivada direccional en \mathbb{R}^n de una función en un punto x y en la dirección del vector v es la siguiente:

$$\frac{\partial f}{\partial v} = \lim_{s \rightarrow 0} \frac{f(x + sv) - f(x)}{s}.$$

En una variedad diferenciable la noción de $x + sv$ la extendemos por la noción de flujo: $\phi_s(v)$. Se tiene un campo vectorial V tal que $V(x) = v$, es decir, el campo vectorial en el punto x es el vector v , la dirección en la cual tomamos la derivada direccional. Así, en una variedad diferenciable la derivada direccional de una función f en un punto x y en la dirección del vector v es:

$$\frac{\partial f}{\partial v} = \lim_{s \rightarrow 0} \frac{f(\phi_s(v)) - f(x)}{s}.$$

Por lo tanto, la derivada direccional del funcional objetivo \mathbb{S} en un punto de la curva $\gamma(t)$ en la dirección del vector $\xi = \delta\gamma(t)$ es la siguiente:

$$\frac{\partial \mathbb{S}(\gamma(t))}{\partial \xi} = \lim_{s \rightarrow 0} \frac{\mathbb{S}(\phi_s(\gamma(t))) - \mathbb{S}(\gamma(t))}{s},$$

donde para todo t , $\gamma(t) \in M$, $\xi(t) \in T_{\gamma(t)}M$, $\xi_{\gamma(t)} = \xi(t)$, cuyo flujo es ϕ_s . Si denotamos por $\mathbb{L} = L + p_i(\dot{x}^i - f^i)$ con las definiciones del hamiltoniano y de la 1-forma siguientes: $H = L - p_i f^i$ y $\theta = p_i dx^i$, como $\dot{\gamma} = \dot{x}^i \partial / \partial x^i + \dot{p}_i \partial / \partial p_i + \dot{u}^a \partial / \partial u^a$, nos queda que $\theta(\dot{\gamma}) = p_i \dot{x}^i$, por lo que el lagrangiano extendido se puede escribir como $\mathbb{L} = \theta_{\gamma}(\dot{\gamma}) - H_{\gamma}$, entonces tenemos que

$$\begin{aligned} \frac{\partial \mathbb{S}(\gamma(t))}{\partial \xi} &= \lim_{s \rightarrow 0} \left(\frac{1}{s} \int_0^T [\mathbb{L}(\phi_s(\gamma)) - \mathbb{L}(\gamma)] dt \right) = \\ &= \int_0^T \lim_{s \rightarrow 0} \frac{1}{s} [\mathbb{L}(\phi_s(\gamma)) - \mathbb{L}(\gamma)] dt = \\ &= \int_0^T \lim_{s \rightarrow 0} \left[\frac{1}{s} \left(\theta_{\phi_s(\gamma)}(\overline{\phi_s(\gamma)}) - H(\phi_s(\gamma)) - \theta_{\gamma}(\dot{\gamma}) + H(\gamma) \right) \right] dt = \\ &= \int_0^T \lim_{s \rightarrow 0} \left[\frac{1}{s} \left(\theta_{\phi_s(\gamma)}(\overline{\phi_s(\gamma)}) - \theta_{\gamma}(\dot{\gamma}) \right) \right] dt - \\ &\quad - \int_0^T \lim_{s \rightarrow 0} \left[\frac{1}{s} \left(H(\phi_s(\gamma)) - H(\gamma) \right) \right] dt. \end{aligned}$$

Veamos qué son cada una de las partes del integrando. Como H es una función diferenciable se tiene que

$$\lim_{s \rightarrow 0} \frac{H(\phi_s(\gamma)) - H(\gamma)}{s} = dH(\gamma)(\xi).$$

Con la definición de la derivada de Lie, tenemos que

$$\begin{aligned} (\mathcal{L}_{\xi}\theta)(\dot{\gamma}) &= \frac{d}{ds} (\phi_s^* \theta)_{\gamma}(\dot{\gamma}) \Big|_{s=0} = \frac{d}{ds} [\theta_{\phi_s(\gamma)}(\phi_{s^*}(\dot{\gamma}))] \Big|_{s=0} = \\ &= \frac{d}{ds} [\theta_{\phi_s(\gamma)}(\overline{\phi_s(\gamma)})] \Big|_{s=0} = \lim_{s \rightarrow 0} \frac{\theta_{\phi_s(\gamma)}(\overline{\phi_s(\gamma)}) - \theta_{\gamma}(\dot{\gamma})}{s}. \end{aligned}$$

Por lo tanto, la derivada direccional de \mathbb{S} queda de la siguiente manera:

$$\begin{aligned} \frac{\partial \mathbb{S}(\gamma(t))}{\partial \xi} &= \int_0^T [(\mathcal{L}_{\xi}\theta)(\dot{\gamma}) - dH(\gamma)(\xi)] dt = \\ &= \int_0^T [(i_{\xi}d\theta_{\gamma})(\dot{\gamma}) + d(i_{\xi}\theta_{\gamma})(\dot{\gamma}) - dH(\gamma)(\xi)] dt, \end{aligned}$$

donde hemos utilizado la siguiente relación de la derivada de Lie de una 1-forma:

$$(\mathcal{L}_\xi\theta) = i_\xi d\theta_\gamma + di_\xi\theta_\gamma.$$

Con la siguiente expresión: $i_\xi d\theta(\dot{\gamma}) = d\theta(\xi, \dot{\gamma}) = -d\theta(\dot{\gamma}, \xi) = \Omega(\dot{\gamma}, \xi)$ y denotando $dH(\gamma)(\xi) = dH_\gamma(\xi)$ la integral anterior queda de la siguiente forma:

$$\begin{aligned} \frac{\partial \mathbb{S}(\gamma(t))}{\partial \xi} &= \int_0^T [-d\theta(\dot{\gamma}, \xi) - dH_\gamma(\xi)] dt + \int_0^T d(i_\xi\theta)(\dot{\gamma}) dt = \\ &= \int_0^T [(i_{\dot{\gamma}}\Omega)|_\gamma - dH_\gamma](\xi) dt + \int_0^T d(i_\xi\theta)(\dot{\gamma}) dt. \end{aligned}$$

Por último, utilizando el teorema de Stokes la segunda integral se convierte en

$$\int_0^T d(i_\xi\theta)_\gamma(t)(\dot{\gamma}) dt = \int_\gamma d(i_\xi\theta) = \int_{\delta\gamma} i_\xi\theta = (i_\xi\theta)_\gamma(T) - (i_\xi\theta)_\gamma(0).$$

Por lo que hemos demostrado el siguiente teorema.

Teorema 4.1.2 *Las soluciones extremales del problema de control óptimo (4.3-4.4), que son C^∞ a trozos, son curvas integrales de los campos vectoriales Γ que satisfacen la ecuación dinámica:*

$$i_\Gamma\Omega = dH. \quad (4.5)$$

Consideraremos en lo que sigue todos nuestros espacios de curvas como subvariedades de la variedad de Hilbert de curvas H^1 en C , denotado como $\mathcal{P}(C)$.

El fibrado tangente de clase H^1 lo denotaremos por $T\mathcal{P}$ o bien por $H^1(TC)$ y el fibrado tangente de clase H^0 se denotará por $H^0(TC)$.

La ecuación de control Γ es un campo vectorial a lo largo de la aplicación π . Tenemos los siguientes espacios:

$$\mathcal{P}_{x_0, x_T}(C) = \{ \gamma \in H^1([0, T]; C) \mid \gamma \in \mathcal{P}(C), \pi\gamma(0) = x_0, \pi\gamma(T) = x_T \}.$$

$$T_\gamma\mathcal{P}_{x_0, x_T}(C) = \{ \delta\gamma \in H^1(u^*(TC)) \mid \tau_C\delta\gamma(t) = \gamma(t), \pi_*(\delta\gamma(0)) = \pi_*(\delta\gamma(T)) = 0 \}.$$

Donde las proyecciones y los espacios vienen representados en los siguientes diagramas:

$$\begin{array}{ccc} & TP & \\ \Gamma \nearrow & \downarrow \tau_P & \\ C & \xrightarrow{\pi} & P \end{array} \qquad \begin{array}{ccc} TC & \xrightarrow{\pi_*} & TP \\ \tau_C \downarrow & & \downarrow \tau_P \\ C & \xrightarrow{\pi} & P \end{array}$$

Definimos la subvariedad dinámica como el conjunto de curvas γ que verifican la ecuación de control, es decir,

$$\mathcal{W} = \left\{ \gamma \in \mathcal{P}(C) \mid \pi_* \frac{d\gamma(t)}{dt} = \Gamma(\gamma(t), t) \right\}. \quad (4.6)$$

Al igual que en las secciones previas tenemos que

$$T_\gamma \mathcal{W} = \left\{ \delta\gamma \in T_\gamma \mathcal{P} \mid \pi_* \frac{d}{dt} \delta\gamma = D\Gamma(\gamma(t), t) \delta\gamma \right\}. \quad (4.7)$$

De nuevo, tenemos que $\mathcal{W} = A^{-1}(\mathbf{0})$, donde $A: \mathcal{P}(C) \longrightarrow H^0(TC)$ definido como

$$A(\gamma) = \pi_* \frac{d\gamma}{dt} - \Gamma(\gamma(t), t).$$

Nótese que con la proyección $\pi: C \longrightarrow P$ tenemos una aplicación inducida $\tilde{\pi}: \mathcal{P}(C) \longrightarrow \mathcal{P}(P)$, $\tilde{\pi}(\gamma) = \pi \circ \gamma$ y podemos tomar el pull-back del fibrado $H^0(TP)$ a $\mathcal{P}(C)$, es decir, $\tilde{\pi}^*(H^0(TP)) \longrightarrow \mathcal{P}(C)$. Por lo tanto, podemos decir que A es una sección de dicho fibrado sobre $\mathcal{P}(C)$. Entonces \mathcal{W} es el conjunto de nivel cero de la sección A . Denotando como anteriormente $TA(\gamma)$ como $\mathfrak{L}_\gamma(\delta\gamma) = \pi_* \frac{d}{dt} \delta\gamma - D\Gamma(\gamma(t), t) \delta\gamma$, se tiene que:

$$T_\gamma \mathcal{W} = \ker \mathfrak{L}_\gamma.$$

y por tanto:

$$\ker \mathfrak{L}_\gamma = \overline{\text{Ran } \mathfrak{L}_\gamma^+}.$$

Aplicando el Teorema de los multiplicadores de Lagrange (3.3.3), tendremos que un punto $\gamma \in \mathcal{P}(C)$ será un punto crítico ordinario de $S: \mathcal{P}(C) \longrightarrow \mathbb{R}$ restringido a \mathcal{W} si y sólo si $\nabla S_\gamma \in \text{Ran } \mathfrak{L}_\gamma^+$, esto es, si $\exists p \in H^0(\gamma^*(TP))$ tal que

$$\langle \nabla S_\gamma, \delta\gamma \rangle_{H^1} = \langle \mathfrak{L}_\gamma^+ p, \delta\gamma \rangle_{H^1} = \langle p, \mathfrak{L}_\gamma \delta\gamma \rangle_{L^2}. \quad (4.8)$$

Por lo tanto, $\gamma \in \mathcal{W}$ es un punto crítico ordinario de $S|_{\mathcal{W}}$ si y sólo si γ es un punto crítico de $\mathbb{S}: \mathcal{P}(\pi^* T^* P) \longrightarrow \mathbb{R}$ definido como sigue:

$$\mathbb{S}(\gamma, p) = S(\gamma) - \langle p, A(\gamma) \rangle_{L^2} = \int_0^T L(x(t), u(t)) - \langle p(t), A(\gamma(t)) \rangle dt.$$

4.2. La geometría del principio de Hamilton-Pontryagin: El fibrado de Pontryagin

Construiremos el espacio de fases de control, que denotaremos de nuevo por M , como el pull-back del fibrado C a lo largo de π_P , la proyección natural del

fibrado cotangente $\pi_P: T^*P \longrightarrow P$, así $M = \pi^*(T^*P)$, o, equivalentemente como el pull-back del fibrado cotangente a P , también llamado espacio de coestados T^*P , a lo largo de π , $M = \pi_P^*(T^*P)$. En ambos casos, el espacio M se puede describir como el conjunto de puntos (ξ, α) , $\xi \in C$, $\alpha \in T^*P$ que verifican $\pi(\xi) = \pi_P(\alpha)$. El espacio M no es más que la suma de Whitney de la variedad C y la variedad cotangente de P , $M = C \oplus T^*P$ y también puede denominarse el fibrado de Pontryagin del problema de control óptimo objeto de estudio. Las coordenadas canónicas en la variedad $(2n + m)$ -dimensional M las denotamos como (x^i, p_i, u^a) . Tenemos, además, dos proyecciones naturales $pr_1: M \longrightarrow C$, $pr_1(x^i, p_i, u^a) = (x^i, u^a)$ y $pr_2: M \longrightarrow T^*P$, $pr_2(x^i, p_i, u^a) = (x^i, p_i)$, y la relación entre todas estas aplicaciones se recoge en el siguiente diagrama conmutativo:

$$\begin{array}{ccc}
 & M = C \oplus T^*P & \\
 pr_1 \swarrow & & \searrow pr_2 \\
 C & & T^*P \\
 \pi \searrow & & \swarrow \pi_P \\
 & P &
 \end{array}$$

Con respecto a pr_2 , M es un fibrado sobre T^*P cuyas fibras son los espacios de controles mientras que, con respecto a pr_1 , M es un fibrado sobre C con fibras los espacios de coestados.

El espacio M está equipado de forma natural con una 1-forma canónica, la 1-forma de Liouville θ , que en coordenadas locales (x^i, p_i, u^a) tiene la expresión $\theta = p_i dx^i$ y que de manera intrínseca está definida por:

$$\theta(\xi, \alpha_x)(U) = \langle \alpha_x, (\pi \circ pr_1)_*(U) \rangle = \langle \alpha_x, (\pi_P \circ pr_2)_*(U) \rangle, \quad \forall U \in T_{(\xi, \alpha_x)}M. \quad (4.9)$$

Donde, en la parte derecha de la ecuación (4.9), $\langle \cdot, \cdot \rangle$ denota la dualidad natural entre TP y T^*P . Asociada a esta 1-forma tenemos la 2-forma presimpléctica canónica Ω en M ,

$$\Omega = -d\theta, \quad (4.10)$$

que en coordenadas locales es:

$$\Omega = dx^i \wedge dp_i.$$

Esta estructura coincide con la estructura presimpléctica Ω en M inducida mediante el pull-back a lo largo de pr_2 de ω_P , es decir,

$$\Omega = pr_2^* \omega_P. \quad (4.11)$$

La expresión en coordenadas locales de Ω es exactamente la misma que la de ω_P , $\Omega = dx^i \wedge dp_i$, y tiene rango constante igual a $2n$. Esta 2-forma cerrada Ω

es degenerada. Su distribución característica $\ker \Omega$, que denotamos por K , está generada por los “campos vectoriales de control”, $\partial/\partial u^a$, $a = 1, \dots, m$,

$$K = \text{Lin} \left\{ \frac{\partial}{\partial u^a}, \quad a = 1, \dots, m \right\}.$$

La ecuación de estado es un campo vectorial en C a lo largo de la proyección $\pi: C \longrightarrow P$. Podemos extender de manera natural dicha ecuación a M .

En el espacio de fases total M de un problema de control óptimo tenemos, por tanto, la función canónica hamiltoniana H definida como

$$H(x, p, u) = \langle \theta_{(x,p,u)}, \Gamma(x, u) \rangle - L(x, u) = p_i f^i(x, u) - L(x, u), \quad (4.12)$$

donde θ denota la 1-forma canónica de Liouville en M . Este hamiltoniano coincide con el definido por Pontryagin en la presentación estándar de la Teoría de Control Óptimo (4.2). La tripleta (M, Ω, H) define un sistema hamiltoniano pre-simpléctico cuya dinámica viene dada por todos los campos vectoriales Γ_C en M que satisfacen la ecuación dinámica:

$$i_{\Gamma_C} \Omega = dH. \quad (4.13)$$

Como hemos visto en el apartado anterior Ω es degenerada y tiene núcleo distinto de cero; existirán soluciones Γ_C a la ecuación (4.13) si y sólo si

$$i_Z(dH) = 0, \quad \forall Z \in \ker \Omega = K.$$

Con la base $Z_a = \{\partial/\partial u^a\}$ para el núcleo K , la condición previa se convierte en:

$$\phi_a^{(1)} := i_{Z_a}(dH) = Z_a(dH) = \frac{\partial H}{\partial u^a} = 0, \quad a = 1, \dots, m,$$

y, por lo tanto, sólo habrá soluciones a la ecuación $i_{\Gamma_C} \Omega = dH$ en el subconjunto M_1 de M definido por:

$$M_1 = \left\{ (x, u, p) \in M \mid \phi_a^{(1)} = \frac{\partial H}{\partial u^a} = p_i \frac{\partial f^i}{\partial u^a} - \frac{\partial L}{\partial u^a} = 0 \right\}. \quad (4.14)$$

A las funciones $\phi_a^{(1)}$ las vamos a denominar ligaduras de primer orden o primarias de acuerdo con la terminología usual.

Puede ocurrir que este subconjunto M_1 no sea una subvariedad diferenciable. Sin embargo, tenemos que:

Lema 4.2.1 M_1 es una subvariedad diferenciable $\iff \mathbf{0}$ es un valor regular de la aplicación $\Phi: M \longrightarrow \mathbb{R}^m$, $\Phi = \{\phi_a^{(1)}\}$, $a = 1, \dots, m \iff T\Phi_\xi$ es una aplicación sobreyectiva $\forall \xi \in \Phi^{(-1)}(\mathbf{0})$, es decir, la matriz

$$T\Phi = \left[\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial p}, \frac{\partial \Phi}{\partial u} \right]$$

tiene rango máximo $\forall \xi \in M_1$.

Gracias al lema de Sard sabemos que si tenemos un valor crítico c_0 de una aplicación diferenciable, $\forall \varepsilon > 0 \exists c$ tal que $|c - c_0| < \varepsilon$ y c es un valor regular de Φ , es decir, que tan cerca como queramos del valor crítico tenemos un valor regular de la aplicación.

Podemos construir la siguiente forma cuadrática a lo largo de la fibración vertical de $pr_2|_{M_1}$ definida localmente como

$$W = W_{ab} du^a \otimes du^b, \quad (4.15)$$

siendo $W_{ab} = \partial^2 H / \partial u^a \partial u^b$. La forma cuadrática W es un objeto intrínseco, es decir, que no depende de las coordenadas, puesto que es la derivada fibrada o vertical segunda de una función [Ar86].

Con esta forma cuadrática establecemos ya la definición de regularidad desde el punto de vista de control óptimo.

Definición 4.2.1 *El problema de control óptimo definido por las ecuaciones $\dot{x}^i(t) = f^i(x(t), u(t))$, $i = 1, \dots, n$, y $S(\gamma) = \int_0^T L(x(t), u(t)) dt$ se dice regular en el punto $\xi_0 = (x_0, p_0, u_0)$ si la forma cuadrática W es regular en dicho punto. Diremos que el problema (P, C, Γ, L) es regular si W es regular en todo punto de la subvariedad M_1 , en otro caso diremos que el problema es singular en el sentido de control óptimo.*

Entonces, tenemos que si $\mathbf{0}$ es un valor crítico de $\Phi \rightarrow W$ es no invertible y, por tanto, hay puntos en M_1 singulares en el sentido de control óptimo. Por otro lado, si los puntos de M_1 no son singulares en el sentido de control óptimo $\rightarrow \mathbf{0}$ es un valor regular para Φ .

