

2



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO



FACULTAD DE CIENCIAS

Algunas Aplicaciones de los Grupos de Lie a la Teoría de la Relatividad General

T E S I S

QUE PARA OBTENER EL TÍTULO DE M A T E M Á T I C A

P R E S E N T A :

ADRIANA BARRERA FRANCO

**DIRECTOR DE TESIS:
DR. PABLO PADILLA LONGORIA**

2002

DIVISION DE ESTUDIOS PROFESIONALES



**FACULTAD DE CIENCIAS
SECCION ESCOLAR**





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

M. EN C. ELENA DE OTEYZA DE OTEYZA

Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunico a usted que hemos revisado el trabajo escrito:

Algunas Aplicaciones de los Grupos de Lie a la Teoría de la Relatividad
General

realizado por Adriana Barrera Franco

con número de cuenta 09550433-1 , quién cubrió los créditos de la carrera de: Matemáticas

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

**Director de Tesis
Propietario**

Dr. Pablo Padilla Longoria

Propietario

Dr. José Alejandro Ayala Mercado

Propietario

Dr. Enrique Javier Elizondo Huerta

Suplente

M. en C. José Antonio Gómez Ortega

Suplente

M. en C. Manuel Cruz López

Pablo Padilla
Javier Elizondo
997
[Signature]

Consejo Departamental de Matemáticas



997
FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL DE
MATEMÁTICAS

M. en C. José Antonio Gómez Ortega

MATEMÁTICAS

*A mis papás,
los quiero mucho.*

Algunas Aplicaciones de los Grupos de Lie a la Teoría de la Relatividad General

Adriana Barrera Franco

24 de julio de 2002

Índice General

Agradecimientos	v
Introducción	vii
1 Grupos de Lie	1
1.1 Variedades.	1
1.2 Grupos de Lie.	8
1.3 Campos Vectoriales.	14
1.4 Foliaciones y el Teorema de Frobenius.	31
1.5 Álgebras de Lie.	35
2 Simetrías en Ecuaciones	41
2.1 Ecuaciones Algebraicas.	41
2.2 Ecuaciones Diferenciales.	52
2.3 La Ecuación de Onda.	66
3 Grupos de Simetría y Leyes de Conservación	73
3.1 Elementos del Cálculo Variacional.	73
3.2 Ejemplos de Leyes de Conservación: Mecánica Clásica.	77
3.3 Simetrías y Cantidades Conservadas.	82
4 Aplicaciones a la Relatividad General	85
4.1 Relatividad Especial.	85
4.2 Relatividad General.	87
4.3 La Solución de Schwarzschild.	89
4.4 Mecánica de Hoyos Negros.	96
A El Método de Bases Ortonormales para Calcular la Curvatura	99
B Teorema de Frobenius (versión dual) y Campos de Killing	107
B.1 Formas Diferenciales y el Teorema de Frobenius	107
B.2 Mapeo Diferencial para Tensores y Campos de Killing	110

Agradecimientos

Quiero agradecer al Dr Pablo Padilla Longoria por el interés y el entusiasmo que mostró al asesorarme en la realización de esta tesis; por su paciencia, motivación y apoyo; por su manera abierta y sencilla de compartir sus ideas. Para mí ha sido todo un gusto trabajar con él.

Al M. en C. José Antonio Gómez Ortega, al Dr. Alejandro Ayala Mercado, al Dr. Javier Elizondo Huerta y al M. en C. Manuel Cruz López por su disposición y ayuda en la revisión de mi trabajo de tesis.

Al Dr. Oscar Palmas Velasco, no sólo por los conocimientos que me ha transmitido, sino por su amistad.

A los Drs. Michael Barot y Max Newman por su apoyo así como al Dr. Enrique López y la M. en C. Marcela Grethler por darme la oportunidad de iniciar mi experiencia en la docencia.

A mis papás, que todo me lo han dado, por quererme tanto y estar conmigo, por la formación que tengo, por la posibilidad de cumplir mis sueños.

A mis hermanos, Gabriela y Oswaldo, por su cariño.

A Breno por la cercanía y el conocimiento mutuo; por lo que me ha enseñado y lo que hemos aprendido juntos; por tantas cosas que no puedo resumir en unas cuantas líneas.

A Diana por la ayuda académica que siempre me ha dado; por ser la gran persona que es y darme el orgullo de decir que es mi amiga.

A todas las chicas: Valeria, Paulina, Inés, Dení, Brenda y Jazmín, por todo lo que significa su amistad para mí.

A Murat por la fuerte amistad que hemos desarrollado.

A mis compañeros en los grupos de Cálculo y Análisis, a Tere, Rita y Carmen. A todas las personas que conocí durante los últimos semestres y que me demostraron que nunca es tarde para hacer amigos.

A mis vecinas, Verónica, Katia y Tania.

Al personal que labora en el IIMAS y a sus estudiantes, a Sara por la enorme ayuda y a Violeta.

A Lourdes que me ha acompañado en este proceso de titulación y no me ha dejado caer, a Javier Torales, a Acolmán y a Humberto por toda la ayuda, el apoyo y las risas, sé muy bien que sin ustedes este trabajo no sería posible.

A la M. en C. Elena de Oteyza y a Adelina, por haberme ayudado tanto a concluir este proceso.

A la Facultad de Ciencias y a la UNAM.

Con todo mi cariño,

Adriana Barrera Franco.

Introducción

Tanto desde el punto de vista geométrico como físico, el análisis de las simetrías es de gran importancia. El suponer una simetría en determinado espacio permite obtener bastante información, no sólo porque, desde el punto de vista operacional, resulta más sencillo resolver algunos problemas, sino porque en el marco de dicha estructura, hay una serie de consecuencias que resultan ser problemas interesantes en sí mismas. Tal es el caso de la relación entre las simetrías de un problema variacional y ciertas cantidades conservadas, como por ejemplo la energía.

En particular, la relatividad general es una teoría donde la estructura geométrica del espacio-tiempo y los aspectos mecánicos se funden por medio de la ecuación de Einstein. Mientras que en la mecánica newtoniana y la Teoría Especial de la Relatividad las características del espacio-tiempo están dadas y forman un marco de referencia para cualquier sistema mecánico, en la Teoría General de la Relatividad uno debe preguntarse cómo es el espacio-tiempo al mismo tiempo que se pregunta por el comportamiento mecánico de determinado objeto. En este sentido, es de suponer que la presencia de simetrías en el espacio-tiempo tenga notables implicaciones en cualquier sistema físico que se estudie. La mecánica de hoyos negros es una rama de la relatividad general que ha tenido gran desarrollo en las últimas décadas y que ilustra muy bien estas ideas; las tres leyes de los hoyos negros, que en principio se deducen bajo ciertas suposiciones de simetría en la geometría del espacio-tiempo, parecen tener un importante contenido físico.

En este trabajo se pretende hacer un estudio detallado del lenguaje de cierto tipo de simetrías, para luego ilustrar una pequeña parte de las consecuencias que tiene suponer algunas de ellas en el marco de la Teoría General de la Relatividad.

Para hablar de simetrías lo natural es usar la teoría de grupos; si se quiere hablar de simetrías continuas, los grupos de Lie resultan una herramienta indispensable. Tomando en cuenta lo anterior, en el capítulo 1 desarrollamos el caso de los grupos monoparamétricos y sus propiedades.

En el capítulo 2, siguiendo la intención original de Lie, hacemos una aplicación de la teoría de grupos de Lie a las ecuaciones diferenciales, estableciendo para ello la definición precisa de un grupo de simetría de un sistema de ecuaciones y algunos criterios.

En el capítulo 3 se establece la relación entre los grupos de simetría y las cantidades

conservadas.

En el último capítulo se hace una breve aplicación de las ideas manejadas en los tres anteriores, tomando a la Teoría General de la Relatividad y, dentro de ella, a la mecánica de hoyos negros.

Dado que este trabajo gira en torno a las simetrías continuas y su lenguaje, se omite la introducción del formalismo tensorial, así como las definiciones precisas relacionadas con las ideas de métrica, curvatura, transporte paralelo y derivada covariante. Todos estos conceptos (que aparecen en el capítulo 4 y los dos apéndices) así como la notación de los mismos, se dan por conocidos y se refiere al lector al libro de Wald [10].

Capítulo 1

Grupos de Lie

En este capítulo desarrollamos las herramientas matemáticas de grupos de Lie que emplearemos más adelante. Para esto, partimos desde la definición de variedad, de modo que gran parte del material representa un breve repaso de la teoría de variedades y grupos de Lie. Dentro de los resultados aquí presentados, destacan el Teorema del Rango y el Teorema de Frobenius, además de los conceptos de grupos de transformaciones, en concreto, los grupos monoparamétricos de difeomorfismos.

1.1 Variedades.

Como una primera aproximación, podemos pensar que las variedades son espacios topológicos obtenidos a partir de “pegar” copias del espacio euclideo \mathbb{R}^n . Hay varias maneras de realizar el “pegado”: a través del uso de homeomorfismos, usando difeomorfismos y mediante el empleo de homeomorfismos analíticos, en cuyo caso llamamos a las variedades resultantes variedades topológicas, variedades diferenciables (o suaves) y variedades analíticas, respectivamente. En nuestro caso, trabajaremos siempre con variedades diferenciables, de modo que, cada vez que hagamos referencia al término “variedad”, entenderemos con ello que se trata de una variedad diferenciable. Demos ahora una definición más precisa de una variedad.

Definición 1.1 Sea M un espacio topológico y \mathbb{R}^m un espacio euclideo m -dimensional. Una *carta* (o *vecindad coordenada*) de M es un par (U, χ) que consiste de un subconjunto abierto U de M y un homeomorfismo χ de U a un subconjunto abierto de \mathbb{R}^m ($\chi : U \subset M \rightarrow \mathbb{R}^m$).

Si para $p \in U \subset M$ escribimos $\chi(p)$ como:

$$\chi(p) = (x^1(p), \dots, x^m(p)),$$

decimos que $x^i(p) : U \rightarrow \mathbb{R}$ es la *i-ésima función coordenada* y que (U, χ) es una *carta* en p . También puede hacerse referencia a una carta en p al denotar las funciones coordenadas $(x^1(p), \dots, x^m(p))$ y llamarlas *coordenadas locales* en p .

Una colección $\{(U_\alpha, \chi_\alpha)\}$ de cartas se llama un *atlas* (o *sistema coordenado*) en M si $\{U_\alpha\}$ forma una cubierta abierta de M . El atlas es *suave* si para cada par de índices α y β , con $U_\alpha \cap U_\beta \neq \emptyset$, el mapeo

$$\chi_\beta \circ \chi_\alpha^{-1} : \chi_\alpha(U_\alpha \cap U_\beta) \rightarrow \chi_\beta(U_\alpha \cap U_\beta)$$

es suave (i.e. de clase C^∞). Esta definición se ilustra en la figura 1.1.

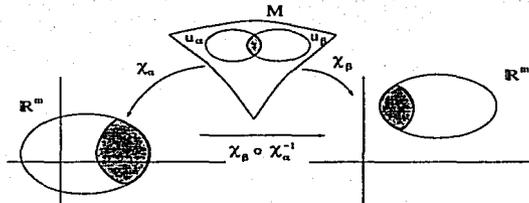


Figura 1.1: Un atlas suave.

Es claro que $\chi_\beta \circ \chi_\alpha^{-1}$ es un difeomorfismo (una función de clase C^∞ de \mathbb{R}^m a \mathbb{R}^m con inversa suave) ya que $(\chi_\beta \circ \chi_\alpha^{-1})^{-1} = \chi_\alpha \circ \chi_\beta^{-1} \in C^\infty$ por definición de atlas suave.

Definición 1.2 Decimos que dos atlas suaves $\{(U_\alpha, \chi_\alpha)\}$ y $\{(V_\beta, \nu_\beta)\}$ en M son *equivalentes* si para cada par de índices α y β , con $U_\alpha \cap V_\beta \neq \emptyset$, el mapeo

$$\nu_\beta \circ \chi_\alpha^{-1} : \chi_\alpha(U_\alpha \cap V_\beta) \rightarrow \nu_\beta(U_\alpha \cap V_\beta)$$

es un difeomorfismo. Una clase de equivalencia de atlas suaves en M se llama *estructura suave* o *diferencial* en M . La estructura suave representada por el atlas $\{(U_\alpha, \chi_\alpha)\}$ se denota como $[(U_\alpha, \chi_\alpha)]$. Una *variedad suave* o *diferenciable de dimensión m* (o *variedad m-dimensional*, o simplemente *variedad*) es un par $(M, [(U_\alpha, \chi_\alpha)])$ que consiste en un espacio Hausdorff M segundo numerable y una estructura suave $[(U_\alpha, \chi_\alpha)]$ en M .

Generalmente denotamos la variedad $(M, [(U_\alpha, \chi_\alpha)])$ como M , entendiendo que existe una estructura suave $[(U_\alpha, \chi_\alpha)]$ en M .

Existen varios ejemplos de variedades, el más sencillo lo constituyen los espacios Euclidianos \mathbb{R}^n . También las n -esferas (S^n) y los espacios proyectivos (P^n) son variedades suaves.

Es posible construir nuevas variedades a partir de variedades anteriores empleando varios métodos, por ejemplo:

- Un subconjunto abierto O de una variedad $(M, [(U_\alpha, \chi_\alpha)])$ constituye una *subvariedad abierta* de M con una estructura dada por $[(O \cap U_\alpha, \chi_\alpha|_{O \cap U_\alpha})]$.

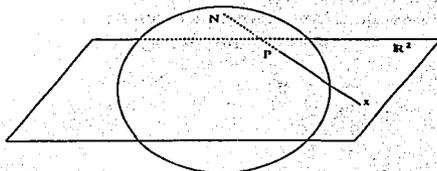


Figura 1.2: La proyección estereográfica como un sistema de coordenadas de la esfera a \mathbb{R}^2 .

- Dadas dos variedades suaves $(M, [(U_\alpha, \chi_\alpha)])$ y $(N, [(V_\beta, \nu_\beta)])$, el producto cartesiano $M \times N$ es una variedad suave con la estructura $[(U_\alpha \times V_\beta, \chi_\alpha \times \nu_\beta)]$.

En términos generales, una subvariedad N de una variedad M ($N \subset M$) es un subespacio topológico de M que adquiere su estructura de variedad a partir de la de M . En particular, las subvariedades regulares (ilustradas en la figura 1.3) son de gran importancia.

Definición 1.3 Sea M una variedad m -dimensional, N un subconjunto de M , y $n \in \mathbb{N}$ tal que $0 \leq n \leq m$. Si para cada $p \in N$ existe una carta (U, χ) de M en p tal que:

$$N \cap U = \{a \in U \mid x^{n+1}(a) = \dots = x^m(a) = 0\}$$

entonces N es una subvariedad regular de M .

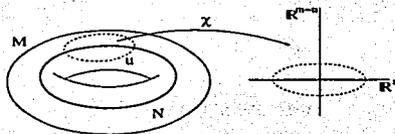


Figura 1.3: Una subvariedad regular.

Muchos conceptos del cálculo en \mathbb{R}^n pueden extenderse a las variedades de manera natural. Así, la noción de suavidad se extiende de la siguiente manera:

Definición 1.4 Un mapeo continuo $F : M \rightarrow N$ entre dos variedades suaves $(M, [(U, \chi_\alpha)])$ y $(N, [(V, \nu_\beta)])$ es suave (o C^∞ o diferenciable) si para cada par de índices α y β , con $U_\alpha \cap F^{-1}(V_\beta) \neq \emptyset$, el mapeo

$$\nu_\beta \circ F \circ \chi_\alpha^{-1} : \chi_\alpha(U_\alpha \cap F^{-1}(V_\beta)) \rightarrow \nu_\beta(V_\beta)$$

es suave o C^∞ como función de \mathbb{R}^m a \mathbb{R}^n (ver figura 1.3). Un mapeo suave $F : M \rightarrow N$ entre variedades suaves es un *difeomorfismo* si es un homeomorfismo y el mapeo inverso F^{-1} también es suave; se dice entonces que M y N son *difeomorfas*.

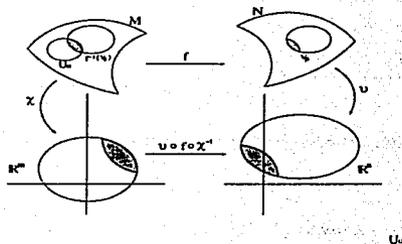


Figura 1.4: Un mapeo suave entre dos variedades.

Pese a que este tipo de definiciones y generalizaciones hacen referencia a un sistema coordenado, es fácil notar que, dado que estamos tratando con variedades diferenciables, no dependen de las coordenadas, lo cual nos da la libertad de elegir el sistema coordenado que resulte conveniente en cada caso.

Definición 1.5 Sea $F : M \rightarrow N$ un mapeo suave entre dos variedades $(M, \{(U, \chi_\alpha)\})$ y $(N, \{(V, \nu_\beta)\})$, con dimensiones m y n respectivamente. El *rango de F* en un punto $p \in M$ es el rango de la matriz jacobiana de F evaluada en el punto p , es decir: dadas cartas (U, χ) en p y (V, ν) en $F(p) = q$, expresadas como:

$$\chi(p) = (x^1(p), \dots, x^m(p))$$

y

$$\nu(q) = (y^1(q), \dots, y^n(q))$$

respectivamente, el rango de F está dado por el rango de la siguiente matriz:

$$J(F)(p) = \begin{pmatrix} \frac{\partial}{\partial x^1} y^1 \circ F \circ \chi^{-1} \Big|_{\chi(p)} & \dots & \frac{\partial}{\partial x^m} y^1 \circ F \circ \chi^{-1} \Big|_{\chi(p)} \\ \vdots & & \vdots \\ \frac{\partial}{\partial x^1} y^n \circ F \circ \chi^{-1} \Big|_{\chi(p)} & \dots & \frac{\partial}{\partial x^m} y^n \circ F \circ \chi^{-1} \Big|_{\chi(p)} \end{pmatrix}$$

El mapeo F es de *rango máximo* en un subconjunto $S \subset M$ si $\forall p \in S$ el rango de F es igual al mínimo entre m y n (el máximo valor que puede alcanzar el rango de la matriz $J(F)(p)$).

Es sencillo ver que esta definición no depende del sistema coordenado que escojamos. Al establecer condiciones sobre el rango de un mapeo suave, podemos garantizar la existencia de un sistema coordenado que nos permita expresar a dicho mapeo en su forma más simple:

Teorema 1.6 (del Rango).

Sea $F : M \rightarrow N$ un mapeo suave entre variedades con dimensiones m y n respectivamente, y sean $p_0 \in M$, $F(p_0) = q_0 \in N$. Si F es de rango máximo en p_0 , entonces existen cartas (U, χ) en p_0 , $\chi(p_0) = (x^1(p_0), \dots, x^m(p_0))$, y (V, ν) en q_0 , $\nu(q_0) = (y^1(q_0), \dots, y^n(q_0))$, tales que, $\forall p \in U$, con $F(p) = q \in V$, se tiene:

$$\begin{aligned} \nu(q) = \nu(F(p)) &= (x^1(p), \dots, x^m(p), 0, \dots, 0) && \text{si } m \leq n && \text{y} \\ \nu(q) = \nu(F(p)) &= (x^1(p), \dots, x^n(p)) && \text{si } n \leq m. && \end{aligned} \quad (1.1)$$

Prueba.

Primero demostraremos el teorema del rango para el caso especial en el que las variedades corresponden a espacios euclidianos \mathbb{R}^m y \mathbb{R}^n . En este caso, lo que dice el teorema es que, si F es de rango máximo en $p_0 \in \mathbb{R}^m$, existen vecindades de p_0 y de $F(p_0) = q_0 \in \mathbb{R}^n$ y difeomorfismos $\chi : U \rightarrow \mathbb{R}^m$, $\nu : V \rightarrow \mathbb{R}^n$, tales que, $\forall p \in U$, con $F(p) = q \in V$, se cumple:

$$\begin{aligned} \nu \circ F \circ \chi^{-1}(p) &= \nu \circ F \circ \chi^{-1}(x^1, \dots, x^m) = (x^1, \dots, x^m, 0, \dots, 0) \\ & \hspace{15em} \text{si } m \leq n && \text{y} \\ \nu \circ F \circ \chi^{-1}(p) &= \nu \circ F \circ \chi^{-1}(x^1, \dots, x^m) = (x^1, \dots, x^n) \\ & \hspace{15em} \text{si } n \leq m. \end{aligned}$$

Es decir, que el mapeo se comporta localmente como una inclusión ($m \leq n$) o como una proyección ($n \leq m$). Comencemos con el caso $m \leq n$.

La matriz jacobiana de $F = (F^1, \dots, F^n)$ en p_0 es:

$$J(F)(p) = \begin{pmatrix} \left. \frac{\partial}{\partial x^1} F^1 \right|_{p_0} & \cdots & \left. \frac{\partial}{\partial x^m} F^1 \right|_{p_0} \\ \vdots & & \vdots \\ \left. \frac{\partial}{\partial x^1} F^n \right|_{p_0} & \cdots & \left. \frac{\partial}{\partial x^m} F^n \right|_{p_0} \end{pmatrix}.$$

Como F es de rango máximo en p_0 , existe una submatriz de $m \times m$ de $J(F)(p_0)$ cuyo determinante es distinto de cero. Sin pérdida de generalidad, tomando en cuenta mapeos lineales que permutan coordenadas, podemos suponer que

$$A = \begin{pmatrix} \left. \frac{\partial}{\partial x^1} F^1 \right|_{p_0} & \cdots & \left. \frac{\partial}{\partial x^m} F^1 \right|_{p_0} \\ \vdots & & \vdots \\ \left. \frac{\partial}{\partial x^1} F^m \right|_{p_0} & \cdots & \left. \frac{\partial}{\partial x^m} F^m \right|_{p_0} \end{pmatrix}$$

es dicha submatriz. Definimos $\chi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ como:

$$\chi(x^1, \dots, x^m) = (F^1(x^1, \dots, x^m), \dots, F^m(x^1, \dots, x^m)),$$

entonces,

$$J(\chi)(p_0) = \begin{pmatrix} \frac{\partial}{\partial x^1} F^1 \Big|_{p_0} & \dots & \frac{\partial}{\partial x^m} F^1 \Big|_{p_0} \\ \vdots & & \vdots \\ \frac{\partial}{\partial x^1} F^m \Big|_{p_0} & \dots & \frac{\partial}{\partial x^m} F^m \Big|_{p_0} \end{pmatrix} = A,$$

de modo que $\det(J(\chi)(p_0)) \neq 0$. Por teorema de la función inversa, existen vecindades abiertas de p_0 y de $\chi(p_0)$ en las que χ es un difeomorfismo y, por lo tanto, existe $\chi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ suave.

Del mismo modo, definimos $v^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ como:

$$v^{-1}(x^1, \dots, x^m, x^{m+1}, \dots, x^n) = (x^1, \dots, x^m, x^{m+1} + F^{m+1}(x^1, \dots, x^m), \dots, x^n + F^n(x^1, \dots, x^m)),$$

cuya matriz jacobiana en cualquier punto es:

$$\begin{pmatrix} \frac{\partial}{\partial x^1} x^1 & \dots & \frac{\partial}{\partial x^m} x^1 & \dots & \frac{\partial}{\partial x^{m+1}} x^1 & \dots & \frac{\partial}{\partial x^n} x^1 \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial}{\partial x^1} x^m & \dots & \frac{\partial}{\partial x^m} x^m & \dots & \frac{\partial}{\partial x^{m+1}} x^m & \dots & \frac{\partial}{\partial x^n} x^m \\ \frac{\partial}{\partial x^1} x^{m+1} + F^{m+1} & \dots & \frac{\partial}{\partial x^m} x^{m+1} + F^{m+1} & \dots & \frac{\partial}{\partial x^{m+1}} x^{m+1} + F^{m+1} & \dots & \frac{\partial}{\partial x^n} x^{m+1} + F^{m+1} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial}{\partial x^1} x^n + F^n & \dots & \frac{\partial}{\partial x^m} x^n + F^n & \dots & \frac{\partial}{\partial x^{m+1}} x^n + F^n & \dots & \frac{\partial}{\partial x^n} x^n + F^n \end{pmatrix},$$

la cual, dado que las funciones F^i , con $i = m+1, \dots, n$, sólo dependen de las primeras m coordenadas, es igual a:

$$\begin{pmatrix} I_{m \times m} & 0 \\ * & I_{n \times n} \end{pmatrix},$$

donde $I_{m \times m}$ denota la matriz identidad de $m \times m$, similar para $I_{n \times n}$ y $*$ es una submatriz cuyo valor no influye en nuestros cálculos. Claramente, $\det(J(v)(q)) = 1 \neq 0 \quad \forall q$, así que, aplicando el teorema de la función inversa, sabemos que existen vecindades de q y de $v^{-1}(q)$ en donde v^{-1} es un difeomorfismo y, por ello, tenemos $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ suave.

Si denotamos por ι a la inclusión de \mathbb{R}^m en \mathbb{R}^n , obtenemos que:

$$\iota \circ \chi(x^1, \dots, x^m) = \iota(F^1, \dots, F^m) = (F^1, \dots, F^m, 0, \dots, 0),$$

de donde:

$$v^{-1} \circ \iota \circ \chi(x^1, \dots, x^m) = v(F^1, \dots, F^m, 0, \dots, 0) = (F^1, \dots, F^m, F^{m+1}, \dots, F^n) = F(x^1, \dots, x^m).$$

Así pues, $F = v^{-1} \circ \iota \circ \chi$ y, componiendo con las funciones inversas χ^{-1} y v , tenemos:

$$v \circ F \circ \chi^{-1}(x^1, \dots, x^m) = \iota(x^1, \dots, x^m) = (x^1, \dots, x^m, 0, \dots, 0).$$

Por lo tanto, existen difeomorfismos χ y v y vecindades U y V (dadas por el teorema de la función inversa) tales que:

$$v \circ F \circ \chi^{-1}(x^1, \dots, x^m) = (x^1, \dots, x^m, 0, \dots, 0).$$

Si $n \leq m$, la Prueba es bastante similar, definiendo $\chi: \mathbb{R}^m \rightarrow \mathbb{R}^n$ como $\chi(x^1, \dots, x^m) = (F^1, \dots, F^n, x^{n+1}, \dots, x^m)$ y a $v: \mathbb{R}^n \rightarrow \mathbb{R}^n$ como la función identidad.

Para el caso general en el que F es un mapeo entre variedades M y N , consideramos que (U, χ) y (V, ψ) son cartas en p_0 y $q_0 = F(p_0)$, respectivamente; entonces, el mapeo $\psi \circ F \circ \chi^{-1}$ es una función entre espacios euclidianos \mathbb{R}^m y \mathbb{R}^n , de modo que podemos aplicar el teorema del rango en el caso que acabamos de demostrar. Así, sabemos que existen difeomorfismos $\tilde{\chi}: \tilde{U} \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ y $\tilde{\psi}: \tilde{V} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, tales que:

$$\begin{aligned} \tilde{\psi} \circ (\psi \circ F \circ \chi^{-1}) \circ \tilde{\chi}^{-1}(x^1, \dots, x^m) &= (x^1, \dots, x^m, 0, \dots, 0) \\ &\quad \text{si } m \leq n \quad \text{y} \\ \tilde{\psi} \circ (\psi \circ F \circ \chi^{-1}) \circ \tilde{\chi}^{-1}(x^1, \dots, x^m) &= (x^1, \dots, x^n) \\ &\quad \text{si } n \leq m. \end{aligned}$$

Por lo que existen cartas en p_0 y en q_0 dadas por $\chi = \tilde{\chi} \circ \tilde{\chi}^{-1}$ y $\psi = \tilde{\psi} \circ \tilde{\psi}^{-1}$, junto con las vecindades U y V correspondientes, que cumplen el teorema del rango. \square

La anterior Prueba nos puede servir para ilustrar cómo se trabaja con coordenadas y el hecho de que las variedades sean espacios localmente homeomorfos a \mathbb{R}^n . En general, se trabaja indistintamente con los puntos en la variedad, $p \in M$, o directamente con sus coordenadas en \mathbb{R}^m , $x = (x^1, \dots, x^m) = \chi(p)$, considerando que "localmente son iguales". Por ejemplo, si tenemos un mapeo $F: M \rightarrow N$, considerar $\left. \frac{\partial}{\partial x^i} F^j \right|_p$ equivale a tomar $\left. \frac{\partial}{\partial x^i} y^j \circ F \circ \chi^{-1} \right|_x$.

Hay que notar que el Teorema del Rango no sólo garantiza la expresión sencilla de un mapeo suave en términos de ciertas coordenadas locales, sino que establece una condición para generar subvariedades regulares:

- Si $m \leq n$, F es un mapeo inyectivo y de rango máximo en M , entonces $F(M)$ es una subvariedad (regular) de N . A F se le conoce como *inmersión* y a $F(M)$ como una *subvariedad inmersa* en N .

- Si $n \leq m$, $q \in N$, F es suprayectiva y de rango máximo en $F^{-1}(q)$, entonces $F^{-1}(q)$ es una subvariedad (regular) de M .

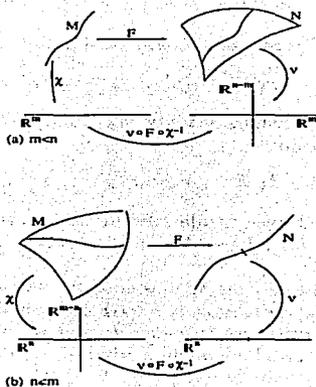


Figura 1.5: Ilustración del Teorema del Rango.

Otro concepto que también se hereda del cálculo es el de una curva sobre una variedad. Así, una *curva suave* en una variedad M es un mapeo suave $\phi : I \rightarrow M$, donde I es un intervalo en \mathbb{R} . Ahora bien, como cualquier espacio topológico, una variedad es *conexa* si no hay dos conjuntos abiertos que la separen. Dado que cualquier variedad es localmente difeomorfa a un espacio euclidiano, entonces las variedades comparten todas las propiedades locales de \mathbb{R}^n , en particular la conexidad y la compacidad locales. Siguiendo este razonamiento, no es difícil probar, mediante vecindades coordinadas, que cualquier variedad conexa es conexa por trayectorias, es decir que, para cualquier par de puntos en la variedad, existe una curva suave que los une. En nuestro caso, consideraremos siempre que nuestras variedades son conexas.

1.2 Grupos de Lie.

En una primera aproximación, los grupos de Lie parecen la simple unión de conceptos provenientes de dos ramas de las matemáticas: por el lado del álgebra tenemos la idea de grupo, por el lado de la geometría tenemos a las variedades. Como veremos más adelante, esta unión, que en principio parece "poco natural", resulta ser una poderosa herramienta en el estudio de las simetrías, permitiendo resolver una gran variedad de problemas, tanto en las cuestiones algebraicas, como en las geométricas. Primero recordemos el concepto de grupo:

Definición 1.7 Un grupo es un par ordenado $(G, *)$, donde G es un conjunto y $*$ es una operación binaria en G , es decir, $*$: $G \times G \rightarrow G$. En general, $\forall g, h \in G$, se escribe $*(g, h) \in G$ como $g * h$, o simplemente gh , y a la operación $*$ se le suele llamar producto en G . Además, deben cumplirse las siguientes condiciones:

- 1) El producto es asociativo: $(gh)f = g(hf) \quad \forall g, h, f \in G$.
- 2) Existe un único elemento e en G , llamado *idéntico* o *identidad*, tal que $ge = eg = g \quad \forall g \in G$.
- 3) $\forall g \in G \quad \exists g^{-1} \in G$ tal que $gg^{-1} = g^{-1}g = e$. A g^{-1} se le conoce como el inverso de g .

Existen varios ejemplos de grupos, desde aquellos que se refieren a sistemas de números (\mathbb{Z} , \mathbb{Q} y \mathbb{R} son grupos con la suma como operación binaria, por mencionar algunos), hasta otros que ilustran claramente cómo la estructura de grupo es la adecuada para hablar de simetrías (el grupo de movimientos de un polígono regular respecto a sus ejes de simetría con la composición como operación, por citar uno).

La característica particular de los grupos de Lie es que son grupos que también tienen una estructura de variedad, de modo que ambas estructuras son compatibles.

Definición 1.8 Un *grupo de Lie* es un conjunto G tal que es una variedad suave y G es también un grupo cuyas operaciones son suaves, es decir que $\forall g, h \in G$, el producto en G

$$\begin{aligned} * : G \times G &\rightarrow G \\ (g, h) &\rightarrow gh \end{aligned}$$

y la inversión

$$\begin{aligned} i : G &\rightarrow G \\ g &\rightarrow g^{-1}, \end{aligned}$$

son mapeos suaves entre variedades (de hecho, la inversión es un difeomorfismo). Aquí, la estructura de variedad de $G \times G$ está dada por la estructura de producto cartesiando de variedades.

Nuevamente, podemos definir y caracterizar a los mapeos entre grupos de Lie, por ejemplo, un *homomorfismo* entre grupos de Lie G y H , es un mapeo suave $\varphi : G \rightarrow H$ que respeta las operaciones de grupo:

$$\varphi(g_1 g_2) = \varphi(g_1) \varphi(g_2) \quad \forall g_1, g_2 \in G.$$

Si φ es un homomorfismo que además es un difeomorfismo (como mapeo entre variedades), entonces φ es un *isomorfismo* entre G y H . Generalmente no se hacen distinciones entre grupos de Lie que son isomorfos.

Hay numerosos ejemplos de grupos de Lie, a pesar de que en principio podríamos pensar que no es así. De entrada, los ejemplos más sencillos vuelven a ser los espacios

euclidianos \mathbb{R}^n , aunque existen también otros que no son tan triviales:
El grupo general lineal con coeficientes reales,

$$\mathbf{GL}(n, \mathbb{R}) = \{A \in \mathbf{M}(\mathbb{R})_{n \times n} \mid \det \neq 0\}.$$

Podemos identificar a las matrices cuadradas de $n \times n$ con coeficientes reales con el espacio euclideo \mathbb{R}^{n^2} , de modo que $\mathbf{M}(\mathbb{R})_{n \times n}$ es una variedad. Claramente, el mapeo determinante $\det : \mathbf{M}(\mathbb{R})_{n \times n} \rightarrow \mathbb{R}$ es continuo. Dado que $\mathbb{R}^* (= \mathbb{R} - \{0\})$ es un conjunto abierto de \mathbb{R} y que $\mathbf{GL}(n, \mathbb{R}) = \det^{-1}(\mathbb{R}^*)$, entonces $\mathbf{GL}(n, \mathbb{R})$ es un subconjunto abierto de $\mathbf{M}(\mathbb{R})_{n \times n}$ y, por lo tanto, una subvariedad de $\mathbf{M}(\mathbb{R})_{n \times n}$. Además, $\mathbf{GL}(n, \mathbb{R})$ es un grupo con la multiplicación de matrices como operación. El producto AB de dos elementos A y $B \in \mathbf{GL}(n, \mathbb{R})$ tiene entradas que son polinomios en las entradas de A y B , las cuales son precisamente las expresiones en coordenadas del mapeo producto, de modo que éste es suave. Las entradas de la matriz inversa de A , A^{-1} , son funciones racionales de las entradas de A cuyos denominadores no se anulan, de modo que son mapeos suaves. Por lo tanto, $\mathbf{GL}(n, \mathbb{R})$ es un grupo de Lie.

Un subgrupo H de un grupo de Lie G se convierte naturalmente en un grupo de Lie cuando está equipado con la estructura de variedad heredada a partir de la de G . Pensando en G como espacio topológico, si H es cerrado, abierto, etc., en G , a H se le llama *subgrupo cerrado*, *subgrupo abierto*, etc., respectivamente. Particularmente, si H es un subgrupo cerrado de G , entonces H es un grupo de Lie.¹ Así, también puede demostrarse, directamente o usando la afirmación anterior, que los siguientes grupos de matrices son grupos de Lie:

$$\begin{aligned} \mathbf{O}(n) &= \{A \in \mathbf{GL}(n, \mathbb{R}) \mid AA^t = I\}, \\ \mathbf{SL}(n, \mathbb{R}) &= \{A \in \mathbf{GL}(n, \mathbb{R}) \mid \det = 1\}, \\ \mathbf{SO}(n) &= \mathbf{O}(n) \cap \mathbf{SL}(n, \mathbb{R}). \end{aligned}$$

Aquí A^t denota la matriz transpuesta de A .

Muchas veces no estamos interesados en todo un grupo de Lie, sino sólo en aquellos elementos que están cerca del elemento idéntico. En este caso, tendremos subvariedades abiertas y conexas de un grupo de Lie que forman vecindades del elemento identidad y únicamente nos preocupamos por verificar que los axiomas usuales de grupo de Lie estén definidos y se cumplan en dichas vecindades, las cuales se conocen como *grupos locales de Lie*. Inversamente, es posible partir de un grupo local de Lie y encontrar un grupo global de Lie, conexo, que contiene al grupo local y preserva su estructura.

¹La Prueba de esta afirmación requiere un poco más de herramienta, en concreto, del mapeo exponencial (que veremos más adelante) y es un poco larga, pero puede consultarse en Kawakubo [5].

Propiedad 1.9

Sea G un grupo de Lie conexo y U una vecindad de la identidad. Sea $U^n = \{g_1 g_2 \dots g_n \mid g_i \in U\}$ el conjunto de productos de n elementos de U , entonces

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

Es decir, que cada elemento de G puede escribirse como un producto finito de elementos de U .

Prueba.

Sea V un subconjunto abierto de U que contiene a la unidad y tal que $V = V^{-1}$ (donde $V^{-1} = \{g^{-1} \mid g \in U\}$), por ejemplo $V = U \cap U^{-1}$. Sea

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

entonces H es un subgrupo de G y es también un subconjunto abierto de G , puesto que $\forall h \in H$, hV es una vecindad abierta de h contenida en H . Además, dado que H es abierto, cada clase lateral gH es un abierto en G , $\forall g \in G$. H es el complemento en G de la unión de todas las clases laterales distintas de H , de modo que H es también cerrado en G . Como G es conexo y $H \neq \emptyset$ (al menos la identidad está en H), entonces $H = G$. Por lo tanto,

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

□

Para un grupo cualquiera, no necesariamente de Lie, se define la *acción* del grupo sobre un conjunto S como sigue:

Definición 1.10 Sea G un grupo y S un conjunto. Definimos una *acción* del grupo G en el conjunto S como un mapeo $\cdot : G \times S_j \rightarrow S$ (denotado como $g \cdot s \forall g \in G$ y $\forall s \in S$), que satisface las siguientes propiedades:

- 1) $g_1 \cdot (g_2 \cdot s) = (g_1 g_2) \cdot s \quad \forall g_1, g_2 \in G$ y $\forall s \in S$.
- 2) $e \cdot s = s \quad \forall s \in S$.

Cuando existe una acción de G en S , se dice que G *actúa* en S . Para cada $s \in S$, se define la órbita de s como el conjunto

$$O_s = \{g \cdot s \mid g \in G\} \subseteq S.$$

Este concepto es importante ya que la herramienta de los grupos de Lie resulta de particular interés cuando tenemos uno de estos grupos actuando sobre una variedad.

Definición 1.11 Sea G un grupo de Lie y M una variedad. Si un mapeo suave $\cdot : G \times M \rightarrow M$ es una acción de G en M , entonces el mapeo es una *acción suave*; donde nuevamente estamos pensando en que la estructura de variedad de $G \times M$ está dada por la estructura de un producto cartesiano de variedades. El triplete (G, M, \cdot) se conoce como un *grupo de transformaciones suaves* en M y de hecho a M puede llamársele también una G -variedad.

Por ejemplo, dado un vector $a \in \mathbb{R}^n$, $a \neq 0$ fijo, y tomando a \mathbb{R} como grupo de Lie, definimos una acción de \mathbb{R} en \mathbb{R}^n como sigue:

$$t \cdot x = x + ta \quad \forall x \in \mathbb{R}^n, \quad \forall t \in \mathbb{R}.$$

Claramente \cdot es una acción suave, ya que:

$$t_1 \cdot (t_2 \cdot x) = t_1 \cdot (x + t_2 a) = x + (t_1 + t_2)a = (t_1 + t_2) \cdot x$$

$$\text{y } 0 \cdot x = x + 0a = x.$$

De hecho, esto se conoce como el *grupo de traslaciones* en \mathbb{R}^n . Para cada $x \in \mathbb{R}^n$, su órbita está dada por una recta paralela al vector a que pasa por x .

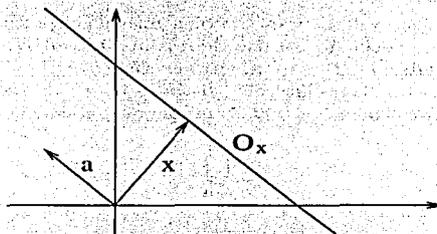


Figura 1.6: Acción del grupo de traslaciones en $x \in \mathbb{R}^2$.

Otro ejemplo interesante de un grupo de transformaciones que actúa en \mathbb{R}^n es $\text{SO}(n)$. En particular, vemos que $\text{SO}(2)$ corresponde al grupo de rotaciones en el plano:

$$\text{SO}(2) = \left\{ \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad 0 \leq \theta \leq 2\pi \right\},$$

con θ el ángulo de rotación. Para cada punto $(x, y) \in \mathbb{R}^2$, su órbita corresponde a un círculo de radio r , ($r^2 = x^2 + y^2$).

Con esto tenemos otra muestra de que los grupos nos proveen del lenguaje adecuado para hablar de simetrías, sólo que ahora podemos trabajar también con las ideas de continuidad y diferenciabilidad.

Podemos tener grupos de Lie actuando únicamente en regiones de la variedad, en cuyo caso estaremos hablando de *grupos locales de transformaciones*, en concreto:

Definición 1.12 Sea M una variedad suave. Un *grupo local de transformaciones* actuando en M está dado por un grupo de Lie G (con idéntico e), un subconjunto U tal que:

$$\{e\} \times M \subseteq U \subseteq G \times M$$

y un mapeo suave $\cdot : U \rightarrow M$ que, $\forall p \in M$, cumple con lo siguiente:

- 1) Si $(h, p), (g, h \cdot p)$ y $(gh, p) \in U \Rightarrow g \cdot (h \cdot p) = gh \cdot p$.
- 2) $e \cdot p = p$.
- 3) Si $(g, p) \in U \Rightarrow (g^{-1}, g \cdot p) \in U$.

Notemos entonces que si tenemos un grupo local de transformaciones actuando en M y para cada $p \in M$ definimos:

$$G_p = \{g \in G \mid (g, p) \in U\} \subseteq G,$$

entonces G_p es un grupo local de Lie que actúa en una vecindad de p . De hecho, a los grupos G_p se les suele llamar grupos locales de transformaciones, puesto que son grupos locales de Lie que actúan en subvariedades definidas por una vecindad del punto p .

Para grupos locales de transformaciones, las órbitas están dadas por productos de elementos del grupo en los cuales está definida la acción:

$$O_p = \{g_1 g_2 \dots g_n \cdot p \mid 1 \leq n, g_i \in G \text{ y } g_1 g_2 \dots g_n \cdot p \text{ está definido}\}.$$

Ciertamente, las órbitas de un grupo de transformaciones son subvariedades de M , no necesariamente regulares y con dimensiones que pueden variar. En referencia a esto, hay dos clases importantes de acciones de grupos de Lie:

Definición 1.13 Sea G un grupo local de transformaciones actuando en M , entonces:

- 1) G actúa semiregularmente si todas las órbitas son de la misma dimensión, como subvariedades de M .
- 2) G actúa regularmente si la acción es semiregular y $\forall p \in M \exists U \subseteq M$ una vecindad con la propiedad de que cada órbita de G intersecta a U en un subconjunto conexo.

Recordando nuestros ejemplos anteriores, el grupo de traslaciones en \mathbb{R}^n actúa regularmente, puesto que todas sus órbitas son rectas; lo mismo ocurre con el grupo de rotaciones en el plano, si quitamos el origen, ya que todas sus órbitas son círculos alrededor del $(0, 0)$. Encontrar un ejemplo de una acción semiregular que no sea regular es más complicado, lo más sencillo es tomar el llamado "flujo irracional" del toro.²

²Ver Olver [7], sección 1.2.

1.3 Campos Vectoriales.

La herramienta principal de los grupos de Lie y de los grupos de transformaciones son las "transformaciones infinitesimales", las cuales están dadas en términos del concepto de campo vectorial. Para entender esta idea, lo primero que necesitamos es conocer lo que significa un vector tangente.

Existen varias formas de definir a los vectores tangentes debido a que estos objetos tienen una estructura verdaderamente multifacética. Por un lado, tienen un aspecto geométrico que se manifiesta como direcciones en el espacio: cuando estamos "parados" en una variedad, podemos movernos sobre ella en diversas direcciones, las cuales pueden ser descritas por vectores tangentes a la variedad en nuestra posición original. Por otro lado, hay un aspecto analítico, según el cual, los vectores tangentes aparecen como "derivadas direccionales" (o bien, operadores diferenciales parciales de primer orden) que, cuando se aplican a funciones suaves, nos dan la razón de cambio de dichas funciones en una dirección determinada. A pesar de que, en cierto sentido, estas nociones parecen lo mismo, tienen un desarrollo distinto: mientras que las direcciones nos proveen de una idea estrictamente geométrica (básicamente la de un vector visto como "flecha"), los operadores diferenciales son objetos del dominio del análisis y, por ejemplo, pueden estar contruidos de tal manera que nos lleven a pensar en otro tipo de operadores cuyo significado geométrico puede, en principio, no ser tan claro.

Para hablar de vectores tangentes, es conveniente hacer una consideración del concepto de variedad: de acuerdo a lo revisado en la sección 1.1, particularmente la definición de subvariedad regular y el Teorema del Rango, siempre podemos pensar que nuestras variedades M son subvariedades de un espacio euclideo mayor \mathbb{R}^n , es decir, podemos imaginarnos a las variedades como superficies (de dimensión m) inmersas en un espacio euclideo de dimensión mayor a m .³

Comencemos con el punto de vista geométrico de los vectores. Supongamos que queremos encontrar un vector tangente a una variedad M en un punto $p \in M$. Hay distintas formas de hacer esto, dependiendo de la manera en que tengamos descrita a nuestra variedad. Si M es definida paramétricamente, es decir, por medio de un sistema de coordenadas, lo que hacemos es tomar una carta (U, χ) en p . Como estamos considerando a nuestra variedad encajada en \mathbb{R}^n , la parametrización de M estará dada por el mapeo χ^{-1} , el cual lo podemos pensar como $\chi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$, $(x^1, \dots, x^m) \rightarrow (y^1, \dots, y^m, y^{m+1}, \dots, y^n)$ y, sin pérdida de generalidad, tomamos $\chi^{-1}(0) = p$, (es decir, $\chi(p) = 0$). Del cálculo diferencial, sabemos que un vector tangente a M en p estará

³Este resultado se conoce como el teorema de Whitney, el cual puede consultarse en Guillemin y Pollack, *Diferential Topology*, Prentice-Hall, 1974.

dado de la siguiente manera: Sea $v = (\xi^1, \dots, \xi^m) \in \mathbb{R}^m$ un vector, entonces

$$(J(\chi^{-1})(0))v = \begin{pmatrix} \frac{\partial}{\partial x^1} y^1 \Big|_0 & \cdots & \frac{\partial}{\partial x^m} y^1 \Big|_0 \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x^1} y^n \Big|_0 & \cdots & \frac{\partial}{\partial x^m} y^n \Big|_0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^m \end{pmatrix},$$

nos da un vector en \mathbb{R}^n que es tangente a M en p . Así, el plano tangente a M en p es:

$$T_p M_{\text{geométrico}} = \{p + (J(\chi^{-1})(0))v \mid v \in \mathbb{R}^m\}.$$

Claramente, dado que $J(\chi^{-1})(0)$ tiene rango m , el plano tangente es un espacio vectorial de dimensión m .

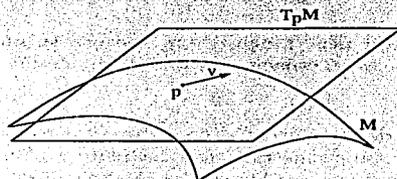


Figura 1.7: Plano tangente a $p \in M$.

De este modo, al tomar una curva en la variedad que pasa por p , $\varphi : \mathbb{R} \rightarrow M$, entonces $\chi \circ \varphi$ es una curva en \mathbb{R}^n que pasa por 0. Si el vector tangente a esta última curva en 0 está dado por v , tendremos que el vector tangente a φ en p está dado por $(J(\chi^{-1})(0))v$.

Ahora bien, nuestra variedad M puede estar descrita como una superficie de nivel, es decir, dado un mapeo suave $F : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$, podemos ver a nuestra variedad definida implícitamente como la preimagen de un valor regular $q \in \mathbb{R}^{n-m}$. Supongamos que $p = (x_0^1, \dots, x_0^n) \in U \subseteq \mathbb{R}^n$, U abierto tal que $U \cap M$ es el conjunto cero del mapeo F y que $F(p) = 0$, es decir, $p \in U \cap M$. Entonces, sabemos del cálculo que un vector tangente a M en p estará dado implícitamente como $v = (\xi^1, \dots, \xi^n) \in \mathbb{R}^n$ tal que $(J(F)(0))(v - p) = 0$, es decir:

$$\begin{pmatrix} \frac{\partial}{\partial x^1} F^1 \Big|_p & \cdots & \frac{\partial}{\partial x^n} F^1 \Big|_p \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x^1} F^{n-m} \Big|_p & \cdots & \frac{\partial}{\partial x^n} F^{n-m} \Big|_p \end{pmatrix} \begin{pmatrix} \xi^1 - x_0^1 \\ \vdots \\ \xi^n - x_0^n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

El plano tangente a M en p quedará definido también implícitamente como:

$$T_p M_{\text{geométrico}} = \{v \in \mathbb{R}^n \mid (J(F)(p))(v - p) = 0\}.$$

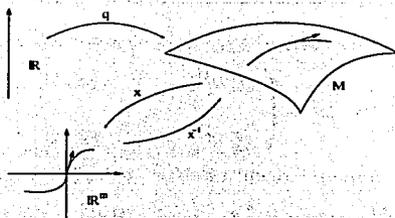


Figura 1.8: Vector tangente a una curva en una variedad M .

En este caso, $J(F)(p)$ tiene rango $n - m$, así que su núcleo tiene dimensión m , por lo que $T_p M_{\text{geométrico}}$ es un espacio vectorial de dimensión m (igual a la de nuestra variedad en cuestión).

Pese a que este tipo de construcciones de vectores tangentes (junto con el correspondiente plano tangente) son geoméricamente bastante explícitas, tienen la desventaja de que dependen del modo en que la variedad M está inmersa en \mathbb{R}^n . La descripción analítica de los vectores tangentes, como derivadas direccionales, permite encontrar un modo intrínseco a la variedad para hablar de ellos y que, además, coincide con el punto de vista geométrico. Empecemos notando que a cada vector $v = (\xi^1, \dots, \xi^n) \in \mathbb{R}^n$ le podemos asociar un operador diferencial

$$D_v = \sum_{i=1}^n \xi^i \frac{\partial}{\partial x^i} ,$$

llamado la *derivada direccional* en la dirección de v . En el caso en el que v sea un vector unitario en \mathbb{R}^n y $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sea una función suave, el significado geométrico de $D_v(f)|_x$ es el incremento de f en la dirección v : $D_v(f)|_x = v \cdot \nabla f|_x$, de modo que este operador nos da la noción que estamos buscando.

Definición 1.14 Una derivación en $p \in M$ es un operador D (de valores reales), definido en las funciones suaves cuyo dominio contenga a una vecindad de p , de modo que $\forall f, g : M \rightarrow \mathbb{R}$ suaves, D satisface las siguientes dos propiedades:

- 1) (Linealidad) $D(\alpha f + \beta g)|_p = \alpha D(f)|_p + \beta D(g)|_p, \quad \forall \alpha, \beta \in \mathbb{R}.$
- 2) (Regla de Leibniz) $D(fg)|_p = g(p)D(f)|_p + f(p)D(g)|_p.$

Consideraremos ahora a los vectores tangentes a M en p como derivaciones en p , y

al plano tangente en p como el espacio vectorial de todas las derivaciones en p :

$$T_p M_{\text{analítico}} = \{D : \mathcal{F}(M) \rightarrow \mathbb{R} \mid D \text{ es una derivación en } p\},$$

donde $\mathcal{F}(M) = \{f : M \rightarrow \mathbb{R} \text{ suaves}\}$.

Antes de mostrar la equivalencia entre ambos planos tangentes, veamos algunas propiedades de $T_p M_{\text{analítico}}$.

observación 1. Si $f = g$, $f, g \in \mathcal{F}(M)$ en una vecindad de p , entonces, por linealidad, $D(f) = D(g)$. También por linealidad se cumple que la derivación de toda $f = \text{constante}$ es cero.

observación 2. Supongamos que las coordenadas en una vecindad U de p están dadas por (x^1, \dots, x^n) , entonces $\frac{\partial}{\partial x^i} \Big|_p$ es un vector tangente a M en p , es decir, que $\forall f, g \in \mathcal{F}(M)$, se tiene:

$$\begin{aligned} \frac{\partial}{\partial x^i}(\alpha f + \beta g) \Big|_p &= \alpha \frac{\partial f}{\partial x^i} \Big|_p + \beta \frac{\partial g}{\partial x^i} \Big|_p, \quad \forall \alpha, \beta \in \mathbb{R} \text{ y} \\ \frac{\partial}{\partial x^i}(fg) \Big|_p &= g(p) \frac{\partial f}{\partial x^i} \Big|_p + f(p) \frac{\partial g}{\partial x^i} \Big|_p \end{aligned}$$

observación 3. Los vectores $\frac{\partial}{\partial x^i} \Big|_p$ son linealmente independientes, pues si tenemos:

$$0 = \sum_{i=1}^n a^i \frac{\partial f}{\partial x^i} \Big|_p, \quad \forall f \in \mathcal{F}(M),$$

entonces se cumple en particular para x^j , de donde

$$0 = \sum_{i=1}^n a^i \frac{\partial x^j}{\partial x^i} \Big|_p = a^j \quad \forall j.$$

Lema 1.15

Si f es una función suave definida en una vecindad de $x_0 \in \mathbb{R}^n$, entonces

$$f(x) = f(x_0) + \sum_{i=1}^n (x^i - x_0^i) \left\{ \frac{\partial f}{\partial x^i} \Big|_{x_0} + a^i(x) \right\},$$

donde las funciones $a^i(x)$ son suaves y se anulan en x_0 .

Prueba.

Usando regla de la cadena, obtenemos:

$$f(x) - f(x_0) = \int_0^1 \frac{\partial}{\partial t} f(x_0 + t(x - x_0)) dt = \sum_{i=1}^n (x^i - x_0^i) \int_0^1 \frac{\partial}{\partial x^i} f(x_0 + t(x - x_0)) dt.$$

Integrando por partes, se tiene que

$$\begin{aligned} \int_0^1 \frac{\partial}{\partial x^i} f(x_0 + t(x - x_0)) dt \\ &= t \frac{\partial}{\partial x^i} f(x_0 + t(x - x_0)) \Big|_0^1 - \int_0^1 t \sum_{j=1}^n (x^j - x_0^j) \frac{\partial^2}{\partial x^i \partial x^j} f(x_0 + t(x - x_0)) dt \\ &= \frac{\partial f}{\partial x^i} \Big|_{x_0} + a^i(x), \end{aligned}$$

donde $a^i(x)$ son funciones suaves que claramente se anulan en x_0 . □

Proposición 1.16

Si U es un subconjunto abierto de \mathbb{R}^n y D es una derivación en $x_0 \in U$, entonces

$$D = \sum_{i=1}^n D(x^i) \frac{\partial}{\partial x^i} \Big|_{x_0}.$$

Prueba.

Para una $f \in \mathcal{F}(M)$ arbitraria, escribimos

$$f(x) = f(x_0) + \sum_{i=1}^n (x^i - x_0^i) \left\{ \frac{\partial f}{\partial x^i} \Big|_p + a^i(x) \right\},$$

como en el lema anterior. Aplicando D a ambos lados y usando la regla de Leibnitz:

$$\begin{aligned} D(f) \Big|_{x_0} &= 0 + \sum_{i=1}^n D(x^i - x_0^i) \Big|_{x_0} \left\{ \frac{\partial f}{\partial x^i} \Big|_{x_0} + a^i(x_0) \right\} + \sum_{i=1}^n (x_0^i - x_0^i) D \left\{ \frac{\partial f}{\partial x^i} \Big|_{x_0} + a^i(x) \right\} \Big|_{x_0} \\ &= \sum_{i=1}^n D(x^i) \Big|_{x_0} \frac{\partial f}{\partial x^i} \Big|_{x_0} \end{aligned}$$

Corolario 1.17

$T_p \mathbb{R}^n_{\text{analítico}}$ tiene dimensión n .

Prueba.

Por la observación 3 y la proposición anterior, $\left\{ \frac{\partial}{\partial x^i} \Big|_{x_0} \right\}$, $i = 1, \dots, n$, es base de $T_p \mathbb{R}^n_{\text{analítico}}$ □

Además, si tenemos un mapeo suave entre variedades $F : M \rightarrow N$ y fijamos un punto $p \in M$, F induce un mapeo lineal entre los espacios tangentes a p y a $F(p)$:

Definición 1.18 Sea $F : M \rightarrow N$ un mapeo suave entre variedades y $p \in M$. La diferencial de F en p (denotada como F_p^* ó dF_p) es una transformación $F_p^* : T_p M_{\text{analítico}} \rightarrow T_{F(p)} N_{\text{analítico}}$ definida como sigue:

$$F_p^*(D)(f) = D(f \circ F) \Big|_p \quad \forall f \in \mathcal{F}(N).$$

Teorema 1.19

i) F_p^* es una transformación lineal.

ii) (Regla de la Cadena) Sean $G : M \rightarrow N$ y $F : N \rightarrow P$ dos mapeos suaves, $p \in M$ y $G(p) \in N$, entonces

$$(F \circ G)_p^* = F_{G(p)}^* \circ G_p^*.$$

Prueba.

i) $\forall \alpha, \beta \in \mathbb{R}$, tomando $f \in \mathcal{F}(N)$ arbitraria, tenemos:

$$\begin{aligned} F_p^*(\alpha D_1 + \beta D_2)(f) &= (\alpha D_1 + \beta D_2)(f \circ F) \\ &= \alpha D_1(f \circ F) + \beta D_2(f \circ F) \\ &= \alpha F_p^*(D_1)(f) + \beta F_p^*(D_2)(f) \end{aligned}$$

ii) Tomemos $f \in \mathcal{F}(P)$ arbitraria, entonces

$$\begin{aligned} (F \circ G)_p^*(D)(f) &= D(f \circ (F \circ G)) = D((f \circ F) \circ G) \\ &= G_p^*(D)(f \circ F) = F_{G(p)}^*(G_p^*(D))(f) \\ &= (F_{G(p)}^* \circ G_p^*)(D)(f). \end{aligned}$$

□

Corolario 1.20

Sea $F : M \rightarrow N$ un mapeo suave y $p \in M$, entonces el rango de F en p es igual al rango de $F_p^* : T_p M_{\text{analítico}} \rightarrow T_{F(p)} N_{\text{analítico}}$.

Prueba.

Dado que esto es un problema local, podemos dar coordenadas (x^1, \dots, x^m) en una vecindad de p y (y^1, \dots, y^n) en una vecindad de $F(p)$. Por el corolario 1.17, $\left\{ \frac{\partial}{\partial x^i} \Big|_p \right\}, i = 1, \dots, m$, es base de $T_p M_{\text{analítico}}$ ($= T_p \mathbb{R}^m$, por el homeomorfismo local entre M y \mathbb{R}^m). Considerando $F = (F^1, \dots, F^n) = (y^1 \circ F, \dots, y^n \circ F)$, tendremos que las entradas de la matriz asociada a F_p^* son:

$$F_p^* \left(\frac{\partial}{\partial x^i} \Big|_p \right) (y^j) = \frac{\partial}{\partial x^i} (y^j \circ F) \Big|_p = \frac{\partial F^j}{\partial x^i} \Big|_p,$$

es decir, localmente la matriz de transformación de F_p^* es igual a la matriz jacobiana de F en p . □

Corolario 1.21

Sea M una variedad conexa y $F: M \rightarrow N$ un mapeo suave. Si $\forall p \in M$, se tiene que $F_p^* = 0$, entonces F es una función constante.

Prueba.

Sea $q \in F(M) \subset N$. De entrada, $F^{-1}(q)$ es un conjunto cerrado en M , por continuidad. Por otro lado, sea $p \in F^{-1}(q)$. Escogemos coordenadas (x^1, \dots, x^m) en una vecindad de p y (y^1, \dots, y^n) en una vecindad de q . Por hipótesis,

$$0 = F_p^* \left(\frac{\partial}{\partial x^j} \Big|_p \right) = \sum_{i=1}^n \frac{\partial}{\partial x^j} (y^i \circ F) \frac{\partial}{\partial y^i} \Big|_q,$$

de donde, $\frac{\partial}{\partial x^j} (y^i \circ F) \equiv 0$, es decir, $y^i \circ F$ es constante en U , $\forall j = 1, \dots, n$ por lo que $F(U) = q$, lo que implica que $F^{-1}(q)$ es abierto como subconjunto de M . Dado que $F^{-1}(q)$ es abierto y cerrado en M conexa, $M = F^{-1}(q)$, por lo que F es constante. □

Ahora que contamos con esta herramienta, podemos mostrar la equivalencia entre $T_p M_{\text{geométrico}}$ y $T_p M_{\text{analítico}}$, de modo que no haremos distinción entre ellos al referirnos al espacio tangente y denotaremos a éste simplemente como $T_p M$.

Supongamos que tenemos una variedad M descrita como el conjunto cero de un mapeo suave $F: \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$, $F = (F^1, \dots, F^{n-m})$, de rango $n - m$. Entonces,

$$T_p M_{\text{geométrico}} = \ker J(F)(p),$$

con $J(F)(p): \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ vista como transformación lineal.

Ahora bien, la inclusión $\iota: M \rightarrow \mathbb{R}^n$ induce el mapeo

$$\iota_p^*: T_p M_{\text{analítico}} \rightarrow T_p \mathbb{R}^n.$$

Si $D \in T_p M_{\text{analítico}}$, se tiene que $\iota_p^*(D) \in T_p \mathbb{R}^n$ y, por la proposición 1.16, podemos escribir

$$\iota_p^*(D) = \sum_{i=1}^n \xi^i \frac{\partial}{\partial x^i}, \quad \text{donde } \xi^i = \iota_p^*(D)(x^i) \Big|_p \in \mathbb{R}.$$

Lo que queremos ver es que $(\xi^1, \dots, \xi^n) \in T_p M_{\text{geométrico}}$. Para esto, notemos que $F^j \circ \iota$ se anula en todo punto de M , $\forall j$. Así:

$$0 = D(F^j \circ \iota) = (F^j \circ \iota)_p^*(D) = F_{(\iota(p))}^j(\iota_p^*(D))$$

$$= F_{i(p)}^{*j} \left(\sum_{i=1}^n \xi^i \frac{\partial}{\partial x^i} \right) = \sum_{i=1}^n \xi^i \frac{\partial F^j}{\partial x^i} \Big|_p, \quad \forall j.$$

Entonces, $(J(F)(p))(\xi^1, \dots, \xi^n) = 0$, así que $(\xi^1, \dots, \xi^n) \in T_p M_{\text{geométrico}}$. Esto nos lleva a un isomorfismo entre espacios vectoriales dado por:

$$\begin{aligned} T_p M_{\text{analítico}} &\leftrightarrow T_p M_{\text{geométrico}} \\ D &\rightarrow (\xi^1, \dots, \xi^n) = \left(i_p^*(D)(x^1) \Big|_p, \dots, i_p^*(D)(x^n) \Big|_p \right). \end{aligned}$$

Una vez que hemos hecho esto, podemos tomar a los vectores tangentes como derivaciones sin dejar de pensar en su sentido geométrico.