Entonces, la pregunta es la siguiente: ¿Existe un sistema de control óptimo próximo al original de manera que $\mathbf{0}$ sea un valor regular y no altere la naturaleza de la singularidad como problema de control óptimo? La respuesta es que sí. Si tenemos el problema de control óptimo siguiente:

$$\begin{aligned} \dot{x} &= f(x, u) \\ S &= \int_0^T L(x, u) dt, \end{aligned}$$

el hamiltoniano de Pontryagin es, como hemos definido previamente,

$$H = pf(x, u) - L(x, u)$$

y, por lo tanto, las ligaduras primarias quedan

$$\Phi = \frac{\partial H}{\partial u} = p \frac{\partial f}{\partial u} - \frac{\partial L}{\partial u}.$$

La derivada de esta aplicación es la siguiente:

$$T\Phi = \left[p \frac{\partial^2 f}{\partial x \partial u} - \frac{\partial^2 L}{\partial x \partial u}, \frac{\partial f}{\partial u}, p \frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 L}{\partial u^2} \right].$$

Si esta aplicación no tiene rango máximo en un punto dado, podemos perturbar el problema de partida de la siguiente forma:

$$\dot{x} = f_\varepsilon(x, u) = f(x, u) + \varepsilon \rho(u) Bu, \quad (4.16)$$

donde B es una matriz de rango máximo de tal forma que $\text{rank}(\partial f_\varepsilon / \partial u)$ sea máximo en un punto dado (y, por tanto, en un entorno de él), además, el funcional objetivo no se modifica y $\rho(u)$ es una función "bump" con dominio en un disco de radio 1 en torno al punto crítico. Es importante señalar que no estamos cambiando la caracterización del problema como problema de control óptimo singular, puesto que la nueva matriz W_ε cumple que $W_\varepsilon = W$, $\forall \varepsilon$, ya que hemos introducido una perturbación lineal en los controles. Hemos probado por tanto el siguiente lema:

Lema 4.2.2 *Dado un problema de control óptimo (P, C, Γ, L) tal que $\xi_0 = (x_0, p_0, u_0)$ es un punto crítico para Φ y dado $\varepsilon > 0$ existe un problema de control óptimo $(P, C, \Gamma_\varepsilon, L)$ tal que ξ_0 es un punto regular para Φ_ε y $\|\Gamma_\varepsilon - \Gamma\|_\infty < \varepsilon$. Además, la naturaleza de ξ_0 como punto singular del problema de control óptimo no cambia.*

Hemos visto en el apartado anterior la condición necesaria para la existencia de dinámica. Sólo va a ocurrir en la subvariedad M_1 . Vamos a ver ahora cómo tienen que ser los campos vectoriales Γ_C para que sus curvas integrales permanezcan en M_1 . Si el campo vectorial lo escribimos de la manera siguiente:

$$\Gamma_C = \dot{x}^i \frac{\partial}{\partial x^i} + \dot{p}_i \frac{\partial}{\partial p_i} + C^a \frac{\partial}{\partial u^a}, \quad (4.17)$$

al imponer que se cumpla la ecuación dinámica (4.13), se obtiene que

$$\dot{x}^i dp_i - \dot{p}_i dx^i = \frac{\partial H}{\partial x^i} dx^i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial u^a} du^a,$$

de donde deducimos que las curvas integrales de Γ_C verifican las ecuaciones del principio del máximo de Pontryagin (4.1) como ya señalamos anteriormente:

$$\dot{x}^i = \frac{\partial H}{\partial p_i} = f^i \quad (4.18)$$

$$\dot{p}_i = -\frac{\partial H}{\partial x^i} = \frac{\partial L}{\partial x^i} - p_j \frac{\partial f^j}{\partial x^i} \quad (4.19)$$

$$\frac{\partial H}{\partial u^a} = p_j \frac{\partial f^j}{\partial u^a} - \frac{\partial L}{\partial u^a} = 0. \quad (4.20)$$

Las soluciones de estas ecuaciones las llamaremos curvas optimales críticas. El campo vectorial Γ_C queda indeterminado al no aparecer las derivadas de los controles y su expresión es:

$$\Gamma_C = f^i \frac{\partial}{\partial x^i} + \left(\frac{\partial L}{\partial x^i} - p_j \frac{\partial f^j}{\partial x^i} \right) \frac{\partial}{\partial p_i} + C^a \frac{\partial}{\partial u^a}. \quad (4.21)$$

Entonces, en M_1 hay solución de la ecuación de la dinámica (4.13), pero no sabemos si los campos vectoriales Γ_C son paralelos a esta subvariedad, donde paralelo significa que cualquier trayectoria que parta de M_1 va a permanecer todo el tiempo posterior en la subvariedad M_1 . Veremos más adelante que esta condición es necesaria para la consistencia de las ecuaciones. Para que esto se cumpla, tenemos que imponer que el campo deje invariante a la subvariedad, es decir, que aplicado a las ecuaciones que definen la subvariedad M_1 se anule:

$$\Gamma_C(\phi_a^{(1)}) = 0 \quad \text{en } M_1, \quad (4.22)$$

por lo tanto, en M_1 :

$$\begin{aligned} 0 &= \Gamma_C(\phi_a^{(1)}) = f^i \frac{\partial^2 H}{\partial x^i \partial u^a} + \left(\frac{\partial L}{\partial x^i} - p_j \frac{\partial f^j}{\partial x^i} \right) \frac{\partial^2 H}{\partial p_i \partial u^a} + C^b \frac{\partial^2 H}{\partial u^b \partial u^a} = \\ &= f^i \left(p_j \frac{\partial^2 f^j}{\partial x^i \partial u^a} - \frac{\partial^2 L}{\partial x^i \partial u^a} \right) + \left(\frac{\partial L}{\partial x^i} - p_j \frac{\partial f^j}{\partial x^i} \right) \frac{\partial f^i}{\partial u^a} + \\ &+ C^b \left(p_i \frac{\partial^2 f^i}{\partial u^b \partial u^a} - \frac{\partial^2 L}{\partial u^b \partial u^a} \right). \end{aligned}$$

Como ya indicamos previamente, diremos que el problema es regular si la siguiente matriz es invertible en cada punto de M_1 :

$$W_{ab} = \left(p_i \frac{\partial^2 f^i}{\partial u^b \partial u^a} - \frac{\partial^2 L}{\partial u^b \partial u^a} \right) = \frac{\partial^2 H}{\partial u^a \partial u^b} = \frac{\partial \phi_a^{(1)}}{\partial u^b}. \quad (4.23)$$

Si W es invertible podemos obtener C^b como:

$$\begin{aligned} C^b &= -W^{ab} \left\{ f^i \left(p_j \frac{\partial^2 f^j}{\partial x^i \partial u^a} - \frac{\partial^2 L}{\partial x^i \partial u^a} \right) + \left(\frac{\partial L}{\partial x^i} - p_j \frac{\partial f^j}{\partial x^i} \right) \frac{\partial f^i}{\partial u^a} \right\} = \\ &= W^{ab} B_a, \end{aligned}$$

donde $W^{ab}W_{bc} = \delta^a_c$. En cuyo caso se tiene solución única y va a existir un campo vectorial dinámico Γ_C , único y bien definido en M_1 , cuyas curvas integrales van a ser trayectorias extremales de S y tangente en todo punto a la subvariedad M_1 . Podemos, además, despejar los controles de las ecuaciones $\phi_a^{(1)} = 0$, (4.14), que nos definen la subvariedad M_1 en función de las variables $(x, p) \in T^*P$, y así obtenemos un "feedback" optimal ó síntesis optimal:

$$u^a = u^a(x, p), \quad (4.24)$$

con esta definición local de los controles, la introducimos en las ecuaciones del problema (4.3) y (4.4) y nos queda localmente un sistema en el fibrado cotangente a P , T^*P . En este trabajo no abordaremos problemas de Control Óptimo Singulares.

Hay una interpretación muy interesante de la noción de la regularidad de un sistema de control óptimo basada en la noción de singularidades de aplicaciones. La variedad M proyecta sobre T^*P mediante la aplicación de proyección $pr_2: M \longrightarrow T^*P$. La subvariedad M_1 , definida mediante las funciones de ligadura $\phi_a^{(1)}$, proyecta sobre T^*P al restringir la aplicación de proyección pr_2 a ésta. Vamos a denotar la restricción de pr_2 a M_1 como p_1 , es decir, $p_1 = pr_2|_{M_1}$, y la inmersión de ésta en M como $i_1: M_1 \longrightarrow M$, por tanto, $p_1 = pr_2 \circ i_1$. La subvariedad M_1 hereda, por tanto, una 2-forma cerrada $\Omega_1 = i_1^*\Omega$. Finalmente, si denotamos por W_1 la restricción de W a M_1 , tenemos (ver [De04]):

$$\ker \Omega_1 = \ker p_{1*} = TM_1 \cap K = \ker W_1.$$

No es difícil probar las siguientes caracterizaciones equivalentes de regularidad.

Teorema 4.2.1 *Las siguientes afirmaciones son equivalentes:*

1. El punto $\xi = (x, p, u) \in M_1$ es regular.
2. $\det W_{ab}(\xi) \neq 0$.
3. $T_\xi M_1 \cap K_\xi = 0$.
4. p_1 es un difeomorfismo local en ξ .
5. pr_2 es transversal a M_1 en ξ .
6. Ω_1 es de rango máximo en ξ .
7. p_1 define un symplectomorfismo local entre (M_1, Ω_1) y (T^*P, w_p) en ξ .
8. Existe una solución única maximal de la ecuación $i_\Gamma \Omega_1 = dH_1$ que pasa por ξ .
9. Existe una única curva optimal crítica a través de ξ .

4.3. Sistemas de Control Óptimo de tipo Lie-Scheffers-Brockett.

Teorema de Lie-Scheffers

Teorema 4.3.1 *Dado un sistema no autónomo de n ecuaciones diferenciales de primer orden:*

$$\frac{dx^i}{dt} = X^i(x^1, \dots, x^n, t), i = 1, \dots, n. \quad (4.25)$$

una condición necesaria y suficiente para que exista un principio de superposición $\Phi: \mathbb{R}^{n(m+1)} \longrightarrow \mathbb{R}^n$ tal que la solución general puede ser escrita como:

$$x = \Phi(x_{(1)}, \dots, x_{(m)}; k_1, \dots, k_n) \quad (4.26)$$

con $\{x_{(a)} \mid a = 1, \dots, m\}$ un conjunto de soluciones particulares independientes del sistema y k_1, \dots, k_n un sistema de n constantes arbitrarias, es que el sistema se puede escribir como:

$$\frac{dx^i}{dt} = Z_1(t)\xi^{1i}(x) + \dots + Z_r(t)\xi^{ri}(x) \quad (4.27)$$

donde Z_1, \dots, Z_r son funciones que dependen solamente de t y $\xi^{\alpha i}, \alpha = 1, \dots, r$ son funciones de $x = (x^1, \dots, x^n)$ tal que los r campos vectoriales en \mathbb{R}^n dados por:

$$Y^{(\alpha)} = \sum_{i=1}^n \xi^{\alpha i}(x^1, \dots, x^n) \frac{\partial}{\partial x^i}, \alpha = 1, \dots, r. \quad (4.28)$$

cierran un álgebra de Lie real de dimensión r , es decir $Y^{(\alpha)}$ son linealmente independientes y existen r^3 constantes de estructura $f_{\gamma}^{\alpha\beta}$, tales que:

$$[Y^{(\alpha)}, Y^{(\beta)}] = \sum_{\gamma=1}^r f_{\gamma}^{\alpha\beta} Y^{(\gamma)} \quad (4.29)$$

Llamaremos sistema de control de Lie-Scheffers-Brockett a la clase de problemas de control que en coordenadas locales de estado x^i , tienen la forma

$$\dot{x}^i = \sum_{a=1}^r c_a(t) X_a^i(x),$$

donde además los campos de de vectores $X_a = X_a^i \partial / \partial x^i$, satisfacen la condición de cierre de Lie

$$[X_a, X_b] = C_{ab}^c X_c,$$

para algunas constantes C_{ab}^c . Resulta natural utilizar este nombre ya que como hemos indicado anteriormente esta clase de sistemas fueron considerados, obviamente no como un problema de control, por Lie and Scheffers [Li84] y mucho más tarde por R. Brockett [Br70], [Br73b], aunque ya en el marco de la teoría de control.

Los sistemas de Lie-Scheffers-Brockett son los ejemplos prototipo de sistemas dinámicos sobre grupos de Lie ó espacios homogéneos. Lie y Scheffers encontraron que para un sistema dinámico de esta forma siempre existían principios de superposición no lineal para la composición de soluciones de la ecuación diferencial, siendo la ecuación de Riccati el ejemplo más conocido y celebrado.

Por otra parte, si M es un espacio homogéneo para el grupo de Lie G y denotamos por ξ_a una base de su álgebra de Lie \mathfrak{g} , podemos considerar la familia de campos de vectores X_a inducidos sobre M por ellos. Un sistema de Brockett sobre M es exactamente un campo vectorial

$$\Gamma = \sum_{a=1}^r u_a(t) X_a$$

donde u_a son las variables de control. Bajo esta forma, la teoría de control de sistemas de Brockett sobre espacios homogéneos es una formulación global de los sistemas considerados por Lie and Scheffers (ver [Ca00] y la sección §4.3 para conocer en detalle el teorema de Lie-Scheffers y sus aplicaciones).

Sea G un grupo de Lie actuando por la derecha sobre una variedad diferenciable P . Denotaremos por \mathfrak{g} al álgebra de Lie de G y por $\xi \in \mathfrak{g}$ a un elemento genérico. Entonces ξ_P denotará el campo de vectores sobre P asociado a ξ por la acción de G , es decir,

$$\xi_P(x) = \left. \frac{d}{dt} (x \cdot \exp(-t\xi)) \right|_{t=0}, \quad x \in P.$$

Si ξ_a es una base dada en \mathfrak{g} , $[\xi_a, \xi_b] = C_{ab}^c \xi_c$, y X_a denota el correspondiente campo vectorial en P , esto es $X_a = (\xi_a)_P$ con la notación anterior, tendremos que

$$\xi_P = \sum_{a=1}^r u^a X_a,$$

donde $\xi = u^a \xi_a$ y,

$$[X_a, X_b] = C_{ab}^c X_c.$$

Si permitimos que las coordenadas u^a sobre el álgebra de Lie \mathfrak{g} sean funciones que dependan del tiempo, podríamos interpretarlas como variables de control que controlan nuestro sistema $\Gamma := \xi_P$.

Consideraremos el sistema dinámico no autónomo sobre P definido por el campo vectorial ξ_P , es decir,

$$\dot{x}^i = \xi_P(x) = \sum_{a=1}^r u^a(t) X_a^i(x), \quad (4.30)$$

y lo interpretaremos como la ecuación de estado para un sistema de control.

Restringiremos la clase de los sistemas que nos interesan, introduciendo una estructura adicional, que nos procurará una realización particular de las variables de control u^a . Supongamos que el grupo de Lie G actúa por la izquierda sobre una variedad diferenciable Q . Asumiremos por simplicidad que la acción es propia y libre, por tanto el espacio cociente Q/G , es una variedad diferenciable N y la

proyección canónica $\pi: Q \longrightarrow N$ es una submersion. Por consiguiente π define una fibración principal sobre N con grupo estructural G .

Sea B una conexión principal sobre el fibrado principal $Q(G, N)$, esto es B es una \mathfrak{g} -valuada 1-forma $B: TQ \longrightarrow \mathfrak{g}$ sobre Q tal que $B(\xi_Q(q)) = \xi$, $\forall \xi \in \mathfrak{g}$, and $L_g^* B = Ad_g B$, $g \in G$, donde Ad_g denota la acción adjunta de G sobre \mathfrak{g} y L_g la acción por la izquierda de G sobre Q . La restricción de una conexión al espacio tangente $T_q Q$ la denotamos por B_q . Entonces podemos definir una aplicación $\xi: TQ \longrightarrow \mathfrak{g}$ por medio de:

$$\xi(q, \dot{q}) = B_q(\dot{q}), \quad (q, \dot{q}) \in T_q Q. \quad (4.31)$$

Si consideramos ahora un sistema de Lie-Scheffers-Brockett de la forma,

$$\dot{x} = \xi(q, \dot{q})_P(x), \quad (4.32)$$

diremos que está controlado por las variables $u = (q, \dot{q}) \in TQ$ y la conexión B . En coordenadas locales (q^i, \dot{q}^i) sobre TQ , la aplicación ξ tendrá la forma:

$$\xi(q, \dot{q}) = \dot{q}^i B_i(q) = \dot{q}^i B_i^a(q) \xi_a,$$

y si describimos los campos de vectores X_a en coordenadas locales x^α sobre P como

$$X_a = \xi_a^\alpha(x) \frac{\partial}{\partial x^\alpha},$$

finalmente obtendremos, para el sistema de control anterior, la expresión en coordenadas locales siguiente:

$$\dot{x}^\alpha = \dot{q}^i B_i^a(q) \xi_a^\alpha(x).$$

Finalmente, si $L: P \times TQ \longrightarrow \mathbb{R}$ es una función sobre $P \times TQ$, podemos construir el funcional objetivo

$$S = \int_0^T L(x, q, \dot{q}) dt, \quad (4.33)$$

definido en un espacio apropiado de curvas sobre $P \times Q$. Restringiremos nuestra atención a funcionales objetivo G -invariantes, es decir, tales que $L: P \times TQ \longrightarrow \mathbb{R}$ es una función G -invariante sobre TQ . Notar que G actúa por la izquierda sobre TQ por levantamiento de la acción de Q a TQ . El espacio cociente $P \times TQ/G$ puede ser identificado (usando una conexión) con el pull-back a TN del fibrado $P \times adQ$ donde adQ es el fibrado adjunto \mathfrak{g} -algebra de $Q \longrightarrow N$ que puede ser identificado como $Q \times_G \mathfrak{g} = Q \times \mathfrak{g}/G$ con coordenadas locales naturales $(n^\alpha, \dot{n}^\alpha, \xi^a)$, donde n^α son coordenadas locales en N y ξ^a en \mathfrak{g} .

Finalmente, si fijamos las condiciones en los extremos

$$x_0 = x(0), \quad x_T = x(T), \quad (4.34)$$

podemos considerar el problema de control óptimo en el espacio de estados P con ecuación de control (4.32), funcional objetivo (4.33), y dichas condiciones extremales (4.34).

Siempre y cuando la variedad Q no tenga bordes, es bien sabido que si la curva $(x(t), q(t), \dot{q}(t))$ es un minimizador normal del problema de control óptimo dado, entonces existe una extensión $(x(t), p(t))$ de la trayectoria de estado $x(t)$ al espacio de coestados T^*P , que satisface las ecuaciones de Pontryagin:

$$\dot{x} = \xi(q, \dot{q})_P(x), \quad \dot{p} = -\frac{\partial H_P}{\partial x}, \quad \frac{\partial H_P}{\partial q} = \frac{\partial H_P}{\partial \dot{q}} = 0,$$

donde H_P es la función Hamiltoniana de Pontryagin,

$$H_P(x, p; q, \dot{q}) = \langle p, \xi(q, \dot{q})_P(x) \rangle - L(x, q, \dot{q}).$$

En términos más geométricos, podemos decir que los minimizadores normales son curvas integrales del sistema presimpléctico (K, Ω_P, H_P) donde $K = T^*P \times TQ$ y Ω_P es la forma presimpléctica obtenida por pull-back a K de la forma canónica simpléctica ω_P sobre T^*P .