A la colección de todos los espacios tangentes que corresponden a todos los puntos en la variedad M se le llama el haz tangente de M y se denota como

$$TM = \bigcup_{p \in M} T_p M.$$

De hecho, TM es una variedad suave de dimensión $2m$ y puede ser visto como

$$TM = \{(p, v_p) \mid p \in M, v_p \in T_p M\}.$$

La idea es tomar vectores tangentes "pegados" de una manera suave, de modo que si tenemos una curva suave $\varphi: \mathbb{R} \rightarrow M$, los vectores tangentes $\varphi'(t) \in T_{\varphi(t)} M$ varíen suavemente de punto a punto. Un campo vectorial v en M puede verse como una función que asigna un vector tangente $v|_p$ a cada $p \in M$. A nosotros nos interesan los campos vectoriales suaves, es decir, aquellos que asignan vectores tangentes que varían suavemente de punto a punto.

Las propiedades de los campos vectoriales se heredan de las propiedades de los vectores tangentes, de modo que el conjunto de todos los campos vectoriales suaves forma un módulo sobre $\mathcal{F}(M)$, con las siguientes operaciones:

$$\begin{aligned} (v+w)|_p &= v|_p + w|_p, \quad y \\ (fv)|_p &= f(p)v|_p. \end{aligned}$$

También tenemos que, en coordenadas locales (x^1, \dots, x^m) , un campo vectorial tiene la forma

$$v|_x = \xi^1(x) \frac{\partial}{\partial x^1} \Big|_x + \dots + \xi^m(x) \frac{\partial}{\partial x^m} \Big|_x,$$

donde cada $\xi^i(x)$ es una función suave de x . Generalmente, hacemos referencia a un campo vectorial indistintamente como v o como (ξ^1, \dots, ξ^m) . Además, el mapeo diferencial inducido por una función suave entre variedades $F: M \rightarrow N$, se puede generalizar

a los campos vectoriales tomando, para cada $p \in M$, el mapeo $F_p^* : T_p M \rightarrow T_{F(p)} N$. Es importante notar que, si v es un campo vectorial en M , entonces, en general, $F^*(v)$ no necesariamente es un campo vectorial bien definido en N : para empezar, $F^*(v)$ puede no estar definido en todo N y, si dos puntos p_1 y $p_2 \in M$ van a dar al mismo punto en $q \in N$ bajo el mapeo F , no hay garantía de que los vectores $F_{p_1}^*(v|_{p_1})$ y $F_{p_2}^*(v|_{p_2})$ sean el mismo en N . Sin embargo, si F es un difeomorfismo, entonces $F^*(v)$ siempre está bien definido, más aún, si tenemos dos campos vectoriales v en M y w en N , diremos que v y w están F -relacionados si $F_p^*(v|_p) = w|_{F(p)} \quad \forall p$.

Una *curva integral* de un campo vectorial v es una curva parametrizada suave $\varphi(t)$ cuyo vector tangente en cualquier punto coincide con el valor de v en el mismo punto:

$$\frac{d\varphi}{dt} \equiv \dot{\varphi}(t) = v|_{\varphi(t)} \quad \forall t.$$

Dada una condición inicial $\varphi(0) = p_0$, el teorema de existencia y unicidad de ecuaciones diferenciales ordinarias nos garantiza que hay una única curva integral "máximal" (con el mayor dominio posible) que pasa por p_0 .

Sean (x^1, \dots, x^m) coordenadas locales para una vecindad en M . Con cada campo vectorial $v|_x = (\xi^1(x), \dots, \xi^m(x))$, definido en esta vecindad, tenemos asociado el siguiente sistema autónomo de ecuaciones diferenciales:

$$\frac{dx^i}{dt} \equiv \dot{x}^i(t) = \xi^i(x(t)), \quad i = 1, \dots, m.$$

Las soluciones $x(t)$ de este sistema son las curvas integrales del campo vectorial v . Demos, además, una condición inicial

$$x^i|_{t=0} = x_0^i.$$

Tomemos ahora el siguiente mapeo:

$$\begin{aligned} \varphi : U \subset \mathbb{R} \times M &\rightarrow M \\ (t, x_0^1, \dots, x_0^m) &\rightarrow (x^1, \dots, x^m), \\ (0, x_0^1, \dots, x_0^m) &\rightarrow (x_0^1, \dots, x_0^m). \end{aligned}$$

La función φ asocia a cada x_0 , en un tiempo dado t , el punto sobre la curva integral de v en dicho t . Por el teorema de existencia y unicidad, hay justamente una única solución para cada x_0 , así que, para cada t fijo, φ define un difeomorfismo φ_t de M a M . Es decir, que al aplicar φ_t a un punto (x_0) , esto nos da la nueva posición de una partícula que inicialmente estaba en x_0 y que se ha movido a lo largo de la curva integral de v correspondiente, transcurrido un tiempo t . A los mapeos $\varphi(t, x)$ y φ_t se les conoce indistintamente como el *flujo* generado por v . Tenemos entonces que, por el

teorema de existencia y unicidad, φ_t cumple lo siguiente:

i) $\varphi_0(x_0) = x_0$

ii) Para valores suficientemente pequeños del parámetro, $\varphi_t(\varphi_s(x_0)) = \varphi_{t+s}(x_0)$.

iii) Para cada t con $|t|$ suficientemente pequeño, $\varphi_{-t}(x) = (\varphi_t(x_0))^{-1}$.

Al comparar las propiedades anteriores con la definición 1.12, nos damos cuenta de que el flujo generado por un campo vectorial es un grupo local de transformaciones, al cual se le conoce como *grupo (local) monoparamétrico de difeomorfismos o de transformaciones*, y al campo vectorial v se le llama *generador infinitesimal del grupo*.

Inversamente, si tenemos un grupo local monoparamétrico de difeomorfismos φ_t , podemos definir su campo de velocidades (su generador infinitesimal) como:

$$v|_x = \left. \frac{d}{dt} \varphi_t(x) \right|_{t=0},$$

el cual, por definición, es tangente a las trayectorias dadas por el grupo.

Por ejemplo, consideremos al grupo de traslaciones en \mathbb{R}^n , en la dirección x^1 . Sea $x_0 = (x_0^1, \dots, x_0^n)$ y $\varphi_t(x_0) = (x_0^1 + t, x_0^2, \dots, x_0^n) = (x^1, \dots, x^n)$, entonces,

$$\begin{aligned} \left. \frac{dx^1}{dt} \right|_{t=0} &= 1 \\ \left. \frac{dx^i}{dt} \right|_{t=0} &= 0 \quad \forall i = 2, \dots, n \end{aligned}$$

por lo que $v|_{x_0} = (1, 0, \dots, 0)$ y en general se tiene el campo vectorial es constante y con curvas integrales dadas por rectas:

$$v = (1, 0, \dots, 0) = \frac{\partial}{\partial x^1}$$

En otro ejemplo, tomemos al grupo monoparamétrico de rotaciones alrededor del origen, por un ángulo θ , en \mathbb{R}^2 . En este caso, para $(x_0, y_0) \neq (0, 0)$, $\varphi_\theta(x_0, y_0)$ está dado por:

$$\begin{aligned} x &= x_0 \cos \theta - y_0 \sin \theta \\ y &= x_0 \sin \theta + y_0 \cos \theta, \end{aligned}$$

de donde

$$\begin{aligned} \left. \frac{dx}{dt} \right|_{t=0} &= -y_0 \\ \left. \frac{dy}{dt} \right|_{t=0} &= x_0 \end{aligned}$$

así que $v|_{(x_0, y_0)} = (-y_0, x_0)$ y el campo vectorial es

$$v = (-y, x) = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y},$$

por lo que las curvas integrales de este campo son círculos $x^2 + y^2 = c$.

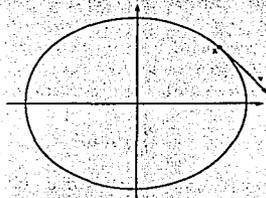


Figura 1.9: Campo vectorial asociado al grupo de rotaciones en \mathbb{R}^2 evaluado en un punto x .

El efecto de un cambio de coordenadas $y = y(x)$ en un campo vectorial v está determinado por su efecto en cada vector tangente $v|_x$, $x \in M$:

Sea

$$v|_{x_0} = \sum_{i=1}^m \xi^i(x_0) \frac{\partial}{\partial x^i} \Big|_{x_0}, \quad \xi^i(x_0) = \frac{dx^i}{dt} \Big|_{t=0}$$

entonces,

$$\frac{dy^i}{dt} \Big|_{t=0} = \sum_{j=1}^m \frac{\partial y^i}{\partial x^j} \Big|_{x_0} \frac{dx^j}{dt} \Big|_{t=0},$$

por lo que

$$v|_{y_0=y(x_0)} = \sum_{i=1}^m \left(\sum_{j=1}^m \frac{\partial y^i}{\partial x^j} \Big|_{x_0} \xi^j(x_0) \right) \frac{\partial}{\partial y^i} \Big|_{y_0},$$

de modo que el campo vectorial v en términos de las coordenadas (y^1, \dots, y^m) es

$$v = \sum_{i=1}^m \left(\sum_{j=1}^m \frac{\partial y^i}{\partial x^j} \xi^j \right) \frac{\partial}{\partial y^i}.$$

El siguiente resultado resalta la conveniencia de escoger un sistema coordenado adecuado. En este caso, tenemos que, localmente, siempre podemos ver a las curvas

integrales de un campo vectorial v como rectas paralelas al eje x^1 .

Teorema 1.22 (Rectificación de coordenadas)

Sea v un campo vectorial definido en una variedad M , con coordenadas (x^1, \dots, x^m) , y un punto $p \in M$ en el cual $v|_p$ no se anula. Entonces, existe un sistema de coordenadas (y^1, \dots, y^m) , llamado *sistema adaptado a v* , en una vecindad de p tal que $v = \frac{\partial}{\partial y^1}$.

Prueba.

Como este resultado es local, al considerar traslaciones y rotaciones alrededor del origen, podemos suponer que nuestras coordenadas (x^1, \dots, x^m) están dadas de modo que a p le corresponde el punto $x_0 = 0$ y $v|_{x_0} = \frac{\partial}{\partial x^1}|_{x_0}$. En nuestras coordenadas, el campo vectorial es:

$$v = \sum_{i=1}^m \xi^i \frac{\partial}{\partial x^i},$$

de modo que $\xi^1(x_0) = 1$ y, por continuidad, $\xi^1(x) > 0$ en una vecindad de $x_0 = 0$. Tenemos entonces que, para los puntos en las curvas integrales $(x^1(t), \dots, x^m(t))$ dentro de dicha vecindad, se tiene que

$$\frac{dx^1}{dt} = \xi^1(x) > 0 \Rightarrow \frac{dx^1}{dt} \neq 0,$$

por lo que las curvas integrales varían en la dirección x^1 , cruzando el hiperplano $\{(0, x^2, \dots, x^m)\}$. Así, podemos señalar a los puntos de la forma $(0, x^2, \dots, x^m)$ como condiciones iniciales del campo v .

Si tenemos un punto (x^1, \dots, x^m) en una vecindad de 0 , construimos sus nuevas coordenadas tomando a la curva integral de v que pasa por dicho punto, regresando sobre esta curva hasta alcanzar su condición inicial $(0, y^2, \dots, y^m)$ en el hiperplano $\{(0, x^2, \dots, x^m)\}$ y asignando entonces las coordenadas (y^1, y^2, \dots, y^m) , donde y^1 está dado por el tiempo que le toma a $(0, y^2, \dots, y^m)$ llegar a (x^1, \dots, x^m) sobre la curva integral. Así, el punto $\varphi_t(y^1, \dots, y^m)$, corresponde a un punto sobre la curva integral de (x^1, \dots, x^m) que, por teorema de existencia y unicidad, es la curva integral con condición inicial $(0, y^2, \dots, y^m)$; de modo que a $(0, y^2, \dots, y^m)$ le toma un tiempo $y^1 + t$ llegar a $\varphi_t(y^1, \dots, y^m)$, por lo que

$$\varphi_t(y^1, \dots, y^m) = \varphi_t(\varphi_{y^1}(0, y^2, \dots, y^m)) = \varphi_{t+y^1}(0, y^2, \dots, y^m) = (y^1 + t, y^2, \dots, y^m)$$

$$y \quad v = \frac{\partial}{\partial y^1}.$$

□

Supongamos que tenemos un grupo monoparamétrico de difeomorfismos $\varphi_t(x)$, junto con su correspondiente campo vectorial asociado v , además de una función $f \in \mathcal{F}(M)$. Definimos la acción del grupo en f como

$$\varphi_t(x) \cdot f = f(\varphi_t(x)).$$

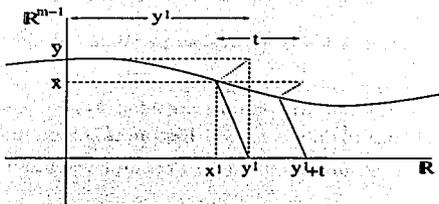


Figura 1.10: Un sistema coordenado adaptado a v .

Para saber cuál es el efecto de esta acción, hay que estudiar cómo cambia f a lo largo del flujo generado por v , conforme varía el parámetro t . Tomando en cuenta que la expresión del campo en coordenadas, $v = \sum_{i=1}^m \xi^i \frac{\partial}{\partial x^i}$, tenemos

$$\begin{aligned} \frac{df}{dt}(\varphi_t(x)) &= \sum_{i=1}^m \xi^i(\varphi_t(x)) \frac{\partial f}{\partial x^i} \Big|_{\varphi(t,x)} \\ &= v(f) \Big|_{\varphi_t(x)}. \end{aligned}$$

En particular, para $t = 0$,

$$\frac{df}{dt}(\varphi_t(x)) \Big|_{t=0} = \sum_{i=1}^m \xi^i(x) \frac{\partial f}{\partial x^i} \Big|_x = v(f) \Big|_x.$$

Si la serie de Taylor de $f(x) \in \mathcal{F}(M)$ converge en una vecindad de x , podemos escribir

$$f(\varphi_t(x)) = f(x) + tv(f) \Big|_x + \frac{t^2}{2!} v^2(f) \Big|_x + \dots$$

donde $v^2 = v(v(f))$ y así sucesivamente. Entonces, como la serie de Taylor de f converge,

$$f(\varphi_t(x)) = \sum_{n=1}^{\infty} \frac{t^n}{n!} v^n(f) \Big|_x.$$

Esta expresión tiene una sugerente semejanza con la serie de Taylor para la función exponencial, lo que nos lleva a definir un mapeo exponencial para un campo v con flujo $\varphi_t(x)$:

Definición 1.23 El *operador exponencial* del un campo vectorial $v = \sum_{i=1}^m \xi^i \frac{\partial}{\partial x^i}$ es el operador:

$$\exp(tv) = 1 + tv + \frac{t^2}{2!} v^2 + \dots = \sum_{n=1}^{\infty} \frac{t^n}{n!} v^n.$$

La acción de este operador en funciones $f(x) \in \mathcal{F}(M)$ se define naturalmente como:

$$\exp(tv)(f) \Big|_x = f(x) + tv(f) \Big|_x + \frac{t^2}{2!} v^2(f) \Big|_x + \dots = \sum_{n=1}^{\infty} \frac{t^n}{n!} v^n(f) \Big|_x.$$

Tomando a t suficientemente pequeña para que las series converjan.

Claramente, para t suficientemente pequeña, la acción del operador exponencial de $v = (\xi^1, \dots, \xi^m)$ en $f \in \mathcal{F}(M)$ coincide con la acción de $\varphi_t(x)$ en f , más aún, si tomamos las funciones coordenadas, obtenemos una expresión en serie del flujo mismo (bajo suposiciones de convergencia):

$$\varphi_t(x) = x + tv \Big|_x + \frac{t^2}{2!} v^2 \Big|_x + \dots = \sum_{n=1}^{\infty} \frac{t^n}{n!} v^n \Big|_x,$$

por lo que también se suele denotar al flujo asociado a v como $\exp(tv)x$.

Generalmente, se conoce a $v(f) \Big|_x = \sum_{i=1}^m \xi^i(x) \frac{\partial f}{\partial x^i} \Big|_x = \frac{df}{dt}(\varphi_t(x)) \Big|_{t=0}$ como la *derivada de Lie* de f (también denotada como $L_v f$). Utilizando la noción de derivada direccional, nos damos cuenta del hecho de que, si f tiene una derivada de Lie nula ($L_v f \equiv 0$), entonces la función f es constante en las curvas integrales del campo vectorial v , lo que nos da cierta idea de "invariancia" bajo la acción del grupo mon-paramétrico asociado. Más adelante retomaremos estas ideas.

A continuación definiremos la operación más importante para campos vectoriales:

Definición 1.24 Sean v y w campos vectoriales en M , entonces, el *paréntesis de Lie* (o *conmutador*) es el campo vectorial, denotado como $[v, w]$, dado por:

$$[v, w](f) = v(w(f)) - w(v(f)) \quad \forall f \in \mathcal{F}(M).$$

Si $v = \sum_{i=1}^m \xi^i \frac{\partial}{\partial x^i}$, $w = \sum_{i=1}^m \eta^i \frac{\partial}{\partial x^i}$, entonces, usando la regla de Leibniz, se tiene que:

$$[v, w] = \sum_{i=1}^m \left(\sum_{j=1}^m \xi^j \frac{\partial \eta^i}{\partial x^j} - \eta^j \frac{\partial \xi^i}{\partial x^j} \right) \frac{\partial}{\partial x^i}$$

Proposición 1.25

El paréntesis de Lie tiene las siguientes propiedades:

i) *IR-Bilinealidad*

$$\begin{aligned} [av_1 + bv_2, w] &= a[v_1, w] + b[v_2, w], \\ [v, aw_1 + bw_2] &= a[v, w_1] + b[v, w_2], \end{aligned}$$

donde a y b son constantes.

ii) *Antisimetría*

$$[v, w] = -[w, v].$$

iii) Identidad de Jacobi

$$[u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0.$$

Prueba.

La \mathbb{R} -bilinealidad es consecuencia directa de la linealidad de los campos vectoriales vistos como derivaciones. La asimetría se sigue directamente de la definición del paréntesis de Lie. En cuanto a la identidad de Jacobi, ésta se resuelve directamente sustituyendo la definición en todos los términos en donde aparezca el conmutador, de modo que, una vez que se desarrolla completamente la expresión, los doce términos se cancelan por pares. □

Proposición 1.26

Sea $F : M \rightarrow N$ un mapeo suave entre variedades y v, w , campos vectoriales en M que están F -relacionados con campos vectoriales bien definidos $F^*(v)$ y $F^*(w)$ en N , entonces, los correspondientes paréntesis de Lie también están F -relacionados, es decir:

$$F^*([v, w]) = [F^*(v), F^*(w)].$$

Prueba.

Primero demosetremos el siguiente resultado: dados w y u campos vectoriales en M y N , respectivamente, y un mapeo $F : M \rightarrow N$, si w y u están F -relacionados, entonces

$$w(f \circ F) = u(f) \circ F.$$

Como w está F -relacionado con u , se tiene $F^*(w) = u$, es decir,

$$\begin{aligned} F_x^*(w|_x) &= u|_{F(x)} \\ \Rightarrow F_x^*(w|_x)(f) &= u|_{F(x)}(f), \quad f \in \mathcal{F}(N) \\ \Rightarrow w|_x(f \circ F)|_x &= u(f)|_{F(x)} \\ \Rightarrow w|_x(f \circ F)|_x &= u(f) \circ F|_x \\ \Rightarrow w(f \circ F) &= u(f) \circ F. \end{aligned}$$

Regresando a la Prueba de la proposición, para $f \in \mathcal{F}(N)$, $x \in M$, tenemos:

$$\begin{aligned} F_x^*([v, w]|_x)(f) &= [v, w]|_x(f \circ F)|_x \\ &= v|_x(w(f \circ F))|_x - w|_x(v(f \circ F))|_x \\ &= v|_x(F^*(w)(f) \circ F) - w|_x(F^*(v)(f) \circ F) \end{aligned}$$

$$\begin{aligned}
 &= F_x^*(v|_x)(F^*(w)(f)) - F_x^*(w|_x)(F^*(v)(f)) \\
 &= (F^*(v))|_{F(x)}(F^*(w)(f)) - (F^*(w))|_{F(x)}(F^*(v)(f)) \\
 &= [F^*(v), F^*(w)]|_{F(x)}(f)
 \end{aligned}$$

□

Hay una caracterización más geométrica del paréntesis de Lie de dos campos vectoriales v y w visto como "conmutador infinitesimal" de los correspondientes grupos monoparamétricos:

Teorema 1.27

Sean v, w , campos vectoriales en una variedad M . Para cada $x \in M$, usando la notación exponencial para los flujos de v y w , el conmutador

$$\psi_t(x) = \exp(-\sqrt{t}v) \exp(-\sqrt{t}w) \exp(\sqrt{t}v) \exp(\sqrt{t}w)x$$

define una curva suave para t suficientemente pequeña, $0 \leq t$. Además, el paréntesis de Lie $[v, w]|_x$ es el vector tangente a esta curva en el punto inicial $\psi_0(x) = x$:

$$[v, w] = \frac{\partial \psi_t(x)}{\partial t} \Big|_{t=0^+}$$

Prueba.

Escribimos $y = \exp(\sqrt{t}w)x$, $z = \exp(\sqrt{t}v)y$, $u = \exp(-\sqrt{t}w)z$, para expresar el conmutador como $\psi_t(x) = \exp(-\sqrt{t}v)u$. Entonces, usando las expansiones de Taylor para los flujos, tenemos:

$$\begin{aligned}
 \psi_t(x) &= \exp(-\sqrt{t}v)u = u + \sqrt{t}v|_u + \frac{t^2}{2}v^2|_u + O(t^{\frac{3}{2}}) \\
 &= z - \sqrt{t}\{w|_z + v|_z\} + t\{\frac{1}{2}w^2|_z + v(w)|_z + \frac{1}{2}v^2|_z\} + O(t^{\frac{3}{2}}) \\
 &= y - \sqrt{t}w|_y + t\{-w(v)|_y + v(w)|_y + \frac{1}{2}w^2|_y\} + O(t^{\frac{3}{2}}) \\
 &= x + t\{v(w)|_x - w(v)|_x\} + O(t^{\frac{3}{2}}).
 \end{aligned}$$

Por lo tanto:

$$\frac{\partial \psi_t(x)}{\partial t} \Big|_{t=0^+} = v(w)|_x - w(v)|_x = [v, w]|_x.$$

□

Como otra ilustración del sentido geométrico del paréntesis de Lie, mostraremos que los flujos generados por dos campos vectoriales conmutan si y sólo si el paréntesis de Lie de los campos se anula en cualquier punto.

Teorema 1.28

Sean v y w campos vectoriales en una variedad M . Entonces

$$\exp(tv) \exp(\tau w)x = \exp(\tau w) \exp(tv)x \quad (1.2)$$

para todo $x \in M$, $\forall t, \tau \in \mathbb{R}$ en donde los flujos estén definidos, si y sólo si

$$[v, w] = 0$$

Prueba.

Por el teorema 1.27, se sigue inmediatamente que, si el flujo conmuta, entonces $[v, w] = 0$. Supongamos ahora que $[v, w] = 0$ y sea $x \in M$. Si ambos campos v y w se anulan en x , entonces los flujos dejan fijo a x y es claro que conmutan en x . De otro modo, al menos uno de los campos es distinto de cero en x , digamos $v|_x \neq 0$. Usando el teorema 1.22, podemos escoger coordenadas locales (y^1, \dots, y^m) , cerca de x , tales que $v = \frac{\partial}{\partial y^1}$ en la vecindad que corresponda a dichas coordenadas. Entonces, si $w = \sum_{i=1}^m \eta^i \frac{\partial}{\partial y^i}$, usando la expresión de $[v, w]$ en coordenadas,

$$0 = [v, w] = \sum_{i=1}^m \frac{\partial \eta^i}{\partial y^1} \frac{\partial}{\partial y^i}.$$

De modo que cada η^i es independiente de y^1 . El flujo generado por v en estas coordenadas es simplemente

$$\exp(tv)(y^1, \dots, y^m) = (y^1 + t, y^2, \dots, y^m),$$

y el flujo generado por w es solución del sistema de ecuaciones diferenciales

$$\frac{dy^i}{d\tau} = \eta^i(y^2, \dots, y^m), \quad i = 1, \dots, m.$$

Consideremos las funciones

$$y(\tau, t) = \exp(\tau w) \exp(tv)y = \exp(\tau w)(y^1 + t, y^2, \dots, y^m)$$

y

$$\begin{aligned} \tilde{y}(\tau, t) &= \exp(tv) \exp(\tau w)y = \exp(tv)y(\tau, 0) \\ &= (y^1(\tau, 0) + t, y^2(\tau, 0), \dots, y^m(\tau, 0)). \end{aligned}$$

dado que y^1 no aparece en el lado derecho del sistema de ecuaciones diferenciales para el flujo generado por w , como funciones de τ , y y \bar{y} son soluciones del sistema y ambas tienen la misma condición inicial:

$$y(0, t) = (y^1 + t, y^2, \dots, y^m) = \bar{y}(0, t).$$

Por teorema de existencia y unicidad, $y(\tau, t) = \bar{y}(\tau, t)$, lo que demuestra la validez del teorema para t y τ suficientemente pequeñas.

Para probar el resultado general, consideremos los siguientes subconjuntos del plano $\tau - t$:

$$V = \{(\tau, t) \mid \text{ambos lados de (1.2) están definidos en } (\tau, t)\}$$

y

$$U = \{(\tau, t) \mid \text{ambos lados de (1.2) están definidos y son iguales en } (\tau, t)\}.$$

Notemos que $U \subset V$, que V es un conjunto conexo en el plano $\tau - t$ y que, por lo que acabamos de probar, U es abierto. Por otro lado, si se cumple (1.2) en una sucesión $(\tau_i, t_i) \in U$ cuyo límite es $(\tau^*, t^*) \in V$, entonces, por continuidad, (1.2) se cumple en (τ^*, t^*) , de modo que U contiene a todos sus puntos de acumulación y, por lo tanto, es cerrado. Así, U es abierto y cerrado como subconjunto de V conexo, por lo que $U = V$.

□

1.4 Foliaciones y el Teorema de Frobenius.

Una foliación es, en términos generales, una descomposición de una variedad M en subvariedades de dimensión menor (conocidas como las hojas de la foliación), que yacen pegadas unas con otras de manera suave. En el caso más sencillo, hemos visto cómo cada campo vectorial v en una variedad M determina una curva integral que pasa por cada punto de M , de modo que v sea tangente a la curva integral en todo punto. El teorema de Frobenius trata con el caso análogo para dimensiones mayores: determinar "subvariedades integrales" de sistemas de campos vectoriales, con la propiedad de que cada campo vectorial sea tangente a la subvariedad en cada punto de ella, de modo que tengamos subespacios vectoriales de los espacios tangentes en cada punto, pegados unos a otros suavemente.

Definición 1.29 Sea M una variedad m -dimensional y $n \in \mathbb{N}$, $0 \leq n \leq m$. Una *distribución n -dimensional* \mathcal{D} en M es una colección de subespacios vectoriales $\mathcal{D}_p \subset T_p M$, uno para cada $p \in M$, que son suaves en el sentido de que pueden ser descritos por n campos vectoriales suaves $\{v_1, \dots, v_n\}$ que generan a \mathcal{D} en cada punto de una vecindad $U \subset M$. Un campo vectorial v está en la distribución \mathcal{D} si $v|_p \in \mathcal{D}_p$ para cada $p \in M$. Una distribución \mathcal{D} se llama *involutiva* o *completamente integrable* si para cada dos campos vectoriales en ella, v y w , se tiene que $[v, w] \in \mathcal{D}$, es decir, dados

$\{v_1, \dots, v_n\}$ que generan a \mathcal{D} , existen funciones suaves $h_{ijk}(x) \in \mathcal{F}(M)$, $i, j, k = 1, \dots, n$ tales que, para cada $i, j = 1, \dots, n$, se tiene:

$$[v_i, v_j] = \sum_{k=1}^n h_{ij}^k v_k.$$

Definición 1.30 Una *subvariedad integrable* de una distribución \mathcal{D} , generada por $\{v_1, \dots, v_n\}$, es una subvariedad, N de M , cuyo espacio tangente, $T_q N$, está generado por los vectores $\{v_1|_q, \dots, v_n|_q\}$, para cada $q \in N$.

Proposición 1.31

Sea \mathcal{D} una distribución suave en M tal que cada punto de M pasa por una subvariedad integral de \mathcal{D} , entonces, \mathcal{D} es involutiva.

Prueba.

Sea N una subvariedad integral de \mathcal{D} que pasa por $p \in M$. Podemos pensar que N es una subvariedad inmersa en M mediante un mapeo F , de modo que $F(q) = p$, para algún $q \in N$, y que, para dos campos vectoriales u y v en \mathcal{D} , existen campos vectoriales \tilde{u} y \tilde{w} en N que están F -relacionados con v y w .

Por la proposición 1.26, $[\tilde{u}, \tilde{w}]|_q \in T_q N$ y $[v, w]|_p$ están F -relacionados y, como N es una subvariedad integral de \mathcal{D} , $F'_q: T_q N \rightarrow \mathcal{D}_{F(q)}$ es un isomorfismo en cada $q \in N$, por lo que $[v, w] = F^*[\tilde{u}, \tilde{w}] \in \mathcal{D}$

□

Teorema 1.32 (Frobenius).

Sea \mathcal{D} una distribución n -dimensional, suave e involutiva en M y $p \in M$. Entonces, existe una subvariedad integral de \mathcal{D} que pasa por p ; de hecho, existe una carta (U, χ) en p cuyas funciones coordenadas x^1, \dots, x^m son tales que, las hojas

$$x^i = \text{constante} \quad \forall i = n+1, \dots, m,$$

son subvariedades integrales de \mathcal{D} y, si N es una subvariedad integral conexa de \mathcal{D} contenida en U , entonces N está en una de estas hojas.

Prueba.

Probaremos la parte del teorema correspondiente a la existencia por inducción sobre n , la dimensión de la distribución. Para el caso $n = 1$, escogemos un campo vectorial v en \mathcal{D} definido en una vecindad de $p \in M$ tal que $v|_p \neq 0$. Entonces, el teorema 1.22 establece coordenadas $v = (y^1, \dots, y^m)$ alrededor de p tales que $v(p) = y_0 = 0$ y $v = \frac{\partial}{\partial y^1}$, por lo que la curva integral de v que pasa por cero es la subvariedad buscada.

Supongamos que el teorema se cumple para $n-1$, queremos demostrar que se cumple para una distribución suave \mathcal{D} de dimensión n . Tomemos campos vectoriales v_1, \dots, v_n

que generan a \mathcal{D} en una vecindad de p . Por el teorema 1.22, existen coordenadas $v = (y^1, \dots, y^m)$ en una vecindad V de p tales que $v(p) = y_0 = 0$ y

$$v_1 \Big|_V = \frac{\partial}{\partial y^1}.$$

En V , sean

$$\begin{aligned} w_1 &= v_1, \\ w_i &= v_i - v_i(y^1)v_1, \quad i = 2, \dots, n. \end{aligned}$$

Entonces, los w_1, \dots, w_n son campos vectoriales independientes que generan a \mathcal{D} en V . Sea $S \subset M$ definida por la hoja en la que $y^1 = 0$ y sean

$$u_i = w_i \Big|_S \quad i = 2, \dots, n.$$

Por definición de los w_i , tenemos que

$$w_i(y^1) = 0 \quad i = 2, \dots, n.$$

de modo que los u_i son campos vectoriales en S , es decir, que $u_i \Big|_y \in T_y S$, siempre que $y \in S$ y, por lo tanto, los campos u_i generan una distribución suave en S de dimensión $n - 1$. Afirmamos que esta distribución es involutiva: los u_i están ι -relacionados (con $\iota : S \rightarrow M$ el mapeo inclusión) a los w_i y, por la proposición 1.26, sus paréntesis de Lie también están ι -relacionados a los paréntesis de Lie correspondientes a los w_i . Como los $[w_i, w_j]$, ($i, j \geq 2$) no tienen componente en la dirección y^1 , existen funciones suaves h_{ij}^k tales que

$$[w_i, w_j] = \sum_{k=2}^n h_{ij}^k w_k$$

en V y, por ello,

$$[u_i, u_j] = \sum_{k=2}^n h_{ij}^k u_k.$$

Esto prueba que la distribución en S es involutiva. Por hipótesis de inducción, existen coordenadas (z^2, \dots, z^m) en una vecindad de $p \in S$ tales que, las hojas definidas por $z^i = \text{constante} \quad \forall i = n+1, \dots, m$, son precisamente subvariedades integrales de la distribución generada por $\{u_2, \dots, u_n\}$ en dicha vecindad.

Las funciones

$$\begin{aligned} x^1 &= y^1, \\ x^j &= z^j \circ \pi, \quad j = 2, \dots, n. \end{aligned}$$

donde $\pi : V \rightarrow S$ es la proyección natural en el sistema coordenado $v = (y^1, \dots, y^m)$, están definidas en una vecindad U de $p \in M$, son independientes y todas se anulan en p , por lo que forman una carta (U, χ) en p . Ahora queremos demostrar que

$$w_i(x^{n+r}) \equiv 0 \quad \text{en } U, \quad \text{con } i = 1, \dots, n; \quad r = 1, \dots, m - n,$$

ya que de esto se sigue que los campos vectoriales $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ forman una base de \mathcal{D} en cada punto de U y las hojas $x^i = \text{constante} \quad \forall i = n + 1, \dots, m$, son subvariedades integrales de \mathcal{D} en dicha vecindad.

Observemos que la definición de las x^i implica que

$$\frac{\partial x^j}{\partial y^i} = \begin{cases} 1 & \text{si } j = 1 \\ 0 & \text{si } j = 2, \dots, m \end{cases}$$

en U . Por definición de los w_i y lo anterior, en U se cumple que

$$w_1 = \frac{\partial}{\partial x^1},$$

así que, claramente, $w_1(x^{n+r}) \equiv 0 \quad \forall r = 1, \dots, m - n$. Al tomar $i = 2, \dots, m; r = 1, \dots, m - n$, se tiene que

$$\frac{\partial}{\partial x^1}(w_i(x^{n+r})) = w_1(w_i(x^{n+r})) = [w_1, w_i](x^{n+r}).$$

El hecho de que \mathcal{D} sea involutiva implica que existen funciones h_i^k tales que

$$[w_1, w_i] = \sum_{k=1}^n h_i^k w_k,$$

por lo que, igualando las expresiones para el paréntesis de Lie, se obtiene:

$$\frac{\partial}{\partial x^1}(w_i(x^{n+r})) = \sum_{k=1}^n h_i^k w_k(x^{n+r}).$$

Fijemos una hoja en U de la forma $x^2 = \text{constante}, \dots, x^m = \text{constante}$. En esta hoja, $w_i(x^{n+r})$ son funciones que sólo dependen de x^1 y la igualdad anterior se convierte en un sistema de $n - 1$ ecuaciones diferencial es lineales y homogneas respecto a x^1 , cuya solución es única, dadas condiciones iniciales, por el teorema de existencia y unicidad. Como el sistema es homogéneo, las funciones 0 son una solución pero, además, cada una de estas hojas tiene un único punto en $S \cap U$, donde, por definición de u_i y de x^j :

$$w_i(x^{n+r}) = u_i(z^{n+r}), \quad i = 2, \dots, n$$

y, tomando en cuenta el hecho de que las subvariedades integrales de la distribución en S , determinada por los u_i , están dadas por las hojas adecuadas en las coordenadas z^i , se tiene:

$$u_i(z^{n+r}) = 0, \quad i = 2, \dots, n$$

$$\Rightarrow w_i(x^{n+r}) = 0, \quad i = 2, \dots, n.$$

De las últimas tres igualdades se sigue que las funciones $w_i(x^{n+r})$ deben ser idénticamente cero en U , lo cual completa la Prueba por inducción.

Finalmente, supongamos que N es una subvariedad integral de \mathcal{D} conexa, vista como subvariedad inmersa a través de un mapeo F , de modo que $F(N) \subset U$. Sea $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^{m-n}$ la proyección sobre las últimas $m - n$ coordenadas, entonces, los vectores en \mathcal{D} se anulan al tomar $(\pi \circ \chi)^*$, así que

$$(\pi \circ \chi \circ F)_q^* \equiv 0$$

para cada q en N conexa, y, por el corolario 1.21, $\pi \circ \chi \circ F$ es un mapeo constante. Por lo tanto, $F(N)$ está contenida en una hoja $x^i = \text{constante} \quad \forall i = n+1, \dots, m$. □

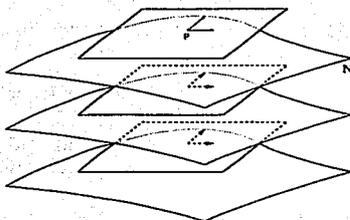


Figura 1.11: Ilustración del Teorema de Frobenius.

1.5 Álgebras de Lie.

Para un grupo de Lie G , existen ciertos campos vectoriales distinguidos, caracterizados por su "invarianza" bajo el producto del grupo. Estos campos vectoriales invariantes forman un espacio vectorial finito dimensional llamado el álgebra de Lie, que puede ser entendida como el "generador infinitesimal" de G ; de hecho, casi toda la información del grupo G está contenida en su álgebra de Lie, lo que constituye el punto medular en la teoría de grupos de Lie.

Empecemos con un grupo de Lie G . Para cada elemento $g \in G$ se define el mapeo *producto por la izquierda* como:

$$I_g : G \rightarrow G$$

$$h \rightarrow gh,$$

el cual es un difeomorfismo con inversa

$$(I_g)^{-1} = I_{g^{-1}}.$$

En este sentido, un campo vectorial es *izquierdo-invariante* si $\forall g$ y $h \in G$, se tiene:

$$(I_g)_h^*(v|_h) = v|_{I_g(h)} = v|_{gh}.$$

Notemos que si v y w son derecho-invariantes, también lo es cualquier combinación lineal suya, por lo que el conjunto de todos los campos vectoriales derecho-invariantes en G forman un espacio vectorial.

Definición 1.33 El *álgebra de Lie de un grupo de Lie G* , denotada como \mathfrak{g} , es el espacio vectorial de todos los campos vectoriales izquierdo-invariantes de G .

Observemos que cualquier campo vectorial izquierdo-invariante puede asociarse a un único vector tangente a G en la identidad, ya que

$$v|_g = v|_{ge} = v|_{I_g(e)} = (I_g)_e^*(v|_e).$$

Igualmente, cualquier vector tangente a G en la identidad determina un único campo vectorial izquierdo-invariante a través de la fórmula anterior:

$$(I_g)_h^*(v|_h) = (I_g)_h^*((I_h)_e^*(v|_e)) = (I_g \circ I_h)_e^*(v|_e) = (I_{gh})_e^*(v|_e) = v|_{gh}.$$

De este modo, podemos identificar a \mathfrak{g} , el álgebra de Lie de G , con el espacio tangente a G en el elemento identidad:

$$\mathfrak{g} \simeq T_e G.$$

Notemos que \mathfrak{g} tiene la misma dimensión que su grupo de Lie.

Además de la estructura de espacio vectorial, un álgebra de Lie está equipada con una operación bilineal y antisimétrica, el paréntesis de Lie de campos vectoriales, puesto que si dos campos vectoriales v y w son izquierdo-invariantes, también lo es su paréntesis de Lie (por la proposición 1.26):

$$(I_g)^*([v, w]) = [(I_g)^*(v), (I_g)^*(w)] = [v, w].$$

Esto motiva una definición general de un álgebra de Lie:

Definición 1.34 Un *álgebra de Lie* es un espacio vectorial sobre \mathbb{R} , \mathfrak{g} , junto con una operación binaria $[,] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, llamada el *paréntesis de Lie* de \mathfrak{g} , que satisface las

siguientes condiciones $\forall v, v_1, v_2, w, w_1, w_2, u \in \mathfrak{g}$:

i) \mathbb{R} -Bilinealidad

$$\begin{aligned} [av_1 + bv_2, w] &= a[v_1, w] + b[v_2, w], \\ [v, aw_1 + bw_2] &= a[v, w_1] + b[v, w_2], \end{aligned}$$

donde a y b son constantes.

ii) Antisimetría

$$[v, w] = -[w, v].$$

iii) Identidad de Jacobi

$$[u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0.$$

Además de los campos vectoriales izquierdo-invariantes (junto con el paréntesis de Lie para campos), existen otros ejemplos de álgebras de Lie, como son: los espacios vectoriales en los que se define que el paréntesis de Lie sea nulo para cualquier par de vectores (a estas álgebras de Lie se les llama *abelianas*); el grupo general lineal con el paréntesis de Lie definido como $[A, B] = AB - BA$; los vectores en \mathbb{R}^3 con el producto cruz como paréntesis de Lie, etc.

El siguiente resultado muestra que hay una correspondencia biyectiva entre los subespacios vectoriales de \mathfrak{g} de dimensión uno (las subálgebras de \mathfrak{g} de una dimensión) y los subgrupos monoparamétricos conexos de G .

Proposición 1.35

Sea $v \neq 0$ un campo vectorial izquierdo-invariante en un grupo de Lie G . Entonces, el flujo generado por v que pasa por la identidad

$$g_t = \exp(tv)e$$

está definido para toda $t \in \mathbb{R}$ y forma un subgrupo monoparamétrico de G con

$$g_{t_1+t_2} = g_{t_1} g_{t_2}, \quad g_0 = e, \quad g_t^{-1} = g_{-t},$$

isomorfo a \mathbb{R} o al grupo circular $SO(2)$. De igual modo, cualquier subgrupo de G , conexo de una dimensión, es generado por un campo vectorial izquierdo-invariante.

Prueba.

Notemos que, en general, para un difeomorfismo F y dos vectores v y $w = F^*(v)$ F -relacionados, F manda a una curva integral de v en una curva integral de w :

$$F(\exp(tv)x) = \exp(tF^*(v))F(x)$$

(1.3)

En nuestro caso, para t_1 y t_2 suficientemente pequeñas se tiene que:

$$\begin{aligned} g_{t_1} g_{t_2} &= I_{g_{t_1}}(g_{t_2}) = I_{g_{t_1}}(\exp(t_2 v) e) \\ &= \exp(t_2 (I_{g_{t_1}})^*(v)) I_{g_{t_1}}(e) \\ &= \exp(t_2 v) g_{t_1} \\ &= \exp(t_2 v) \exp(t_1 v) e \\ &= \exp(t_1 + t_2 v) e \\ &= g_{t_1 + t_2}. \end{aligned}$$

Por lo que g_t es al menos un subgrupo monoparamétrico local. Más aún, g_t está definido al menos para una vecindad $-\frac{1}{2}t_0 \leq t \leq \frac{1}{2}t_0$, para algún $t_0 > 0$, así que podemos definir inductivamente:

$$g_{nt_0+t} = g_{nt_0} g_t, \quad -\frac{1}{2}t_0 \leq t \leq \frac{1}{2}t_0,$$

con n un número entero. Esto último muestra que g_t es una curva suave en G que satisface las propiedades de grupo antes mencionadas $\forall t_1, t_2$, por lo que el flujo está bien definido globalmente y forma un subgrupo de G . Si $g_{t_1} = g_{t_2}$ para algún $t_1 \neq t_2$, entonces $g_{t_0} = e$ para un valor positivo mínimo $t_0 = |t_2 - t_1| > 0$ y $g_{t+t_0} = g_t$, de donde g_t es periódico de periodo t_0 . En este caso $\{g_t\}$ es isomorfo a $SO(2)$. De otro modo, $g_{t_1} \neq g_{t_2}$ para todos $t_1 \neq t_2$, y $\{g_t\}$ es isomorfo a \mathbb{R} .

De igual modo, si $H \subset G$ es un subgrupo unidimensional, tomamos $v|_e$ un vector tangente a H en la identidad distinto de cero. Usando el isomorfismo entre \mathfrak{g} y $T_e G$, podemos extender a v en un campo izquierdo-invariante en G . Dado que H es un subgrupo, se sigue que $v|_h$ es tangente a H en cualquier $h \in H$ y, por ello, H es la curva integral a v que pasa por e .

□

En general, una subálgebra \mathfrak{h} , de un álgebra de Lie \mathfrak{g} , es un subespacio vectorial cerrado bajo el paréntesis de Lie, de modo que $[v, w] \in \mathfrak{h}$, $\forall v, w \in \mathfrak{h}$. Si H es un subgrupo de un grupo de Lie G , cualquier campo vectorial izquierdo-invariante v en H se puede extender a un campo vectorial izquierdo-invariante en G tomando $v|_g = (I_g)_* (v|_e)$. En este sentido, el álgebra de Lie \mathfrak{h} de H se entiende como una subálgebra del álgebra \mathfrak{g} de G . De manera que, usando el teorema de Frobenius, la correspondencia entre los subgrupos monoparamétricos conexos de G y los subespacios de dimensión uno de su álgebra de Lie, \mathfrak{g} , se generaliza a una correspondencia entre los subgrupos conexos de G y las subálgebras de \mathfrak{g} .

El mapeo exponencial $\exp : \mathfrak{g} \rightarrow G$ se define como la función que asocia a cada $v \in \mathfrak{g}$ el flujo generado por v que pasa por el idéntico, al tiempo $t = 1$:

$$\exp : \mathfrak{g} \rightarrow G$$

$$v \rightarrow \exp(v)e \equiv \exp(v).$$

Al analizar la diferencial del mapeo exponencial en 0:

$$\exp_{*0} : T_0\mathfrak{g} \simeq \mathfrak{g} \rightarrow T_e G \simeq \mathfrak{g}$$

observamos que se trata del mapeo identidad. Entonces, por el teorema de la función inversa, \exp determina un difeomorfismo local entre \mathfrak{g} y una vecindad del elemento idéntico en G . En consecuencia, cada elemento $g \in G$, lo suficientemente cerca de la identidad, puede ser escrito como una exponencial: $g = \exp(v)$ para algún v en \mathfrak{g} . En general, el mapeo exponencial no es globalmente inyectivo ni sobre, sin embargo, usando la propiedad 1.9, siempre podemos escribir cualquier elemento del grupo g como un producto finito de exponentiales:

$$g = \exp(v_1) \exp(v_2) \dots \exp(v_k)$$

para algunos v_1, \dots, v_k en \mathfrak{g} . Como consecuencia de esta observación, la prueba de invarianza de algunos objetos bajo la acción de un grupo de Lie completo, G , se reduce a la prueba de invarianza bajo la acción de subgrupos monoparamétricos de G , escribiendo a cualquier $g \in G$ en términos de una base $\{v_1, \dots, v_n\}$ de \mathfrak{g} :

$$g = \exp(t_1 v_1) \exp(t_2 v_2) \dots \exp(t_n v_n) \tag{1.4}$$

lo que se conoce como "invarianza infinitesimal" bajo los generadores infinitesimales de \mathfrak{g} .

Recordando cómo se construyen los grupos locales de Lie a partir de conceptos globales, podemos hacer lo mismo y definir álgebras de Lie para grupos locales de Lie, las cuales seguirán las mismas propiedades que ya hemos visto, tomando en cuenta las regiones en donde el grupo local, junto con sus operaciones, están bien definidos.

Capítulo 2

Simetrías en Ecuaciones

En este capítulo nos enfocamos a obtener los grupos de simetría (conexos) de un sistema de ecuaciones diferenciales. Para esto, desarrollamos la teoría de un caso más simple (las ecuaciones algebraicas) y, una vez que hemos planteado el problema de las ecuaciones diferenciales en términos geométricos similares a los de las ecuaciones algebraicas, establecemos los criterios de simetría equivalentes. Finalmente, analizamos el caso de la ecuación de onda como un ejemplo ilustrativo que nos será útil más adelante.

2.1 Ecuaciones Algebraicas.

Nos interesa estudiar los grupos de simetría de ecuaciones algebraicas, ya que serán una herramienta fundamental en la siguiente sección, donde estudiaremos los grupos de simetría de ecuaciones diferenciales. De acuerdo con lo que hemos venido desarrollando, a partir de ahora vamos a obviar la referencia a sistemas coordenados, en el sentido de que, si tenemos un punto $p \in M$ y una carta (U, χ) en p , de modo que $\chi(p) = x_0$, entonces consideraremos directamente al punto " x_0 en la variedad M ".

Entendemos un "sistema de ecuaciones algebraicas" como un sistema de ecuaciones $F^i(x) = 0$, $i = 1, \dots, n$, donde las F^i son funciones suaves de valores reales definidas en una variedad M ($F^i : M \rightarrow \mathbb{R}$, $F^i \in \mathcal{F}(M)$, $i = 1, \dots, n$). Una *solución* es un punto $x \in M$ tal que $F^i(x) = 0$, $\forall i = 1, \dots, n$. Un grupo de simetría de este sistema será un grupo local de transformaciones que actúa en M transformando soluciones del sistema en otras soluciones; es decir, que si x es solución de $F^i(x) = 0$, $i = 1, \dots, n$ y G es el grupo de simetría, entonces $g \cdot x$ es solución de $F^i(x) = 0$ ($F^i(g \cdot x) = 0$) para todo $g \in G$ donde $g \cdot x$ esté definido.

Definición 2.1 Sea G un grupo local de transformaciones que actúa en una variedad M y $S \subset M$ un subconjunto. S es G -invariante y G es el grupo de simetría de S si $\forall x \in S$ y $g \in G$ donde $g \cdot x$ esté definido, $g \cdot x \in S$.

Por ejemplo, tomemos a $M = \mathbb{R}^2$ y a G_a el grupo de traslaciones en \mathbb{R}^2 :

$$g_t \cdot (x, y) = (x + at, y + t), \quad t \in \mathbb{R}, \quad a = \text{constante}$$

Sea $S_d = \{(x, y) \mid x = ay + d\} \subset M$ el conjunto de las rectas con pendiente $\frac{1}{a}$ y ordenada al origen $-\frac{d}{a}$. Tomamos $(x, y) \in S$, entonces

$$g_t \cdot (x, y) = (x + at, y + t) = (ay + d + at, y + t) = (a(y + t) + d, y + t) \in S$$

Por lo tanto, S es G_a -invariante. Si determinamos las órbitas de la acción, observamos que éstas coinciden con los conjuntos G_a -invariantes; esto se cumple en general, por definición de las órbitas.

Usualmente, el conjunto S corresponde al conjunto solución o la subvariedad determinada por los ceros de una colección de funciones suaves en la variedad, $F = (F^1, \dots, F^n)$, $F: M \rightarrow \mathbb{R}^n$:

$$S = S_F \equiv \{x \mid F^i(x) = 0, \quad i = 1, \dots, n\}.$$

Además de buscar simetrías a los conjuntos solución de un sistema de ecuaciones algebraicas, también podemos buscar simetrías a las funciones $F(x)$ que les dan origen.

Definición 2.2 Sea G un grupo local de transformaciones que actúa en una variedad M . Un mapeo entre variedades $F: M \rightarrow N$ se llama *función G -invariante* si $\forall x \in M$ y $\forall g \in G$ donde $g \cdot x$ esté definido,

$$F(g \cdot x) = F(x),$$

es decir, si la función F manda a cada órbita a un solo punto en N . Una función de valores reales G -invariante $f: M \rightarrow \mathbb{R}$ se llama también *invariante de G* . Notemos que $F: M \rightarrow \mathbb{R}^n$ es G -invariante si y sólo si, cada componente F^i de $F = (F^1, \dots, F^n)$ es una invariante de G .

Retomando el ejemplo anterior ($M = \mathbb{R}^2$, G_a el grupo de traslaciones), podemos ver que $f(x, y) = x - ay$ es invariante de G_a :
Sea $(x, y) \in \mathbb{R}^2$, entonces

$$f(g_t \cdot (x, y)) = f(x + at, y + t) = x + at - a(y + t) = x - ay = f(x, y).$$

Por lo que f es G_a -invariante. Más aún, si $\tilde{f}(x, y) = \phi(x - cy)$, con ϕ función de $x - cy$ suave, se tiene:

$$\tilde{f}(g_t \cdot (x, y)) = \tilde{f}(x + at, y + t) = \phi(x + at - a(y + t)) = \phi(x - ay) = \tilde{f}(x, y).$$

Por lo que \tilde{f} es invariante de G_a .

Si $F : M \rightarrow \mathbb{R}^n$ es una función G -invariante, es claro que todo conjunto de nivel de F es un subconjunto de M G -invariante, en particular el conjunto solución. Sin embargo, no necesariamente se cumple que, si el conjunto solución es G -invariante, entonces F es G -invariante; la condición necesaria y suficiente es, pues, que todo conjunto de nivel sea G -invariante.

Proposición 2.3

Si G actúa en M y $F : M \rightarrow \mathbb{R}^n$ es una función suave, entonces F es una función G -invariante si y sólo si todo conjunto de nivel $\{x \mid F(x) = c\}$, $c \in \mathbb{R}^n$, es un subconjunto G -invariante de M .

Prueba.

En la dirección " \Rightarrow ", sean F una función suave G -invariante y $S = \{x \in M \mid F(x) = c\}$, $c \in \mathbb{R}^n$, $c = \text{constante}$. Tomamos $x \in S$ y $g \in G$, entonces, como F es G -invariante, $F(g \cdot x) = F(x) = c$ por lo que $g \cdot x \in S$ y, por lo tanto, S es G -invariante.

En la dirección " \Leftarrow ", tenemos que $S_c = \{x \in M \mid F(x) = c\}$ es invariante $\forall c \in \mathbb{R}^n$, $c = \text{constante}$. Sean $x \in M$ y $g \in G$, entonces $x \in S_c$ para alguna $c \in \mathbb{R}^n$, por lo que $g \cdot x \in S_c$, es decir, $F(g \cdot x) = c = F(x)$, por lo tanto, F es G -invariante. □

Otro modo de entender este resultado es que el grupo de simetría del conjunto solución S_F de un sistema de ecuaciones algebraicas es, en general, más grande que el grupo de simetría de la función que lo determina.

Gran parte de la utilidad de la teoría de grupos de Lie radica en el hecho de que es posible cambiar condiciones complicadas, no lineales, para la invariancia de conjuntos o funciones, por condiciones lineales equivalentes en los generadores infinitesimales de la acción de un grupo. Este criterio de invarianza nos ayudará más adelante a determinar explícitamente los grupos de simetría de sistemas de ecuaciones diferenciales. En el caso de funciones invariantes, el criterio se sigue al describir cómo cambian las funciones a lo largo de un flujo generado por un campo vectorial.

Proposición 2.4

Sea G un grupo de transformaciones conexo que actúa en una variedad M . Una función suave $f : M \rightarrow \mathbb{R}$ es invariante de G si y sólo si, para todo $x \in M$ y todo v generador infinitesimal de G , se cumple

$$v(f) = 0$$

Es decir, que f será invariante de G si y sólo si no cambia a lo largo de las órbitas de la acción, que son las curvas integrales de los generadores de la acción correspondiente.

Prueba.

Supongamos que f es una función invariante de G y v un generador infinitesimal de G . Sabemos que

$$v(f)(x) = \left. \frac{d}{dt} f(\exp(tv)x) \right|_{t=0}$$

y como f es invariante de G , $f(\exp(tv)x) = f(x)$, de donde

$$\frac{d}{dt} f(\exp(tv)x) = \frac{d}{dt} f(x) = 0$$

Por lo que $v(f) = 0$.

Por otro lado, si $v(f) = 0 \quad \forall x \in M$ y $\forall v$ generador infinitesimal de G , entonces:

$$\begin{aligned} v(f)(\exp(tv)x) &= 0 \\ \Rightarrow \frac{d}{dt} f(\exp(tv)x) &= 0 \\ \Rightarrow f(\exp(tv)x) &= \text{constante} = f(x), \end{aligned}$$

de modo que $f(\exp(tv)x)$ es una constante para el subgrupo local monoparamétrico definido por el flujo de v , que está contenido en $G_x \equiv \{g \mid g \cdot x \text{ está definido}\}$. Por (1.4), sabemos que todo $g \in G_x$ puede escribirse como producto finito de exponenciales de los generadores infinitesimales de G , v_1, \dots, v_n , de donde

$$f(g \cdot x) = f(\exp(t_1 v_1) \exp(t_2 v_2) \dots \exp(t_n v_n) x) = \text{constante} = f(x) \quad \forall g \in G_x, \quad \forall x \in M.$$

Por lo tanto, f es invariante de G . □

Si $\{v_1, \dots, v_n\}$ forman una base para \mathfrak{g} , el álgebra de Lie de los generadores infinitesimales de G , lo que dice la proposición anterior es que $f(x)$ es invariante si y sólo si $v_k(f) = 0, \quad k = 1, \dots, n$. De manera que, en coordenadas locales, f debe ser solución del siguiente sistema de ecuaciones diferenciales parciales lineales de primer orden:

$$v_k(f) = \sum_{i=1}^n \xi_k^i(x) \frac{\partial f}{\partial x^i} = 0, \quad k = 1, \dots, n.$$

Retomando el ejemplo del grupo de traslaciones G_a en \mathbb{R}^2 , tenemos que el generador infinitesimal de la acción es: $v = \xi^1 \frac{\partial}{\partial x} + \xi^2 \frac{\partial}{\partial y}$, con

$$\begin{aligned} \xi^1 &= \frac{\partial}{\partial t} x + at = a; & \xi^2 &= \frac{\partial}{\partial t} y + t = 1 \\ \Rightarrow v &= a \frac{\partial}{\partial x} + \frac{\partial}{\partial y}. \end{aligned}$$

Si tomamos $f(x, y) = x - ay$, entonces

$$v(f)(x, y) = v(x - ay) = \left(a \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) (x - ay) = a - a = 0$$

de donde $v(f) = 0$ y, por la proposición anterior, f es invariante de G_a , lo cual coincide con lo que habíamos obtenido anteriormente.

Para el caso de soluciones a sistemas de ecuaciones algebraicas $F(x) = 0$, el criterio de invarianza requiere una condición más, a saber, la condición de rango máximo.

Teorema 2.5

Sea G un grupo de transformaciones conexo que actúa en una variedad M de dimensión n . Sea $F : M \rightarrow \mathbb{R}^n$, $n \leq m$ una función que define un sistema de ecuaciones algebraicas

$$F^i(x) = 0, \quad i = 1, \dots, n.$$

Spongamos que el sistema es de rango máximo, es decir, que la matriz jacobiana $\frac{\partial F^i}{\partial x^j}$ es de rango $n \quad \forall x$ solución del sistema. Entonces, G es el grupo de simetría del sistema (el grupo de simetría del conjunto solución) si y sólo si, para todo v generador infinitesimal de G , se cumple

$$v(F^i)(x) = 0,$$

siempre que $F(x) = 0$ (F^i no cambia en la intersección del conjunto solución y las órbitas de la acción).

Prueba.

Spongamos que G es el grupo de simetría del sistema, entonces G es el grupo de simetría de S_F , así que $\forall g \in G$ y $x \in S_F$ donde $g \cdot x$ esté definido,

$$\begin{aligned} g \cdot x \in S_F &\Rightarrow F(g \cdot x) = 0 \\ &\Rightarrow F^i(g \cdot x) = 0 \end{aligned}$$

En particular, para v generador de G

$$\begin{aligned} F^i(\exp(tv)x) &= 0 \\ \Rightarrow v(F^i)(x) &= \frac{d}{dt} F^i(\exp(tv)x) = 0. \end{aligned}$$

Por lo que $v(F^i)(x) = 0$, $x \in S_F$.

Consideremos ahora que $v(F^i)(x) = 0 \quad \forall i = 1, \dots, n$, $x \in S_F$. Por la condición de rango máximo en el conjunto solución y el teorema 1.6, podemos suponer que existen coordenadas $y = (y^1, \dots, y^m)$ en una vecindad de $x_0 \in M$, tales que $x_0 = 0$ y

$F(y) = (y^1, \dots, y^n)$. Tomemos $v = \sum_{i=1}^n \frac{\partial}{\partial y^i}$, un generador infinitesimal de G expresado en las nuevas coordenadas. Como $v(F^i)(y) = 0$, sucede que

$$v(y^i) = 0, \quad i = 1, \dots, n, \quad \Rightarrow \quad \xi^i(y) = 0, \quad i = 1, \dots, n,$$

en la vecindad de x_0 donde tenemos definidas nuestras coordenadas. Ahora bien, el flujo de v que pasa por $x_0 = 0$, $\varphi(t, 0) = \exp(tv)x_0$, cumple con el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned} \frac{\partial \varphi^i}{\partial t} &= \xi^i = 0, \quad \varphi^i(0) = 0, \quad i = 1, \dots, m \\ &\Rightarrow \varphi(t, 0) = 0 \\ &\Rightarrow \exp(tv)x_0 = x_0 \\ &\Rightarrow F(\exp(tv)x_0) = F(x_0) = 0 \end{aligned}$$

para t suficientemente pequeña, lo que implica que $\exp(tv)x_0$ es solución del sistema si t es lo suficientemente chica. Por continuidad de F , S_F es cerrado y, como $\exp((t + \tau)v)x_0 = \exp(tv)\exp(\tau v)x_0$ y el mapeo exponencial es continuo, podemos decir que, para cualquier $g = \exp(tv)$ en el subgrupo monoparamétrico conexo generado por v contenido en G_{x_0} , $g \cdot x_0$ es solución del sistema. Finalmente, si aplicamos un argumento similar al empleado en la demostración de la proposición 2.4, se tiene que $F(g \cdot x_0) = F(x) \quad \forall g \in G, x \in S_F$ donde $g \cdot x$ esté definido. Por lo tanto, G es grupo de simetría del sistema. □

Tomemos como ejemplo el grupo de rotaciones en \mathbb{R}^2 , $\mathbf{SO}(2)$, cuyo generador infinitesimal es $v = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$. El círculo unitario $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$ es el conjunto solución de $f(x, y) = x^2 + y^2 - 1$; de hecho,

$$v(f) = \left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right)(x^2 + y^2 - 1) = -2xy + 2xy = 0.$$

Además, $\nabla f|_{(x,y)} = (2x, 2y) \neq 0 \quad \forall (x, y) \in S^1$, por lo que se cumplen las condiciones del teorema 2.5 y tenemos que $\mathbf{SO}(2)$ es el grupo de simetría de f .

Otro ejemplo menos trivial consiste en tomar la función $\tilde{f}(x, y) = x^4 + x^2y^2 + y^2 - 1$. Se tiene que:

$$\nabla \tilde{f}|_{(x,y)} = (4x^3 + 2xy^2, 2x^2y + 2y) = 0 \quad \Leftrightarrow \quad x = y = 0$$

y

$$v(\tilde{f}) = -4x^3y - 2xy^3 + 2x^3y + 2xy = \frac{-2xy}{x^2+1}(x^4 + x^2y^2 + y^2 - 1) = \frac{-2xy}{x^2+1}\tilde{f}(x, y).$$

De modo que $\tilde{f}(x, y) = 0 \Rightarrow v(\tilde{f})(x, y) = 0$ y, como el $(0, 0)$ no está en el conjunto solución, \tilde{f} es de rango máximo en $S_{\tilde{f}} = \{(x, y) \mid x^4 + x^2 y^2 + y^2 = 1\}$. Por el teorema 2.5, $\text{SO}(2)$ es el grupo de simetría de \tilde{f} y $S_{\tilde{f}}$ es rotacionalmente invariante. Cabe mencionar que en este caso \tilde{f} no es $\text{SO}(2)$ -invariante, de hecho, cualquier otro conjunto de nivel de \tilde{f} no es rotacionalmente invariante.