4.4. Un teorema de dualidad para problemas de Control Óptimo de Lie-Scheffers-Brockett y sistemas Lagrangianos con simetría

4.4.1. Representación de Clebsch de sistemas de Lie-Scheffers-Brocket de control óptimo

En este capítulo establecemos que los extremales normales para el problema de control óptimo presentado anteriormente pueden ser caracterizados como caminos críticos del funcional objetivo (4.33) en el espacio de curvas $\Omega(P \times Q; x_0, x_T)$ sujeto a la restricción impuesta por la ecuación (4.32) ó, alternativamente, usando una versión apropiada del teorema de los multiplicadores de Lagrange (ver Thm. 2.3.1 y Thm. 3.3.2 del capítulo anterior para los detalles) podemos considerarlos como caminos críticos $(x(t), p(t); q(t))$ del funcional extendido

$$S_L = \int_0^T (L(x, q, \dot{q}) + \langle p, \dot{x} - \xi(q, \dot{q})_P(x) \rangle) dt \quad (4.35)$$

en un espacio apropiado de curvas $\gamma(t) = (x(t), p(t), q(t)) \in \Omega_{x_0, x_T}(T^*P \times Q)$ con extremos fijos x_0, x_T .

4.4.2. Puntos críticos del Lagrangiano de Clebsch

Un simple cálculo nos permite escribir la densidad Lagrangiana $\mathbb{L}: T(T^*P) \times TQ \longrightarrow \mathbb{R}$ del funcional (4.35), como

$$\begin{aligned} \mathbb{L}(x, p, \dot{x}, \dot{p}, q, \dot{q}) &= L(x, q, \dot{q}) - \langle p, (B_q(\dot{q}))_P(x) \rangle + \langle p, \dot{x} \rangle = \\ &= L(x, q, \dot{q}) + \langle J(x, p), B_q(\dot{q}) \rangle + \langle \theta_P(x, p), \dot{x} \rangle. \end{aligned} \quad (4.36)$$

La aplicación $J: T^*P \longrightarrow \mathfrak{g}^*$ en (4.36) denota la aplicación momento asociada al levantamiento cotangente de la acción de G a P , $\langle J(x, p), \xi \rangle = \langle \xi_P(x), p \rangle$, para todo $\xi \in \mathfrak{g}$, $(x, p) \in T^*P$. La 1-forma θ_P denota la 1-forma canónica de Liouville sobre T^*P con expresión en coordenadas locales $\theta_P = p_\alpha dx^\alpha$.

Demostraremos en las siguientes secciones que la expresión en la parte derecha de la ecuación (4.36) tiene la forma de un Lagrangiano de Clebsch. Mientras tanto, resumiremos la discusión en el siguiente teorema:

Teorema 4.4.1 *Sea x_0, x_T dos puntos fijos en P y q_0 en Q , entonces las siguientes afirmaciones son equivalentes:*

i.- La curva diferenciable $(x(t), q(t))$ es un punto crítico del funcional objetivo:

$$S_L[(x, q)] = \int_0^T L(x, q, \dot{q}) dt,$$

sobre el espacio de curvas que satisface la ecuación $\dot{x} = \xi(q, \dot{q})_P(x)$, y $x(0) = x_0, x(T) = x_T, q(0) = q_0$.

*ii.- Existe un levantamiento $p(t)$ de la curva $x(t)$ a T^*P tal que la curva diferenciable $(x(t), p(t); q(t))$ es un punto crítico del funcional:*

$$S_{\mathbb{L}}[(x, p, q)] = \int_0^T (L(x, q, \dot{q}) + \langle J(x, p), B_q(\dot{q}) \rangle + \langle \theta_P(x, p), \dot{x} \rangle) dt.$$

Llamaremos al Lagrangiano \mathbb{L} el Lagrangiano de Clebsch para el problema de control óptimo de Lie-Scheffers-Brockett.

4.4.3. Equivalencias entre las distintas representaciones del problema de Lie-Scheffers-Brockett

Y ahora, reuniendo los resultados obtenidos en el Teorema 4.4.1 de la sección anterior y el Teorema 3.3.1 previo, podemos establecer las siguientes equivalencias que resultan de las diversas representaciones del problema de Lie-Scheffers-Brockett ó del sistema Lagrangiano con simetría.

Teorema 4.4.2 *Dados $g_T \in G$, $q_0 \in Q$, $x_0 \in P$, $q_T = g_T q_0 \in Q$, $x_T = x_0 g_T^{-1} \in P$, hay una correspondencia uno-uno entre los puntos críticos de los siguientes funcionales en sus correspondientes espacios de caminos.*

i.- (Sistema Lagrangiano invariante) *La curva $q(t) \in \Omega_{q_0, x_0 G q_T}(Q)$ es un punto crítico del funcional $S_L : \Omega_{q_0, x_0 G q_T}(Q) \rightarrow \mathbb{R}$ definido por*

$$S_L[q] = \int_0^T L(q, \dot{q}) dt. \quad (4.37)$$

ii.- (Sistema Lagrangiano de Clebsch) *Hay una curva $(x(t), p(t)) \in \Omega_{x_0, x_T}(T^*P)$ tal que la curva $(x(t), p(t); q(t))$ es un punto crítico del funcional $S_{\mathbb{L}} : \Omega_{x_0, x_T; q_0}(T^*P \times Q) \rightarrow \mathbb{R}$ definido por:*

$$\begin{aligned} S_{\mathbb{L}}[x, p; q] &= \int_0^T \mathbb{L}(x(t), p(t), \dot{x}(t), \dot{p}(t); q(t), \dot{q}(t)) dt \\ &= \int_0^T (L(q, \dot{q}) + \langle J(x, p), B_q(\dot{q}) \rangle + \theta_P(x, p)(\dot{x}, \dot{p})) dt. \end{aligned} \quad (4.38)$$

iii.- (Sistema de control óptimo de Lie-Scheffers-Brockett) *Dados x_0, x_T en P y q_0 en Q , la curva $(x(t), q(t))$ en el subespacio $\Omega_{x_0, x_T; q_0}(P \times Q)$ definido por la ecuación $\dot{x} = \xi(q, \dot{q})(x)$, es un punto crítico del funcional objetivo:*

$$S[q] = \int_0^T L(q, \dot{q}) dt.$$

iv.- (Sistema Lagrangiano reducido de Clebsch) *La curva $(x(t), p(t); n(t), \zeta(t)) \in \Omega_{x_0, x_T}(T^*P) \times \Omega_{n_0, n_T}(N) \times \Omega(\mathfrak{g})$, con $x_T = i_q^{-1}(x_0 q_h(T))$, es un punto crítico del funcional \bar{S} inducido por $S_{\mathbb{L}}$ y definido por*

$$\bar{S}[x, p; n, \zeta] = \int_0^T (l(n, \dot{n}, \zeta) + \langle J(x, p), B_q(\dot{q}) \rangle + \theta_P(x, p)(\dot{x}, \dot{p})) dt \quad (4.39)$$

Corolario 4.4.1 *Con la notación anterior, las curvas críticas en $\Omega_{q_0, x_0 G q_T}(Q)$ del funcional acción $S_L[q] = \int_0^T L(q, \dot{q}) dt$ del Lagrangiano L con grupo de simetría G , están en correspondencia uno-uno con las curvas críticas del problema de control óptimo de Lie-Scheffers-Brockett definidas por la ecuación de control $\dot{x} = B_q(\dot{q})_P(x)$ y el funcional objetivo S con extremos x_0, x_T y $x_T = x_0 g_T^{-1}$, $q_T = g_T q_0$.*

Capítulo 5

Estructuras de Dirac y la Teoría de Control Óptimo

Recientemente se ha comenzado a explorar el uso de diversas estructuras geométricas para obtener una comprensión más profunda de diversos problemas dinámicos y del cálculo de variaciones. Una vía prometedora consiste en utilizar la noción de algebroides (ver E. Martínez [Mr04]) y estrechamente relacionado con ella, el uso de estructuras de Dirac. En un reciente conjunto de artículos Yoshimura y Marsden [Yo07a], [Yo09], han presentado un conjunto de ideas para el estudio de la reducción de sistemas Lagrangianos ó Hamiltonianos implícitos utilizando estructuras de Dirac. En este capítulo aplicaremos este punto de vista al problema de la reducción de un problema de control óptimo con simetría.

5.1. Estructuras de Dirac

En esta sección discutiremos la reducción de Lie-Dirac de una estructura de Dirac sobre el fibrado cotangente T^*M de cierta variedad diferenciable M .

Recordemos primero la definición de una estructura de Dirac sobre un espacio vectorial V . Sea V^* el espacio dual de V , y $\langle \cdot, \cdot \rangle$ el pairing natural entre V y V^* . Definimos el pairing simétrico $\langle\langle \cdot, \cdot \rangle\rangle$ sobre $V \oplus V^*$ por:

$$\langle\langle (v, \alpha), (\tilde{v}, \tilde{\alpha}) \rangle\rangle = \langle \alpha, \tilde{v} \rangle + \langle \tilde{\alpha}, v \rangle,$$

para $(v, \alpha), (\tilde{v}, \tilde{\alpha}) \in V \oplus V^*$. Una estructura de Dirac sobre V es un subespacio $D \subset V \oplus V^*$ tal que $D = D^\perp$, donde D^\perp es el ortogonal de D relativo al pairing $\langle\langle \cdot, \cdot \rangle\rangle$.

Ahora sea M una variedad dada y sea $TM \oplus T^*M$ el fibrado suma de Whitney sobre M , es decir el fibrado sobre la base M y con fibra sobre el punto $m \in M$ igual a $T_m M \times T_m^* M$. En esta memoria llamaremos a un subfibrado $D \subset TM \oplus T^*M$ una estructura de Dirac fibrada sobre M cuando es una estructura de Dirac en el sentido de espacios vectoriales en cada punto $m \in M$.

Una 2-forma dada Ω sobre M junto con una distribución $\Delta \subset TM$ sobre M determina una estructura de Dirac sobre M como sigue: sea $m \in M$, y definimos

$$D(m) = \{(v_m, \alpha_m) \in T_m M \times T_m^* M \mid v_m \in \Delta(m), \alpha_x(\omega_x) = \Omega_\Delta(v_x, \omega_x), \forall \omega_x \in \Delta(x)\}$$

donde Ω_Δ es la restricción de Ω a Δ .

Diremos que una estructura de Dirac es integrable si la condición

$$\langle \mathcal{L}_{X_1} \alpha_2, X_3 \rangle + \langle \mathcal{L}_{X_2} \alpha_3, X_1 \rangle + \langle \mathcal{L}_{X_3} \alpha_1 \rangle = 0,$$

es satisfecha para todas las parejas de vectores y 1-formas (X_1, α_1) , (X_2, α_2) y (X_3, α_3) que toman valores en D , donde \mathcal{L}_X denota la derivada de Lie a lo largo del campo vectorial X sobre M . En el conjunto de secciones del fibrado D podemos introducir un corchete siguiendo la definición de Courant [?]. Así, si $(X, \alpha), (Y, \beta) \in D$, se define

$$[(X, \alpha), (Y, \beta)] = ([X, Y], L_X \beta - L_Y \alpha + d()).$$

Se puede comprobar que son equivalentes la anterior noción de integrabilidad y que la distribución D sea integrable respecto al corchete de Courant.

Una de las más importantes e interesantes estructuras de Dirac en mecánica es la inducida por ligaduras cinemáticas no holónomas junto a la estructura (pre)simpléctica definida por un lagrangiano. Tales ligaduras están generalmente dadas por una distribución sobre una variedad de configuración

Sea M una variedad de configuración. Sea TM el fibrado tangente y T^*M el fibrado cotangente. Sea $\Delta_M \subset TM$ una distribución regular sobre M y definimos una distribución levantada sobre T^*M por

$$\Delta_{T^*M} = (T\pi_M)^{-1}(\Delta_M) \subset TT^*M.$$

donde $\pi_M: T^*M \rightarrow M$ es la proyección canónica luego su tangente será $T\pi_M: TT^*M \rightarrow TM$. Sea Ω la 2-forma canónica sobre T^*M . La estructura de Dirac inducida D_{Δ_M} sobre T^*M , es el subfibrado de $TT^*M \oplus T^*T^*M$, cuya fibra es dada para cada $p_m \in T^*M$ como:

$$D_{\Delta_M}(p_m) = \{(v_{p_m}, \alpha_{p_m}) \in T_{p_m} T^*M \times T_{p_m}^* T^*M \mid v_{p_m} \in \Delta_{T^*M}(p_m),$$

$$\alpha_{p_m}(\omega_{p_m}) = \Omega(p_m)(v_{p_m}, \omega_{p_m}), \forall \omega_{p_m} \in \Delta_{T^*M}(p_m)\}.$$

Discutiremos la aparición de dicho concepto en problemas de control óptimo más adelante.

5.1.1. La estructura de Dirac de un problema de control óptimo

Después del marco presimpléctico de los problemas de control óptimo iremos un paso más allá y describiremos su geometría usando la noción de estructuras de Dirac.

Hay dos maneras de introducir una estructura de Dirac asociada a un problema de control óptimo. Por un lado podemos considerar el espacio de Pontryagin (ó fibrado de Pontryagin) $M = T^*P \times_P C$ de un problema de control, equipado con su estructura presimpléctica natural Ω . Dicha estructura presimpléctica define, como indicamos en la sección 5.1, una estructura de Dirac $D_\Omega \subset TM \oplus T^*M$. De hecho podemos escribir explícitamente:

$$D_\Omega(x, p, u) = \{(v_x, v_p, v_u; p_x, p_p, p_u) \in T_{x,p,u}M \oplus T_{x,p,u}^*M \mid v_x = p_p; v_p = -p_x; p_u = 0\}.$$

La dinámica definida por dicha estructura de Dirac es simplemente $(X, dH) \in D_\Omega$ donde H es el Hamiltoniano de Pontryagin en M . Podemos por tanto concluir con la proposición-definición:

Proposición 5.1.1 *Consideremos el problema de control óptimo definido en el fibrado de controles $C \xrightarrow{\pi} P$ con ecuación de control Γ y densidad Lagrangiana $L: C \rightarrow \mathbb{R}$, entonces hay una estructura de Dirac canónica sobre $M = T^*P \times_P C$ dada por D_Ω donde Ω es la estructura presimpléctica canónica sobre M . Además si $H(x, p, u) = \langle p, \Gamma(x, u) \rangle - L(x, u)$ define el Hamiltoniano de Pontryagin, entonces la dinámica definida por D_Ω y dH describe las curvas extremales para el problema de control óptimo.*

$$(X, dH) \in D_\Omega$$

Hay, sin embargo otra manera de proporcionar una descripción en términos de estructuras de Dirac de un sistema de control óptimo. Podemos inspirarnos para ello en la descripción de Marsden y Yoshimura [Yo09] de un sistema lagrangiano implícito. Por tanto denominaremos esta descripción como descripción implícita de problemas de control óptimo en términos de estructuras de Dirac.

$$\begin{array}{ccc} M & \xrightarrow{X} & T(T^*P) \\ & \searrow p_P & \downarrow \tau_{T^*P} \\ & & T^*P \end{array} \qquad \begin{array}{ccc} M & \xrightarrow{p_P} & T^*P \\ & \searrow \Phi & \downarrow \pi_P \\ & & P \end{array}$$

Consideraremos la estructura de Dirac canónica $D_P \subset T(T^*P) \oplus T^*(T^*P)$ definida por la estructura simpléctica canónica ω_P en T^*P . La dinámica del sistema estará descrita por un campo vectorial X a lo largo de la aplicación canónica $p_P: M = T^*P \times_P C \rightarrow T^*P$, esto es, $X: M \rightarrow T(T^*P)$ tal que

$\tau_{T^*P} \circ X = p_P$, ó en términos de los puntos de los espacios involucrados involucrados, si $(x, p, u) \in M$, entonces $X(x, p, u) \in T_{x,p}(T^*P)$. El hamiltoniano de Pontryagin $H: M \rightarrow \mathbb{R}$, nos permite definir $dH: M \rightarrow T^*M$, pero podemos tomar la restricción de dH al subespacio TM_1 , donde $i_1: M_1 \rightarrow M$ es la subvariedad definida por las condiciones $\phi_a = \frac{\partial H}{\partial u^a} = 0$. Nótese que $i_{1*}: TM_1 \hookrightarrow TM$ permite considerar de manera natural la restricción de dH a TM_1 . Por lo tanto la formulación de Dirac implícita del problema de control óptimo (Γ, L, C) está dada por las ecuaciones:

$$(X, dH \Big|_{TM_1}) \in D_P \text{ en } M_1$$

ó, más explícitamente, si $(x, p, u) \in M$, entonces

$$\left(X(x, p, u), dH(x, p, u) \Big|_{TM_1} \right) \in D_P(x, p), \forall (x, p, u) \in M_1.$$

Nótese que si tomamos el pull-back de la estructura de Dirac D_P a lo largo de la aplicación p_P recuperaremos la descripción presimpléctica anterior.

5.1.2. Simetrías y Sistemas de Control Óptimo

Estudiaremos la reducción de sistemas de control óptimo con simetría desde el punto de vista de la reducción del problema variacional que los definen. Para ello comenzaremos precisando la noción de simetría de un problema de control óptimo. Sea G un grupo de Lie actuando sobre el fibrado de controles C . Asumiremos que la acción de G en C desciende a una acción de G en P . Denotaremos la acción de G en C por $\Psi: G \times C \rightarrow C$ y la acción de G en P por $\Phi: G \times P \rightarrow P$. Por tanto tendremos que $\pi \circ \Psi_g = \Phi_g \circ \pi$, $\forall g \in G$ donde $\Phi_g(x) = \Phi(g, x)$ y $\Psi_g(x, u) = \Psi(g, (x, u))$, $\forall g \in G$, $(x, u) \in C$. ó en otras palabras, el siguiente diagrama es conmutativo.

$$\begin{array}{ccc} C & \xrightarrow{\Psi_g} & C \\ \pi \downarrow & & \downarrow \pi \\ P & \xrightarrow{\Phi_g} & P \end{array}$$

El sistema de control $(C \rightarrow P, \Gamma, L)$ se dirá que es G -invariante, ó que tiene G como grupo de simetrías, si verifica que tanto la ecuación de control Γ como el Lagrangiano L son invariantes con respecto a la acción de G , esto es:

1. $L(gx, gu) = L(x, u)$, $\forall (x, u) \in C$, $\forall g \in G$,
2. $T\Phi_g \circ \Gamma = \Gamma \circ \Psi_g$, $\forall g \in G$,

Nótese en primer lugar que la acción del grupo G se levanta naturalmente tanto a TP como a T^*P respectivamente. Denotaremos dichas acciones por $g \cdot v = T\Phi_g(v)$ y $g \cdot p = T\Phi_{g^{-1}}^*(p)$, $\forall v \in TP$, $p \in T^*P$, $g \in G$. Por lo tanto el grupo G también actuará sobre el fibrado de Pontryagin $M = T^*P \times_P C$ a través de:

$$g \cdot (x, p, u) = (gx, gp, gu) = (\Phi_g(x), T\Phi_{g^{-1}}^*(x)p, \Psi_{(g,x)}(u))$$

Segundo, observamos que el marco presimpléctico del sistema de control es también G -invariante y el grupo G es representado (pre)simplécticamente sobre tal sistema.

Proposición 5.1.2 *El grupo G actúa (pre)simplécticamente sobre (M, Ω, H) . Además, hay una aplicación momento $J_M : M \rightarrow \mathfrak{g}^*$ tal que:*

$$i_{\xi_M} \Omega = \langle dJ_M, \xi \rangle$$

donde $\xi \in \mathfrak{g}$ y ξ_M es el campo vectorial de killing definido por la acción de G en M asociada a ξ .

Demostración. Primero comprobamos que $H(x, p, u) = H(gx, gp, gu)$, $\forall g \in G$

$$\begin{aligned} H(gx, gp, gu) &= \langle gp, \Gamma(gx, gu) \rangle - L(gx, gu) = \\ &= \langle T\Phi_{g^{-1}}^*(x)p, T\Phi_g(x)\Gamma(x, u) \rangle - L(x, u) = \\ &= \langle p, \Gamma(x, u) \rangle - L(x, u) = H(x, p, u) \end{aligned}$$

Ahora verificamos que $g^*\Omega = \Omega$. De hecho, como el siguiente diagrama conmuta:

$$\begin{array}{ccc} M & \xrightarrow{g} & M \\ p_P \downarrow & & \downarrow p_P \\ T^*P & \xrightarrow{g} & T^*P \end{array}$$

ya que es fácil comprobar las siguientes igualdades

$$\begin{aligned} p_P \circ g(x, p, u) &= p_P(gx, gp, gu) = \\ &= (gx, gp) = g(x, p) = g \circ p_P(x, p, u) \end{aligned}$$

y por tanto, obtenemos

$$g^*\Omega = g^*p_P^*\omega_P = (p_P \circ g)^*\omega_P = (g \circ p_P)^*\omega_P = p_P^*(g^*\omega_P) = p_P^*\omega_P = \Omega.$$

Ahora comprobemos que la aplicación $J_M = p_P^*J_P : M \rightarrow \mathfrak{g}^*$, donde $J_P : T^*P \rightarrow \mathfrak{g}^*$ es la aplicación momento canónica definida por la acción cotangente del grupo de Lie G en T^*P . La aplicación J_P está definida como sigue:

$$\langle J_P(x, p), \xi \rangle = \langle p, \xi_P(x) \rangle, \quad \forall (x, p) \in T^*P,$$

donde ξ_P denota el campo vectorial de killing definido por la acción de G en P , (recordemos que $\xi_P(x) = \frac{d}{ds} (\exp(s\xi) \cdot x) |_{s=0}$).

Ahora obtenemos

$$\begin{aligned} i_{\xi_M} \Omega(\cdot) &= \Omega(\xi_M, \cdot) = \\ &= p_P^* \omega_P(\xi_M, \cdot) = \omega_P(p_{P*} \xi_M, p_{P*} \cdot) = \\ &= \omega_P(\xi_{T^*P}, p_{P*} \cdot) = i_{\xi_{T^*P}} \omega_P \circ p_{P*} = \\ &= d\langle J_P, \xi \rangle \circ p_{P*} = p_{P*} d\langle J_P, \xi \rangle = \\ &= d\langle p_P^* J_P, \xi \rangle = d\langle J_M, \xi \rangle. \end{aligned}$$

Notar que la aplicación J_M es K -invariante, donde $K = \text{Ker}(\Omega)$, por construcción. □

Como establecimos anteriormente, el problema que estamos enfrentando consiste en la explotación de las simetrías del problema para conseguir simplificarlo. Un modo de hacerlo es reducir el sistema deshaciéndonos de los grados "redundantes" de libertad del problema. Esto podría hacerse de varios modos. El primero y más obvio sería reducir el sistema presimpléctico (M, Ω, H) usando reducción (pre)simpléctica. Sin embargo, preferiremos reducir el principio variacional de Lagrange-Pontryagin equivalente a (M, Ω, H) , por mantener el formalismo variacional que indicamos al inicio de la memoria.

5.1.3. Reducción de Lie-Dirac de sistemas presimplécticos.

Consideremos un sistema hamiltoniano presimpléctico (M, Ω, H) . Supongamos que la distribución $K = \text{Ker} \Omega$ es regular y el espacio de hojas de la foliación R definida por K es una variedad $M/K = \tilde{M}$ tal que la proyección $p: M \rightarrow M/K$ es una submersión. Supongamos que G es un grupo de Lie actuando a izquierda en M tal que H es G -invariante, y con aplicación momento $J: M \rightarrow \mathfrak{g}^*$, esto es, $\xi_M \lrcorner \Omega = dJ_\xi$ con $J_\xi(x) = \langle J(x), \xi \rangle$, $\forall \xi \in \mathfrak{g}$.

Observemos en primer lugar que la distribución K es G -invariante, es decir $g_* Z \in K$, $\forall Z \in K$. De hecho $\Omega(g_* Z, Y) = (g^* \Omega)(Z, g_*^{-1} Y) = \Omega(Z, g_*^{-1} Y) = 0$. Por tanto la acción de G pasa al cociente M/K . Por otro lado la forma presimpléctica Ω desciende a M/K y define una forma simpléctica $\tilde{\Omega}$ en $\tilde{M} = M/K$. Notemos finalmente que la aplicación momento satisface $\mathcal{L}_Z J_\xi = 0$, $\forall \xi \in \mathfrak{g}$ y $\forall Z \in K$.

De hecho $\langle dJ_\xi, Z \rangle = i_Z i_{\xi_M} \Omega = 0$, $\forall Z \in K$, $\xi \in \mathfrak{g}$. La aplicación J_ξ induce una aplicación \tilde{J}_ξ en \tilde{M} . Es fácil comprobar que $\tilde{J}: \tilde{M} \rightarrow \mathfrak{g}^*$, definida por $\langle \tilde{J}(x), \xi \rangle = \tilde{J}_\xi(x)$ es la aplicación momento de la acción de G en \tilde{M} .

En efecto, si denotamos por $\xi_{\tilde{M}}$ el campo asociado a ξ en \tilde{M} , entonces $p_*\xi_M = \xi_{\tilde{M}}$. Por tanto

$$\langle i_{\xi_{\tilde{M}}}\tilde{\Omega}, \tilde{Y} \rangle = \tilde{\Omega}(p_*\xi_M, \tilde{Y}) = (p^*\tilde{\Omega})(\xi_M, Y) = \Omega(\xi_M, Y) \circ p = \langle dJ_\xi, Y \rangle \circ p = \langle d\tilde{J}_\xi, \tilde{Y} \rangle,$$

donde Y es cualquier campo vectorial en M proyectable tal que $p_*Y = \tilde{Y}$. Hemos probado así que:

Proposición 5.1.3 *Sea un sistema hamiltoniano presimpléctico (M, Ω, H) y G un grupo de Lie actuando en M tal que la acción es hamiltoniana con aplicación momento $J: M \rightarrow \mathfrak{g}^*$ y H es G -invariante. Si el espacio cociente \tilde{M} de la distribución característica de Ω es una variedad diferenciable y la proyección canónica $p: M \rightarrow \tilde{M}$ es una submersión, entonces el sistema hamiltoniano presimpléctico induce un G -sistema hamiltoniano $(\tilde{M}, \tilde{\Omega}, \tilde{H})$ con aplicación momento $\tilde{J}: \tilde{M} \rightarrow \mathfrak{g}^*$, inducida por la aplicación momento J .*

La conclusión de la pequeña proposición anterior implica que podemos proceder a la reducción de la variedad presimpléctica (M, Ω, H) bien reduciendo directamente los niveles de la aplicación momento y obteniendo de esa manera espacios presimplécticos reducidos, ó cocientando directamente la distribución característica de Ω y procediendo a la reducción simpléctica ordinaria a la Marsden-Weinstein del correspondiente espacio de hojas. Así tenemos:

Proposición 5.1.4 *En las mismas condiciones de la proposición anterior. Sea un G -espacio presimpléctico hamiltoniano (M, Ω, H, G, J) . Si $\mu \in \mathfrak{g}^*$ es un valor regular de la aplicación momento J y el grupo de isotropía G_μ de μ bajo la acción coadjunta actúa libre y propiamente en $J^{-1}(\mu)$, entonces el espacio cociente $J^{-1}(\mu)/G_\mu = M_\mu$ hereda una estructura presimpléctica Ω_μ . La distribución característica K_μ de Ω_μ es el cociente respecto a la acción de G_μ de la restricción de la distribución característica de Ω a $J^{-1}(\mu)$. Además el espacio cociente M_μ/K_μ es simplécticamente isomorfo al espacio simpléctico reducido $\tilde{J}^{-1}(\mu)/G_\mu = \tilde{M}_\mu$ obtenido por reducción simpléctica de Marsden-Weinstein de $(\tilde{M}, \tilde{\Omega})$ con respecto a μ .*

$$\begin{array}{ccc} J^{-1}(\mu) & \longrightarrow & \tilde{J}^{-1}(\mu) \\ \pi_\mu \downarrow & & \downarrow \tilde{\pi}_\mu \\ J^{-1}(\mu)/G_\mu = M_\mu & \xrightarrow{p_\mu} & \tilde{J}^{-1}(\mu)/G_\mu = \tilde{M}_\mu \end{array} \qquad \begin{array}{ccc} M & \xrightarrow{p} & \tilde{M} = M/K \\ & \searrow J & \downarrow \tilde{J} \\ & & \mathfrak{g}^* \end{array}$$

El esquema de reducción presimpléctica anterior tiene su análogo en términos de estructuras de Dirac. Dado que $D_\Omega \subset TM \oplus T^*M$, y dado que la acción de G en M se levanta de manera canónica tanto a TM como a T^*M , es inmediato

comprobar que la estructura de Dirac D_Ω es G -invariante, esto es, si $(X, \alpha) \in D_\Omega$, entonces $(g.X, g.\alpha) \in D_\Omega$, donde $g.X$ y $g.\alpha$ denota el levantamiento natural de la acción de G a TM y T^*M respectivamente. La prueba es simple. En efecto, si $(X, \alpha) \in D_\Omega$, entonces $\Omega(X, Y) = \langle \alpha, Y \rangle$, $\forall Y \in TM$ y por tanto $(g^*\Omega)(X, Y) = g^*\langle \alpha, Y \rangle = \langle g^{-1}\alpha, gY \rangle$. Por otro lado, $(g^*\Omega)(X, Y) = \Omega(g_*X, g_*Y) = \Omega(g.X, g.Y)$ y por tanto $(g.X, g.Y) \in D_\Omega$.

Si denotamos por $[D_\Omega]_G$ el cociente del subfibrado D_Ω por la acción de G , es decir, $[D_\Omega]_G = D/G$, que será un fibrado sobre M/G . Tendremos que $[D_\Omega]_G \subset TM/G \oplus T^*M/G$ será una estructura de Dirac fibrada, llamada reducción de Lie-Dirac de D_Ω . Dado que $TM/G \cong T(M/G) \oplus \tilde{\mathfrak{g}}$ y $T^*M/G \cong T^*(M/G) \oplus \tilde{\mathfrak{g}}^*$, tendremos que $[D_\Omega]_G \subset T(M/G) \oplus T^*(M/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*$.

Por otro lado si denotamos como anteriormente $M/K = \tilde{M}$, entonces podemos considerar la estructura de Dirac sobre \tilde{M} asociada a $\tilde{\Omega}$, esto es $D_{\tilde{\Omega}} \subset T\tilde{M} \oplus T^*\tilde{M}$. La aplicación $p: M \rightarrow \tilde{M}$ induce una aplicación $Tp: TM \rightarrow T\tilde{M}$, entonces D_Ω es el pull-back a lo largo de Tp de $D_{\tilde{\Omega}}$. En efecto utilizando la notación en [Yo09], si denotamos por $BTp(D_{\tilde{\Omega}}) = \{(X, p^*\tilde{\alpha}) \in TM \oplus T^*M \mid (p_*X, \tilde{\alpha}) \in D_{\tilde{\Omega}}\}$ se ve inmediatamente que $BTp(D_{\tilde{\Omega}}) = D_\Omega$. También denotaremos la relación indicando que $D_\Omega/K := D_{\tilde{\Omega}}$ ya que $K = \text{Ker}Tp$. También diremos que la estructura de Dirac presimpléctica D_Ω "proyecta" sobre la estructura de Dirac presimpléctica $D_{\tilde{\Omega}}$.

Finalmente observamos que la reducción de Lie-Dirac de la estructura $D_{\tilde{\Omega}}$ proporciona otra estructura fibrada de Dirac $[D_{\tilde{\Omega}}]_G$ sobre $T(\tilde{M}/G)$, esto es:

$$[D_{\tilde{\Omega}}]_G := D_{\tilde{\Omega}}/G \subset T(\tilde{M}/G) \oplus T^*(\tilde{M}/G) \oplus \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}^*.$$

La aplicación $p: M \rightarrow \tilde{M}$ induce una aplicación $p_G: M/G \rightarrow \tilde{M}/G$. Se obtiene fácilmente que $Bp_G[D_{\tilde{\Omega}}]_G = [D_\Omega]_G$, ó dicho en otras palabras, la reducción de Lie-Dirac de la estructura de Dirac definida por la forma simpléctica Ω es el pull-back de la reducción de Lie-Dirac de la estructura de Dirac definida por la forma simpléctica $\tilde{\Omega}$, ó todavía de otra forma: la estructura de Dirac fibrada $[D_\Omega]_G$ "proyecta" a la estructura de Dirac fibrada $[D_{\tilde{\Omega}}]_G$.

Podemos aplicar todas estas ideas al sistema presimpléctico definido por un problema de control óptimo $M = T^*P \times_p C$ como veremos en la siguiente sección.

5.1.4. Reducción de la estructura de Dirac D_Ω de un problema de Control Óptimo

Alternativamente, podemos proceder a la reducción de la estructura de Dirac definida por la forma presimpléctica Ω . Dado que $D_\Omega \subset TM \oplus T^*M$, podemos por tanto proceder a la reducción de D_Ω considerando el espacio cociente D_Ω/G que será un subfibrado de $(TM \oplus T^*M)/G = TM/G \oplus T^*M/G$. Esta reducción propuesta por Marsden y Yoshimura la denominaremos reducción de Lie-Dirac de D_Ω y la denotaremos por $[D_\Omega]_G$, es decir $[D_\Omega]_G := D_\Omega/G$.

Para proceder al cálculo de D_Ω/G , supondremos que el grupo G actúa libre y propiamente en P , esto es, que $P \rightarrow P/G$ es un fibrado principal con grupo estructural G . Supondremos también por simplicidad y sin merma de generalidad, que el fibrado de control $C \xrightarrow{\pi} P$ es el pullback a P de un fibrado asociado a $P \xrightarrow{\pi_1} P/G$ con fibra standard U , esto es tenemos una acción de G en el espacio lineal (ó afín) U y consideramos el fibrado sobre P/G definido por $P \times U/G = P \times_G U$. Denotaremos tal fibrado por \tilde{U} como en ocasiones anteriores. Así se tiene que $C = \pi_1^* \tilde{U}$ aunque abusaremos de la notación denotando por \tilde{U} los diversos pull-backs de dicho fibrado. Un sencillo cálculo nos muestra que $TM/G \cong T(T^*P \times \tilde{U})/G \cong TT^*(P/G) \oplus (\tilde{\mathfrak{g}}^* \oplus \tilde{U}) \times (\tilde{U} \oplus \tilde{V})$ y $T^*M/G \cong T^*(T^*P \times \tilde{U})/G \cong T^*T^*(P/G) \oplus (\tilde{\mathfrak{g}}^* \oplus \tilde{U}) \times (\tilde{U}^* \oplus \tilde{V}^*)$.

Por tanto

$$[D_\Omega]_G \subset TT^*(P/G) \oplus T^*T^*(P/G) \oplus (\tilde{\mathfrak{g}}^* \oplus \tilde{U}) \times (\tilde{U} \oplus \tilde{U}^* \oplus \tilde{V} \oplus \tilde{V}^*).$$

Es fácil probar que la proyección de $[D_\Omega]_G$ a lo largo del fibrado $\tilde{U} \oplus \tilde{U}^*$, proyecta sobre la reducción cotangente de la estructura de Dirac canónica sobre T^*P dado que los factores \tilde{U} y \tilde{U}^* no aparecen en la definición de Ω . Denotaremos tal proyección por $\Pi: [D_\Omega]_G \longrightarrow [D_{T^*P}]_G$, con

$$[D_{T^*P}]_G \subset TT^*(P/G) \oplus T^*T^*(P/G) \oplus (\tilde{\mathfrak{g}}^* \times (\tilde{V} \oplus \tilde{V}^*)),$$

la estructura de Dirac cotangente reducida. Nótese finalmente que:

$$\begin{aligned} [D_{T^*P}]_G &= \{(z, p_z, \mu; v_z, v_{p_z}; \pi_z, \pi_{p_z}; \xi, v_\mu; \alpha_\xi, \alpha_{v_\mu}) \mid \\ &|\langle \pi_z, \delta z \rangle + \langle \delta p_z, \pi_{p_z} \rangle + \langle \alpha_\xi, \xi \rangle + \langle \alpha_{v_\mu}, v_\mu \rangle = [\Omega]_G(z, p_z, \mu)(v_z, v_{p_z}, \xi, v_\mu)(\delta z, \delta p_z, \xi, \delta \mu), \\ &\forall (\delta z, \delta p_z, \xi, \delta \mu) \in T_{(z, p_z)} T^*(P/G) \times \tilde{V}\}. \end{aligned}$$

5.1.5. Descripción de los fibrados reducidos

Para hacer esto necesitaremos describir varios fibrados reducidos que destacarán a lo largo del proceso. Notar que el principio variacional de Lagrange-Pontryagin

$$\mathbb{S}(x, p, u) = \int_0^T [L(x, u) + \langle p, \dot{x} - \Gamma(x, u) \rangle] dt$$

es definido sobre las curvas $\gamma(t) = (x(t), p(t), u(t))$ sobre el fibrado de Pontryagin M , por tanto el problema reducido estará definido sobre el espacio de curvas sobre el espacio cociente M/G . Primero, describiremos este espacio de un modo conveniente. Asumiremos por simplicidad que G actúa libre y propiamente sobre P , es decir $P \xrightarrow{\pi} P/G$ es un G -fibrado principal sobre la variedad P/G .

La mayoría de estas reducciones ya han sido discutidas en la literatura, por ejemplo:

- i. El fibrado cotangente reducido T^*P/G ó el espacio de Weinstein [We78]. El espacio T^*P/G tiene una estructura de fibrado canónica sobre $T^*(P/G)$. La proyección $\tilde{\pi}_P : T^*P/G \rightarrow T^*(P/G)$ es definida como:

$$\tilde{\pi}_P([x, p]) = [x] = G \cdot x.$$

Si denotamos por $\tilde{\mathfrak{g}} \rightarrow P/G$ el fibrado asociado a $P \rightarrow P/G$ con fibra \mathfrak{g} y acción de G la acción adjunta, es decir $\tilde{\mathfrak{g}} = P \times_G \mathfrak{g}$, y $g(x, \xi) = (gx, Ad_g \xi)$, $\forall (x, \xi) \in P \times \mathfrak{g}$. El fibrado adjunto P es también denotado como:

$$AdP \equiv \tilde{\mathfrak{g}} = P \times_G \mathfrak{g}.$$

Similarmente definiremos el fibrado coadjunto $\tilde{\mathfrak{g}}^* \equiv Ad^*P = P \times_G \mathfrak{g}^*$, que es el fibrado asociado a $P \rightarrow P/G$ con respecto a la acción coadjunta de G sobre \mathfrak{g}^* . Por tanto $g(x, \mu) = (gx, Ad_{g^{-1}}^* \mu)$, $(g, \mu) \in P \times \mathfrak{g}^*$, $g \in G$. El fibrado $T^*P/G \rightarrow T^*(P/G)$ es el pull-back del fibrado coadjunto $\tilde{\mathfrak{g}}^*$ a $T^*(P/G)$ a lo largo de la proyección $\pi_{P/G} : T^*(P/G) \rightarrow P/G$. Con algún ligero abuso de notación lo denotaremos por $T^*(P/G) \oplus \tilde{\mathfrak{g}}^*$.

Los elementos sobre T^*P/G se denotarán por $[x, p] = (z, p_z, \mu)$ donde $(z, p_z) \in T^*(P/G)$ y z^a , $a = 1, \dots, m$ son coordenadas locales sobre P/G , y $\mu \in \mathfrak{g}^*$.