La condición de rango máximo en el conjunto solución es muy importante puesto que, si tomamos $h(x, y) = (y - 1)^2$, el conjunto solución estará dado por la recta $y = 1$, la cual es claramente no invariante bajo rotaciones. Por otro lado, al tomar $v = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}$ y aplicarlo a h tenemos, $v(h) = 2x(y - 1)$, de modo que $h(x, y) = 0 \Rightarrow v(h)(x, y) = 0$. Se cumplen, pues, casi todas las condiciones del teorema 2.5 y uno podría pensar que dicho teorema nos está llevando a la conclusión errónea de que la recta $y = 1$ es invariante bajo rotaciones. Sin embargo, no se cumple la condición de rango máximo en el conjunto solución, puesto que

$$\nabla h \Big|_{(x,y)} = (0, 2y - 2) = 0 \quad \text{en } y = 1$$

Proposición 2.6

Sea $F : M \rightarrow \mathbb{R}^n$ de rango máximo en la subvariedad S_F , el conjunto solución de F . Una función de valores reales $f : M \rightarrow \mathbb{R}$ se anula en S_F si y sólo si existen funciones suaves Q^1, \dots, Q^n tales que

$$f(x) = Q^1(x)F^1(x) + \dots + Q^n(x)F^n(x) \quad \forall x \in M.$$

Prueba.

Supongamos que $f : M \rightarrow \mathbb{R}$ se anula en S_F y sea $x_0 \in S_F$. Como S_F es el conjunto solución del sistema de ecuaciones algebraicas determinado por $F = (F^1, \dots, F^n)$, entonces $F(x_0) = 0 \Rightarrow F^i(x_0) = 0, \quad i = 1, \dots, n$. Tenemos también que $f(x_0) = 0$, así que en x_0 podemos encontrar una factorización de f en términos de las F^i :

$$f(x_0) = Q^1(x_0)F^1(x_0) + \dots + Q^n(x_0)F^n(x_0) = \sum_{i=1}^n Q^i(x_0)F^i(x_0) = 0,$$

donde las funciones $Q^i(x), \quad i = 1, \dots, n$ son suaves y no todas se anulan en x_0 .

Como F es de rango máximo en S_F , existen coordenadas (y^1, \dots, y^m) en una vecindad V de $x_0 = y_0$ tales que

$$y = (y^1, \dots, y^n, 0, \dots, 0), \quad \forall y \in V \cap S_F.$$

En estas coordenadas, la matriz jacobiana de F , evaluada en y_0 , es

$$\begin{pmatrix} \left. \frac{\partial}{\partial y^1} F^1 \right|_{y_0} & \dots & \left. \frac{\partial}{\partial y^m} F^1 \right|_{y_0} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial}{\partial y^1} F^n \right|_{y_0} & \dots & \left. \frac{\partial}{\partial y^m} F^n \right|_{y_0} \end{pmatrix}$$

y podemos suponer que hay una submatriz invertible de $n \times n$:

$$A = \begin{pmatrix} \left. \frac{\partial}{\partial y^1} F^1 \right|_{y_0} & \cdots & \left. \frac{\partial}{\partial y^n} F^1 \right|_{y_0} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial}{\partial y^1} F^n \right|_{y_0} & \cdots & \left. \frac{\partial}{\partial y^n} F^n \right|_{y_0} \end{pmatrix}.$$

Notemos que, como f se anula en S_F , en las nuevas coordenadas sucede que

$$f(y) = f(y^1, \dots, y^n, 0, \dots, 0) = 0, \quad \forall y \in V \cap S_F \Rightarrow \left. \frac{\partial f}{\partial y^i} \right|_{y \in V \cap S_F} = 0, \quad i = 1, \dots, n,$$

así que, en $V \cap S_F$, f depende sólo de las coordenadas y^{n+1}, \dots, y^m .

Definimos $g: M \rightarrow \mathbb{R}$ como

$$g(y) = f(y) - \sum_{i=1}^n Q^i(y) F^i(y) = 0.$$

claramente, $g(y_0) = 0$ y

$$\begin{aligned} \nabla g \Big|_{y_0} &= \nabla f \Big|_{y_0} - \left(\frac{\partial}{\partial y^1} \sum_{i=1}^n Q^i(y) F^i(y), \dots, \frac{\partial}{\partial y^m} \sum_{i=1}^n Q^i(y) F^i(y) \right) \Big|_{y_0} \\ &= \left(0, \dots, 0, \left. \frac{\partial f}{\partial y^{n+1}} \right|_{y_0}, \dots, \left. \frac{\partial f}{\partial y^m} \right|_{y_0} \right) - \left(\sum_{i=1}^n Q^i(y_0) \left. \frac{\partial F^i}{\partial y^1} \right|_{y_0}, \dots, \sum_{i=1}^n Q^i(y_0) \left. \frac{\partial F^i}{\partial y^m} \right|_{y_0} \right) \\ &\quad - \left(\sum_{i=1}^n F^i(y_0) \left. \frac{\partial Q^i}{\partial y^1} \right|_{y_0}, \dots, \sum_{i=1}^n F^i(y_0) \left. \frac{\partial Q^i}{\partial y^m} \right|_{y_0} \right) \\ &= \left(\sum_{i=1}^n Q^i \left. \frac{\partial F^i}{\partial y^1} \right|_{y_0}, \dots, \sum_{i=1}^n Q^i \left. \frac{\partial F^i}{\partial y^n} \right|_{y_0}, \left. \frac{\partial f}{\partial y^{n+1}} \right|_{y_0}, \dots, \left. \frac{\partial f}{\partial y^m} \right|_{y_0} - \sum_{i=1}^n Q^i \left. \frac{\partial F^i}{\partial y^{n+1}} \right|_{y_0}, \dots, \left. \frac{\partial f}{\partial y^m} \right|_{y_0} - \sum_{i=1}^n Q^i \left. \frac{\partial F^i}{\partial y^m} \right|_{y_0} \right) \Big|_{y_0}. \end{aligned}$$

Notemos que si tomamos $Q = (Q^1, \dots, Q^n)$, la primeras n coordenadas de $\nabla g \Big|_{y_0}$ pueden escribirse como

$$A^t Q = \begin{pmatrix} \left. \frac{\partial}{\partial y^1} F^1 \right|_{y_0} & \cdots & \left. \frac{\partial}{\partial y^n} F^1 \right|_{y_0} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial}{\partial y^1} F^n \right|_{y_0} & \cdots & \left. \frac{\partial}{\partial y^n} F^n \right|_{y_0} \end{pmatrix} \begin{pmatrix} Q^1(y_0) \\ \vdots \\ Q^n(y_0) \end{pmatrix}.$$

Dado que A es invertible y escogimos las $Q^i(x)$ de modo que no todas fueran cero en y_0 , tenemos que

$$\begin{aligned} \sum_{i=1}^n Q^i(y_0) \left. \frac{\partial F^i}{\partial y^j} \right|_{y_0} &\neq 0 \quad \forall j = 1, \dots, n. \\ \Rightarrow \nabla g \Big|_{y_0} &\neq 0. \end{aligned}$$

De este modo, se cumplen las condiciones del teorema de la función implícita y, por ello, podemos tomar una vecindad abierta U que contiene a $y_0 = x_0$ donde $g(y) = 0$, $\forall y \in U$. Por lo tanto,

$$f(y) = Q^1(y)F^1(y) + \dots + Q^n(y)F^n(y) \quad \forall y \in U.$$

Con esto hemos probado localmente la dirección " \Rightarrow " del teorema. Una vez hecho esto, es posible obtener la versión global usando particiones de la unidad.

En la dirección " \Leftarrow ", supongamos que existen funciones suaves Q^1, \dots, Q^n tales que

$$f(x) = Q^1(x)F^1(x) + \dots + Q^n(x)F^n(x) \quad \forall x \in M.$$

Sea $x_0 \in S_f \Rightarrow F(x_0) = 0 \Rightarrow F^i(x_0) = 0$, $\forall i = 1, \dots, n$. Por lo tanto,

$$f(x_0) = Q^1(x_0)F^1(x_0) + \dots + Q^n(x_0)F^n(x_0) = 0,$$

es decir, f se anula en S_f . □

Lo que nos dice la proposición anterior es que podemos remplazar el criterio de invarianza dado por el teorema 2.5, por la condición equivalente

$$v(F^i) = \sum_{j=1}^n Q^j(x)F^j(x), \quad i = 1, \dots, n, \quad x \in M,$$

donde las funciones suaves Q^j están por determinarse.

Nuevamente, para un grupo de transformaciones, podemos tomar las ideas de invarianza de funciones o conjuntos y reducirlas al caso local. Así, únicamente se requiere cumplir la invarianza para un grupo de transformaciones suficientemente cercano a la identidad.

Definición 2.7 Sea G un grupo local de transformaciones que actúa en una variedad M . Un subconjunto $S \subset M$ se llama *localmente G -invariante* si $\forall x \in S$ existe una vecindad de la identidad en G , $\tilde{G}_x \subset G_x$, tal que $g \cdot x \in S \quad \forall g \in \tilde{G}_x$. Una función suave entre variedades $F: U \rightarrow N$, donde U es un subconjunto abierto de M , se llama *localmente G -invariante* si $\forall x \in U$ existe una vecindad de la identidad en G , $\tilde{G}_x \subset G_x$, tal que $F(g \cdot x) = F(x) \quad \forall g \in \tilde{G}_x$. F se llama *globalmente G -invariante* si $F(g \cdot x) = F(x)$, $\forall x \in U$ y $\forall g \in G$ tal que $g \cdot x \in U$.

En el caso de una subvariedad $N \in M$, N es localmente G -invariante si y sólo si los generadores infinitesimales v de G son tangentes a N en cualquiera de sus puntos.

Muchas veces es interesante determinar el "número de invariantes" que tiene un grupo de transformaciones. En concreto, el problema consiste en encontrar funciones

de valores reales $f^1(x), \dots, f^n(x)$ que sean invariantes de un grupo de transformaciones. Si se tiene un mapeo suave $F(z^1, \dots, z^n)$, entonces $F(f^1(x), \dots, f^n(x))$ también será un invariante, el cual no aporta información adicional al problema y se le conoce como "funcionalmente dependiente de las invariantes anteriores $f^1(x), \dots, f^n(x)$ ".

Definición 2.8 Sean $f^1(x), \dots, f^n(x)$ funciones suaves de valores reales, definidas en una variedad M .

i) f^1, \dots, f^n son funcionalmente dependientes si $\forall x \in M$ existe una vecindad $U \subset M$ de x y una función $F: V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ no idénticamente cero, tal que

$$F(f^1(x), \dots, f^n(x)) = 0, \quad \forall x \in U$$

ii) f^1, \dots, f^n son funcionalmente independientes si no son funcionalmente dependientes cuando se les restringe a un subconjunto abierto $U \subset M$.

Teorema 2.9

Sea G un grupo de transformaciones que actúa semirregularmente en M una variedad de dimensión m . Consideremos que las órbitas de la acción son de dimensión n y sea $x_0 \in M$. Entonces, existen $m - n$ invariantes f^1, \dots, f^{m-n} funcionalmente independientes definidas en una vecindad de x_0 . Además, cualquier otra invariante local de la acción del grupo definida cerca de x_0 es de la forma:

$$f = F(f^1(x), \dots, f^n(x))$$

para alguna F , función suave. Si la acción de G es regular, entonces los invariantes pueden considerarse globalmente invariantes en una vecindad de x_0 .

Prueba.

Sea \mathfrak{g} el sistema de generadores infinitesimales de G . De la definición de álgebra de Lie, tenemos que \mathfrak{g} es una distribución suave e involutiva en x_0 , por lo que, aplicando el teorema de Frobenius, podemos dar coordenadas locales y^1, \dots, y^m cerca de x_0 tales que las hojas

$$y^i = \text{constante}, \quad \forall i = n + 1, \dots, m,$$

son las órbitas de \mathfrak{g} . Así, estas nuevas coordenadas $y^{n+1} = f^1(x), \dots, y^m = f^{m-n}(x)$ son en sí mismas invariantes locales de G . Además, cualquier otro invariante de G deberá ser constante en estas hojas, de modo que depende solamente de las coordenadas y^{n+1}, \dots, y^m . Ahora bien, si G actúa regularmente, podemos escoger las cartas coordenadas de modo que cada órbita interseque sólo a una hoja; en este caso, y^{n+1}, \dots, y^m son invariantes globales.

□

Por lo general se conoce a los invariantes del teorema como *conjunto completo de invariantes funcionalmente independientes*. En el caso de subvariedades, existe un

resultado análogo que nos dice que, si f^1, \dots, f^{m-n} son un conjunto completo de invariantes funcionalmente independientes definidas en $U \subset M$ abierto y S_F es G -invariante, entonces, para cada solución $x_0 \in S_F$ existe una vecindad \tilde{U} y una función $\tilde{F} = \tilde{F}(f^1(x), \dots, f^{m-n}(x))$ tal que su conjunto solución coincide con el de F en \tilde{U} .

El problema de encontrar los invariantes de un grupo de transformaciones puede ser bastante complejo, sin embargo, podemos resolverlo de manera sencilla para el caso en que tenemos un grupo monoparamétrico de transformaciones G , actuando en una variedad M . Tomemos $v = \sum_{i=1}^m \xi^i \frac{\partial}{\partial x^i}$ el generador infinitesimal de la acción, expresado en coordenadas locales. De acuerdo con la proposición 2.4, una función f , invariante local de G , debe ser solución de la siguiente ecuación diferencial lineal y homogénea, de primer orden:

$$v(f) = \sum_{i=1}^m \xi^i \frac{\partial f}{\partial x^i} = 0 \quad (2.1)$$

Por el teorema 2.9, existen $m - 1$ invariantes funcionalmente independientes, es decir, existen $m - 1$ soluciones a (2.1) funcionalmente independientes en una vecindad de x_0 .

Ahora bien, de acuerdo a la teoría clásica de ecuaciones diferenciales parciales, si $v|_{x_0} \neq 0$ en $x_0 \in M$, podemos buscar las curvas integrales de v usando el método de características:

$$\frac{dx^i}{dt} = \xi^i \Rightarrow \frac{dx^i}{\xi^i} = dt, \quad i = 1, \dots, m,$$

por lo que tenemos el siguiente sistema característico:

$$\frac{dx^1}{\xi^1} = \dots = \frac{dx^m}{\xi^m},$$

cuayas soluciones generales pueden ser escritas de la forma:

$$f^1(x^1, \dots, x^m) = c^1, \dots, f^{m-1}(x^1, \dots, x^m) = c^{m-1},$$

donde c^1, \dots, c^{m-1} son constantes de integración y las f^i son funciones independientes de las c^j , $i, j = 1, \dots, m - 1$, $i \neq j$. Es claro que las f^i son las soluciones a (2.1) funcionalmente independientes que estamos buscando.

Por ejemplo, tomemos el grupo de rotaciones en \mathbb{R}^2 , $\text{SO}(2)$, cuyo generador infinitesimal es $v = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}$. El correspondiente sistema característico es:

$$\frac{dx}{-y} = \frac{dy}{x}.$$

Resolviendo el sistema, encontramos que la solución general es $x^2 + y^2 = c$, donde c es una constante arbitraria. Entonces, $f(x, y) = x^2 + y^2$ o cualquier otra función de f es una invariante de $\text{SO}(2)$.

Otro ejemplo consiste en tomar el campo vectorial

$$v(f) = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} + (1 + z^2) \frac{\partial}{\partial z}$$

en \mathbb{R}^3 . De acuerdo con el teorema 2.9, existen $3 - 1 = 2$ invariantes del grupo monoparamétrico generado por v en una vecindad del un punto en \mathbb{R}^3 . En este caso, el sistema característico es:

$$\frac{dx}{-y} = \frac{dy}{x} = \frac{dz}{(1 + z^2)}$$

Podemos obtener una invariante usando la primera igualdad, la cual ya resolvimos en el ejemplo anterior: $r = \sqrt{x^2 + y^2}$. Para obtener la segunda invariante, sustituimos x por $\sqrt{r^2 + y^2}$ en la segunda igualdad e integramos:

$$\begin{aligned} \frac{dy}{\sqrt{r^2 + y^2}} &= \frac{dz}{(1 + z^2)} \\ \Rightarrow \arcsin\left(\frac{y}{r}\right) &= \arctan z + k, \end{aligned}$$

con k constante. Así, $\arctan z - \arcsin\left(\frac{y}{r}\right) = \arctan z - \arctan\left(\frac{y}{x}\right)$ es una segunda invariante funcionalmente independiente, la cual puede expresarse de un modo más sencillo si sacamos su tangente, de modo que

$$f^1 = x^2 + y^2 \quad \text{y} \quad f^2 = \frac{xz - y}{x + yz}$$

forman un conjunto completo de invariantes funcionalmente independientes del grupo monoparamétrico generado por v .

2.2 Ecuaciones Diferenciales.

Un sistema de ecuaciones diferenciales S , plantea un problema que puede ser abordado por la teoría de grupos de Lie. Para hacer esto, es necesario hacer un replanteamiento que nos permita analizar la situación desde un enfoque geométrico, de manera que podamos hacer uso de la teoría desarrollada en el sección anterior. Digamos que el sistema S involucra p variables independientes, x^1, \dots, x^p y q variables dependientes u^1, \dots, u^q , entonces podemos considerar al espacio de las variables independientes como $X \simeq \mathbb{R}^p$, con coordenadas $x = (x^1, \dots, x^p)$, y al espacio de las variables dependientes como $U \simeq \mathbb{R}^q$, con coordenadas $u = (u^1, \dots, u^q)$. Las soluciones del sistema son de la forma $u = f(x)$ y el grupo de simetría de S es el grupo local de transformaciones G que actúa en un conjunto abierto $M \subset X \times U$, transformando soluciones de S en soluciones de S .

Si queremos dar un significado más preciso a la noción de grupo de simetría de un sistema de ecuaciones diferenciales, debemos empezar por entender cómo un determinado $g \in G$ transforma una función: $u = f(x)$. Lo primero es identificar la función con su gráfica:

$$\Gamma_f = \{(x, f(x)) \mid x \in \Omega\} \subset X \times U,$$

donde $\Omega \subset X$ es el dominio de la función f . Notemos que Γ_f es una subvariedad m -dimensional de $X \times U$. Consideremos además que $\Gamma_f \subset M_g$, el dominio donde la transformación determinada por g está definida, entonces entendemos la transformación de Γ_f por g como:

$$g \cdot \Gamma_f = \{(\tilde{x}, \tilde{u}) = g \cdot (x, f(x)) \mid (x, u) \in \Gamma_f\}.$$

El conjunto $g \cdot \Gamma_f$ no necesariamente forma la gráfica de una función $\tilde{u} = \tilde{f}(\tilde{x})$, sin embargo, dado que la acción es suave y que el idéntico de G no cambia a Γ_f , si reducimos el dominio de definición de f , podemos garantizar que, para una vecindad cerca de la identidad, la transformación $g \cdot \Gamma_f$ es igual a la gráfica de una función \tilde{u} ($g \cdot \Gamma_f = \Gamma_{\tilde{f}}$, $\tilde{u} = \tilde{f}(\tilde{x})$). Llamamos a la función $\tilde{f} = g \cdot f$ la *transformada de f por medio de g* .

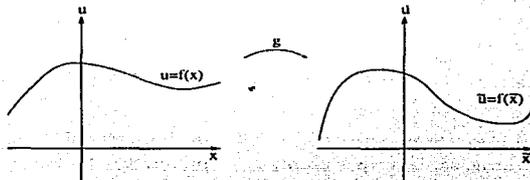


Figura 2.1: Transformada de f por medio de g .

Por ejemplo, sea $u = f(x) = ax + b$ y consideremos a $G = \text{SO}(2)$, el grupo de rotaciones en \mathbb{R}^2 . La gráfica de f es una línea recta así que, si la rotamos un ángulo $\theta \neq 0$, seguimos teniendo una línea recta que, a menos de que sea vertical, podemos considerar como la gráfica de otra función $\theta \cdot f = \tilde{f}$, que es la transformación de f por la rotación de un ángulo θ . Concretamente, para encontrar $\theta \cdot f = \tilde{f}$, tomamos un punto en Γ_f , $(x, u) = (x, ax + b)$ y la transformación de (x, u) por θ :

$$\begin{aligned} \theta \cdot (x, u) &= (x \cos \theta - u \sin \theta, x \sin \theta + u \cos \theta) \\ &= (x \cos \theta - (ax + b) \sin \theta, x \sin \theta + (ax + b) \cos \theta) = (\tilde{x}, \tilde{u}) \end{aligned}$$

de donde

$$\tilde{x} = x \cos \theta - ax \sin \theta - b \sin \theta$$

$$\Rightarrow x = \frac{\bar{x} + b \sin \theta}{\cos \theta - a \sin \theta},$$

que está bien definido si $\cot \theta \neq a$, en particular para θ suficientemente cerca de 0. Así

$$\begin{aligned} \bar{u} &= x \sin \theta + ax \cos \theta - b \cos \theta \\ &= \left(\frac{\bar{x} + b \sin \theta}{\cos \theta - a \sin \theta} \right) (\sin \theta + a \cos \theta) + b \cos \theta \\ &= \frac{\sin \theta + a \cos \theta}{\cos \theta - a \sin \theta} \bar{x} + \frac{b}{\cos \theta - a \sin \theta} \end{aligned}$$

Por lo tanto, $\bar{u} = \bar{f}(\bar{x})$ y nuevamente tenemos una recta bien definida como función, bajo la condición que hemos impuesto a θ .

En general, el procedimiento para encontrar $\bar{f} = g \cdot f$ es muy similar a lo que se hizo en el ejemplo anterior. Si la transformación definida por g está dada como

$$(\bar{x}, \bar{u}) = g \cdot (x, u) = (\Psi_g(x, u), \Phi_g(x, u)),$$

donde Ψ_g y Φ_g son funciones suaves, la gráfica $\Gamma_f = g \cdot \Gamma_f$ está dada paramétricamente por las ecuaciones

$$\begin{aligned} \bar{x} &= \Psi_g(x, f(x)) = \Psi_g \circ (id \times f)(x) \\ \bar{u} &= \Phi_g(x, f(x)) = \Phi_g \circ (id \times f)(x), \end{aligned}$$

con $x \in \Omega$, un dominio adecuado e $id(x) = x$ la función identidad en X . Para encontrar $\bar{f} = g \cdot f$ explícitamente, hay que eliminar el parámetro x de estos dos sistemas de ecuaciones. Si $g = e$, se tiene $\Psi_e \circ (id \times f) = id$, por lo tanto, la matriz jacobiana de $\Psi_e \circ (id \times f)$ es no singular, así que, si tomamos g suficientemente cerca de la identidad, por el teorema de la función inversa, podemos resolver para x localmente:

$$x = (\Psi_e \circ (id \times f))^{-1}(\bar{x}).$$

Sustituyendo en el segundo sistema de ecuaciones, encontramos $g \cdot f$:

$$g \cdot f = (\Psi_g \circ (id \times f)) \circ (\Psi_g \circ (id \times f))^{-1}. \quad (2.2)$$

Ahora podemos dar una definición precisa del concepto de grupo de simetría de un sistema de ecuaciones diferenciales:

Definición 2.10 Sea S un sistema de ecuaciones diferenciales. Un grupo de simetría del sistema S es un grupo local de transformaciones G que actúa en un subconjunto M del espacio de variables independientes y dependientes del sistema, con la propiedad

de que, si $u = f(x)$ es una solución de \mathcal{S} y $g \cdot f$ está definida para $g \in G$, entonces $\tilde{u} = g \cdot f$ también es solución del sistema.

Como ejemplo, demos $u_{xx} = \frac{d^2 u}{dx^2} = 0$. Sabemos que las soluciones a esta ecuación son funciones lineales y ya vimos que el grupo de rotaciones en \mathbb{R}^2 transforma rectas en rectas, por lo que $\text{SO}(2)$ es un grupo de simetría de $u_{xx} = 0$. Otro grupo de simetría de la misma ecuación está dado por las traslaciones en \mathbb{R}^2 .

La idea es poder establecer un criterio para distinguir si un grupo de transformaciones es o no un grupo de simetría de un sistema de ecuaciones diferenciales. De hecho, una vez que hallamos establecido la herramienta geométrica apropiada, podremos usar los criterios de la sección 2.1, no sólo para saber si un grupo es un grupo de simetría de un sistema de ecuaciones diferenciales, sino también para calcular el grupo de simetría más general del sistema.

Comencemos por remplazar la noción de "sistema de ecuaciones diferenciales" por un objeto geométrico concreto, determinado por los puntos donde se anulan ciertas funciones. Para hacer esto, debemos "ampliar" el espacio básico $X \times U$, que representa a las variables independientes y dependientes involucradas, y considerar un espacio donde también estén representadas todas las derivadas parciales que aparecen en el sistema.

Dada una función de p variables independientes que toma valores reales $u = f(x) = f(x^1, \dots, x^p) \in \mathbb{R}$, hay

$$p_k \equiv \binom{p+k-1}{k}$$

diferentes derivadas parciales de orden k , que denotamos como

$$u_J = \partial_J f \Big|_x = \frac{\partial^k f}{\partial x^{j_1} \partial x^{j_2} \dots \partial x^{j_k}} \Big|_x,$$

donde $J = (j_1, \dots, j_k)$ es una colección de k números enteros cuyos valores están entre 1 y m e indican qué derivadas se están considerando; el orden de J ($|J| = k$) indica cuántas derivadas se están tomando en cuenta. En el caso más general, si $f: X \rightarrow U$ es una función suave de $X \simeq \mathbb{R}^p$ a $U \simeq \mathbb{R}^q$, tenemos qp_k números de la forma $u_J^l = \partial_J f^l \Big|_x$, que representan todas las derivadas parciales de orden k de las q componentes de f , evaluadas en el punto x . Así, podemos tomar $U^k \equiv \mathbb{R}^{qp_k}$ como el espacio euclidiano con coordenadas u_J^l , $l = 1, \dots, q$, $J = (j_1, \dots, j_k)$ en todos los multi-índices de orden k , para representar todas las derivadas de orden k de la función $u = f(x)$. Más aún, tomemos el espacio $U^{(n)} = U \times U^1 \times \dots \times U^n$, cuyas coordenadas representan todas las derivadas de u de todos los órdenes, desde 0 hasta n , evaluadas en el punto x . La dimensión de $U^{(n)}$ es

$$q + qp_1 + \dots + qp_n = q \binom{p+n}{n} \equiv qp^{(n)},$$

por lo que un punto en $U^{(n)}$ se denota como $u^{(n)}$ y tiene $qp^{(n)}$ componentes distintas $u_l^j, l = 1, \dots, q, J = (j_1, \dots, j_k)$ en todos los multi-índices con $1 \leq j_k \leq p$ y $0 \leq k \leq n$ (u_0^j se refiere a la componente u^j de u).

Finalmente, dada una función suave $u = f(x), f : X \rightarrow U$, hay una función inducida, denotada como $\text{pr}^{(n)}f$, de X a $U^{(n)}$, llamada la *enésima prolongación* de f :

$$\begin{aligned} \text{pr}^{(n)}f : X &\rightarrow U^{(n)} \\ x &\rightarrow u^{(n)}. \end{aligned}$$

El espacio $X \times U^{(n)}$, cuyas coordenadas representan las variables independientes, las variables dependientes y las derivadas de las variables dependientes hasta orden n se conoce como el *enésimo espacio jet* del espacio $X \times U$. Muchas veces no estamos interesados en ecuaciones diferenciales definidas en todo $X \times U$, sino sólo en un subconjunto abierto $M \subset X \times U$; en este caso, definimos el *enésimo espacio jet* de M como

$$M^{(n)} \equiv M \times U^1 \times \dots \times U^{(n)}.$$

Si $u = f(x)$ es una función cuya gráfica está en M , la *enésima prolongación* de f , $\text{pr}^{(n)}f(x)$, es una función cuya gráfica está en $M^{(n)}$.

Tomemos el caso en que $X \simeq \mathbb{R}^3$, con coordenadas (t, x, y) y $U \simeq \mathbb{R}$, con coordenada $u = f(t, x, y)$ ($p = 3, q = 1$). El espacio U^1 es isomorfo a \mathbb{R}^3 , con coordenadas $(u_t, u_x, u_y) = (\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$, que representan todas las derivadas de primer orden de f , evaluadas en (t, x, y) . Del mismo modo, $U^2 \simeq \mathbb{R}^6$ tiene coordenadas $(u_{tt}, u_{tx}, u_{ty}, u_{xx}, u_{xy}, u_{yy})$, que representan todas las derivadas de segundo orden de f , evaluadas en (t, x, y) . Por último, $U^{(2)} = \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^6 \simeq \mathbb{R}^{10}$ tiene coordenadas $u^{(2)} = (u; u_t, u_x, u_y; u_{tt}, u_{tx}, u_{ty}, u_{xx}, u_{xy}, u_{yy})$, que corresponden a todas las derivadas de f , hasta orden 2, evaluadas en (t, x, y) , de manera que la función inducida $\text{pr}^{(2)}f$, de $X \simeq \mathbb{R}^3$ a $U^{(n)} \simeq \mathbb{R}^{10}$, asigna a cada (t, x, y) el vector $(u; u_t, u_x, u_y; u_{tt}, u_{tx}, u_{ty}, u_{xx}, u_{xy}, u_{yy})$.

Un sistema \mathcal{S} de m ecuaciones diferenciales de orden n en p variables independientes y q variables dependientes, está dado como un sistema de ecuaciones

$$\Delta^r(x, u^{(n)}) = 0, \quad r = 1, \dots, m.$$

Vamos a suponer que las funciones $\Delta(x, u^{(n)}) = (\Delta^1(x, u^{(n)}), \dots, \Delta^m(x, u^{(n)}))$ son suaves en sus argumentos, es decir, que $\Delta : X \times U^{(n)} \rightarrow \mathbb{R}^m$ es una función suave. De este modo, las ecuaciones diferenciales del sistema \mathcal{S} indican los puntos en $X \times U^{(n)}$ donde el mapeo Δ se anula, por lo que determinan una subvariedad del enésimo espacio jet:

$$S_{\Delta} = \{(x, u^{(n)}) \mid \Delta(x, u^{(n)}) = 0\} \subset X \times U^{(n)}.$$

Al identificar el sistema de ecuaciones diferenciales con su subvariedad correspondiente, damos una reinterpretación de las relaciones "abstractas" entre las distintas derivadas

de u , determinadas por el sistema, como un subconjunto geométrico concreto S_Δ del enésimo espacio jet. Desde este punto de vista, una *solución* suave del sistema de ecuaciones diferenciales es una función suave $u = f(x)$ tal que

$$\Delta^r(x, u^{(n)}) = 0, \quad r = 1, \dots, m,$$

siempre que x se encuentre en el dominio de f . Esta condición es equivalente al hecho de que la gráfica de $\text{pr}^{(n)}f$ esté completamente incluida en la subvariedad S_Δ , determinada por el sistema:

$$\Gamma_f^{(n)} \equiv \{(x, \text{pr}^{(n)}f(x))\} \subset S_\Delta = \{(x, u^{(n)}) \mid \Delta(x, u^{(n)}) = 0\}.$$

De este modo, hemos conseguido reformular el problema inicial en términos geométricos adecuados para utilizar la herramienta de los grupos de Lie.