- ii. El fibrado $TP/G \cong T(P/G) \oplus \tilde{\mathfrak{g}}$ puede ser identificado con el pull-back a $T(P/G)$ del fibrado adjunto $\tilde{\mathfrak{g}} \rightarrow P/G$, y otra vez con un ligero abuso de notación lo llamaremos $T(P/G) \oplus \tilde{\mathfrak{g}}$. Sus elementos serán de la forma $[x, v_x] = (z, v_z, \zeta)$ con $(z, v_z) \in T(P/G)$ y $\zeta \in \mathfrak{g}$.

Si escogemos una conexión principal A sobre $P \rightarrow P/G$, entonces tenemos una identificación explícita de $T(P/G) \oplus \tilde{\mathfrak{g}}$ y TP/G . La aplicación dada por:

$$\alpha : TP/G \rightarrow T(P/G) \oplus \tilde{\mathfrak{g}}, \quad \alpha([x, v_x]) = (T\rho_x(v_x), A_x(v_x))$$

donde $T\rho_x : T_xP \rightarrow T_{G \cdot x}(P/G)$ es la aplicación tangente de la proyección $\rho : P \rightarrow P/G$. El dual de esta aplicación nos permite también hacer la identificación $T^*P/G \cong T^*(P/G) \oplus \tilde{\mathfrak{g}}^*$.

- iii. El fibrado de control reducido C/G . El fibrado $\tilde{\pi} : C/G \rightarrow P/G$ estará definido por el push-forward del fibrado C a lo largo de la proyección ρ . Notar que la aplicación $\tilde{\pi} : C/G \rightarrow P/G$ dada por $\tilde{\pi}([x, u]) = [x]$ está bien definida $\tilde{\pi}^{-1}[x] \cong \pi^{-1}(x)$ gracias a las aplicaciones $\beta_x : \pi^{-1}(x) \rightarrow \tilde{\pi}^{-1}[x]$, $\beta_x(u) = [x, u]$ y $\gamma_x : \tilde{\pi}^{-1}[x] \rightarrow \pi^{-1}(x)$, $\gamma_x([x, u]) = (x, u)$. Por lo tanto, $\beta_x \circ \gamma_x = id_{\pi^{-1}(x)}$. Los elementos sobre $\tilde{C} = C/G$ serán de la forma $[x, u] = (z, u)$ con $z \in P/G$.
- iv. El fibrado reducido de Pontryagin $M/G = (T^*P \times_P C)/G$. El espacio cociente M/G es entonces un fibrado sobre P/G , y puede ser identificado con el fibrado $T^*(P/G) \oplus \tilde{\mathfrak{g}}^* \oplus \tilde{C}$ sobre P/G . Entonces los elementos de fibrado reducido de Pontryagin serán de la forma: $[x, p, u] = ((z, p_z), \mu, u)$ donde $(z, p_z) \in T^*(P/G)$, $\mu \in \mathfrak{g}^*$ y $u \in \tilde{C}$.

5.1.6. El principio variacional reducido de Lagrange-Pontryagin

Ahora podemos escribir la forma explícita del principio variacional reducido de Lagrange-Pontryagin. La invarianza del lagrangiano L nos permite definir la función $l : \tilde{C} \rightarrow \mathbb{R}$, esto es, $L(x, u) = l([x, u]) = l(z, u)$, $(z, u) \in \tilde{C}$.

El campo vectorial dinámico Γ toma la forma $\Gamma(x, u) = \tilde{\Gamma}(x, u) \oplus \tilde{\gamma}(z, u)$ donde $\tilde{\Gamma}(z, u) \in T(P/G)$ y $\tilde{\gamma}(z, u) \in \mathfrak{g}$. De hecho, $T_x P = T_z(P/G) \oplus TG$, por lo tanto $\Gamma(x, u) \in T_x P$ se descompone como un elemento de $T_z(P/G)$ y otro elemento en \mathfrak{g} . Notar que si $\gamma(x, u) = A_x(\Gamma(x, u))$ donde A es una conexión principal en P , entonces $\gamma(gx, gu) = A_{gx}(\Gamma(gx, gu)) = A_{gx}(g_*\Gamma(x, u)) = (L_g^* A)_x(\Gamma(x, u)) = Ad_g(A_x(\Gamma(x, u))) = Ad_g \gamma(x, u)$, entonces $\tilde{\gamma}(z, u) = [\gamma(x, u)]$. Podemos describir entonces la densidad $\mathbb{L}(x, u, p) = L(x, u) + \langle p, \dot{x} - \Gamma(x, u) \rangle$ del principio variacional de Lagrange-Pontryagin como:

$$\mathbb{L}(x, u, p) = l(z, u) + \langle p_z, \dot{z} - \tilde{\Gamma}(z, u) \rangle + \langle \mu, \zeta - \tilde{\gamma}(z, u) \rangle$$

donde otra vez hemos usado $T^*P = T^*(P/G) \oplus T^*G$. Luego $(x, p) = (z, p_z, g, \mu)$ y $\dot{x} = (\dot{z}, \zeta) \in T(P/G) \oplus \tilde{\mathfrak{g}}$.

Observaciones

En lo que sigue vamos a establecer muchas notaciones e identidades para TP y T^*P . Hemos demostrado que $TP/G \cong T(P/G) \oplus \tilde{\mathfrak{g}}$ como un fibrado sobre P/G . Además como G actúa libremente sobre P , también actúa libremente sobre TP y $TP \rightarrow TP/G$ es un fibrado principal con fibra G . Además también es un fibrado principal con fibra TG . Podemos verlo en dos pasos:

1. Consideramos la aplicación $T_\rho : TP \rightarrow T(P/G)$, la aplicación tangente de la proyección canónica $\rho : P \rightarrow P/G$. Mapea $(x, v_x) \in T_x P$ en (z, v_x) donde $v_x = v_z + v_g$.

$$\begin{array}{ccc} TP & \xrightarrow{T_\rho} & T(P/G) \\ \tau_P \downarrow & & \downarrow \tau_{P/G} \\ P & \xrightarrow{\rho} & P/G \end{array}$$

2. Consideramos el grupo de Lie TG . Podemos usar traslaciones a izquierdas para escribirlo como $G \times \mathfrak{g}$. Entonces la identificación natural es $(g, v_g) \rightsquigarrow (g, \xi)$ donde $\xi = L_{g*}^{-1} v_g$, $(g, v_g) \in TG$.

Recordemos que la ley de composición natural de TG es

$$(g, v_g) \cdot (h, v_h) = (g \cdot h, v_g \cdot v_h) \text{ donde } v_g \cdot v_h = \frac{d}{dt} g(t) \cdot h(t)|_{t=0} \text{ y respectivamente}$$

$$v_g = \frac{d}{dt} g(t)|_{t=0}, \quad v_h = \frac{d}{dt} h(t)|_{t=0}.$$

Ahora

$$\begin{aligned} L_{g.h^*}^{-1}v_g.v_h &= L_{g.h^*}^{-1}\frac{d}{dt}g(t)h(t)|_{t=0} = \frac{d}{dt}(g.h)^{-1}g(t)h(t)|_{t=0} = \\ &= \frac{d}{dt}h^{-1}g^{-1}g(t)h(t)|_{t=0} = \frac{d}{dt}(h^{-1}g^{-1}g(t)h)(h^{-1}h(t))|_{t=0} = Ad_h\xi + \zeta, \end{aligned}$$

con $\xi = L_{g^*}^{-1}v_g$ y $\zeta = L_{h^*}^{-1}v_h$.

Luego $(g, \xi).(h, \zeta) = (g.h, Ad_h\xi + \zeta)$, $\forall (g, \xi), (h, \zeta) \in G \times \mathfrak{g}$. Hay una acción natural de TG en TP dada como:

$$(g, v_g).(x, v_x) = (g.x, \frac{d}{dt}g(t)x(t)|_{t=0}) \in T_{gx}P.$$

Notar que $\frac{d}{dt}g(t)x(t)|_{t=0} = \frac{d}{dt}g(t)g^{-1}gx(t)|_{t=0} = (Ad_g\xi)_P(gx) + g_*v_x$ ó en otras palabras:

$$(g, \xi).(x, v_x) = (g.x, g_*v_x + (Ad_g\xi)_P(g.x))$$

Por tanto el principio variacional reducido de Lagrange-Pontryagin está definido sobre un espacio de curvas en M/G , sin embargo estas curvas están directamente relacionadas con curvas en M , luego las variaciones de tales curvas no son completamente independientes.

Esto es causado por la anterior discusión donde obtuvimos que el cociente $M/G = T^*P \times_G C = T^*(P/G) \oplus \tilde{\mathfrak{g}}^* \oplus \tilde{C}$ con elementos de la forma (z, p_z, μ, u) y variaciones $\delta z, \delta p_z, \delta \mu, \delta u$. Notar que $\delta z(0) = 0 = \delta z(T)$, ya que las curvas $x(t)$ proyectan a curvas $z(t) \in P/G$ y $x(0) = x_0, x(T) = x_T$. Luego $z(0) = \rho(x_0)$ y $z(T) = \rho(x_T)$.

Además, dada una conexión principal A en P , podemos siempre escribir $x(t) = g(t)x^h(t)$ donde $x^h(t)$ es la única curva horizontal sobre P comenzando en x_0 y proyectándose sobre $z(t) = \rho(x(t))$. Por tanto $x_0 = g(0)x^h(0) = g(0)x_0$, entonces $g(0) = e$ y similarmente $g(T) = e$. Cuando escribimos $\dot{x}(t) = \dot{g}(t)x^h(t) + g(t)\dot{x}^h(t) = \dot{g}(t)g^{-1}(t)g(t)x^h(t) + g(t)\dot{x}^h(t) = \zeta(t)_P(x(t)) + g(t)\dot{x}^h(t)$.

Por tanto usando la identificación $T_xP \cong T_z(P/G) \oplus \mathfrak{g}$, proporcionada por $\dot{x} = (\dot{z}, \zeta)$, donde $\dot{x} = \dot{x}^h + \dot{x}^v$, $\rho_*\dot{x}^h = \dot{z}$, $\dot{x}^v = \zeta_P(x)$, $\zeta \in \mathfrak{g}$ obtenemos que la curva $\zeta(t) \in \mathfrak{g}$ corresponde a la componente vertical de la curva $\dot{x}(t)$. Además $\zeta(t) = \dot{g}(t)g^{-1}(t)$, luego:

$$\begin{aligned} \delta\zeta &= \delta\dot{g}g^{-1} + \dot{g}\delta g^{-1} = \delta\dot{g}g^{-1} - \dot{g}g^{-1}\delta g g^{-1} = \widehat{\delta g g^{-1}} - \delta g \widehat{g^{-1}} - \zeta\delta g g^{-1} = \\ &= \widehat{\delta g g^{-1}} + \delta g g^{-1}\dot{g}g^{-1} - \zeta\delta g g^{-1} = \dot{\eta} + [\eta, \zeta] \end{aligned}$$

y las variaciones de la curva $\xi(t)$ tendrán la forma:

$$\delta\zeta = \dot{\eta} + [\eta, \zeta] \in \mathfrak{g}$$

para un arbitrario $\eta(t) \in \mathfrak{g}$

5.1.7. El principio del Maximo de Pontryagin Reducido

Las ecuaciones de Euler-Lagrange definidas por el principio variacional reducido de Lagrange-Pontryagin nos proporcionara las condiciones diferenciales del PMP para el sistema reducido. Podemos calcularlas entonces de:

$$\mathcal{S}_G = \int_0^T [l(z, u) + \langle p_z, \dot{z} - \tilde{\Gamma}(z, u) \rangle + \langle \mu, \xi - \gamma(z, u) \rangle] dt$$

Por tanto:

$$\begin{aligned} \delta \mathcal{S}_G &= \int_0^T \left[\frac{\partial l}{\partial z} \delta z + \frac{\partial l}{\partial u} \delta u + \langle \delta p_z, \dot{z} - \tilde{\Gamma}(z, u) \rangle + \langle \delta \mu, \xi - \gamma(z, u) \rangle + \right. \\ &\quad \left. + \langle p_z, \frac{d}{dt} \delta z - \frac{\partial \tilde{\Gamma}}{\partial z} \delta z - \frac{\partial \tilde{\Gamma}}{\partial u} \delta u \rangle + \langle \mu, \delta \xi - \frac{\partial \gamma}{\partial z} \delta z - \frac{\partial \gamma}{\partial u} \delta u \rangle \right] dt = \\ &= \int_0^T \left(\frac{\partial l}{\partial z} - \dot{p}_z - \frac{\partial \tilde{\Gamma}}{\partial z} p_z - \mu \frac{\partial \gamma}{\partial z} \right) \delta z dt + \int_0^T \left(\frac{\partial l}{\partial u} - p_z \frac{\partial \tilde{\Gamma}}{\partial u} - \mu \frac{\partial \gamma}{\partial u} \right) \delta u dt + \\ &+ \int_0^T \langle \delta p_z, \dot{z} - \tilde{\Gamma}(z, u) \rangle dt + \int_0^T \langle \delta \mu, \xi - \gamma(z, u) \rangle dt + p_z \delta z \Big|_0^T + \int_0^T \langle \mu, \dot{\eta} + [\eta, \xi] \rangle dt \end{aligned}$$

El ultimo termino da:

$$\begin{aligned} \int_0^T \langle \mu, \dot{\eta} + [\eta, \xi] \rangle dt &= \int_0^T \langle \dot{\mu}, \eta \rangle + \langle \mu, [\eta, \xi] \rangle dt + \langle \mu, \eta \rangle \Big|_0^T = \\ &= \int_0^T [\langle \dot{\mu}, \eta \rangle + \langle ad_{\xi}^* \mu, \eta \rangle] dt + \langle \mu, \eta \rangle \Big|_0^T = \\ &= \int_0^T \langle \dot{\mu} + ad_{\xi}^* \mu, \eta \rangle dt + \langle \mu, \eta \rangle \Big|_0^T. \end{aligned}$$

Ahora, fijando las condiciones frontera para $\eta(t)$:

$$\begin{aligned} \eta(0) &= \eta(T) = 0, \\ \dot{\eta}(0) &= \xi(0), \quad \dot{\eta}(T) = \xi(T). \end{aligned}$$

Concluimos:

$$\left\{ \begin{array}{l} \dot{z} = \tilde{\Gamma}(z, u), \\ \dot{p}_z = -\frac{\partial \tilde{H}}{\partial z} = \frac{\partial l}{\partial z} - \frac{\partial \tilde{\Gamma}}{\partial z} p_z - \mu \frac{\partial \gamma}{\partial z}, \\ \dot{\mu} = -ad_{\xi}^* \mu, \\ \xi = \gamma(z, u), \\ \frac{\partial \tilde{H}}{\partial u} = \frac{\partial l}{\partial u} - \frac{\partial \tilde{\Gamma}}{\partial u} p_z - \mu \frac{\partial \gamma}{\partial u} = 0. \end{array} \right. \quad (5.1)$$

donde el Hamiltoniano reducido de Pontryagin $\tilde{H} : M/G \rightarrow \mathbb{R}$ es definido como:

$$\tilde{H}(z, p_z, u, \mu) = \langle p_z, \tilde{\Gamma}(z, u) \rangle + \langle \mu, \gamma(z, u) \rangle - l(z, u) \quad (5.2)$$

Por tanto las ecuaciones reducidas anteriores (5.1) pueden ser escritas también como ecuaciones de Hamilton-Poisson:

$$\left\{ \begin{array}{l} \dot{z} = \frac{\partial \tilde{H}}{\partial p_z}, \\ \dot{p}_z = -\frac{\partial \tilde{H}}{\partial z}, \\ \dot{\mu} = -ad_{\xi}^* \mu, \\ \xi = \frac{\partial \tilde{H}}{\partial \mu}, \\ \frac{\partial \tilde{H}}{\partial u} = 0. \end{array} \right. \quad (5.3)$$

ó pueden ser escritas como ecuaciones de Lagrange-Poincaré como en [Yo09].

5.2. Reducción de Lie-Dirac y el Principio del Máximo de Pontryagin reducido

En la sección §5.1.7 obtuvimos las ecuaciones del Principio del Máximo de Pontryagin reducido para un sistema de Control Óptimo con simetría, a base de calcular la diferencial de la reducción del principio de Hamilton-Pontryagin. Por

otro lado en las secciones anteriores, mostramos cómo un problema de control óptimo podía describirse de dos maneras distintas a través de estructuras de Dirac: como un problema presimpléctico ó como un problema implícito con estructura de Dirac $D_{T^*P} = D_P$. Vimos también que ambas estructuras de Dirac se hallan relacionadas entre sí, siendo la segunda una proyección de la primera. Finalmente también vimos que en presencia de simetría podemos reducir ambas estructuras de Dirac y las correspondientes estructuras de Dirac reducidas se hallan de nuevo relacionadas por proyección. Así $[D_\Omega]_G$ y $[D_P]_G$ son tales que $[D_P]_G$ es la proyección de $[D_\Omega]_G$ con respecto a $\tilde{V} \oplus \tilde{V}^*$. La estructura de Dirac reducida $[D_P]_G$ no es más que la reducción de Lie-Dirac de la estructura de Dirac cotangente en T^*P . Por tanto si denotamos por H_G el Hamiltoniano de Pontryagin inducido en $M/G = T^*(P/G) \oplus \tilde{\mathfrak{g}}^* \oplus \tilde{U}$ el espacio cociente del fibrado de Pontryagin, tendremos que la subvariedad $M_1 = \{\frac{\partial H}{\partial u^a} = 0\}$ proyecta a la subvariedad $[M_1]_G = \{\frac{\partial H_G}{\partial u^a} = 0\}$ y $M_1/G = [M_1]_G$. Por lo tanto definiremos la ecuación implícita de Pontryagin a la dinámica definida por la ecuación:

$$\left([X]_G(z, p_z, \mu, u), [dH_G] \Big|_{T[M_1]_G} (z, p_z, \mu, u) \right) \in [D_P]_G(z, p_z; \mu), \quad (5.4)$$

$$(z, p_z) \in T^*(P/G), \quad \mu \in \mathfrak{g}^*.$$

Las ecuaciones así obtenidas coinciden con las ecuaciones reducidas del Principio del Máximo de Pontryagin obtenidas anteriormente.

Teorema 5.2.1 *Dado un sistema de control óptimo regular $(L, \Gamma, C \rightarrow P)$ con grupo de simetría G ; las ecuaciones del Principio del Máximo de Pontryagin reducido obtenidas como reducción del principio variacional de Hamilton-Pontryagin, coinciden con las ecuaciones de la reducción de Lie-Dirac de las ecuaciones de la descripción de Dirac implícita del problema de control óptimo, ecuaciones 5.4.*

Demostración. La demostración del teorema se apoya en la descripción de la reducción de Lie-Dirac de la estructura de Dirac cotangente D_P obtenida en [Yo09]. Utilizaremos dicho resultado sin probarlo de nuevo. Así obtenemos que la estructura $[D_P]_G$ para cada $(z, p_z; \mu) \in T^*(P/G) \oplus \mathfrak{g}^*$ tiene la forma (Prop. 4 [Yo09]):

$$[D_P]_G(z, p_z; \mu) = \{(v_z, v_{p_z}; \xi, v_\mu), (\pi_z, \pi_{p_z}; \alpha_\xi, \alpha_{v_\mu}) \in T_{(z, p_z)} T^*(P/G) \times \tilde{V} \oplus \oplus T_{(z, p_z)}^* T^*(P/G) \times \tilde{V}^* \mid v_z = \pi_{p_z}, \quad v_{p_z} + \pi_z = -F_A(v_z, \cdot), \quad \xi = \alpha_{v_\mu}, \quad v_\mu + \alpha_\xi = ad_\xi^* \mu\}.$$

Entonces si denotamos $[X]_G$ como el campo vectorial reducido tal que para cada curva $(z(t), p_z(t); \mu(t), u(t)) \in T^*(P/G) \oplus \tilde{\mathfrak{g}}^* \oplus \tilde{U}$, toma el valor

$$[X]_G(z(t), p_z(t); \mu(t), u(t)) = (z(t), p_z(t), u(t); \frac{dz(t)}{dt}, \frac{Dp_z(t)}{Dt}, \mu(t), \xi(t), \frac{D\mu(t)}{Dt})$$

y teniendo en cuenta que $[dH]_G = dH_G$ restringido a $T[M_1]_G = T(M_1/G)$ tiene la forma:

$$[dH]_G \Big|_{T[M_1]_G} = (z, p_z; \mu, u; -\frac{\partial \tilde{H}}{\partial z}, \frac{\partial \tilde{H}}{\partial p_z}, \frac{\partial \tilde{H}}{\partial \mu}, 0)$$

y finalmente:

$$\left\{ \begin{array}{l} \frac{dz}{dt} = \frac{\partial \tilde{H}}{\partial p_z}, \\ \frac{Dp_z}{Dt} = -\frac{\partial \tilde{H}}{\partial z} - F_A\left(\frac{dz}{dt}, \cdot\right), \\ \frac{D\mu}{Dt} = \text{ad}_\xi^* \mu, \\ \xi = \frac{\partial \tilde{H}}{\partial \mu}. \end{array} \right. \quad (5.5)$$

junto con $\frac{\partial \tilde{H}}{\partial u} = 0$.

□

Capítulo 6

Aplicaciones y ejemplos.

6.1. Las ecuaciones de Euler para el sólido rígido

Empezaremos discutiendo el caso más sencillo, como son las ecuaciones de Euler para el sólido rígido, como un problema en control óptimo, ejemplo que sirvió de guía para la discusión en [B100].

6.1.1. Ecuaciones para el grupo $SO(3)$

Sea $G = SO(3)$ el grupo de rotaciones. Como fibrado principal Q consideraremos el grupo $SO(3)$ actuando por la izquierda sobre sí mismo. Por otra parte, consideraremos como espacio auxiliar P de nuevo el grupo de Lie $SO(3)$, pero esta vez, actuando sobre sí mismo por la derecha. El espacio de control para el problema de control óptimo será el fibrado tangente $TQ = TSO(3) \cong_R SO(3) \times \mathfrak{so}(3)$ donde el subíndice R indica que hemos utilizado traslaciones por la derecha sobre $SO(3)$ para la identificación. La densidad Lagrangiana es la energía cinética $L(q, \omega) = 1/2 \langle \omega, J(\omega) \rangle$, $(q, \omega) \in SO(3) \times \mathfrak{so}(3)$ definida por una métrica J invariante por la derecha en $SO(3)$. Consideraremos por simplicidad la conexión B sobre Q como la 1-forma canónica invariante por la derecha de Maurer-Cartan (en la representación de $TSO(3)$ anterior, B es la matriz identidad). Por tanto, calculando el campo de vectores $B(\omega)$ sobre P en el punto x , obtenemos $(B(\omega))_P(x) = d/dt(\exp(tB(\omega))x) |_{t=0} = \omega x$ y la ecuación de control (4.32) será:

$$\dot{x} = (B(\omega))_P(x) = \omega x.$$

Además el Hamiltoniano de Pontryagin será de la siguiente forma:

$$H(q, \omega; x, p) = \langle p, \omega x \rangle - \frac{1}{2} \langle \omega, J(\omega) \rangle,$$

donde $(q, \omega) \in TQ = TSO(3)$ y $(x, p) \in T^*P \cong TP$.

Las ecuaciones de Hamilton serán entonces

$$\dot{x} = \omega x, \quad \dot{p} = \omega p,$$

junto con la condición de ligadura siguiente:

$$px^T - xp^T - J(\omega) = 0.$$

6.1.2. Ecuaciones para el grupo $SU(2)$

Ahora, consideremos como en el ejemplo anterior $G = Q = P = SU(2, \mathbb{C})$ donde hemos reemplazado el grupo ortogonal $SO(3)$ por su cubrimiento universal, el grupo especial unitario $SU(2, \mathbb{C})$. El grupo de Lie

$$SU(2, \mathbb{C}) = \left\{ g = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C}, \quad a\bar{a} + b\bar{b} = 1 \right\}$$

tiene como álgebra de Lie

$$\mathfrak{su}(2, \mathbb{C}) = \{ \xi_+ e_+ + \xi_- e_- + \xi_0 e_0 \mid \xi_+, \xi_-, \xi_0 \in \mathbb{C} \},$$

con

$$e_+ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad e_- = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad e_0 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

Identificaremos TQ con $SU(2, \mathbb{C}) \times \mathfrak{su}(2, \mathbb{C})$ por traslaciones por la derecha, es decir, $(g, \xi) \in SU(2, \mathbb{C}) \times \mathfrak{su}(2, \mathbb{C})$, corresponde al elemento $(g, \dot{g}) \in TSU(2, \mathbb{C})$, con

$$\dot{g} = \xi g.$$

El campo de vectores $\xi_P(x)$ toma la forma $\xi_P(x) = \xi x$. Y por tanto la ecuación de control se convierte en

$$\dot{x} = \xi x$$

El funcional objetivo es

$$S = \frac{1}{4} \int_0^T \langle \xi, J(\xi) \rangle dt, \quad (6.1)$$

donde hemos usado la forma de Killing $\langle \xi, \zeta \rangle = 4\text{Tr}(\xi\zeta)$. Las ecuaciones de Euler-Lagrange para el Lagrangiano

$$L(g, \dot{g}) = \frac{1}{2} \langle \dot{g}g^{-1}, J(\dot{g}g^{-1}) \rangle \quad (6.2)$$

están dadas por:

$$\dot{g} = \xi g, \quad \dot{J}(\xi) = [\xi, J(\xi)]. \quad (6.3)$$

Las ecuaciones del movimiento dadas por el principio del máximo de Pontryagin son:

$$\dot{x} = \xi x, \quad \dot{p} = \xi p,$$

y la relación entre ambos conjuntos de ecuaciones está dada por:

$$px^* - xp^* - J(\xi) = 0,$$

y tomando derivadas con respecto a t obtenemos

$$\dot{J}(\xi) = [\xi, J(\xi)].$$

6.2. Ecuaciones de Ricatti

6.2.1. Ecuaciones para el grupo $SL(2, \mathbb{R})$

Consideremos ahora el grupo $G = SL(2, \mathbb{R})$ y como en el ejemplo del sólido rígido, tomaremos como espacio de control Q el mismo grupo $SL(2, \mathbb{R})$. Para mantener las mismas convenciones que hemos utilizado a lo largo de esta tesis, consideraremos que el grupo G actúa por la izquierda sobre Q por multiplicación por la izquierda. El grupo de Lie

$$SL(2, \mathbb{R}) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R}, \quad ad - bc = 1 \right\}$$

tiene álgebra de Lie

$$\mathfrak{sl}(2, \mathbb{R}) = \{ \xi_+ e_+ + \xi_- e_- + \xi_0 e_0 \mid \xi_+, \xi_-, \xi_0 \in \mathbb{R} \},$$

con

$$e_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

y las siguientes relaciones de conmutatividad distintas de cero,

$$[e_+, e_-] = e_0, \quad [e_+, e_0] = -2e_+, \quad [e_-, e_0] = 2e_-.$$

Identificaremos TQ con $SL(2, \mathbb{R}) \times \mathfrak{sl}(2, \mathbb{R})$ por traslaciones por la izquierda, es decir, $(g, \xi) \in SL(2, \mathbb{R}) \times \mathfrak{sl}(2, \mathbb{R})$, corresponde al elemento $(g, \dot{g}) \in TSL(2, \mathbb{R})$, con

$$\dot{g} = g\xi.$$

Ahora consideremos el espacio de estados $P = \mathbb{R}$ y $SL(2, \mathbb{R})$ actúa sobre él por la transformación de Moebius, es decir,

$$x \cdot g^{-1} = \frac{dx - b}{a - cx}, \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{R}).$$

Si A es una 1-forma de conexión diagonal, es decir, $A = \text{diag}(A_+, A_-, A_0)$ en la base e_+, e_-, e_0 , entonces el campo de vectores $\xi_P(x) = A_g(\dot{g})_P(x)$ toma la forma:

$$\xi_P(x) = (\xi_+ A_+ + 2\xi_0 A_0 x + \xi_- A_- x^2) \frac{\partial}{\partial x}.$$

Por tanto la ecuación de control se convierte en la ecuación de Ricatti:

$$\dot{x} = \alpha + \beta x + \gamma x^2, \quad (6.4)$$

con $\alpha = \xi_+ A_+$, $\beta = 2\xi_0 A_0$, $\gamma = -\xi_- A_-$.

Para definir el problema de control óptimo, consideraremos otra vez una "energía" tipo funcional objetivo como la que sigue:

$$S = \int_0^T \frac{1}{2} \langle \xi, I\xi \rangle dt, \quad (6.5)$$

donde hemos usado el producto escalar $\langle \xi, \zeta \rangle = 4\text{Tr}(\xi^T \zeta)$.

Los resultados discutidos a lo largo de esta tesis, demuestran que hay una relación bien definida entre las soluciones del problema de control óptimo, dadas por las ecuaciones (6.4), (6.5) y los puntos críticos del Lagrangiano

$$L(g, \dot{g}) = \frac{1}{2} \langle g^{-1} \dot{g}, I g^{-1} \dot{g} \rangle \quad (6.6)$$

en $TSL(2, \mathbb{R})$.

Las ecuaciones de Euler-Lagrange para el Lagrangiano (6.6) están dadas por:

$$\dot{g} = g\xi, \quad I\dot{\xi} = [\xi^T, I\xi].$$

Las ecuaciones del movimiento resultantes de aplicar el principio del máximo de Pontryagin son:

$$\dot{x} = \alpha + \beta x + \gamma x^2, \quad \dot{p} = -(\beta + 2\gamma x)p,$$

y la relación entre ambos conjuntos de ecuaciones está dada por:

$$pA_+ - I_+ \xi_+ = 0, \quad -pA_- x^2 - I_- \xi_- = 0, \quad 2pA_0 x - I_0 \xi_0 = 0.$$

6.2.2. Ecuaciones para el grupo $SU(2)$

En este ejemplo, si sustituimos $SL(2, \mathbb{R})$ por $SU(2)$, la ecuación de control es ahora la siguiente:

$$\dot{x} = (\alpha - i\gamma) + i\beta x + (\alpha + i\gamma)x^2, \quad (6.7)$$

con $\alpha = \xi_+ A_+$, $\beta = 2\xi_0 A_0$, $\gamma = -\xi_- A_-$. Las ecuaciones de Euler-Lagrange para el Lagrangiano (6.6), donde hemos utilizado la forma de Frobenius están dadas por:

$$\dot{g} = g\xi, \quad I\dot{\xi} = [\xi^*, I\xi] = [I\xi, \xi],$$

y usando la forma de Killing obtenemos:

$$\dot{g} = g\xi, \quad \dot{J}(\xi) = [J(\xi), \xi],$$

con

$$J(\xi) = I\xi + \xi I.$$

6.2.3. Ecuaciones para el grupo $SO(2, 1)$

De nuevo, si reemplazamos $SL(2, \mathbb{R})$ por $SO(2, 1)$ donde

$$\mathfrak{so}(2, 1) = \{ \xi_+ e_+ + \xi_- e_- + \xi_0 e_0 \mid \xi_+, \xi_-, \xi_0 \in \mathbb{C} \},$$

con

$$e_+ = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad e_- = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad e_0 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

la ecuación de control sería:

$$\dot{x} = i(2\alpha - \gamma(1 + x)), \quad (6.8)$$

con $\alpha = \xi_+ A_+$, $\gamma = -\xi_- A_-$. Las ecuaciones de Euler-Lagrange para el Lagrangiano (6.6), donde usamos la forma de Frobenius están dadas por:

$$\dot{g} = g\xi, \quad I\dot{\xi} = [\xi^*, I\xi],$$

y si usamos la forma de Killing son:

$$\dot{g} = g\xi, \quad \dot{J}(\xi) = [J(\xi), \xi],$$

con $J(\xi) = I\xi + \xi I$.

Figura 6.1: Elroy's beanie

6.3. Las ecuaciones simétricas de "Elroy's beanie"

En realidad, Elroy's beanie se refiere tanto a Elroy como a su gorra (propulsor). Asumimos que Elroy está flotando en medio del espacio, libre de los efectos gravitatorios, de la fricción y de cualquier cosa que complique el problema. Elroy y su gorra se mueven por separado, cada uno girando independientemente. Por esto, es realmente un sistema de dos cuerpos rígidos girando alrededor del mismo eje.

Consideramos dos cuerpos rígidos en el plano unidos en sus centros de masas. Las ecuaciones simétricas de Elroy's beanie como un problema en teoría de control óptimo pueden ser discutidas como sigue:

Sea G un grupo de Lie actuando por la izquierda sobre una variedad diferenciable Q representando el espacio de estados de un sistema de control. Supongamos además que G actúa por la derecha sobre una variedad diferenciable Q . Puesto que tanto la gorra como Elroy son libres de moverse en un círculo, y sus movimientos son independientes, como fibrado principal Q consideraremos el grupo $\mathbb{S}^1 \times \mathbb{S}^1$ para el cual usaremos coordenadas (θ, ψ) , donde θ es la posición angular de Elroy con respecto a un sistema inercial y ψ es el ángulo del propulsor relativo a Elroy. Por otra parte, consideraremos como espacio auxiliar el espacio de configuraciones P que es de nuevo el grupo de Lie $\mathbb{S}^1 \times \mathbb{S}^1$. El espacio de controles para el problema de control óptimo será el fibrado tangente TQ . Las coordenadas correspondientes para TQ serán $(\theta, \psi, v_\theta, v_\psi)$. Puesto que el momento angular no es medido sobre un círculo, el espacio de fases es $T^*P = \mathbb{S}^1 \times \mathbb{R} \times \mathbb{S}^1 \times \mathbb{R}$.

Entonces, en ausencia de potencial, la densidad Lagrangiana es simplemente la energía cinética, así que:

$$L(\theta, \psi, v_\theta, v_\psi) = (1/2)(I_1 + I_2)v_\theta^2 + (1/2)I_2v_\psi^2 + I_2v_\theta v_\psi, \quad (6.9)$$

definida por la métrica Riemanniana

$$\langle \cdot, \cdot \rangle = (I_1 + I_2)d\theta \otimes d\theta + I_2d\theta \otimes d\psi + I_2d\psi \otimes d\theta + I_2d\psi \otimes d\psi$$

sobre Q , donde I_1 e I_2 son los momentos de inercia de Elroy y su gorra.

Por tanto, al no haber potencial el grupo G será $\mathbb{S}^1 \times \mathbb{S}^1$. Lo parametrizaremos por $\begin{pmatrix} e^{i\phi_1} & 0 \\ 0 & e^{i\phi_2} \end{pmatrix}$. También identificaremos $\mathfrak{g} = \mathfrak{u}(1) \times \mathfrak{u}(1)$ con $\mathbb{R} \times \mathbb{R}$ de la forma usual.

Consideremos la siguiente acción de G sobre P :

$$\Phi \left(\begin{pmatrix} e^{i\phi_1} & 0 \\ 0 & e^{i\phi_2} \end{pmatrix}, (\theta, \psi) \right) = (\theta + \phi_1, \psi + \phi_2).$$

Puesto que podemos identificar \mathfrak{g} con $\mathbb{R} \times \mathbb{R}$, la aplicación exponencial será $\exp(\omega_1, \omega_2) = (\omega_1 \bmod 2\pi, \omega_2 \bmod 2\pi)$.

Por tanto, calculando el campo de vectores $\xi = \omega_1 \frac{\partial}{\partial \phi_1} + \omega_2 \frac{\partial}{\partial \phi_2}$ sobre P en el punto $(\theta, \psi, v_\theta, v_\psi)$, obtenemos

$$\xi_P(\theta, \psi, v_\theta, v_\psi) = \left. \frac{d}{dt} \left(\Phi(\exp(t\xi), (\theta, \psi)) \right) \right|_{t=0} = \omega_1 \frac{\partial}{\partial \theta} + \omega_2 \frac{\partial}{\partial \psi},$$

y la ecuación de control será

$$(\dot{\theta}, \dot{\psi}) = \xi_P(\theta, \psi) = (\omega_1(t), \omega_2(t)).$$

El funcional objetivo es

$$S = \int_0^T L(\theta, \psi, v_\theta, v_\psi) dt. \quad (6.10)$$

Las ecuaciones de Euler-Lagrange para el lagrangiano (6.9) estarán dadas por:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial v_\theta} \right) - \frac{\partial L}{\partial \theta} = 0, \quad \frac{d}{dt} \left(\frac{\partial L}{\partial v_\psi} \right) - \frac{\partial L}{\partial \psi} = 0. \quad (6.11)$$

Por tanto, las ecuaciones de Lagrange para el propulsor serán

$$\ddot{\theta} = 0, \quad \ddot{\psi} = 0. \quad (6.12)$$

Obtenemos velocidad angular constante, como es de esperar por la ausencia de fuerzas.

Y el campo vectorial lagrangiano será

$$\Gamma_L = v_\theta \frac{\partial}{\partial \theta} + v_\psi \frac{\partial}{\partial \psi} \quad (6.13)$$

Además el hamiltoniano de Pontryagin $H : T^*P \times TQ \longrightarrow \mathbb{R}$ tomará la forma:

$$H(p_\theta, p_\psi, \theta, \psi, v_\theta, v_\psi) = \langle (p_\theta, p_\psi), (\omega_1(t), \omega_2(t)) \rangle - L(\theta, \psi, v_\theta, v_\psi),$$

donde $(\theta, \psi, v_\theta, v_\psi) \in TQ$ y $(\theta, \psi, p_\theta, p_\psi) \in T^*P$.

Las ecuaciones del movimiento dadas por el Principio del Máximo de Pontryagin serán entonces:

$$(\dot{\theta}, \dot{\psi}) = (\omega_1(t), \omega_2(t)), \quad (\dot{p}_\theta, \dot{p}_\psi) = (0, 0), \quad p_\theta \dot{\omega}_1(t) + p_\psi \dot{\omega}_2(t) = 0$$

junto con las siguientes ligaduras:

$$p_\psi = I_2 v_\theta + I_2 v_\psi \quad (6.14)$$

y

$$p_\theta = (I_1 + I_2)v_\theta + I_2 v_\psi \quad (6.15)$$

Tomando derivadas con respecto a t en la ecuación (6.14) y (6.15) obtenemos:

$$0 = \dot{p}_\theta = (I_1 + I_2)\dot{v}_\theta + I_2\dot{v}_\psi, \quad 0 = \dot{p}_\psi = I_2\dot{v}_\theta + I_2\dot{v}_\psi.$$

Por lo tanto, tendremos

$$\dot{v}_\theta(t) = 0 = \dot{v}_\psi(t).$$

Entonces, $\omega_1 = \omega_1(t) = v_\theta(t) = v_\theta(0)$ y $\omega_2 = \omega_2(t) = v_\psi(t) = v_\psi(0)$.