Tomemos un grupo local de transformaciones G , actuando en un subconjunto abierto del espacio de variables independientes y dependientes $M \subset X \times U$. Entonces, hay una acción local inducida en el enésimo espacio jet $M^{(n)}$, conocida como la *enésima prolongación de la acción de G en M* y denotada como $\text{pr}^{(n)}G$. Esta acción está definida de manera que transforma las derivadas de las funciones $u = f(x)$ en las correspondientes derivadas de las funciones transformadas $\tilde{u} = \tilde{f}(\tilde{x})$. Para ser más precisos, sea $(x_0, u_0^{(n)})$ un punto en $M^{(n)}$ y $u = f(x)$ definida en una vecindad de x_0 , cuya gráfica está contenida en M y cuyas derivadas evaluadas x_0 están dadas por

$$u_0^{(n)} = \text{pr}^{(n)}f(x_0).$$

Si $g \in G$ está suficientemente cerca de la identidad en G , la función transformada $g \cdot f$ dada por (2.2) está definida en una vecindad del punto $(\tilde{x}_0, \tilde{u}_0) = g \cdot (x_0, u_0)$, con $u_0 = f(x_0)$. Determinamos la enésima prolongación de la transformación dada por g $\text{pr}^{(n)}g$, en el punto $(x_0, u_0^{(n)})$, evaluando las derivadas de la función transformada $g \cdot f$ en \tilde{x}_0 :

$$\text{pr}^{(n)}g \cdot (x_0, u_0^{(n)}) = (\tilde{x}_0, \tilde{u}_0^{(n)}),$$

donde

$$\tilde{u}_0^{(n)} \equiv \text{pr}^{(n)}(g \cdot f)(\tilde{x}_0). \quad (2.3)$$

Retomando el ejemplo en el que el grupo de rotaciones $\text{SO}(2)$ actúa en $\mathbb{R}^2 \simeq X \times U$, calculemos la primera prolongación $\text{pr}^1\text{SO}(2)$. Dada una función $u = f(x)$, su primera prolongación es:

$$\text{pr}^1f(x) = (f(x), f'(x)).$$

Ahora bien, dado un punto $(x_0, u_0, u_{x0}) \in X \times U^{(1)}$ y una rotación por un ángulo $\theta \neq 0$, queremos encontrar $\text{pr}^1\theta \cdot (x_0, u_0, u_{x0}) = (\tilde{x}_0, \tilde{u}_0, \tilde{u}_{x0})$. Demos una recta en $X \times U$ que pase por (x_0, u_0) , con pendiente u_{x0} :

$$u = f(x) = u_{x0}(x - x_0) + u_0, \quad (f(x_0) = u_0, f'(x_0) = u_{x0})$$

De acuerdo a lo que habíamos encontrado antes,

$$\begin{aligned}\bar{u} &= \bar{f}(\bar{x}) = \theta \cdot f(\bar{x}) = \frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta} \bar{x} + \frac{u_0 - u_{x_0} x_0}{\cos \theta - u_{x_0} \sin \theta} \\ \Rightarrow \bar{f}'(\bar{x}) &= \frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta},\end{aligned}$$

por lo que

$$\text{pr}^{(1)}\theta \cdot f(\bar{x}) = \left(\frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta} \bar{x} + \frac{u_0 - u_{x_0} x_0}{\cos \theta - u_{x_0} \sin \theta}, \frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta} \right),$$

de donde

$$\begin{aligned}\text{pr}^{(1)}\theta \cdot (x_0, u_0, u_{x_0}) &= \left(x_0, \frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta} \bar{x} + \frac{u_0 - u_{x_0} x_0}{\cos \theta - u_{x_0} \sin \theta}, \frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta} \right) \\ &= \left(x_0 \cos \theta - u_0 \sin \theta, x_0 \sin \theta + u_0 \cos \theta, \frac{\sin \theta + u_{x_0} \cos \theta}{\cos \theta - u_{x_0} \sin \theta} \right)\end{aligned}$$

y para cualquier punto (x, u, u_x) ,

$$\text{pr}^{(1)}\theta \cdot (x, u, u_x) = \left(x \cos \theta - u \sin \theta, x \sin \theta + u \cos \theta, \frac{\sin \theta + u_x \cos \theta}{\cos \theta - u_x \sin \theta} \right).$$

Podemos observar en el ejemplo anterior que $\text{pr}^{(1)}G$ actúa en las variables (x, u) exactamente igual que G , sólo la acción en la derivada u_x nos da nueva información. Este hecho se cumple en general: dada $\text{pr}^{(n)}G$ que actúa en las variables $(x, u^{(n)})$, si restringimos nuestra atención a las derivadas hasta orden $k \leq n$, de modo que sólo consideremos las variables $(x, u^{(k)})$, entonces la acción de $\text{pr}^{(n)}G$ coincide con la prolongación anterior $\text{pr}^{(k)}G$; en particular, para $k = 0$, $\text{pr}^{(0)}G$ coincide con la acción de G en M . Este resultado puede establecerse de manera más precisa mediante la proyección natural

$$\begin{aligned}\pi_k^n : X \times U^{(n)} &\rightarrow X \times U^{(k)} \\ (x, u^{(n)}) &\rightarrow (x, u^{(k)}),\end{aligned}$$

donde $u^{(k)}$ consiste en las primeras componentes u_j^k , $|J| \leq k$, de $u^{(n)}$. Así, tenemos que

$$\pi_k^n \circ \text{pr}^{(n)}g = \text{pr}^{(k)}g, \quad k \leq n, \quad (2.4)$$

para cualquier elemento $g \in G$. Otro modo de ver esto es que, si ya conocemos la k -ésima prolongación de la acción $\text{pr}^{(k)}G$, entonces, para encontrar $\text{pr}^{(n)}G$, sólo necesitamos calcular la transformación de las derivadas u_j^k de ordenes $k < |J| \leq n$.

Dado un sistema de ecuaciones diferenciales de orden n o, equivalentemente, una subvariedad S_Δ del espacio jet $M^{(n)} \subset X \times U^{(n)}$, definimos a un grupo de simetría del sistema como un grupo local de transformaciones G que actúa en $M \subset X \times U$, transformando soluciones del sistema en otras soluciones. Lo que haremos a continuación es establecer una relación entre esta condición de simetría y la condición geométrica de que la subvariedad correspondiente S_Δ sea invariante ante la prolongación de la acción del grupo $\text{pr}^{(n)}G$. Esto nos permitirá reducir el problema de encontrar el grupo de simetría de un sistema de ecuaciones diferenciales al problema de determinar cuándo una subvariedad es invariante bajo la acción de un grupo local de transformaciones, de manera que podremos hacer uso de la herramienta desarrollada en la sección anterior, pues estamos considerando a los sistemas de ecuaciones diferenciales como sistemas de ecuaciones algebraicas en el n -ésimo espacio jet.

Teorema 2.11

Sea $M \subset X \times U$ un subconjunto abierto y $\Delta(x, u^{(n)}) = 0$ un sistema de ecuaciones diferenciales de orden n definidas sobre M , junto con su subvariedad correspondiente $S_\Delta \subset M^{(n)}$. Sea G un grupo local de transformaciones tal que la prolongación de la acción de G en M deja a S_Δ invariante, es decir, que $\forall (x, u^{(n)}) \in S_\Delta$ se cumple que $\text{pr}^{(n)}g \cdot (x, u^{(n)}) \in S_\Delta \forall g \in G$ donde la acción está definida. Entonces G es un grupo de simetría del sistema de ecuaciones diferenciales.

Prueba.

Sea $u = f(x)$ una solución local de $\Delta(x, u^{(n)}) = 0$, es decir, que

$$\Gamma_f^{(n)} = \{(x, \text{pr}^{(n)}f(x))\} \subset S_\Delta.$$

Si $g \in G$ es tal que $g \cdot f$ está bien definida, entonces la gráfica de su prolongación $\Gamma_{g \cdot f}^{(n)}$ es igual a la transformación de la gráfica $\Gamma_f^{(n)}$ por $\text{pr}^{(n)}g$:

$$\Gamma_{g \cdot f}^{(n)} = \text{pr}^{(n)}g(\Gamma_f^{(n)}).$$

Como S_Δ es invariante bajo $\text{pr}^{(n)}g$, entonces $\text{pr}^{(n)}g(\Gamma_f^{(n)}) \subset S_\Delta$, de donde $\Gamma_{g \cdot f}^{(n)} \subset S_\Delta$ y, por lo tanto, $\bar{u} = g \cdot f$ es una solución del sistema Δ .

□

Al igual que con los grupos de transformaciones, también podemos definir la prolongación de los correspondientes generadores infinitesimales de la acción:

Definición 2.12 Sea $M \subset X \times U$ abierto y v un campo vectorial en M , con su correspondiente grupo local monoparamétrico $\exp(tv)$. La n -ésima prolongación de v , denotada como $\text{pr}^{(n)}v$, es un campo vectorial en el n -ésimo espacio jet $M^{(n)}$ definido

como el generador infinitesimal de la correspondiente prolongación de la acción del grupo monoparamétrico $\text{pr}^{(n)} \exp(tv)$, es decir:

$$\text{pr}^{(n)}v \Big|_{(x, u^{(n)})} = \frac{d}{dt} \text{pr}^{(n)} \exp(tv)(x, u^{(n)}) \Big|_{t=0},$$

para todo $(x, u^{(n)}) \in M^{(n)}$.

Notemos que, dado que las coordenadas $(x, u^{(n)})$ en $M^{(n)}$ representan a todas las variables independientes (x^1, \dots, x^m) , las variables dependientes (u^1, \dots, u^p) y todas las derivadas u_j^i de las variables dependientes hasta orden n , un campo vectorial en $M^{(n)}$ tiene la forma

$$v^* = \sum_{i=1}^p \xi^i \frac{\partial}{\partial x^i} + \sum_{i=1}^q \sum_J \phi_i^J \frac{\partial}{\partial u_i^J},$$

donde J corre en todos los multi-índices con orden $0 \leq |J| \leq n$ y las funciones coeficientes ξ^i , ϕ_i^J pueden depender de todas las variables $(x, u^{(n)})$. En el caso en que v^* es la prolongación $\text{pr}^{(n)}v$ de un campo vectorial

$$v = \sum_{i=1}^p \xi^i(x, u) \frac{\partial}{\partial x^i} + \sum_{i=1}^q \phi_i(x, u) \frac{\partial}{\partial u^i},$$

los coeficientes ξ^i , ϕ_i^J de $v^* = \text{pr}^{(n)}v$ quedan determinados por los coeficientes de v : de acuerdo con (2.4), la prolongación de la acción $\text{pr}^{(n)} \exp(tv)$ restringida a las variables x, u de $M^{(0)} = M$, debe coincidir con la acción del grupo $\exp(tv)$ en M , así que los coeficientes ξ^i , $\phi_i^0 = \phi_i$ de $v^* = \text{pr}^{(n)}v$ deben coincidir con los coeficientes ξ^i , ϕ_i de v , entonces

$$\text{pr}^{(n)}v = \sum_{i=1}^p \xi^i \frac{\partial}{\partial x^i} + \sum_{i=1}^q \sum_J \phi_i^J \frac{\partial}{\partial u_i^J},$$

donde ξ^i , $\phi_i^0 = \phi_i$ provienen de v . De hecho, si $|J| = k$, los coeficientes ϕ_i^J de $\frac{\partial}{\partial u_i^J}$ sólo dependen de las derivadas de u de orden menor o igual a k , $\phi_i^J = \phi_i^J(x, u^{(k)})$ ya que, por (2.4), la acción del grupo correspondiente sólo involucra derivadas de orden menor o igual a k . Esta observación nos permite construir las distintas prolongaciones de un campo vectorial de manera recursiva. Lo que necesitamos ahora es encontrar una fórmula general de los coeficientes ϕ_i^J .

Para el ejemplo anterior del grupo de rotaciones $\text{SO}(2)$ que actúa en $\mathbb{R}^2 \simeq X \times U$, sabemos que el generador de la acción es

$$v = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u},$$

y el flujo está dado por

$$\exp(\theta v)(x, u) = (x \cos \theta - u \sin \theta, x \sin \theta + u \cos \theta).$$

Según lo que habíamos encontrado, la primera prolongación de la acción de $\exp(\theta v)$ es

$$\text{pr}^{(1)} \exp(\theta v)(x, u, u_x) = \left(x \cos \theta - u \sin \theta, x \sin \theta + u \cos \theta, \frac{\sin \theta + u_x \cos \theta}{\cos \theta - u_x \sin \theta} \right).$$

Para obtener la primera prolongación de v sólo necesitamos derivar la última componente de $\text{pr}^{(1)} \exp(\theta v)(x, u, u_x)$ respecto a θ y evaluarla en $\theta = 0$:

$$\frac{\partial}{\partial \theta} \left(\frac{\sin \theta + u_x \cos \theta}{\cos \theta - u_x \sin \theta} \right) \Big|_{\theta=0} = \frac{1 + u_x^2}{(\cos \theta - u_x \sin \theta)^2} \Big|_{\theta=0} = 1 + u_x^2,$$

por lo que

$$\text{pr}^{(1)} v = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u} + (1 + u_x^2) \frac{\partial}{\partial u_x}.$$

Para poder aplicar los criterios de la sección 2.1, necesitamos establecer una condición equivalente a la condición de rango máximo de sistemas de ecuaciones algebraicas.

Definición 2.13 Sea

$$\Delta^r(x, u^{(n)}) = 0, \quad r = 1, \dots, m,$$

un sistema de ecuaciones diferenciales. Se dice que el sistema es de *rango máximo* si la matriz jacobiana, de $p + qp^{(n)} \times m$, de Δ respecto a las variables $(x, u^{(n)})$:

$$J_{\Delta}(x, u^{(n)}) = \left(\frac{\partial \Delta^r}{\partial x^i}, \frac{\partial \Delta^r}{\partial u_j^{(n)}} \right),$$

es de rango m siempre que $\Delta(x, u^{(n)}) = 0$.

Teorema 2.14

Sea

$$\Delta^r(x, u^{(n)}) = 0, \quad r = 1, \dots, m,$$

un sistema de ecuaciones diferenciales de rango máximo en $M \subset X \times U$. Si G es un grupo local de transformaciones que actúa en M y

$$\text{pr}^{(n)} v(\Delta^r(x, u^{(n)})) = 0, \quad r = 1, \dots, m, \quad \text{siempre que } \Delta^r(x, u^{(n)}) = 0, \quad (2.5)$$

para todo v generador infinitesimal de G , entonces G es un grupo de simetría del sistema.

Prueba.

Como el sistema es de rango máximo y

$$\text{pr}^{(n)} v(\Delta^r(x, u^{(n)})) = 0,$$

entonces, por el teorema 2.5, la prolongación de la acción de G deja invariante a S_{Δ} y, por el teorema 2.11, G es un grupo de simetría del sistema de ecuaciones diferenciales.

□

Considerando que se cumple la condición de rango máximo, tomando en cuenta la proposición 2.6, podemos remplazar (2.5) por la condición equivalente de que existan funciones suaves $Q^s(x, u^{(n)})$, $s = 1, \dots, m$, tales que

$$\text{pr}^{(n)}u(\Delta^r(x, u^{(n)})) = Q^1(x, u^{(n)})\Delta^1(x, u^{(n)}) + \dots + Q^m(x, u^{(n)})\Delta^m(x, u^{(n)}) \quad \forall (x, u^{(n)}) \in M^{(n)}. \quad (2.6)$$

Una vez que hemos establecido una relación entre un grupo de simetría de un sistema de ecuaciones diferenciales y un criterio de invarianza del sistema bajo la prolongación de los generadores del grupo, falta encontrar una fórmula explícita para la prolongación de un campo vectorial.

Definición 2.15 Sea $P(x, u^{(n)})$ una función suave de x, u y las derivadas de u hasta orden n , en un subconjunto abierto $M^{(n)} \subset X \times U^{(n)}$. Para $u = f(x)$ suave, definimos la *derivada total* de P con respecto a x^i (o la *i -ésima derivada total* de P), como la función suave $D_i P(x, u^{(n)})$, dada por

$$D_i P(x, \text{pr}^{(n+1)}f(x)) = \frac{\partial}{\partial x^i} P(x, u^{(n)}).$$

Es decir, $D_i P$ se obtiene a partir de P al derivar P con respecto a x^i , tomando todas las funciones u^l y sus derivadas como funciones de x .

Aplicando la regla de la cadena, nos damos cuenta de que la i -ésima derivada total de P tiene una forma general

$$D_i P(x, \text{pr}^{(n+1)}f(x)) = \frac{\partial}{\partial x^i} P(x, u^{(n)}) = \frac{\partial P}{\partial x^i} + \sum_{l=1}^q \sum_J u_{J,i}^l \frac{\partial P}{\partial u_{J,i}^l}, \quad (2.7)$$

donde, para $J = (j_1, \dots, j_k)$,

$$u_{J,i}^l = \frac{\partial u_{J,i}^l}{\partial x^i} = \frac{\partial^{k+1} u^l}{\partial x^i \partial^{j_1} \dots \partial^{j_k}}. \quad (2.8)$$

La suma en (2.7) es sobre todos los J con orden $0 \leq |J| \leq n$.

Para las derivadas totales de órdenes mayores, si $J = (j_1, \dots, j_k)$ con $1 \leq j_k \leq p$ para cada k , entonces la derivada J -ésima es:

$$D_J = D_{j_1} D_{j_2} \dots D_{j_k}.$$

Teorema 2.16

Sea

$$v = \sum_{i=1}^p \xi^i(x, u) \frac{\partial}{\partial x^i} + \sum_{i=1}^q \phi_i(x, u) \frac{\partial}{\partial u^i},$$

un campo vectorial definido en $M \subset X \times U$ cuya n -ésima prolongación es el campo vectorial dado por

$$\text{pr}^{(n)}v = v + \sum_{i=1}^q \sum_{j=1}^n \phi_i^j(x, u^{(n)}) \frac{\partial}{\partial u_{j,i}^i},$$

definido en el correspondiente espacio $\text{jet } M^{(n)}$, con $J = (j_1, \dots, j_k), 1 \leq j_k \leq p$ y $1 \leq k \leq n$. Entonces, las funciones coeficientes ϕ_i^j de $\text{pr}^{(n)}$ están dadas por la siguiente fórmula:

$$\phi_i^j(x, u^{(n)}) = D_J \left(\phi_i - \sum_{i=1}^p \xi^i u_i^j \right) + \sum_{i=1}^p \xi^i u_{j,i}^i, \quad (2.9)$$

donde $u_i^j = \frac{\partial u^i}{\partial x^j}$ y $u_{j,i}^i$ está dada como en (2.8).

La demostración de este teorema se hace por inducción en el orden del operador de prolongación, n , y puede consultarse en Olver [1]. Para $n = 1$, la idea es tomar $g_t = \exp(tv)$ como el correspondiente grupo monoparamétrico de v , cuya transformación está dada por

$$(\bar{x}, \bar{u}) = g_t \cdot (x, u) = (\Psi_t(x, u), \Phi_t(x, u)).$$

Notemos que

$$\begin{aligned} \xi^i(x, u) &= \left. \frac{d}{dt} \Psi_t^i(x, u) \right|_{t=0}, \quad i = 1, \dots, p, \\ \phi_i(x, u) &= \left. \frac{d}{dt} \Phi_t^i(x, u) \right|_{t=0}, \quad i = 1, \dots, q, \end{aligned}$$

donde Ψ_t^i y Φ_t^i son las funciones componentes de Ψ_t y Φ_t . Sea $u = f(x)$, de manera que $u^{(1)} = \text{pr}^{(1)}f(x)$:

$$u^i = f^i(x), \quad u_i^j = \left. \frac{\partial f^i}{\partial x^j} \right|_x.$$

De acuerdo con (2.1), para t suficientemente pequeña, se tiene

$$\bar{u} = g_t \cdot f(\bar{x}) = (\Phi_t \circ (id \times f)) \circ (\Psi_t \circ (id \times f))^{-1}(\bar{x}).$$

Tomando las derivadas esta expresión con respecto a x^i , $i = 1, \dots, p$, obtenemos de forma explícita a $\text{pr}^{(1)}g \cdot (x, u^{(1)})$. Una vez hecho esto, considerando la definición 2.12,

al derivar las componentes de $\text{pr}^{(1)}g \cdot (x, u^{(1)})$ respecto a t y evaluar en $t = 0$, obtenemos las funciones coeficientes ϕ^t de $\frac{\partial}{\partial u_k}$ en $\text{pr}^{(1)}v$:

$$\begin{aligned}\phi^t(x, \text{pr}^{(1)}f(x)) &= \frac{\partial}{\partial x^k} \phi_i(x, f(x)) - \sum_{i=1}^p \frac{\partial f^i}{\partial x^t} \frac{\partial}{\partial x^k} \xi^i(x, f(x)) \\ &= D_k \left(\phi_i - \sum_{i=1}^p \xi^i u_i^t \right) + \sum_{i=1}^p \xi^i u_{k,i}^t,\end{aligned}$$

donde $u_{k,i}^t = \frac{\partial^2 u^i}{\partial x^k \partial x^t}$. Una vez que se ha mostrado esto, hay que realizar la inducción sobre n , haciendo algunas consideraciones respecto a los $(n+1)$ espacios jet , $M^{(n+1)}$, como subespacios del primer espacio jet del n -ésimo espacio jet , $(M^{(n)})^{(1)}$.

Consideremos el caso especial $p = 3$, $q = 1$ en la fórmula de prolongación, de modo que estamos tratando con una ecuación diferencial parcial que involucra a la función $u = f(t, x, y)$. Un campo vectorial que actúa en $X \times U \simeq \mathbb{R}^3 \times \mathbb{R}$, tiene la forma:

$$v = \tau(t, x, y, u) \frac{\partial}{\partial t} + \xi(t, x, y, u) \frac{\partial}{\partial x} + \eta(t, x, y, u) \frac{\partial}{\partial y} + \phi(t, x, y, u) \frac{\partial}{\partial u}. \quad (2.10)$$

La primera prolongación de v es el campo vectorial

$$\text{pr}^{(1)}v = v + \phi^t \frac{\partial}{\partial u_t} + \phi^x \frac{\partial}{\partial u_x} + \phi^y \frac{\partial}{\partial u_y}.$$

Para calcular las funciones ϕ^t , ϕ^x , ϕ^y usamos la fórmula (2.9), en particular:

$$\begin{aligned}\phi^t &= D_t \left(\phi - \tau \frac{\partial u}{\partial t} - \xi \frac{\partial u}{\partial x} - \eta \frac{\partial u}{\partial y} \right) + \tau \frac{\partial u_t}{\partial t} + \xi \frac{\partial u_t}{\partial x} + \eta \frac{\partial u_t}{\partial y} \\ \Rightarrow \phi^t &= D_t \left(\phi - \tau u_t - \xi u_x - \eta u_y \right) + \tau u_{tt} + \xi u_{tx} + \eta u_{ty}.\end{aligned} \quad (2.11)$$

Por (2.7), para $P(t, x, y, u^{(1)}) = \phi - \tau u_t - \xi u_x - \eta u_y$, tenemos:

$$\begin{aligned}D_t P &= \frac{\partial}{\partial t} \left(\phi - \tau u_t - \xi u_x - \eta u_y \right) + u_t \frac{\partial}{\partial u} \left(\phi - \tau u_t - \xi u_x - \eta u_y \right) \\ &+ \frac{\partial u_t}{\partial t} \frac{\partial}{\partial u_t} \left(\phi - \tau u_t - \xi u_x - \eta u_y \right) + \frac{\partial u_x}{\partial t} \frac{\partial}{\partial u_x} \left(\phi - \tau u_t - \xi u_x - \eta u_y \right) \\ &+ \frac{\partial u_y}{\partial t} \frac{\partial}{\partial u_y} \left(\phi - \tau u_t - \xi u_x - \eta u_y \right) \\ \Rightarrow D_t P &= \phi_t - \tau_t u_t - \xi_t u_x - \eta_t u_y + u_t (\phi_u - \tau_u u_t - \xi_u u_x - \eta_u u_y) \\ &- \tau u_{tt} - \xi u_{tx} - \eta u_{ty}.\end{aligned} \quad (2.12)$$

Sustituyendo (2.12) en (2.11) tenemos:

$$\phi^t = \phi_t + (\phi_u - \tau_t)u_t - \xi_t u_x - \eta_t u_y - \tau_u u_t^2 - \xi_u u_t u_x - \eta_u u_t u_y.$$

Las funciones ϕ^x y ϕ^y se calculan de igual manera, por lo tanto, las funciones coeficientes de $\text{pr}^{(1)}v$ están dadas por los coeficientes de v y por

$$\begin{aligned}\phi^t &= \phi_t + (\phi_u - \tau_t)u_t - \xi_t u_x - \eta_t u_y - \tau_u u_t^2 - \xi_u u_t u_x - \eta_u u_t u_y, \\ \phi^x &= \phi_x - \tau_x u_t + (\phi_u - \xi_x)u_x - \eta_x u_y - \tau_u u_t u_x - \xi_u u_x^2 - \eta_u u_x u_y, \\ \phi^y &= \phi_y - \tau_y u_t - \xi_y u_x + (\phi_u - \eta_y)u_y - \tau_u u_t u_y - \xi_u u_x u_y - \eta_u u_y^2.\end{aligned}$$

Similarmente, la segunda prolongación de v es:

$$\text{pr}^{(2)}v = \text{pr}^{(1)}v + \phi^{tt} \frac{\partial}{\partial u_{tt}} + \phi^{tx} \frac{\partial}{\partial u_{tx}} + \phi^{ty} \frac{\partial}{\partial u_{ty}} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xy} \frac{\partial}{\partial u_{xy}} + \phi^{yy} \frac{\partial}{\partial u_{yy}}.$$

Si buscamos ahora ϕ^{tt} , se tiene:

$$\begin{aligned}\phi^{tt} &= D_t^2 \left(\phi - \tau \frac{\partial u}{\partial t} - \xi \frac{\partial u}{\partial x} - \eta \frac{\partial u}{\partial y} \right) + \tau \frac{\partial u_{tt}}{\partial t} + \xi \frac{\partial u_{tt}}{\partial x} + \eta \frac{\partial u_{tt}}{\partial y} \\ &= D_t \left(\phi_t - \tau_t u_t - \xi_t u_x - \eta_t u_y + u_t (\phi_u - \tau_u u_t - \xi_u u_x - \eta_u u_y) \right. \\ &\quad \left. - \tau_{tu} - \xi_{tu} - \eta_{tu} \right) + \tau_{tuu} + \xi_{tuu} + \eta_{tuu}.\end{aligned}$$

Finalmente, al calcular la derivada total respecto a t , obtenemos:

$$\begin{aligned}\phi^{tt} &= \phi_{tt} + (2\phi_{tu} - \tau_{tt})u_t - \xi_{tt}u_x - \eta_{tt}u_y + (\phi_{uu} - 2\tau_{tu})u_t^2 - 2\xi_{tu}u_t u_x - 2\eta_{tu}u_t u_y \\ &\quad - \tau_{uu}u_t^3 - \xi_{uu}u_t^2 u_x - \eta_{uu}u_t^2 u_y + (\phi_u - 2\tau_t)u_{tt} - 2\xi_{tu}u_{tx} - 2\eta_{tu}u_{ty} \\ &\quad - 3\tau_{tu}u_t u_{tt} - \xi_{tu}u_x u_{tt} - \eta_{tu}u_y u_{tt} - 2\xi_{tu}u_t u_{tx} - 2\eta_{tu}u_t u_{ty} \\ &\quad + \xi_{tuu} + \eta_{tuu} + \tau_{tuu}.\end{aligned}\tag{2.13}$$

Para ϕ^{xx} y ϕ^{yy} se obtienen expresiones similares:

$$\begin{aligned}\phi^{xx} &= \phi_{xx} - \tau_{xx}u_t + (2\phi_{xu} - \xi_{xx})u_x - \eta_{xx}u_y - 2\tau_{xu}u_t u_x + (\phi_{uu} - 2\xi_{xu})u_x^2 - 2\eta_{xu}u_x u_y \\ &\quad - \tau_{uu}u_t u_x^2 - \xi_{uu}u_x^3 - \eta_{uu}u_x^2 u_y - 2\tau_{xu}u_{tx} + (\phi_u - 2\xi_x)u_{xx} - 2\eta_{xu}u_{xy} \\ &\quad - \tau_{tu}u_t u_{xx} - 3\xi_{tu}u_x u_{xx} - \eta_{tu}u_y u_{xx} - 2\tau_{tu}u_x u_{tx} - 2\eta_{tu}u_x u_{xy} \\ &\quad + \tau_{tuu} + \xi_{tuu} + \eta_{tuu}.\end{aligned}\tag{2.14}$$

y

$$\begin{aligned}\phi^{yy} &= \phi_{yy} - \tau_{yy}u_t - \xi_{yy}u_x + (2\phi_{yu} - \eta_{yy})u_y - 2\tau_{yu}u_t u_y - 2\xi_{yu}u_x u_y + (\phi_{uu} - 2\eta_{yu})u_y^2 \\ &\quad - \tau_{uu}u_t u_y^2 - \xi_{uu}u_x u_y^2 - \eta_{uu}u_y^3 - 2\tau_{yu}u_{ty} - 2\xi_{yu}u_{xy} + (\phi_u - 2\eta_y)u_{yy} \\ &\quad - \tau_{tu}u_t u_{yy} - \xi_{tu}u_x u_{yy} - 3\eta_{tu}u_y u_{yy} - 2\tau_{tu}u_y u_{ty} - 2\xi_{tu}u_x u_{ty} \\ &\quad + \tau_{tuu} + \xi_{tuu} + \eta_{tuu}.\end{aligned}\tag{2.15}$$

Usando el teorema 2.16, es posible demostrar que, para dos campos vectoriales suaves v y w , se cumplen las siguientes propiedades¹:

$$\text{pr}^{(n)}(av + bw) = a\text{pr}^{(n)}v + b\text{pr}^{(n)}w$$

donde a y b son constantes, y

$$\text{pr}^{(n)}[v, w] = [\text{pr}^{(n)}v, \text{pr}^{(n)}w].$$

Con esto se tiene que, para un sistema de ecuaciones de rango máximo definido en $M \subset X \times U$, el conjunto de todos los generadores infinitesimales de las simetrías del sistema forma un álgebra de Lie de campos vectoriales en M .

2.3 La Ecuación de Onda.

En la sección anterior hemos construido un método para encontrar el grupo de simetrías conexo más general de un sistema de ecuaciones diferenciales parciales. En este método consideramos a un generador infinitesimal v de un supuesto grupo monoparamétrico de simetría del sistema, cuyos coeficientes ξ^i y ϕ_i están por determinarse. Los coeficientes ϕ_i^j de la prolongación del generador infinitesimal, $\text{pr}^{(n)}v$, son expresiones que involucran explícitamente a las derivadas parciales de los coeficientes ξ^i y ϕ^i , con respecto a x y u . Usando los criterios de invarianza (2.5) y (2.6), que involucran a x , u y a todas las derivadas de u con respecto a x , así como a ξ^i y ϕ_i y sus derivadas parciales respecto a x y a u , podemos eliminar todas las dependencias en las derivadas de u que involucran al sistema mismo e igualar varios de los coeficientes restantes a cero, lo que nos permite establecer ecuaciones diferenciales parciales para las funciones ξ^i y ϕ_i llamadas *ecuaciones determinantes* para el grupo de simetría del sistema. Finalmente, hay que resolver dichas ecuaciones, pues su solución general determina el grupo de simetría conexo más general del sistema.

En concreto, tomemos la ecuación de onda en dos dimensiones espaciales:

$$-u_{tt} + u_{xx} + u_{yy} = 0.$$

Aquí tenemos tres variables independientes, t , x y y , y una variable dependiente, u , así que $p = 3$ y $q = 1$. La ecuación de onda es de segundo orden, $n = 2$, y puede identificarse como una subvariedad lineal en el segundo espacio jet $X \times U^{(2)}$, determinada por los puntos donde se anula $\Delta(t, x, y, u^{(2)}) = u_{tt} - u_{xx} - u_{yy}$.

¹La demostración puede consultarse en la sección 2.3 de Olver [7]

Un campo vectorial típico en el espacio de las variables independientes y dependientes tiene la forma dada por (2.10):

$$v = \tau \frac{\partial}{\partial t} + \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y} + \phi \frac{\partial}{\partial u},$$

donde τ , ξ , η y ϕ dependen de t , x , y y u . Lo que queremos hacer es determinar estas funciones coeficientes, de modo que el correspondiente grupo monoparamétrico sea un grupo de simetría de la ecuación de onda.

En este caso vamos a emplear el criterio (2.6), según el cual necesitamos tomar la segunda prolongación de v , aplicarla a la ecuación de onda e igualar esto al producto de una función suave Q por la ecuación de onda:

$$\text{pr}^{(2)}v(-u_{tt} + u_{xx} + u_{yy}) = Q(-u_{tt} + u_{xx} + u_{yy})$$

Donde $Q(t, x, y, u^{(2)})$ depende de las variables independientes, de u y de sus derivadas hasta segundo orden. Así, tenemos que

$$\begin{aligned} (\text{pr}^{(1)}v + \phi^{tt} \frac{\partial}{\partial u_{tt}} + \phi^{tx} \frac{\partial}{\partial u_{tx}} + \phi^{ty} \frac{\partial}{\partial u_{ty}} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xy} \frac{\partial}{\partial u_{xy}} + \phi^{yy} \frac{\partial}{\partial u_{yy}})(-u_{tt} + u_{xx} + u_{yy}) \\ = -Q_{u_{tt}} + Q_{u_{xx}} + Q_{u_{yy}} \\ \Rightarrow -\phi^{tt} + \phi^{xx} + \phi^{yy} = -Q_{u_{tt}} + Q_{u_{xx}} + Q_{u_{yy}} \end{aligned} \quad (2.16)$$

Sustituyendo (2.13), (2.14) y (2.15) en (2.16) y tomando además la igualdad dada por la ecuación de onda misma, obtenemos lo siguiente:

$$\begin{aligned} -\phi_{tt} + \phi_{xx} + \phi_{yy} + (\tau_{tt} - \tau_{xx} - \tau_{yy} - 2\phi_{tu})u_t + (\xi_{tt} - \xi_{xx} - \xi_{yy} + 2\phi_{xu})u_x + (\eta_{tt} - \eta_{xx} - \eta_{yy} + 2\phi_{yu})u_y + (2\tau_{tu} - \phi_{uu})u_t^2 + 2(-\tau_{xu} + \xi_{tu})u_t u_x + 2(-\tau_{yu} + \eta_{tu})u_t u_y + (-2\xi_{xu} + \phi_{uu})u_x^2 - 2(\xi_{yu} + \eta_{xu})u_x u_y + (-2\eta_{yu} + \phi_{uu})u_y^2 + \tau_{uu}u_t^3 + \xi_{uu}u_t^2 u_x + \eta_{uu}u_t^2 u_y - \tau_{uu}u_t u_x^2 - \xi_{uu}u_x^3 - \eta_{uu}u_x^2 u_y - \tau_{uu}u_t u_y^2 - \xi_{uu}u_x u_y^2 - \eta_{uu}u_y^3 + 2\tau_{tu}u_t - 2\xi_{xu}u_{xx} - 2\eta_{yu}u_{yy} + 2(-\tau_x + \xi_t)u_{tx} + 2(-\tau_y + \eta_t)u_{ty} - 2(\xi_y + \eta_x)u_{xy} + 2\tau_{tu}u_t u_{tt} - 2\xi_{xu}u_x u_{xx} - 2\eta_{yu}u_y u_{yy} - 2\tau_{tu}u_x u_{tx} + 2\xi_{xu}u_t u_{tx} - 2\tau_{tu}u_y u_{ty} + 2\eta_{yu}u_t u_{ty} - 2\xi_{xu}u_y u_{xy} - 2\eta_{yu}u_x u_{xy} = -Q_{u_{tt}} + Q_{u_{xx}} + Q_{u_{yy}}. \end{aligned}$$

Para resolver (2.16), observemos que, del lado izquierdo de la ecuación, sólo hay seis términos que involucran a u_{tt} , u_{xx} y u_{yy} , por lo que los coeficientes que multiplican al resto de las derivadas de u deben ser igualados a cero. En particular, anulemos primero los coeficientes de las derivadas de primer orden multiplicadas por las derivadas cruzadas de segundo orden, es decir, los coeficientes de $u_x u_{tx}$, $u_t u_{tx}$, $u_y u_{ty}$, $u_t u_{ty}$, $u_y u_{xy}$ y $u_x u_{xy}$:

$$\tau_u = 0, \quad \xi_u = 0, \quad \eta_u = 0,$$

lo que indica que τ , ξ y η no dependen de u , de donde

$$-\phi_{tt} + \phi_{xx} + \phi_{yy} + (\tau_{tt} - \tau_{xx} - \tau_{yy} - 2\phi_{tu})u_t + (\xi_{tt} - \xi_{xx} - \xi_{yy} + 2\phi_{xu})u_x + (\eta_{tt} - \eta_{xx} - \eta_{yy} + 2\phi_{yu})u_y - \phi_{uu}u_t^2 + \phi_{uu}u_x^2 + \phi_{uu}u_y^2 + 2\tau_{tu}u_t - 2\xi_{tx}u_{tx} - 2\eta_{ty}u_{ty} + 2(-\tau_x + \xi_t)u_{tx} + 2(-\tau_y + \eta_t)u_{ty} - 2(\xi_y + \eta_x)u_{xy} = -Q u_{tt} + Q u_{xx} + Q u_{yy}.$$

De los términos donde aparecen u_{tt} , u_{xx} y u_{yy} , tenemos las siguientes igualdades:

$$\begin{aligned} 2\tau_t &= -Q, & -2\xi_x &= Q, & -2\eta_y &= Q \\ \Rightarrow \tau_t &= \xi_x = \eta_y \end{aligned} \quad (2.17)$$

y, de los coeficientes de las derivadas cruzadas, u_{tx} , u_{ty} y u_{xy} , vemos que:

$$-\tau_x + \xi_t = 0, \quad -\tau_y + \eta_t = 0, \quad \xi_y + \eta_x = 0. \quad (2.18)$$

Desarrollemos a τ , ξ y a η en sus series de Taylor:

$$\begin{aligned} \tau(t, x, y) &= a_{000} + a_{100}t + a_{010}x + a_{001}y + a_{110}tx + a_{101}ty + a_{011}xy + a_{111}txy + \dots, \\ \xi(t, x, y) &= b_{000} + b_{100}t + b_{010}x + b_{001}y + b_{110}tx + b_{101}ty + b_{011}xy + b_{111}txy + \dots, \\ \eta(t, x, y) &= c_{000} + c_{100}t + c_{010}x + c_{001}y + c_{110}tx + c_{101}ty + c_{011}xy + c_{111}txy + \dots \end{aligned} \quad (2.19)$$

Derivando (2.17) dos veces respecto a t , dos veces respecto a x y dos veces respecto a y , obtenemos:

$$\begin{aligned} \tau_{tt} &= \xi_{tx} = \eta_{ty}, \\ \tau_{txx} &= \xi_{xxx} = \eta_{xxy}, \\ \tau_{tyy} &= \xi_{xyy} = \eta_{yyy}, \end{aligned} \quad (2.20)$$

y, derivando la primera ecuación de (2.18) respecto a t y a x , la segunda respecto a t y a y y la tercera respecto a x y a y , se tiene:

$$\tau_{txx} = \xi_{ttx}, \quad \tau_{tyy} = \eta_{tty}, \quad \xi_{xyy} = -\eta_{xxy}. \quad (2.21)$$

De (2.20) y (2.21) podemos ver que:

$$\tau_{ttt} = \xi_{ttt} = \tau_{txx} = \eta_{xxy} = -\xi_{xyy} = -\tau_{tyy} = -\eta_{tty} = -\tau_{ttt}.$$

Por lo tanto, todas estas derivadas de tercer orden se anulan y de igual manera se demuestra que todas las derivadas de tercer orden de τ , ξ y η son iguales a cero. Esto quiere decir que los polinomios en (2.19) son, a lo más, de segundo grado. Si usamos las distintas igualdades en (2.17) y (2.18), así como sus derivadas respecto a las variables t , x y y , llegamos a que τ , ξ y η son de la siguiente forma:

$$\begin{aligned} \tau(t, x, y) &= c_1 + c_4t + c_6x + c_6y - 2c_8tx - 2c_9ty + c_{10}(-t^2 - x^2 - y^2), \\ \xi(t, x, y) &= c_2 + c_5t + c_4x - c_7y + c_8(-t^2 - x^2 + y^2) - 2c_9xy - 2c_{10}tx, \end{aligned}$$

$$\eta(t, x, y) = c_3 + c_6 t + c_7 x + c_4 y - 2c_8 xy + c_9(-t^2 + x^2 - y^2) - 2c_{10} ty. \quad (2.22)$$

Observando el coeficiente de u_t^2 (o de u_x^2 o de u_y^2), tenemos que $\phi_{uu} = 0$, por lo que, como función de u , ϕ debe ser lineal:

$$\phi(t, x, y, u) = \beta(t, x, y)u + \alpha(t, x, y). \quad (2.23)$$

Derivando (2.23) respecto a u y considerando los coeficientes de los términos de primer orden u_t , u_x y u_y obtenemos:

$$\begin{aligned} 2\beta_t &= -\tau_{tt} + \tau_{xx} + \tau_{yy}, \\ 2\beta_x &= \xi_{tt} - \xi_{xx} - \xi_{yy}, \\ 2\beta_y &= \eta_{tt} - \eta_{xx} - \eta_{yy}. \end{aligned} \quad (2.24)$$

Así que, tomando las correspondientes derivadas en (2.22), vemos que:

$$\beta = c_{11} - c_{10}t - c_8 x - c_9 y \quad (2.25)$$

Finalmente, usando los términos independientes, ϕ_{tt} , ϕ_{xx} y ϕ_{yy} , tenemos que:

$$-\beta_{tt} - \alpha_{tt} + \beta_{xx} + \alpha_{xx} + \beta_{yy} + \alpha_{yy} = 0.$$

Como sabemos que las derivadas de tercer orden de τ , ξ y η son nulas, por (2.24), llegamos a:

$$-\alpha_{tt} + \alpha_{xx} + \alpha_{yy} = 0. \quad (2.26)$$

Por lo que α es una solución cualquiera de la ecuación de onda.

De este modo, hemos determinado la solución más general al sistema de ecuaciones determinantes, por lo tanto, el grupo infinitesimal de simetría conexo está generado por los siguientes campos vectoriales:

$$\begin{aligned} &\frac{\partial}{\partial t}, \quad \frac{\partial}{\partial x}, \quad \frac{\partial}{\partial y}; \\ &d = t \frac{\partial}{\partial t} + x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}; \\ &r_{tx} = x \frac{\partial}{\partial t} + t \frac{\partial}{\partial x}, \quad r_{ty} = y \frac{\partial}{\partial t} + t \frac{\partial}{\partial y}, \quad r_{xy} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}; \\ &i_x = -2tx \frac{\partial}{\partial t} + (-t^2 - x^2 + y^2) \frac{\partial}{\partial x} - 2xy \frac{\partial}{\partial y} - xu \frac{\partial}{\partial u}, \\ &i_y = -2ty \frac{\partial}{\partial t} - 2xy \frac{\partial}{\partial x} + (-t^2 + x^2 - y^2) \frac{\partial}{\partial y} - yu \frac{\partial}{\partial u}, \\ &i_t = (-t^2 - x^2 - y^2) \frac{\partial}{\partial t} - 2tx \frac{\partial}{\partial x} - 2ty \frac{\partial}{\partial y} - tu \frac{\partial}{\partial u}; \end{aligned}$$

$$u \frac{\partial}{\partial u}, \quad v_\alpha = \alpha(t, x, y) \frac{\partial}{\partial u}, \quad (2.27)$$

donde α es una solución arbitraria de la ecuación de onda. Notemos que, en correspondencia con el teorema 2.9, dado que $\dim M^{(n)}=13$ y todos los campos vectoriales en (2.27) tienen curvas integrales de dimensión uno, que corresponden a las órbitas de la acción de cada grupo monoparamétrico asociado a cada campo, el total de invariantes funcionalmente independientes de la ecuación de onda (viendo a las órbitas como funciones de las coordenadas en lugar de verlas como curvas parametrizadas) es $13-1=12$.

Lo que nos resta ahora es encontrar los grupos monoparamétricos de transformaciones generados por estos campos vectoriales. Notemos que $u \frac{\partial}{\partial u}$ y v_α generan grupos que únicamente actúan en la variable dependiente, por lo que no nos dan información acerca de cuáles son las simetrías de las soluciones de la ecuación de onda con respecto a los cambios en las variables independientes, así que sólo determinaremos los diez primeros campos vectoriales.

Comencemos con los primeros tres que, de acuerdo a lo que vimos en el capítulo anterior, corresponden a traslaciones en las direcciones t , x y y , respectivamente, por ejemplo:

$$\exp\left(\varepsilon \frac{\partial}{\partial t}\right)(t, x, y, u) = (t + \lambda, x, y, u).$$

Para la acción determinada por d , tomemos el sistema característico:

$$\frac{dt}{t} = \frac{dx}{x} = \frac{dy}{y},$$

de donde se tiene que

$$kt = \bar{k}x = \hat{k}y,$$

que es la ecuación de una recta que pasa por el origen en el espacio $TX Y$. La ecuación paramétrica de una recta así en la dirección (t_0, x_0, y_0) es:

$$\mu(t_0, x_0, y_0)$$

$$\Rightarrow \exp(\mu d)(t, x, y, u) = (\mu t, \mu x, \mu y, u).$$

A esta transformación se le conoce como *rescalamiento* o *dilatación*.

En cuanto a r_{tx} , el sistema característico nos dice que:

$$\frac{dt}{x} = \frac{dx}{t},$$

de modo que las curvas integrales son hipérbolas, $-t^2 + x^2 = c$, en el plano TX . Si queremos parametrizar la hipérbola que pasa por el (t_0, x_0) , tenemos la curva dada por:

$$(t_0 \cosh \vartheta + x_0 \sinh \vartheta, t_0 \sinh \vartheta + x_0 \cosh \vartheta),$$

así que el grupo monoparamétrico que corresponde a r_{tx} es:

$$\exp(\vartheta r_{tx})(t, x, y, u) = (t \cosh \vartheta + x \sinh \vartheta, t \sinh \vartheta + x \cosh \vartheta, y, u).$$

Lo mismo puede hacerse para r_{ty} , por lo que estos dos campos vectoriales generan transformaciones hiperbólicas en los planos TX y TY . Para r_{xy} , sabemos reconocer que se trata de una rotación en el plano XY :

$$\exp(\theta r_{xy})(t, x, y, u) = (t, x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta, u).$$

En el caso de i_x es más difícil integrar el sistema característico pero podemos construir la transformación y verificar que efectivamente corresponde al campo vectorial. Por el momento nos restringimos al espacio de las variables independientes y adoptamos la siguiente notación:

$$x = (x^0, x^1, x^2) \equiv (t, x, y) \quad (2.28)$$

Introducimos, además, la métrica plana de Lorentz:

$$|x|^2 = -(x^0)^2 + (x^1)^2 + (x^2)^2 \quad (2.29)$$

y definimos el mapeo inversión como

$$I(x) = \frac{x}{|x|^2}$$

Tomemos la composición $I \circ \exp(\varepsilon \frac{\partial}{\partial x^1}) \circ I$ en un punto x :

$$\begin{aligned} I \circ \exp(\varepsilon \frac{\partial}{\partial x^1}) \circ I(x) &= I \circ \exp(\varepsilon \frac{\partial}{\partial x^1}) \left(\frac{x}{|x|^2} \right) \\ &= I \left(\frac{x + |x|^2 \varepsilon (0, 1, 0)}{|x|^2} \right) \\ &= \left(\frac{1}{1 + 2\varepsilon x^1 + \varepsilon^2 |x|^2} \right) (x^0, x^1 + \varepsilon |x|^2, x^2). \end{aligned}$$

Al derivar esta transformación respecto al parámetro ε y evaluar en $\varepsilon = 0$, tenemos:

$$\begin{aligned} \frac{d}{d\varepsilon} I \circ \exp(\varepsilon \frac{\partial}{\partial x^1}) \circ I(x) \Big|_{\varepsilon=0} &= \left(\frac{-x^0(2x^1 + 2\varepsilon|x|^2)}{(1 + 2\varepsilon x^1 + \varepsilon^2|x|^2)^2}, \frac{|x|^2(1 + 2\varepsilon x^1 + \varepsilon^2|x|^2) - (x^1 + \varepsilon^2|x|^2)^2(2x^1 + 2\varepsilon|x|^2)}{(1 + 2\varepsilon x^1 + \varepsilon^2|x|^2)^3} \right), \\ \left. \frac{-x^2(2x^1 + 2\varepsilon|x|^2)}{(1 + 2\varepsilon x^1 + \varepsilon^2|x|^2)^2} \right|_{\varepsilon=0} &= (-2x^0 x^1, -(x^0)^2 - (x^1)^2 + (x^2)^2, -2x^1 x^2), \end{aligned}$$

que en nuestra notación anterior equivale al vector $(-2tx, -t^2 - x^2 + y^2, -2xy)$, así que:

$$\exp(\varepsilon i_x)(t, x, y) = I \circ \exp(\varepsilon \frac{\partial}{\partial x^1}) \circ I(t, x, y).$$

Por otro lado, la transformación que genera i_x en la variable dependiente está dada por:

$$u \rightarrow \sqrt{1 + 2\varepsilon x^1 + \varepsilon^2 |x|^2} u.$$

Para construir las acciones generadas por i_y e i_t procedemos de manera similar a lo que hemos hecho con i_x , entendiendo entonces que las acciones generadas por estos campos vectoriales transforman a las variables independientes mediante la composición de inversiones y traslaciones.

Para finalizar, notemos que todas las transformaciones que hemos encontrado son mapeos conformes respecto a la métrica plana de Lorentz en el espacio de variables independientes. Es decir, si en el espacio $TX\bar{Y}$ introducimos la notación y la métrica dadas por (2.28) y (2.29), cualquiera de los grupos monoparamétricos anteriores transforma a cada punto x en otro punto \bar{x} cuya norma es proporcional a la norma de x :

$$|x| = \lambda(x)|\bar{x}|.$$

De hecho, las traslaciones, la rotación y las transformaciones hiperbólicas que encontramos son isometrías ($|x| = |\bar{x}|$). Se dice entonces que la ecuación de onda es invariante respecto al grupo de transformaciones conformes de Lorentz.

Capítulo 3

Grupos de Simetría y Leyes de Conservación

En este capítulo nos preocupamos por establecer la relación que hay entre los grupos de simetría de los sistemas mecánicos y las leyes de conservación, analizando primero algunos ejemplos, que nos lleven a establecer las simetrías propias de la mecánica clásica, y teniendo como principal resultado el teorema de Noether.

3.1 Elementos del Cálculo Variacional.

De acuerdo con lo que vimos en el primer capítulo, para puntos $x(t) = (x^1(t), \dots, x^n(t))$ en una variedad, podemos construir su campo de velocidades de la siguiente manera:

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

de modo que \dot{x} es un vector tangente a x al tiempo t .

Dentro de la formulación lagrangiana de la mecánica, tomamos parejas de puntos (x, \dot{x}) que se encuentran en el haz tangente o *espacio fase*, cuyas $2n$ coordenadas están dadas por (x^1, \dots, x^n) y $(\dot{x}^1, \dots, \dot{x}^n)$ y especifican el estado de movimiento de un sistema mecánico que cambia con respecto al tiempo. Además, definimos una función en el espacio fase que toma valores reales $L(x, \dot{x})$, la cual no depende explícitamente de t .

Por ejemplo, si $x(t)$ denota la posición de una partícula al tiempo t , \dot{x} corresponde a la velocidad de la partícula en el mismo tiempo. Dada una métrica definida en nuestro espacio, al tomar $L = |\dot{x}|$ estamos considerando un lagrangiano que representa a la norma de la velocidad de la partícula.

Ahora tomamos dos puntos fijos $x_1 = (x_1^1, \dots, x_1^n)$ y $x_2 = (x_2^1, \dots, x_2^n)$ y consideramos el conjunto de todas las curvas suaves que unen a x_1 y x_2 :

$$\gamma = x(t), \quad a \leq t \leq b, \quad x^i(a) = x_1^i, \quad x^i(b) = x_2^i \quad \forall \gamma.$$

Para cada γ definimos

$$S[\gamma] = \int_{x_1}^{x_2} L(x(t), \dot{x}(t)) dt.$$

S es un funcional que asocia a cada curva γ , de x_1 a x_2 , el número $S[\gamma]$ y se conoce como la *acción*.

Siguiendo con nuestro ejemplo, tenemos $L = |\dot{x}|$ y

$$S[\gamma] = \int_{x_1}^{x_2} |\dot{x}| dt$$

es la longitud de arco de γ . Notemos que, si encontramos una curva γ para la cual el valor de $S[\gamma]$ es mínimo, estaremos encontrando la trayectoria de menor longitud entre los puntos x_1 y x_2 .

La idea central de esta formulación se basa en conocer cuál es la curva para la que $S[\gamma]$ alcanza un valor extremo (mínimo), a lo que se conoce como *principio de mínima acción*; así, en un espacio euclideo, si $L = \frac{m}{2} |\dot{x}|^2 - U(x)$, donde m es una constante y $U(x)$ es una función escalar, entonces la acción es:

$$S[\gamma] = \int_{x_1}^{x_2} \left(\sum_i \frac{m}{2} (\dot{x}^i)^2 - U(x) \right) dt.$$

La mecánica clásica nos dice cuál es la curva γ para la que $S[\gamma]$ tiene un valor mínimo es la trayectoria de una partícula puntual de masa m en un campo de fuerzas $f_i = -\frac{\partial U}{\partial x^i}$.

El siguiente teorema da una condición necesaria para que una curva γ minimice a $S[\gamma]$.

Teorema 3.1 (Ecuaciones de Euler - Lagrange).

Si $S[\gamma] = \int_{x_1}^{x_2} L(x, \dot{x}) dt$ alcanza su mínimo sobre todas las curvas γ en una curva suave $\gamma_0 : x^i = x^i(t)$, es decir, si $S[\gamma_0] \leq S[\gamma]$ para toda curva suave γ de x_1 a x_2 , entonces, a lo largo de γ_0 , se cumplen las siguientes ecuaciones, conocidas como ecuaciones de Euler - Lagrange:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} = 0, \quad i = 1, \dots, n. \quad (3.1)$$

Las soluciones a estas ecuaciones se llaman curvas extremas.

Aquí tomamos $L = L(x^1, \dots, x^n, \xi^1, \dots, \xi^n)$, de manera que vemos a las variables x^i y ξ^i como independientes; sólo después de que se calculan las derivadas, se evalúa $\xi^i = \dot{x}^i$, es decir, se considera a $\xi = (\xi^1, \dots, \xi^n) \Big|_x$ como un vector tangente a γ_0 en x . Así, tenemos:

$$\frac{\partial L}{\partial \dot{x}^i} = \frac{\partial L(x, \xi)}{\partial \xi^i} \Big|_{\xi=\dot{x}}$$

Además,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) = \frac{d}{dt} \left(\frac{\partial L(x, \xi)}{\partial \xi^i} \right) \Big|_{\xi=\dot{x}}$$

$$\begin{aligned}
 &= \sum_j \left(\frac{\partial}{\partial x^j} \left(\frac{\partial L}{\partial \xi^i} \right) \frac{dx^j}{dt} + \frac{\partial}{\partial \xi^j} \left(\frac{\partial L}{\partial \xi^i} \right) \frac{d\xi^j}{dt} \right) \Big|_{\xi=x} \\
 &= \sum_j \left(\frac{\partial^2 L}{\partial \xi^i \partial \xi^j} \frac{dx^j}{dt} + \frac{\partial^2 L}{\partial \xi^i x^j} \dot{x}^j \right) \Big|_{\xi=x} \\
 &= \sum_j \left(\frac{\partial^2 L}{\partial \xi^i \partial \xi^j} \dot{x}^j + \frac{\partial^2 L}{\partial \xi^i x^j} x^j \right) \Big|_{\xi=x}.
 \end{aligned}$$

Prueba.

Sea $\alpha^i = \alpha^i(t)$, $a \leq t \leq b$, una función suave cualquiera que satisface $\alpha^i(a) = 0 = \alpha^i(b)$. Consideremos la expresión

$$\lim_{\varepsilon \rightarrow 0} \frac{S[\gamma_0 + \varepsilon \alpha] - S[\gamma_0]}{\varepsilon} = \frac{d}{d\varepsilon} S[\gamma_0 + \varepsilon \alpha] \Big|_{\varepsilon=0},$$

donde $\gamma_0 + \varepsilon \alpha$ es una curva que une a x_1 y x_2 cercana a γ_0 para una ε pequeña. Por hipótesis, $S[\gamma_0 + \varepsilon \alpha]$ toma un valor mínimo cuando $\varepsilon = 0$, por lo que

$$\frac{d}{d\varepsilon} S[\gamma_0 + \varepsilon \alpha] \Big|_{\varepsilon=0} = 0.$$

Por otro lado,

$$\begin{aligned}
 0 &= \frac{d}{d\varepsilon} S[\gamma_0 + \varepsilon \alpha] \Big|_{\varepsilon=0} = \int_a^b \left\{ \frac{d}{d\varepsilon} L(x + \varepsilon \alpha, \dot{x} + \varepsilon \dot{\alpha}) \Big|_{\varepsilon=0} \right\} dt \\
 &= \int_a^b \sum_i \frac{\partial L}{\partial x^i} \alpha^i + \frac{\partial L}{\partial \dot{x}^i} \dot{\alpha}^i dt. \tag{3.2}
 \end{aligned}$$

Integrando por partes, encontramos que

$$\int_a^b \frac{\partial L}{\partial \dot{x}^i} \dot{\alpha}^i dt = \left(\alpha^i \frac{\partial L}{\partial \dot{x}^i} \right) \Big|_{t=b} - \left(\alpha^i \frac{\partial L}{\partial \dot{x}^i} \right) \Big|_{t=a} - \int_a^b \alpha^i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) dt$$

y, como $\alpha^i(a) = 0 = \alpha^i(b)$, obtenemos

$$\int_a^b \frac{\partial L}{\partial \dot{x}^i} \dot{\alpha}^i dt = \int_a^b \alpha^i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) dt.$$

Sustituyendo esto en (3.2), tenemos

$$\frac{d}{d\varepsilon} S[\gamma_0 + \varepsilon \alpha] \Big|_{\varepsilon=0} = \int_a^b \sum_i \left(\frac{\partial L}{\partial x^i} + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) \right) \alpha^i dt = 0. \tag{3.3}$$

Esta ecuación es válida para cualquier función suave $\alpha(t)$, por lo tanto,

$$\psi^i(t) = \frac{\partial L}{\partial x^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) = 0, \quad i = 1, \dots, n,$$

pues, si $\psi^i(t) \neq 0$ para alguna i y una $t = t_0$, $a \leq t_0 \leq b$, entonces $\psi^i(t) \neq 0$ en un intervalo de valores de t , así que, si escogemos una $\alpha^i(t)$ adecuada (por ejemplo, $\alpha^i(t) = \psi^i(t)f(t)$ donde $f(a) = f(b) = 0$ y $f(t) > 0$ para $a < t < b$, de modo que el integrando en (3.3) es positivo en un subintervalo de $[a, b]$), podemos obtener que la integral en (3.3) es distinta de cero.

□

Convencionalmente, definimos las siguientes cantidades:

i) La función

$$L = L(x, \dot{x}) \quad (3.4)$$

cuya integral respecto al tiempo queremos minimizar se llama *lagrangiano*.

ii) La *energía* E se conoce como

$$E = E(x, \dot{x}) = \sum_i \dot{x}^i \frac{\partial L}{\partial \dot{x}^i} - L. \quad (3.5)$$

iii) El *momento generalizado* está definido como

$$p_i = \frac{\partial L}{\partial \dot{x}^i}. \quad (3.6)$$

iv) La *fuerza generalizada* está dada por la expresión

$$f_i = \frac{\partial L}{\partial x^i}. \quad (3.7)$$

Siguiendo esto, las ecuaciones de Euler - Lagrange se pueden escribir como:

$$\dot{p}_i = f_i. \quad (3.8)$$

Esta terminología se utiliza teniendo como referencia a la mecánica clásica puesto que, si $L = \frac{m}{2} \sum_i (\dot{x}^i)^2 - U(x)$, las expresiones del momento lineal y la fuerza toman su forma usual para una partícula en movimiento de masa m , de manera que la ecuación

$$\dot{p}_i = f_i \quad \text{ó} \quad m\ddot{x}^i = -\frac{\partial U}{\partial x^i} \quad (3.9)$$

es equivalente a la segunda ley de Newton.

Notemos que las curvas extremas de un lagrangiano $L = L(x, \dot{x})$ también corresponden a un lagrangiano que depende explícitamente del tiempo si este último tiene la forma $\tilde{L}(t, x, \dot{x}) = L + \frac{d}{dt}f(x, t)$, donde f es una función suave, puesto que

$$\frac{d}{d\varepsilon} \int_a^b \frac{d}{dt} f(x + \varepsilon \alpha, t) dt \Big|_{\varepsilon=0} = \int_a^b \frac{d}{dt} \left\{ \frac{d}{d\varepsilon} f(x + \varepsilon \alpha, t) \Big|_{\varepsilon=0} \right\} dt$$

$$\begin{aligned}
 &= \int_a^b \frac{d}{dt} \left\{ \sum_i \frac{\partial f}{\partial x^i} \alpha^i \right\} dt \\
 &= \left[\sum_i \frac{\partial f}{\partial x^i} \alpha^i \right]_a^b = 0,
 \end{aligned}$$

pues $\alpha^i(a) = 0 = \alpha^i(b)$. La energía y el momento generalizado correspondientes están dados por $\hat{E} = E - \frac{d}{dt}f(x, t)$ y $\hat{p} = p + \frac{\partial L}{\partial x}$, respectivamente.

3.2 Ejemplos de Leyes de Conservación: Mecánica Clásica.

Durante el movimiento de un sistema mecánico, las $2n$ cantidades x^i y \dot{x}^i , $i = 1, \dots, n$, que especifican el estado del sistema, varían con el tiempo. Sin embargo, existen funciones de estas cantidades, llamadas *integrales de movimiento*, cuyos valores permanecen constantes, es decir, siguen *leyes de conservación*, y sólo dependen de las condiciones iniciales.

No todas las integrales de movimiento son de igual importancia, hay algunas que tienen gran utilidad debido a que son aditivas, esto es: si un sistema compuesto por varias partículas, cuya interacción es despreciable, tiene una integral de movimiento aditiva, entonces el valor de esta última es igual a la suma de los valores de las integrales de movimiento correspondientes a cada partícula. Por ejemplo, para dos cuerpos que interactúan durante cierto intervalo de tiempo, dado que cada integral de movimiento aditiva de todo el sistema es igual a la suma de los dos valores correspondientes a los cuerpos por separado, antes y después de la interacción, entonces, si conocemos los estados anteriores a que los cuerpos interactúen, las leyes de conservación nos permiten deducir varias conclusiones respecto al estado de ambos cuerpos después de la interacción.

Algunas leyes de conservación tienen un significado profundo en virtud de que se deducen a partir de ciertas suposiciones de simetría en el espacio-tiempo donde se mueve el sistema mecánico que estamos considerando.

Para revisar algunos ejemplos, tomemos el análisis de la mecánica clásica al caso del movimiento de una partícula de masa m en el espacio euclideo plano \mathbb{R}^3 ($(x = (x^1, x^2, x^3), \dot{x} = (\dot{x}^1, \dot{x}^2, \dot{x}^3), |x|^2 = \sum_{i=1}^3 (x^i)^2)$), caracterizado por el lagrangiano $L(x, \dot{x}) = \frac{m}{2} |\dot{x}|^2 - U(x)$. Como ya habíamos comentado, en este caso las ecuaciones de Euler - Lagrange corresponden a la segunda ley de Newton ($\frac{d}{dt}m\dot{x}^i = \frac{\partial U}{\partial x^i} = f_i$) y la energía está dada por $E = \frac{m}{2} |\dot{x}|^2 + U(x)$.