Finalmente, la curva solución en Q con condiciones iniciales $(\theta(0), \psi(0), v_\theta(0), v_\psi(0))$ están dadas por

$$t \mapsto \left(\theta(0) + v_\theta(0)t, \psi(0) + v_\psi(0)t \right)$$

6.3.1. Ecuaciones reducidas de Elroy's beanie

Calcularemos la aplicación momento para la acción Φ dada para el Elroy's beanie. Tenemos $\xi_P^c = \omega_1 \frac{\partial}{\partial \theta} + \omega_2 \frac{\partial}{\partial \psi}$ y

$$\omega_L = d\theta_L = d\left(\frac{\partial L}{\partial v_\theta} d\theta + \frac{\partial L}{\partial v_\psi} d\psi\right) = (I_1 + I_2)dv_\theta \wedge d\theta + I_2 dv_\psi \wedge d\theta + I_2 dv_\psi \wedge d\psi + I_2 dv_\theta \wedge d\psi.$$

Luego, como

$$dJ = \xi_{P \lrcorner} \omega_L = ((I_1 + I_2)dv_\theta + I_2 dv_\psi) e^1 + (I_2 dv_\theta + I_2 dv_\psi) e^2$$

tenemos que

$$J(\theta, \psi, v_\theta, v_\psi) = ((I_1 + I_2)v_\theta + I_2 v_\psi, I_2 v_\theta + I_2 v_\psi).$$

Como $\mu = m_1 e^1 + m_2 e^2 \in \mathfrak{g}^*$, vemos fácilmente que

$$J^{-1}(\mu) = \{(\theta, \psi, v_\theta, v_\psi) \in TQ \mid (I_1 + I_2)v_\theta + I_2 v_\psi = m_1, I_2 v_\theta + I_2 v_\psi = m_2\}. \quad (6.16)$$

Puesto que $Q_\mu = Q/G$ es un único punto, $(\theta(0), \psi(0))$, el campo vectorial $\Gamma_L^\mu \equiv 0$ que tiene curvas integrales $t \mapsto (\theta(0), \psi(0))$ donde $(\theta(0), \psi(0))$ es la condición inicial. Por tanto, $q_\mu(t)$ es la curva sobre Q_μ dada por $t \mapsto (\theta(0), \psi(0))$.

El levantamiento horizontal de $q_\mu(t)$, la curva $q_h(t)$, a través de $(\theta(0), \psi(0), v_\theta(0), v_\psi(0))$ está dada por $t \mapsto (\theta(0), \psi(0))$.

Ahora el uso de la relación algebraica

$$\langle \xi(t)_Q(q_h(t)), \eta_Q(q_h(t)) \rangle = \mu \cdot \eta, \forall \eta \in \mathfrak{g}_\mu$$

determina $\xi(t)$.

Supongamos que $\xi(t) = \omega_1(t)e_1 + \omega_2(t)e_2$ y $\eta = \delta_1 e_1 + \delta_2 e_2$. Entonces,

$$\xi(t)_Q(\theta, \psi, v_\theta, v_\psi) = \omega_1(t) \frac{\partial}{\partial \theta} + \omega_2(t) \frac{\partial}{\partial \psi}$$

y

$$\eta_Q(\theta, \psi, v_\theta, v_\psi) = \delta_1 \frac{\partial}{\partial \theta} + \delta_2 \frac{\partial}{\partial \psi}$$

Calculamos

$$\langle \xi(t)_Q(q_h(t)), \eta_Q(q_h(t)) \rangle = (I_1 + I_2)\omega_1(t)\delta_1 + I_2\omega_1(t)\delta_2 + I_2\omega_2(t)\delta_1 + I_2\omega_2(t)\delta_2.$$

Entonces

$$\left((I_1 + I_2)\omega_1(t) + I_2\omega_2(t) \right) \delta_1 + \left(I_2\omega_1(t) + I_2\omega_2(t) \right) \delta_2 = m_1\delta_1 + m_2\delta_2, \forall \delta_1, \delta_2 \in \mathbb{R}.$$

Y por consiguiente

$$\omega_1(t) = \frac{m_1}{I_1} - \frac{m_2}{I_1} = \omega_1, \quad \omega_2(t) = -\frac{m_1}{I_1} + \frac{(I_1 + I_2)m_2}{I_1 I_2} = \omega_2. \quad (6.17)$$

Aquí, si usamos (6.16) tendremos: $\omega_1 = v_\theta(0)$ y $\omega_2 = v_\psi(0)$. La ecuación diferencial

$$g'(t) = T_e R_{g(t)} \xi(t)$$

en G_μ con condición inicial $g(0) = e$ se lee

$$\begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix} = \begin{pmatrix} i\dot{\phi}_1 e^{i\phi_1} & 0 \\ 0 & i\dot{\phi}_2 e^{i\phi_2} \end{pmatrix} \begin{pmatrix} e^{-i\phi_1} & 0 \\ 0 & e^{-i\phi_2} \end{pmatrix} \quad (6.18)$$

y entonces tenemos:

$$\dot{\phi}_1(t) = -i\omega_1, \quad \phi_1(0) = 0$$

y

$$\dot{\phi}_2(t) = -i\omega_2, \quad \phi_2(0) = 0$$

la cual tiene la solución

$$\phi_1(t) = -i\omega_1 t$$

y

$$\phi_2(t) = -i\omega_2 t.$$

Finalmente, la curva solución en Q con condición inicial $(\theta(0), \psi(0), v_\theta(0), v_\psi(0))$ estará dada por

$$\begin{aligned} t \mapsto \Phi(g(t), q_h(t)) &= \Phi\left(\begin{pmatrix} e^{i(-i\omega_1 t)} & 0 \\ 0 & e^{i(-i\omega_2 t)} \end{pmatrix}, (\theta(0), \psi(0))\right) = \\ &= (\theta(0) + \omega_1 t, \psi(0) + \omega_2 t) = (\theta(0) + v_\theta(0)t, \psi(0) + v_\psi(0)t) \end{aligned}$$

como calculamos en primer lugar.

Ahora consideraremos lo que ocurre cuando ponemos un muelle en el sistema Elroy's beanie de forma que haya una fuerza restauradora, siempre y cuando Elroy y su propulsor no estén alineados. Esto significa que tenemos un potencial de la forma $V = 1/2k\psi^2$. Ahora, si hacemos los cálculos de la sección (6.3) con $G = \mathbb{S}^1$ obtenemos la ecuación $V'(\psi) = 0$. Por tanto, no obtenemos todas las posibles soluciones, y esto falla porque requerimos que el grupo G actúe invariante sobre toda la variedad Q , es decir se corresponde con que tenemos un $\mathbb{S}^1 \times \mathbb{S}^1$ espacio de configuración con sólo una \mathbb{S}^1 simetría, debida al hecho de que la orientación total de Elroy y su propulsor no importa.

6.4. Un ejemplo simple en dimensión finita

6.4.1. Hamiltonianos de Control Cuántico de tipo Lie-Scheffers-Brockett.

Podemos considerar un sistema dependiente del tiempo de la forma:

$$\Gamma_G = \sum_{\alpha=1}^r u^\alpha(t) \xi_\alpha \quad (6.19)$$

El flujo de Γ_G está dado por una familia uniparamétrica $\varphi_t: G \longrightarrow G$ de difeomorfismos del grupo satisfaciendo:

$$\frac{d}{dt}\varphi_t(g) = \Gamma_G(\varphi_t(g)), \forall g \in G$$

ó en otras palabras:

$$\left(\frac{d}{dt}\varphi_t\right) \circ \varphi_t^{-1}(g) = \sum_{\alpha} u^{\alpha}(t)\xi_{\alpha}$$

Si denotamos, dado g , por $g_t = \varphi_t(g)$, la ecuación anterior toma la forma:

$$\frac{d}{dt}g_t \circ g_t^{-1} = \sum_{\alpha} u^{\alpha}(t)\xi_{\alpha}$$

Los coeficientes $u^{\alpha}(t)$ pueden ser considerados como variables de control y la evolución toma lugar en el grupo G .

Un ejemplo particular de esta situación ocurre en la descripción de Heisenberg de sistemas cuánticos. Sea \mathcal{H} un espacio de Hilbert complejo separable que describe (normalizados) los estados puros de un sistema cuántico con la dinámica definida por un operador autoadjunto H . La evolución de un estado dado $|\psi_0\rangle \in \mathcal{H}$ se obtiene integrando la ecuación de Schrödinger:

$$i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle, \quad |\psi(0)\rangle = |\psi_0\rangle \quad (6.20)$$

El teorema de Stone garantiza la existencia de una familia uniparamétrica de operadores unitarios U_t en \mathcal{H} tales que la curva

$$|\psi(t)\rangle = U_t|\psi_0\rangle$$

es la solución de la ecuación de Schrödinger para cada $|\psi_0\rangle \in \mathcal{H}$. Por tanto la ecuación (6.20) puede escribirse como:

$$\frac{d}{dt}U_t = \left(-\frac{i}{\hbar}H\right)U_t \quad (6.21)$$

que puede ser considerada como una ecuación de evolución sobre el grupo $G = \mathcal{U}(\mathcal{H})$ de todos los operadores unitarios sobre \mathcal{H} . Técnicamente hablando $\mathcal{U}(\mathcal{H})$ no es un grupo de Lie, sin embargo muchas aplicaciones en la teoría de control cuántico tratan con sistemas cuánticos finito dimensionales. En tal caso, si $\mathcal{H} = \mathbb{C}^N$, con N la dimensión del sistema cuántico, entonces $\mathcal{U}(\mathcal{H}) = U(N)$, que es un grupo de Lie compacto de dimensión N^2 . Continuaremos trabajando en lo que sigue en el marco general, precisando cuando sea necesario. El operador $-\frac{i}{\hbar}H$ puede ser identificado con un elemento del álgebra de Lie de $\mathcal{U}(\mathcal{H})$, que consiste en

el conjunto de operadores autoadjuntos sobre \mathcal{H} . Típicamente, el Hamiltoniano H en los problemas de control cuántico tienen la forma

$$H = \sum_{\alpha=0}^r u^\alpha H_\alpha = H_0 + \sum_{\alpha=1}^r u^\alpha H_\alpha$$

donde H_0 denota el Hamiltoniano "libre", es decir, el Hamiltoniano del sistema no controlado y los restantes H_α son los Hamiltonianos de control. En general no cierran el álgebra de Lie, sin embargo en este trabajo asumiremos que los Hamiltonianos de control satisfacen:

$$[H_\alpha, H_\beta] = ic_{\alpha\beta}^\gamma H_\gamma$$

es decir, tenemos a nuestra disposición para controlar el sistema cuántico H_0 , a todos los Hamiltonianos del álgebra de Lie dinámica del sistema (recordemos que el álgebra de Lie dinámica de un sistema de control cuántico es el álgebra de Lie generada por los Hamiltonianos de control). Por tanto escribiendo la ecuación de evolución de nuestro sistema de control cuántico en la representación de grupos, tenemos:

$$\frac{dU_t}{dt} \circ U_t^{-1} = -\frac{i}{\hbar} \left(H_0 + \sum_{\alpha} u^\alpha H_\alpha \right)$$

que tiene exactamente la forma de un sistema de Lie-Scheffers-Brockett sobre el grupo $G = \mathcal{U}(\mathcal{H})$.

6.4.2. Control Óptimo de Sistemas Cuánticos y representación de Clebsch.

Consideraremos un problema de control cuántico con Hamiltoniano $H(u) = H_0 + \sum_{\alpha=1}^r u^\alpha(t) H_\alpha$ sobre el espacio de Hilbert \mathcal{H} . Intentaremos entender para qué funciones de control $u^\alpha(t)$, la solución de la ecuación de Schrödinger lleva el estado inicial $|\psi_0\rangle$ a algún estado final $|\psi_T\rangle$ a tiempo T , es decir $|\psi_T\rangle = U_T |\psi_0\rangle$, con U_T la solución de:

$$\frac{dU_t}{dt} \circ U_t^{-1} = -\frac{i}{\hbar} \left(H_0 + \sum_{\alpha} u^\alpha H_\alpha \right); \quad U_0 = I \quad (6.22)$$

Entre todas las posibles soluciones $u^\alpha(t)$ a la ecuación (6.22) estamos interesados en aquellas que además minimizan la energía, ó que se aproximan mucho a cierto camino preestablecido $|\psi_t\rangle$, etc..., es decir, estamos interesados en minimizar un funcional que típicamente tiene la forma:

$$S(u, \psi) = \frac{\delta}{2} \int_0^T u^\alpha(t) u_\alpha(t) dt + \frac{\rho}{2} \int_0^T \|\psi(t) - \psi_t\|^2 dt$$

Esto constituye el problema de control óptimo para el sistema cuántico anterior. En general se sabe muy poco de las soluciones de este problema en el caso infinito dimensional. Sin embargo, el caso finito-dimensional es bien conocido y se han implementado muchas aplicaciones ([Al08], [Mi04]).

En este ejemplo estamos principalmente interesados en problemas de control óptimo generales en los controles, pero sin el término de tracking, es decir:

$$S(u) = \int_0^T L(u) dt$$

aunque incluso el caso general puede también considerarse y el funcional objetivo S tendría la siguiente forma (hablaremos sobre ello más tarde):

$$S(u, \psi) = \int_0^T L(u, \psi) dt$$

Las técnicas usuales para el control óptimo no pueden aplicarse fácilmente. Una idea básica consiste en utilizar la aplicación Φ del espacio de controles $\{u^\alpha\}$ en el espacio de soluciones $\{\psi(t)\}$ para eliminar la dependencia de ψ en el funcional S y obtener:

$$\tilde{S}(u) = \int_0^T L(u, \Phi(u)) dt$$

y entonces aplicar métodos directos para demostrar la existencia de soluciones. Esta idea, sin embargo, no funciona debido a que el sistema (6.19) no permite probar condiciones apropiadas de regularidad de la aplicación Φ , que haría posible analizar el funcional \tilde{S} propiamente. Tomaremos una aproximación diferente al problema. En vez de intentar resolverlo directamente lo consideraremos desde la perspectiva de las variables de Clebsch, es decir lo miraremos como la representación de otro sistema que debería ser más simple de integrar. Para este propósito escribiremos el problema de control óptimo cuántico como un problema de multiplicadores de Lagrange y posteriormente lo manipularemos para obtener la forma que estamos buscando.

Consideraremos entonces el problema de control óptimo definido por el sistema Hamiltoniano $H(u) = H_0 + \sum_{\alpha=1}^r u^\alpha H_\alpha$ sobre el grupo $\mathcal{U}(\mathcal{H})$ de operadores unitarios sobre el espacio de Hilbert \mathcal{H} ; el funcional objetivo $\mathbb{S} = \int_0^T L(\psi, u) dt$ y condiciones iniciales $\psi_0 = \psi(0)$. En lugar del principio del Máximo de Pontryagin, usaremos una aproximación de multiplicadores de Lagrange, y escribiremos el funcional objetivo extendido:

$$\mathbb{S}(\psi, u, \zeta) = \int_0^T \left(L(\psi, u) + \langle \zeta, \dot{\psi} - iH(u)\psi \rangle \right) dt$$

que puede escribirse como:

$$\mathbb{S}(\psi, u, \zeta) = \int_0^T \left(L(\psi, u) - \langle \dot{\zeta}, \psi \rangle + \langle \zeta, -iH(u)\psi \rangle \right) dt + \langle \zeta, \psi \rangle \Big|_0^T$$

El vector $\zeta \in \mathcal{H}$ (ó al dual de \mathcal{H} para este problema) y es la variable de co-estado del sistema. Las variaciones del funcional \mathbb{S} dejarán condiciones de primer orden para la existencia de puntos críticos de S .

No obstante, nos interesaremos por una aproximación distinta al problema que tomará ventaja a la formulación dinámica del grupo. Consideraremos ahora como variable de estado el operador unitario U , por lo que nuestro funcional tendrá la forma:

$$S(U, u) = \int_0^T L(U, u) dt$$

y la ecuación de estado será la ecuación (6.21):

$$\frac{dU}{dt} U^+ = -iH(u), \quad U(0) = I.$$

Con la aplicación del teorema de los multiplicadores de Lagrange obtenemos:

$$\mathbb{S}(U, u, P) = \int_0^T \left(L(U, u) + \langle P, \dot{U}U^+ - iH(u) \rangle \right) dt$$

donde $P \in \mathcal{U}(\mathcal{H})^*$, el dual del álgebra de Lie del grupo unitario $\mathcal{U}(\mathcal{H})$. En el caso finito dimensional, podemos identificar $\mathcal{U}(\mathcal{H}) \cong \mathcal{U}(\mathcal{H})^*$ usando la forma de Killing-Cartan $\langle \cdot, \cdot \rangle$ del grupo unitario ($\langle A, B \rangle = \text{tr} A^+ B$).

Para obtener una descripción apropiada de las ecuaciones introduciremos un formalismo homogéneo, es decir, introduciremos una variable de control auxiliar u_0 que multiplicará a H_0 , luego nuestro Hamiltoniano será:

$$H(u) = u_0 H_0 + \sum_{\alpha=1}^r u^\alpha H_\alpha = \sum_{\alpha=0}^r u^\alpha H_\alpha$$

$$\text{y } [H_i, H_j] = i c_{ij}^k H_k.$$

Denotaremos por G al grupo dinámico generado por H_k . Consideraremos ahora que hay un espacio auxiliar Q tal que el grupo unitario G actúa sobre él. Podemos incluso asumir que G actúa libremente sobre Q y que el espacio cociente Q/G es una variedad y la aplicación de proyección $p: Q \longrightarrow Q/G$ es una submersion,

es decir, que Q es un fibrado principal con grupo de estructura G . Escogeremos una conexión A sobre Q , y consideraremos el siguiente campo vectorial sobre \mathcal{H} :

$$\Gamma_A(\psi) = \widehat{A_q(\dot{q})}(\psi)$$

donde $\widehat{\xi}$ si $\xi \in \mathfrak{g}$, denota el operador sobre \mathcal{H} definido por el elemento ξ , es decir si usamos una base lineal de \mathfrak{g} dada por iH_0, iH_1, \dots, iH_r , entonces para cualquier $\xi \in \mathfrak{g}$; $\xi = i\xi_0 H_0 + \dots + i\xi_r H_r$.

El operador $\widehat{A_q(\dot{q})}$ será

$$\widehat{A_q(\dot{q})} = iA_q^k(\dot{q})H_k = i(A^k(q)_l \dot{q}^l) H_k = iu^k H_k.$$

donde las funciones de control $u^k = u^k(q, \dot{q})$ son funciones sobre las variables auxiliares (q, \dot{q}) . Ahora la ecuación de control de estado se puede escribir como:

$$\frac{dU}{dt} U^+ = \widehat{A_q(\dot{q})}$$

Y el término que contribuye en el funcional extendido tiene la forma:

$$\langle P, \frac{dU}{dt} U^+ - \widehat{A_q(\dot{q})} \rangle = \langle PU, \frac{dU}{dt} \rangle - \langle P, \widehat{A_q(\dot{q})} \rangle = \Theta_{(U,P)}(\dot{U}) - \langle J(U, P), \widehat{A_q(\dot{q})} \rangle$$

donde Θ denota la 1-forma canónica sobre $T^*\mathcal{U}(\mathcal{H})$ definida como :

$$\Theta_{(U,P)}(\dot{U}) = \langle P, \dot{U} U^+ \rangle,$$

y J denota la aplicación momento canónica sobre $T^*\mathcal{U}(\mathcal{H})$ definida por la acción del grupo G , es decir, $J: T^*\mathcal{U}(\mathcal{H}) \longrightarrow \mathfrak{g}^*$,

$$\langle J(U, P), \xi \rangle = \langle P U, \widehat{\xi} U \rangle = \langle P, \widehat{\xi} \rangle$$

La forma anterior del funcional extendido coincide con la representación en términos de variables de Clebsch de los principios variacionales con simetría, que es la razón por la que se le ha llamado la representación de Clebsch de un sistema de control cuántico definido por $H(u)$.

Capítulo 7

Conclusiones y problemas abiertos.

Recapitemos brevemente alguna de las conclusiones y resultados obtenidos en este trabajo, así como algunas cuestiones que habiéndose suscitado, no han encontrado una respuesta total ó parcial. También comentaremos brevemente algunos problemas que, aún no formando parte de esta tesis, se hallan ó bien muy próximos a ella, ó podemos preveer que su solución puede beneficiarse de los resultados aquí expuestos.

7.1. Conclusiones y resultados

0. Los capítulos 1 y 2 se han dedicado a recapitular diversos resultados sobre el problema del cálculo de variaciones con simetría, así como a establecer los elementos analíticos adecuados para el tratamiento posterior. Podemos resaltar aquí el desarrollo sistemático del análisis global de los espacios de curvas en variedades diferenciales y una extensión no completamente del Teorema de los multiplicadores de Lagrange.
1. En el capítulo 3 se ha procedido a una formulación geométrica de la teoría de variables de Clebsch como variables en un espacio auxiliar simpléctico sobre el que actúa el grupo de Lie de simetría de la teoría. Se han discutido también las ligaduras de Lin como condiciones de horizontalidad en curvas en un par de Borel $P \times_G Q$. Finalmente se ha establecido de manera general un teorema sobre la equivalencia entre tres espacios de curvas: curvas en Q , curvas horizontales en $P \times Q$ y pares de curvas en P y $G \backslash Q$. Esta equivalencia se usa para probar un teorema sobre diversas formulaciones equivalentes del principio variacional definido por un Lagrangiano invariante $L: TQ \longrightarrow \mathbb{R}$. Se extienden de esta manera los resultados obtenidos por Cendra, Ibort y Marsden, y Cendra y Marsden [Ce87a], [Ce87b].

2. En segundo lugar en el capítulo 4 se han analizado las soluciones extremales de un problema de control óptimo de Lie-Scheffers-Brockett con grupo G y lagrangiano G -invariante definido en un espacio de controles TQ . Se ha probado que dichas curvas extremales están en correspondencia 1-1 con puntos críticos del sistema lagrangiano con simetría L . La correspondencia que establece dicha relación constituye la descripción geométrica general de la dualidad observada por Bloch, Crouch, Holm y Marsden [Bl00], entre diversas realizaciones de un mismo sistema, bien como sistema lagrangiano con simetría, ó como problema de control óptimo. Dicha relación de dualidad se ha descrito explícitamente para sistemas de Lie-Scheffers-Brockett sobre los grupos $SO(3)$, $SU(2)$, $SL(2, \mathbb{R})$ y $SU(1, 1)$.
3. La teoría de reducción de problemas de control óptimo con simetría se ha discutido en el capítulo 5 desde el punto de vista de la reducción del principio variacional tanto de Lagrange-Pontryagin, como de Hamilton-Pontryagin, siguiendo ideas anticipadas por Yoshimura y Marsden [Yo07a], [Yo09] en el contexto de sistemas lagrangianos invariantes y la reducción cotangente. En el caso de la reducción del principio variacional de Hamilton-Pontryagin, el procedimiento es equivalente a la reducción de Lie-Dirac de la imagen presimpléctica del problema de control óptimo, mientras que la reducción del principio variacional de Lagrange-Pontryagin extiende la reducción de Euler-Poincaré desarrollada por Cendra, Marsden y Ratiu [Ce01a], [Ce01b]. Se muestra que ambas formulaciones son equivalentes, utilizando para ello diversas caracterizaciones de la reducción cotangente de Dirac y una formulación novedosa de los sistemas de control óptimo inspirada en la formulación de Dirac de sistemas lagrangianos implícitos.
4. Varios ejemplos, entre ellos una aplicación a problemas de control cuántico se han desarrollado con detalle.

7.2. Algunos problemas abiertos

1. El análisis de la reducción de problemas de control óptimo utilizando estructuras de Dirac ha quedado esbozado pero no completado. De la misma manera, la relación con la descripción de problemas de control óptimo por medio de algebroides desarrollada por E. Martínez ha quedado apuntada pero no desarrollada, en particular la posibilidad de desarrollar la teoría de pares de Borel como una teoría en un bialgebriode.
2. Desarrollar algoritmos eficaces para la construcción de soluciones aprovechando los principios de superposición no lineales proporcionados por el teorema de Lie-Scheffers para los sistemas de Lie-Scheffers-Brockett.

-
3. Desarrollar esquemas numéricos que utilicen las variables de Clebsch como una herramienta eficaz para estabilizar los algoritmos.
 4. La aplicación de todas estas ideas en dimensión infinita, por ejemplo el caso de fluidos compresibles ó incompresibles, etc.

Bibliografía

- [Ab78] R. Abraham, J. E. Marsden. *Foundations of Mechanics*. Benjamin-Cummings, Reading, MA, 2nd edition, (1978).
- [Ag90] A. A. Agrachev, R. V. GamKrelidze. *Simplectic Geometry for Optimal Control, in Nonlinear Controllability and Optimal Control*. Ed. H.J. Sussmann. Marcel & Dekker Inc., N. Y. 263-277, (1990).
- [Al08] D'Alessandro D. *Introduction to Quantum Control and Dynamics*, (Boca Raton: Chapman & Hall/CRC), (2008).
- [Ar76] V. I. Arnol'd. *Les méthodes mathématiques de la Mécanique classique*. Éditions Mir, Moscou (1976).
- [Ar86] V. I. Arnol'd, A. Varchenco, S. Goussein-Zadé. *Singularités des Applications Différentiables*. Vol 1. Éditions Mir, Moscou (1986).
- [Aru94] A. V. Arutyunov, S. M. Aseev. *The maximum principle in optimal control problems with phase constraints. Nondegeneracy and stability*. Russian Acad. Sci. Dokl. Math. **49**, 38-42, (1994).
- [Ba78] Balachandran, A.P. Marmo, G. *Gauge theories Symmetries and Fiber Bundles*. Lecture Notes in Physics, vol **188**. Springer-Verlag, N.Y., (1978).
- [Ba08] Barbero Liñán, M. Muñoz Lecanda, M. *Geometric approach to Pontryagin's Maximum Principle*. Acta Appl. Math., **104**, (2008) arXiv:0805.1169.
- [Bl96] A.M. Bloch and P.E. Crouch. *Optimal control and geodesic flows*. Systems Control Lett., **28**, 65-72, (1996).
- [Bl97] A.M. Bloch, R. Brockett, and P.E. Crouch. *Double bracket equations and geodesic flows on symmetric spaces*. Comm. Math. Phys. **187**, 357-373, (1997).
- [Bl98] A.A. Bloch, P.E. Crouch, J.E. Marsden, T. Ratiu. *The symmetric representation of the rigid body equations and their discretization*. Nonlinearity, **15**, 1309-1341, (2002).

- [Bl99] A. M. Bloch, P. E. Crouch. *Optimal control, optimization and analytical mechanics*. Mathematical Control Theory. J. Baillieul and J. C. Willems Eds., Springer-Verlag, 265-321 (1999).
- [Bl00] A. A. Bloch, P. E. Crouch, D. D. Holm, J.E. Marsden. *An optimal control formulation for inviscid incompressible fluid flow*. Proc. CDC, **39**, 1273-79, (2000).
- [Br70] R.W. Brockett. *Systems theory on group manifolds and coset spaces*. SIAM J. Control Optim., **10**, 265-284, (1970).
- [Br73a] R. W. Brockett. *Lie theory and control systems defined on spheres*. SIAM J. Appl. Math., **25**, 213-225, (1973).
- [Br73b] R. W. Brockett. *Lie Algebras and Lie groups in Control Theory*. Geometrical Methods in System Theory. D. Q. Mayne, R. W. Brockett Eds., Reidel, Dordrecht, 43-82, (1973).
- [Br78] R. W. Brockett. *Feedback invariants for nonlinear systems*. Proceedings of IFAC Congress, Helsinki, (1978).
- [Br83] R. W. Brockett. *Non linear Control theory and differential geometry*. Proceedings of the ICM Warszawa, august 16-24, 1357-1368, (1983).
- [Bu04] F. Bullo, A. D. Lewis. *Geometric Control of Mechanical Systems*. Modeling, analysis and design for simple mechanical control, Springer, New York, (2004).
- [Ca89] J. F. Cariñena, C. López, E. Martínez. *A new approach to the converse of Noether's Theorem*. J. Phys. A: Math.& Gen., **22**, 4486-4777, (1989).
- [Ca91] J. F. Cariñena, C. López, E. Martínez. *Sections along a map applied to higher-order Lagrangian Mechanics*. Acta Apli. Mathematicae, **25**, 127-151, (1991).
- [Ca00] J.F. Cariñena, J. Grabowski, G. Marmo. *Lie-Scheffers Systems: A Geometric Approach*, Bibliopolis, Napoli, (2000).
- [Ce01a] Cendra H., Marsden J.E. and Ratiu T. *Lagrangian reduction by stages*, Mem. Amer. Math. Soc. 152 **722**, 1-108, (2001).
- [Ce01b] H. Cendra, J.E. Marsden, T. Ratiu. *Geometric Mechanics, Lagrangian reduction, and Nonholonomic systems*. In Mathematics unlimited- 2001 and beyond, B. Engquist and W. Schmid eds., Springer-Verlag, New York, 221-273, (2001).
- [Ce87a] H. Cendra and J.E. Marsden. *Lin Constraints, Clebsch potentials and variational principles*, Physica D, **27**, 63-89, (1987).

- [Ce87b] H. Cendra, J.E. Marsden, A. Ibort. *Variational principles on fiber bundles: a geometric theory of Clebsch potentials and Lin constraints*, J. Geom. Phys. **4**, 183-206, (1987).
- [Ce09] H. Cendra, M. Delgado-Téllez, A. Ibort, T. Rodríguez de la Peña, *Optimal control realizations of lagrangian systems with symmetry*, (Preprint), (2009).
- [De02] Delgado-Tellez M., Ibort A. *Proceedings of the Fourth International Conference on Dynamical Systems and Differential Equations, On the Geometry and Topology of Singular Optimal Control Problems and their solutions, Wilmington, NC, USA.*, (2002).
- [De03a] M. Delgado-Téllez, A. Ibort. *On the Geometry and Topology of Singular Optimal Control Problems and their solutions*. Discrete and Continuous Dynamical Systems, a suplement volume, 223-333, (2003).
- [De03b] M. Delgado-Téllez, A. Ibort. *A Panorama of Geometrical Optimal Control Theory*. Extracta Mathematicae, Vol. **18** (2), 129-151, (2003).
- [De08] M. Delgado-Téllez, A. Ibort. *A Numerical Algorithm for Singular Optimal LQ Control Systems*, Aceptado en Numerical Algorithms, (2008).
- [De04] M. Delgado-Téllez *Métodos Geométricos en Problemas de Control Óptimo Singulares: Fundamentos y Aplicaciones*, Tesis doctoral, (2004).
- [Di49] Dirac., P.A.M.: Rev. Mod. Phys., **21**, 392, (1949).
- [Di50] P. A. M. Dirac. *Generalized Hamiltonian dynamics*. Canad. J. Math., **2**, 129-148 (1950).
- [Di64] P. A. M. Dirac. *Lectures on Quantum Mechanics*. Belfer Graduate School of Science Monograph Series, Vol. **2**, Yeshiva Univ. (1964).
- [Fl89] A. Floer, *Symplectic fixed points and holomorphic spheres*, Commun. Math. Phys., **120**, 575-611, (1989).
- [Go78] M. J. Gotay, J. M. Nester, G. Hinds. *Presymplectic manifolds and the Dirac-Bergmann Theory of constraints*. J. Math. Phys., **19**, 2388-2399 (1978).
- [Go79] M. J. Gotay, J. M. Nester. *Presymplectic Lagrangian Systems I: The constraint algorithm and the equivalence problem*. Ann. Inst. Henri Poincaré, A **30**, 129-142 (1979).
- [Go80] M. J. Gotay, J. M. Nester. *Presymplectic Lagrangian Systems II: The second order equation problem*. Ann. Inst. Henri Poincaré, A **32**, 1-13 (1980).

- [Gr91] X. Gràcia, J. M. Pons. *Constrained systems: a unified approach*. Int. J. Theor. Phys., **30**, 511-516 (1991).
- [Gr92] Gracia, X.; Pons, J.M.: *A generalized framework for constrained systems*. *Diff. Geom. Appl.*, **2**, 223-247, (1992).
- [Gr00] X. Gràcia, J. Marín-Solano, M. C. Muñoz-Lecanda. *Some geometric aspects of variational calculus in constrained systems*. arXiv: math-ph/0004019 v1 13 Apr 2000.
- [Gr04] X. Gràcia, M. Muñoz-Lecanda, N. Román-Roy. *On some aspects of the Geometry of Differential equations in Physics*. arXiv: math-ph/0402030 v1 12 Feb 2004.
- [Gu84] V. Guillemin, S. Sternberg. *Symplectic techniques in Physics*. Cambridge Univ. Press, (1984).
- [Ib85] A. Ibort. *Estructura geométrica de los sistemas con simetría en Mecánica Clásica y Teoría Clásica de Campos*. Publicaciones del seminario García de Galdeano. Serie II, **7** (1985).
- [Ib91] A. Ibort, J. Marín-Solano, *On the inverse problem of the calculus of variations for a class of coupled dynamical systems*. Inverse Problems, **7**, 1-13 (1991).
- [Ib96] A. Ibort, M. de León, G. Marmo, D. Martín de Diego. *Non-holonomic constrained systems as implicit differential equations*. Rend. Sem. Mat. Univ. Pol. Torino, **54**, 295-317 (1996).
- [Ib99] A. Ibort, M. de León, J. C. Marrero, D. Martín de Diego. *Dirac brackets in constrained dynamics*. *Fortschritte der Physik*. **47** (5), 459-492 (1999).
- [Ib05] A. Ibort, T. Rodríguez. *A family of 3D Lie-Scheffers-Brockett optimal control systems*, comunicación en el congreso CEDYA 2005. Leganés, Madrid, 20 de septiembre de (2005).
- [Ib08] A. Ibort, T. Rodríguez de la Peña. *Clebsch variables, Lin constraints and Quantum Control*, Física Teórica, Julio Abad, **1-11** ,(2008).
- [Ib09] A. Ibort, T. Rodríguez, R. Salmoni *Dirac structures and reduction of optimal control problems*, En preparación (2009).
- [Ju96] V. Jurdjevic. *Geometric Control Theory*. Cambridge Univ. Press, Cambridge (1996).
- [Kl78] W. Klingenberg. *Lectures on closed geodesics*. Springer-Verlag, (1978).
- [La32] H. Lamb. *Hydrodynamics*, 6th ed., Cambridge Univ. Press, (1932).

- [La71] S. Lang. *Differentiable manifolds*, Springer-Verlag, N.Y., (1985).
- [Le96a] M. de León, D. Martín de Diego. *On the geometry of non-holonomic Lagrangian systems*. Journal of Mathematical Physics, **37** (7), 3389-3414 (1996).
- [Le96b] M. de León, J.C. Marrero y D. Martín de Diego. *Time-dependent constrained hamiltonian systems and Dirac brackets*. Journal of Physics A: Math. Gen. **29**, 6843-6859(1996).
- [Le04] M. de León, J. Cortés, D. Martín de Diego, S. Martínez. *General symmetries in optimal control*, Reports on Mathematical Physics 53 **1** 55-78, (2004).
- [Li84] S. Lie, *Vorlesungen über continuierlichen Gruppen mit Geometrischen und anderen Anwendungen*, Edited and revised by G. Scheffers, Teubner, Leipzig, (1893).
- [Li63] C.C.Lin, *Hydrodynamics of helium II*, Proc. Int. Sch. Phys. **21**, 93-146, (1963).
- [Lo89] J C. López. *Estudio Geométrico de Sistemas con Ligaduras*, Tesis doctoral, Departamento de Física Teórica, Universidad de Zaragoza, (1989).
- [Lo99] C. López. Conferencia en: *I Encuentros de Invierno: Física Matemática y Teoría Geométrica de Control*, Zaragoza, (1999).
- [Lo00] C. López, E. Martínez. *Subfinslerian metrics associated to an optimal control problem*, SIAM J. Control Opt. 39(3), 798-811, (2000).
- [Ma74] J.E. Marsden, A. Weinstein. *Reduction of symplectic manifolds with symmetry*. Rep. Math. Phys., **5**, 121-130 , (1974).
- [Ma83] J.E. Marsden, A. Weinstein. *Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids*, Physica D **7**, 305-323, (1983).
- [Ma90] E. Marsden, R. Montgomery, T. Ratiu. *Reduction, symmetry, and phases in mechanics*, Memoirs of the A. M. S., **436**, (1990).
- [Ma95] G. Marmo, G. Mendella, W. Tulczyjew. *Integrability of implicit differential equations*. J. Phys. A: Math. Gen. **28**, 149-163 (1995).
- [Me95] Mendella, G., Marmo, G., Tulczyjew, W. *Integrability of implicit differential equations*, *J. Phys. A: Math. Gen.* **28**, 149-163, (1995).
- [Mi04] M. Mirrahimi, P. Rouchon. *Controllability of Quantum Oscillators*, IEEE Automat. Contr., vol. 49, **5**, 745-747, (2004).

- [Mo90] H.K. Moffatt. *The energy spectrum of knots and links*, Nature, **347**, 359-378, (1990).
- [Mr89] E. Martínez. *Geometría de Ecuaciones Diferenciales aplicada a la Mecánica*, Publ. Sem. García Galdeano, Sec 2, **36**, (1989).
- [Mr93] E. Martínez, J. F. Cariñena, W. Sarlet. *Derivations of differential forms along the tangent bundle projection*. Diff. Geometry and its applications **3** 1-29, (1993).
- [Mr99] E. Martínez. Conferencia en: *I Encuentros de Invierno: Física Matemática y Teoría Geométrica de Control*, Zaragoza, (1999).
- [Mr04] E. Martínez. *Reduction in optimal control theory*, Rep. Math. Phys. **53** 79-90, (2004).
- [Mrd02] S. Martínez Díaz. *Geometric Methods in Nonlinear Control Theory with Applications to Dynamic Robotic Systems*. Ph. D. dissertation, Univ. Carlos III de Madrid, (2002).
- [Mu89] M. C. Muñoz-Lecanda, N. Román-Roy. *Hamiltonian systems with constraints: a geometric approach*. Int. J. Theor. Phys. **28** (11), 1405-1417 (1989).
- [Mu92] M. C. Muñoz-Lecanda, N. Román-Roy. *Lagrangian Theory for presymplectic systems*. Ann. Inst. Henri Poincaré **57** (1), 27-45, (1992).
- [Pe98] Petit P. *On the Optimal Control of Implicit Systems*, ESAIM: Control Optimization and Calculus of Variation , **3**, 49-81, (1998).
- [Po62] L. Pontryagin, V.G. Boltyanski, R.V. Gamkrelidze, E.R. Mishchenko. *The Mathematical Theory of Optimal Processes*. Interscience, N.Y., (1962).
- [Ra94] Rabier, P.J., Rheinboldt, W.C. *A geometric treatment of implicit differential-algebraic equations*, J. Diff. Eqs., **109**, 110-146, (1994).
- [Sk83] R. Skinner, R. Rusk. *Generalized Hamiltonian dynamics I. Formulation on $T^*Q \oplus TQ$* . J. Math. Phys., **24** (11), 2589-2594, (1983).
- [Su99] H. J. Sussmann. *Geometry and Optimal Control*. Mathematical Control Theory. J. Baillieul and J. C. Willems Eds., Springer Verlag, 140-198 (1999).
- [We78] A. Weinstein. *An universal construction for particles in Yang-Mills fields*. Lett. Math. Phys., **2**, 417-420, (1978).

-
- [Yo07a] Yoshimura H., and J. E. Marsden. *Reduction of Dirac Structures and the Hamilton-Pontryagin Principle*, Reports on Mathematical Physics, **60**, 381-426, (2007).
- [Yo07b] Yoshimura, H. and J. E. Marsden. *Dirac structures and the Legendre transformation for implicit Lagrangian and Hamiltonian systems*, Lecture Notes in Control and Inform. Sci., **366**, 233-247, (2007).
- [Yo09] Yoshimura, H. and J. E. Marsden, *Dirac cotangent bundle reduction*, J. Geom. Mech., Vol. 1, **1**, 87–158, (2009).