Ahora bien, el hecho de escoger un lagrangiano que tenga precisamente esa forma (en cuya expresión no aparece el tiempo), es una manera implícita de suponer homogeneidad temporal, es decir, en un intervalo de tiempo dado $[a, b]$, el movimiento queda

completamente descrito por las cantidades $x(t)$ y $\dot{x}(t)$ independientemente de que nos traslademos o no a otro intervalo $[a + t_0, b + t_0]$. Así, el cambio total del lagrangiano respecto al tiempo está dado por:

$$\frac{dL}{dt} = \sum_{i=1}^3 m\dot{x}^i \ddot{x}^i - \frac{\partial U}{\partial x^i} \dot{x}^i.$$

Al aplicar las ecuaciones de Euler - Lagrange y la regla de Leibniz tenemos:

$$\begin{aligned} \frac{dL}{dt} &= \sum_{i=1}^3 m\dot{x}^i \ddot{x}^i + \dot{x}^i \frac{d}{dt} m\dot{x}^i = \sum_{i=1}^3 \frac{d}{dt} \dot{x}^i m\dot{x}^i \\ &\Rightarrow \frac{d}{dt} \sum_{i=1}^3 m(\dot{x}^i)^2 - L = 0. \end{aligned}$$

Por lo tanto, la energía

$$E = \sum_{i=1}^3 m(\dot{x}^i)^2 - L = \frac{m}{2} |\dot{x}|^2 - U(x) \quad (3.10)$$

se mantiene constante durante el movimiento de una partícula. Como podemos ver, la expresión de la energía está dada por la suma de la energía cinética ($\frac{m}{2} |\dot{x}|^2$) más la energía potencial ($U(x)$), cuyo gradiente es igual a la fuerza que actúa sobre la partícula ($f_i = -\frac{\partial U}{\partial x^i}$). La suposición de la homogeneidad del tiempo se ve reflejada justamente en que la fuerza no depende explícitamente del tiempo.

Una segunda ley de conservación se sigue de suponer homogeneidad en el espacio o bien, que las propiedades del movimiento de una partícula no cambian si realizamos una traslación espacial de nuestro sistema de referencia, lo que trae como consecuencia que el lagrangiano no varíe bajo dicha traslación. Podemos plantear la homogeneidad espacial tomando la acción de un grupo monoparamétrico (el de las traslaciones) en el espacio, $\varphi_s(x) = (x^1 + as, x^2 + bs, x^3 + cs)$, donde (a, b, c) es un vector constante no nulo que nos da la dirección de la traslación, y considerando que, al trasladar la curva que describe la posición de nuestra partícula $x(t)$, la nueva curva $x(t) + s(a, b, c)$ también describe la posición de la misma bajo un sistema de referencia trasladado, por lo que el lagrangiano no cambia bajo esta acción:

$$\begin{aligned} &\frac{d}{ds} L(x^1 + as, x^2 + bs, x^3 + cs, \dot{x}^1, \dot{x}^2, \dot{x}^3) \Big|_{s=0} = 0 \\ \Rightarrow &\frac{d}{ds} \left(\frac{m}{2} |\dot{x}|^2 - U(x^1 + as, x^2 + bs, x^3 + cs) \right) \Big|_{s=0} = 0 \\ \Rightarrow &-a \frac{\partial U}{\partial x^1} - b \frac{\partial U}{\partial x^2} - c \frac{\partial U}{\partial x^3} = 0 \end{aligned} \quad (3.11)$$

y, por las ecuaciones de Euler - Lagrange, vemos que

$$a \frac{d}{dt} m \dot{x}^1 + b \frac{d}{dt} m \dot{x}^2 + c \frac{d}{dt} m \dot{x}^3 + 0 \\ \Rightarrow (a, b, c) \cdot \frac{d}{dt} m \dot{\mathbf{x}} = 0.$$

Por lo tanto, el momento de una partícula

$$p = m \dot{\mathbf{x}} \quad (3.12)$$

es una cantidad que se conserva.

Observando (3.11), notemos que

$$-a \frac{\partial U}{\partial x^1} - b \frac{\partial U}{\partial x^2} - c \frac{\partial U}{\partial x^3} = 0 \Rightarrow (a, b, c) \cdot \left(-\frac{\partial U}{\partial x^1}, -\frac{\partial U}{\partial x^2}, -\frac{\partial U}{\partial x^3} \right) = (a, b, c) \cdot (-\nabla U) = 0 \\ \Rightarrow (a, b, c) \cdot (f_1, f_2, f_3) = 0. \quad (3.13)$$

De modo que el suponer que el espacio es homogéneo nos lleva a imponer la condición de considerar un sistema cerrado en el que la fuerza total sobre la partícula es nula; una vez hecho esto, la conservación del momento se verifica usando directamente las ecuaciones de Euler - Lagrange y la partícula se mueve siguiendo un movimiento rectilíneo uniforme ($\frac{d}{dt} m \dot{\mathbf{x}} = m \ddot{\mathbf{x}} = 0 \Rightarrow \dot{\mathbf{x}} = cte = \dot{x}_0 \Rightarrow x(t) = \dot{x}_0 t + x_0$). De otro modo, debemos considerar que en el sistema hay una fuerza ejerciéndose sobre la partícula que afecta su movimiento. Sin embargo, también en este caso puede aplicarse la ley de conservación del momento: una de las componentes del momento puede permanecer constante aún si hay una fuerza neta no nula ejerciéndose sobre la partícula, siempre que la energía potencial no dependa de todas las coordenadas. Por ejemplo, si tenemos una fuerza constante en la dirección x^3 , esto quiere decir que $U(x) = U(x^3)$, la homogeneidad del espacio nos permite realizar traslaciones del sistema de referencia en las direcciones x^1 y x^2 (las soluciones a las ecuaciones de Euler - Lagrange no se ven afectadas por traslaciones en el plano formado por estas dos direcciones) y las componentes p_1 y p_2 del momento se conservan.

Hay una tercera ley de conservación que puede derivarse a partir de suponer la isotropía del espacio; esto quiere decir que las propiedades del movimiento de nuestra partícula no cambian por la rotación del sistema de referencia. De manera similar al caso anterior, planteamos la isotropía del espacio tomando la acción del grupo de rotaciones en el espacio y considerando que el lagrangiano no varía bajo esta acción. Para hacer esto, resulta más fácil tomar coordenadas esféricas (r, ϕ, θ) , puesto que las rotaciones se ven como simples traslaciones en las coordenadas angulares. La expresión del lagrangiano $L(x, \dot{x}) = \frac{m}{2} |\dot{x}|^2 - U(x)$ en coordenadas esféricas está dada por:

$$L(r, \phi, \theta, \dot{r}, \dot{\phi}, \dot{\theta}) = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\phi}^2 + r^2 \sin^2 \phi \dot{\theta}^2) - U(r, \phi, \theta)$$

y las ecuaciones de Euler - Lagrange son:

$$\frac{d}{dt}mr\dot{r} = -\frac{\partial U}{\partial r}, \quad \frac{d}{dt}mr^2\dot{\phi} = mr^2 \sin\phi \cos\phi \dot{\theta}^2 - \frac{\partial U}{\partial \phi}, \quad \frac{d}{dt}mr^2 \sin^2\phi \dot{\theta} = -\frac{\partial U}{\partial \theta}. \quad (3.14)$$

Tomemos la rotación $\varphi_s(r, \phi, \theta) = (r, \phi + as, \theta + bs)$. De la isotropía del espacio y las ecuaciones de Euler - Lagrange en (3.14), tenemos:

$$\begin{aligned} & \left. \frac{d}{ds}L(r, \phi + as, \theta + bs, \dot{r}, \dot{\phi}, \dot{\theta}) \right|_{s=0} = 0 \\ \Rightarrow & \left[amr^2 \sin(\phi + as) \cos(\phi + as) \dot{\theta}^2 - a \frac{\partial U}{\partial \phi} - b \frac{\partial U}{\partial \theta} \right]_{s=0} = 0 \\ \Rightarrow & amr^2 \sin\phi \cos\phi \dot{\theta}^2 - a \frac{\partial U}{\partial \phi} - b \frac{\partial U}{\partial \theta} = 0 \\ \Rightarrow & a \frac{d}{dt}mr^2\dot{\phi} + b \frac{d}{dt}mr^2 \sin^2\phi \dot{\theta} = 0. \end{aligned} \quad (3.15)$$

Para entender cuál es la cantidad que se conserva, supongamos primero que que la rotación del sistema de referencia sólo se lleva a cabo en ϕ ($b = 0$). La ecuación (3.15) dice entonces que $a \frac{d}{dt}mr^2\dot{\phi} = 0$, es decir, que $mr^2\dot{\phi}$ permanece constante, pero esta cantidad corresponde a la norma del vector de momento angular $(0, 0, mr^2\dot{\phi})$ de la partícula, cuya dirección permanece constante:

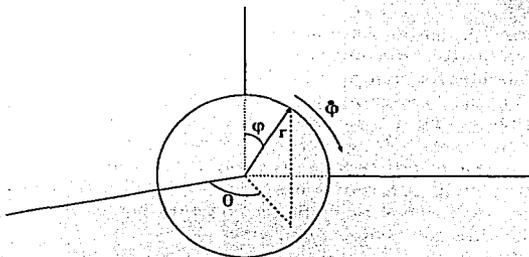


Figura 3.1: Conservación de momento angular, caso $b = 0$.

por lo tanto, el momento angular se conserva.

Ahora consideremos que la rotación es sólo en θ ($a = 0$). En este caso, según (3.15), $b \frac{d}{dt}mr^2 \sin^2\phi \dot{\theta} = 0$, y $mr^2 \sin^2\phi \dot{\theta}$ se conserva. El momento angular es $(0, -mr^2 \sin\phi \dot{\theta}, 0)$, así que la cantidad que permanece constante corresponde al tamaño de la componente del momento angular de la partícula cuya dirección es igual a aquella en la que la fuerza no varía:

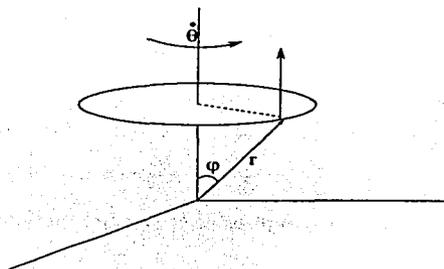


Figura 3.2: Conservación de momento angular, caso $\alpha = 0$.

De modo que hemos encontrado una ley de conservación del momento angular que dice que, en un sistema cerrado donde la fuerza total sobre la partícula es nula, el momento angular se conserva. Por otro lado, si tenemos una fuerza ejerciéndose sobre la partícula que cambia con el tiempo, la componente del momento angular en la dirección en la que la fuerza no varía, permanece constante.

Finalmente, la aditividad de la energía, el momento y el momento angular es una propiedad que se hereda de la aditividad del lagrangiano mismo y de la linealidad de las derivadas.

Lo que hemos hecho en esta sección ha sido establecer parte de la estructura del espacio-tiempo en la mecánica clásica, junto con su grupo de simetrías. En resumen, se tiene que el espacio-tiempo es euclideo de cuatro dimensiones (una temporal y tres espaciales), homogéneo en el tiempo y en el espacio y espacialmente isotrópico. La métrica en el espacio es plana y está dada por $|x| = \sum_{i=1}^3 (x^i)^2$. El estado inicial de un sistema mecánico, es decir, el total de las posiciones y las velocidades de sus puntos en un tiempo dado t_0 , determina de manera única al movimiento, mediante las ecuaciones de Euler - Lagrange.

A esto le añadimos un principio más, conocido como el principio galileano de la relatividad, que dice que los sistemas coordenados de referencia (llamados inerciales) deben cumplir las siguientes propiedades:

- i) Todas las leyes de la naturaleza en cualquier momento del tiempo son iguales en todos los sistemas coordenados inerciales.
- ii) Todos los sistemas coordenados en movimiento rectilíneo uniforme respecto a un sistema inercial son también inerciales.

En consecuencia, hay una transformación más de nuestras coordenadas bajo la cual se siguen cumpliendo las ecuaciones de Euler - Lagrange: el considerar que tenemos un nuevo sistema de referencia, que se mueve en línea recta con velocidad constante v , respecto a nuestro sistema de referencia inicial. De modo que, si tomamos la siguiente

transformación:

$$\begin{aligned} x(t) &\rightarrow x(t) + vt \\ \Rightarrow \dot{x}(t) &\rightarrow \dot{x}(t) + v, \end{aligned} \quad (3.16)$$

para las nuevas coordenadas, el lagrangiano es:

$$\begin{aligned} L(x(t) + vt, \dot{x}(t) + v) &= \frac{m}{2} |\dot{x} + v|^2 - U(x + vt) \\ \Rightarrow \frac{\partial L}{\partial x^i} &= \frac{\partial U}{\partial x^i}, \quad \frac{\partial L}{\partial \dot{x}^i} = m(\dot{x}^i + v^i). \end{aligned}$$

En las coordenadas originales, se cumplen las ecuaciones de Euler - Lagrange, $\frac{d}{dt}m\dot{x}^i = \frac{\partial U}{\partial x^i} \Rightarrow m\ddot{x}^i = f_i$, así que:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) = \frac{d}{dt} m(\dot{x}^i + v^i) = m\ddot{x}^i = f_i = \frac{\partial U}{\partial x^i} = \frac{\partial L}{\partial x^i}.$$

Por lo tanto, se cumplen las ecuaciones de Euler - Lagrange.

Al grupo de transformaciones formado por las traslaciones en el tiempo, las traslaciones en el espacio, las rotaciones en el espacio y el movimiento rectilíneo uniforme se le conoce como el grupo de transformaciones galileanas. Éste es el grupo de simetrías propio de la mecánica clásica, puesto que es el que determina cómo deben ser los sistemas de referencia válidos para esta teoría.

3.3 Simetrías y Cantidades Conservadas.

En términos generales, para un lagrangiano $L(x, \dot{x})$, se puede demostrar que tanto la energía como el momento generalizado se conservan.

Teorema 3.2 (*Ley de la conservación de la Energía*).

Sea $L(x, \dot{x})$ un lagrangiano. A lo largo de una curva extrema, la energía E es constante.

Prueba.

$$\begin{aligned} \frac{dE}{dt} &= \frac{d}{dt} \left(\sum_i \dot{x}^i \frac{\partial L}{\partial \dot{x}^i} - L \right) = \sum_i \ddot{x}^i \frac{\partial L}{\partial \dot{x}^i} + \dot{x}^i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \left(\dot{x}^i \frac{\partial L}{\partial x^i} + \ddot{x}^i \frac{\partial L}{\partial \dot{x}^i} \right) \\ &= \sum_i \dot{x}^i \left(\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} \right) = 0, \end{aligned}$$

pues $L = L(x, \dot{x})$ sólo depende implícitamente del tiempo y, a lo largo de las curvas extremas, se cumplen las ecuaciones de Euler - Lagrange. □

Los sistemas mecánicos en los que la energía se conserva son conocidos como *conservativos*. Estos sistemas pueden estar bajo la influencia de un campo externo de fuerzas, siempre que éste no involucre al tiempo explícitamente. Podemos interpretar el resultado del teorema anterior como que, en sistemas conservativos, las curvas $x(t)$ forman un grupo monoparamétrico de simetrías para la energía, cuyo generador infinitesimal es el campo de velocidades \dot{x}

La ley de conservación del momento generalizado se puede demostrar usando la idea de invariancia del lagrangiano bajo la acción de un grupo monoparamétrico de transformaciones. Como ya vimos en la sección anterior, esto nos lleva a plantear simetrías intrínsecas al espacio-tiempo en el que estamos situando nuestro sistema mecánico.

Definición 3.3 Sea $\varphi_s(x)$ un grupo monoparamétrico de transformaciones. Se dice que φ_s preserva al lagrangiano $L(x, \dot{x})$ si

$$\left. \frac{d}{ds} L(\varphi_s(x), (\varphi_s(\dot{x}))) \right|_{s=0} = 0. \quad (3.17)$$

donde $(\varphi_s(\dot{x})) = \frac{d\varphi_s(x)}{dt}$

Notemos que, si $\gamma^i(t) = x^i(t)$ es una curva extrema de L y φ_s preserva al lagrangiano, la curva $\varphi_s(\gamma(t)) = \varphi_s(x)$ también es extremo de L , ya que

$$\left. \frac{d}{ds} S[\varphi_s(\gamma(t))] \right|_{s=0} = \int_a^b \left\{ \left. \frac{d}{ds} L(\varphi_s(x), (\varphi_s(\dot{x}))) \right|_{s=0} \right\} dt = 0.$$

Según la sección 1.3, sabemos que todo grupo monoparamétrico de transformaciones tiene asociado un campo vectorial; para φ_s , está dado por:

$$\left. \frac{d\varphi_s^i(x)}{ds} \right|_{t=0}, \quad i = 1, \dots, n.$$

Teorema 3.4 (Noether).

Si un grupo monoparamétrico de transformaciones φ_s preserva al lagrangiano L , entonces, en cualquier curva extrema de L , la componente del momento en dirección del generador infinitesimal de φ_s se conserva, es decir:

$$\frac{d}{dt} \left(\sum_i \frac{\partial L}{\partial \dot{x}^i} \frac{d\varphi_s^i(x)}{ds} \right) \Big|_{s=0} = \frac{d}{dt} \left(\sum_i p_i \frac{d\varphi_s^i(x)}{ds} \right) \Big|_{s=0} = 0 \quad (3.18)$$

Prueba.

Como φ_s preserva a L , tenemos que

$$0 = \left. \frac{d}{ds} L(\varphi_s(x), (\varphi_s(\dot{x}))) \right|_{s=0} = \sum_i \frac{\partial L}{\partial x^i} \frac{d\varphi_s^i(x)}{ds} \Big|_{s=0} + \frac{\partial L}{\partial \dot{y}^i} \frac{d(\varphi_s^i(\dot{y}))}{ds} \Big|_{s=0} \quad (3.19)$$

Intercambiando el orden de las derivadas en el segundo término de la suma en (3.19), obtenemos

$$\sum_i \frac{\partial L}{\partial x^i} \frac{d\varphi_s^i(x)}{ds} \Big|_{s=0} + \frac{\partial L}{\partial \dot{x}^i} \frac{d}{dt} \left(\frac{d\varphi_s^i(x)}{ds} \Big|_{s=0} \right) = 0. \quad (3.20)$$

Como estamos en una curva extrema, se cumplen las ecuaciones de Euler - Lagrange, de manera que $\frac{\partial L}{\partial x^i} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right)$, por lo tanto, sustituyendo en el primer término de la suma en (3.20) y usando la regla de Leibniz, se cumple que

$$\sum_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) \frac{d\varphi_s^i(x)}{ds} \Big|_{s=0} + \frac{\partial L}{\partial \dot{x}^i} \frac{d}{dt} \left(\frac{d\varphi_s^i(x)}{ds} \Big|_{s=0} \right) = \frac{d}{dt} \sum_i \left(\frac{\partial L}{\partial \dot{x}^i} \frac{d\varphi_s^i(x)}{ds} \Big|_{s=0} \right) = 0.$$

□

Aunque el teorema de Noether se presenta como una ley de conservación del momento generalizado, de la cual se derivan la conservación del momento y del momento angular como casos particulares, este resultado tiene distintas aplicaciones que van más allá de una idea puramente mecánica, de hecho, es un principio que relaciona los grupos de simetría con las leyes de conservación en general.

85

Capítulo 4

Aplicaciones a la Relatividad General

Se hacen algunas aplicaciones de las ideas manejadas en los capítulos anteriores. Se ilustra el surigimiento de la Teoría Especial de la Relatividad desde el punto de vista de la incompatibilidad entre las simetrías propias de las teorías de Maxwell y Newton. En cuanto a la Teoría General de la Relatividad, se deduce una solución exacta para la ecuación de Einstein (solución de Schwarzschild) usando suposiciones de simetría estática y esférica. Finalmente, se habla de la determinación de la "entropía" de un hoyo negro usando el teorema de Noether ¹.

4.1 Relatividad Especial.

La mecánica newtoniana, con la estructura del espacio-tiempo y las simetrías que plantea, es una teoría bastante exitosa que explica gran cantidad de fenómenos. Sin embargo, existen ciertos hechos experimentales que no pueden entenderse en términos de esta teoría, algunos de ellos relacionados con el movimiento de partículas inestables a grandes velocidades. Además, la mecánica de Newton es incompatible con la teoría electromagnética de Maxwell desde el punto de vista de sus simetrías. La inconsistencia no proviene de suponer homogeneidad y/o isotropía en el espacio-tiempo, sino de la transformación de las velocidades a partir de considerar dos sistemas de referencia, uno moviéndose con respecto a otro a una velocidad constante v , que surge como consecuencia del principio galileano de la relatividad. En efecto, las ecuaciones de Maxwell:

$$\begin{aligned}\nabla \cdot E &= \rho \\ \nabla \cdot B &= 0\end{aligned}$$

¹La introducción del lenguaje tensorial, así como las definiciones precisas relacionadas con los conceptos de métrica, curvatura, transporte paralelo y derivada covariante se dan por conocidos (ver Wild [10]).

$$\begin{aligned} \frac{\partial B}{\partial t} + \nabla \times E &= 0 \\ \frac{\partial E}{\partial t} + \nabla \times B &= j, \end{aligned} \quad (4.1)$$

donde ρ es la densidad de carga, j la corriente eléctrica y E y B corresponden a los campos eléctrico y magnético, respectivamente, no son invariantes bajo la transformación dada por (3.16). En cambio, si introducimos las transformaciones de Lorentz²:

$$\begin{aligned} t &\rightarrow \frac{t - vx}{\sqrt{1 - v^2}} \\ x &\rightarrow \frac{x - vt}{\sqrt{1 - v^2}}, \end{aligned} \quad (4.2)$$

(la velocidad de la luz en el vacío se considera normalizada, es decir $c=1$), podemos observar que las ecuaciones de Maxwell son invariantes.

Esta diferencia entre las transformaciones de coordenadas bajo las cuales la ecuación de Newton y las ecuaciones de Maxwell son invariantes nos lleva a escoger entre la validez de una y otra teoría. En 1905, Albert Einstein decidió escoger a la teoría de Maxwell como aquella que debía considerarse válida y propuso una nueva teoría para la mecánica, conocida como la *Teoría Especial de la Relatividad*, la cual se basa en dos principios fundamentales:

- 1) Las leyes de la física son las mismas desde cualquier sistema de referencia inercial.
- 2) La velocidad de la luz es independiente del movimiento de la fuente.

A partir de estos principios puede deducirse que las transformaciones de coordenadas adecuadas para dos sistemas de referencia entre los que existe un movimiento rectilíneo uniforme con velocidad v , están dadas por (4.2), de modo que la Teoría Especial de la Relatividad es compatible con la teoría electromagnética de Maxwell, además de que explica exitosamente aquellos hechos experimentales en los que la mecánica newtoniana falla.

Por otro lado las ecuaciones de Maxwell en el vacío (de hecho en medios donde la densidad de carga y la corriente electromagnética son constantes, lo que, en el fondo, involucra homogeneidad e isotropía en el espacio-tiempo) se reducen a una ecuación de onda para el campo eléctrico y otra ecuación de onda para el campo magnético. Recordando la sección 2.3, sabemos que la ecuación de onda es invariante ante las transformaciones conformes de Lorentz. Dichas transformaciones introducen una métrica en el espacio-tiempo dada por

$$|x| = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2, \quad (4.3)$$

llamada la métrica de Lorentz plana³ o métrica de Minkowsky. De este modo, podemos pensar que la Teoría Especial de la Relatividad tiene asociada una métrica dada por

²Estas transformaciones toman en cuenta a la coordenada temporal, mientras que todas las que presentamos en la sección 3.2 sólo involucran a las coordenadas espaciales.

³Esta métrica corresponde con la ecuación (2.29) en un espacio-tiempo con tres coordenadas espaciales en lugar de dos.

(4.3) (distinta a la euclidea) y que simetrías del espacio-tiempo como la homogeneidad y la isotropía están íntimamente relacionadas con esta nueva métrica puesto que, como se menciona en la sección 2.3, las traslaciones y las rotaciones (tanto hiperbólicas como esféricas) son isometrías, considerando la métrica de Minkowsky. Entonces, para la Teoría Especial de la Relatividad, el espacio-tiempo es la variedad \mathbb{R}^4 con una métrica de Lorentz plana.

4.2 Relatividad General.

Dado que la teoría electromagnética de Maxwell quedaba perfectamente incorporada dentro del marco de la relatividad especial, el siguiente paso era desarrollar una nueva teoría que incorporara la otra fuerza clásica (la gravedad), de manera que fuera una generalización de las ideas newtonianas, compatible con la Teoría Especial de la Relatividad. De hecho se realizaron varios esfuerzos en este sentido, pero no fue posible encontrar una teoría que fuera satisfactoria respecto a la invarianza de las leyes físicas ante algunas transformaciones de coordenadas admisibles. Finalmente, Einstein optó por un camino totalmente distinto y desarrolló una nueva idea del espacio-tiempo y la gravedad, llamada *Teoría General de la Relatividad*, motivado por dos consideraciones fundamentales. La primera es que todos los cuerpos son influenciados por la gravedad y de hecho caen precisamente de la misma manera en presencia de un campo gravitacional. Este hecho, conocido como *principio de equivalencia*, está implícito en la teoría newtoniana de la gravitación al establecer la proporcionalidad entre la fuerza de gravedad ejercida sobre un cuerpo y la masa inercial del mismo. Dado que el movimiento es independiente de la naturaleza de los cuerpos, las trayectorias de los objetos en caída libre definen un conjunto de curvas preferenciales en el espacio-tiempo, lo que sugiere la posibilidad de asociar propiedades de los campos gravitacionales con la estructura del espacio-tiempo mismo. En cuanto a la segunda idea, ésta consiste en suponer que la estructura del espacio-tiempo se ve influenciada por la presencia de materia y es llamada *principio de Mach*.

Con base en el principio de equivalencia, en la Teoría General de la Relatividad se considera imposible construir marcos de referencia inerciales en el sentido de la relatividad especial y medir la fuerza gravitacional. Esto es acompañado por la hipótesis adicional de que la métrica en espacio-tiempo no es plana como se supone en la relatividad especial, de modo que las trayectorias que siguen los cuerpos en caída libre en un campo gravitacional son simplemente las geodésicas de la métrica del espacio-tiempo. Así, los "observadores de fondo" (las geodésicas de la métrica del espacio-tiempo) automáticamente coinciden con lo que antes era percibido como un movimiento en un campo gravitacional, es decir, lo que se percibe como un campo gravitacional no es en realidad un campo, sino una desviación de la geometría plana del espacio-tiempo propio de la relatividad especial. Como resultado, no hay una manera de describir la fuerza gravitacional, en lugar de ello la gravedad se entiende como un aspecto de la estruc-

tura del espacio-tiempo. Aunque no tiene sentido hablar de una fuerza gravitacional absoluta, es posible hablar de la fuerza de gravedad relativa entre dos objetos cercanos y puede ser medida estudiando la aceleración relativa entre dos cuerpos en caída libre; esta aceleración está directamente relacionada con la curvatura del espacio-tiempo y la desviación de un campo vectorial a lo largo de una geodésica.

En la Teoría General de la Relatividad no hay una restricción *a priori* en la estructura de variedad del espacio-tiempo. A diferencia de la relatividad especial, no se afirma que el espacio-tiempo tiene la estructura de \mathbb{R}^4 con una métrica plana definida en él; ésta es una posibilidad que se manifiesta cuando no hay un campo gravitacional presente, pero no es la única. El marco de la relatividad general permite una métrica de Lorentz⁴, g_{ab} , no necesariamente plana, de hecho, exige que el espacio-tiempo sea curvo en todas las situaciones en las que, físicamente, un campo gravitacional esté presente. Dado que se permiten geodésicas curvas, resulta natural suponer que la estructura del espacio-tiempo es la de una variedad distinta a simplemente tomar \mathbb{R}^4 .

Por otro lado, en relatividad general las leyes de la física están gobernadas por dos principios básicos:

- 1) El principio general de covarianza, que establece que la métrica g_{ab} y todas las cantidades derivadas de ella son las únicas cantidades del espacio-tiempo que pueden aparecer en las ecuaciones de la física.
- 2) Todas las ecuaciones deben reducirse a las ecuaciones satisfechas por la Teoría Especial de la Relatividad en el caso en el que la métrica g_{ab} es plana.

Siguiendo el principio de Mach, en Teoría General de la Relatividad la geometría del espacio-tiempo está influenciada por la distribución de la materia en el universo. De este modo, la métrica del espacio-tiempo no sólo es una referencia en la que se establecen las leyes de la física, sino también una variable dinámica que responde al contenido de materia en el espacio-tiempo, como es el caso de la geometría del espacio-tiempo que describe a la gravedad.

Para deducir la ecuación que relaciona la geometría del espacio-tiempo y la distribución de materia, se hace una comparación entre la gravedad newtoniana, representada por un potencial ϕ , y la relatividad general. Después de considerar la relación de la aceleración con el tensor de curvatura y con el potencial, además de la ecuación de Poisson, que relaciona a ϕ con la densidad de masa y, por lo tanto, con el tensor de energía-momento, se llega a la siguiente ecuación:

$$R_{ab} - \frac{1}{2}Rg_{ab} = 8\pi T_{ab}, \quad (4.4)$$

llamada la ecuación de Einstein, quien la dio a conocer en 1915. Tomando la traza de (4.4), se encuentra que $R = -8\pi T$, de donde

$$R_{ab} = 8\pi(T_{ab} - \frac{1}{2}Rg_{ab}). \quad (4.5)$$

⁴Una métrica semiriemanniana con signos $- + + +$ en la diagonal cuando es aplicada a una base ortonormal.

En resumen, la Teoría General de la Relatividad establece que el espacio-tiempo es una variedad M en la cual se define una métrica de Lorentz g_{ab} . La curvatura de g_{ab} está relacionada con la distribución de la materia en el espacio-tiempo por medio de la ecuación de Einstein.

4.3 La Solución de Schwarzschild.

Actualmente las pruebas cuantitativas de las predicciones que son consecuencia de la Teoría General de la Relatividad se reducen a aquéllas provistas por los campos gravitacionales que ocurren dentro de nuestro Sistema Solar, donde pueden realizarse mediciones precisas. Tomando en cuenta este hecho, resulta de particular interés determinar las soluciones a la ecuación de Einstein que corresponden a un campo gravitacional estático, exterior de un cuerpo esféricamente simétrico, como es el campo producido por el Sol. Este problema fue resuelto por primera vez por Karl Schwarzschild, quien en 1916 dio una solución exacta a la ecuación de Einstein en el vacío.

En general, las predicciones de la Teoría General de la Relatividad tienden a coincidir con las de la mecánica newtoniana en los límites donde los campos son débiles, sin embargo, la solución de Schwarzschild predice pequeñas diferencias con la teoría de Newton para el movimiento de los planetas en el Sistema Solar (en concreto, la precesión del perihelio de Mercurio), además de la deflexión y el corrimiento al rojo de los rayos de luz, debidos a la presencia de cuerpos masivos. Estas predicciones han sido confirmadas con medidas precisas y, junto con las mediciones hechas a pulsares binarios, constituyen la esencia de la evidencia cuantitativa en favor de la Teoría General de la Relatividad.

Por otro lado, la solución Schwarzschild tiene más consecuencias que las predicciones anteriores. Se sabe que los cuerpos suficientemente masivos (estrellas con una masa no mayor a unas cuantas veces la masa solar) no pueden mantenerse y sufren un colapso gravitacional⁵; después de que el colapso de un cuerpo esférico ha ocurrido, la geometría del espacio-tiempo queda completamente descrita por la solución de Swarzschild, la cual plantea una singularidad en el espacio-tiempo asociada a la presencia de un hoyo negro.

Se dice que un espacio-tiempo es *estacionario* si existe un grupo monoparamétrico de isometrías, φ_t , cuyas órbitas son curvas tipo temporal, es decir, que para cada punto, la curva obtenida al variar el parámetro t es de tipo temporal. Otra manera de decir esto es que el espacio-tiempo tiene un campo de Killing ξ^a de tipo temporal.

Un espacio-tiempo es *estático* si es estacionario y existe una hipersuperficie (de tipo espacial), Σ , que es ortogonal a las órbitas de la isometría φ_t . Esto puede entenderse como que existe una foliación del espacio-tiempo tal que las hojas (de tipo espacial) son ortogonales a la curva integral de ξ^a (de hecho, las hojas son ortogonales a ξ^a mismo).

⁵Para una descripción más detallada de este proceso puede consultarse D'Eath [1]

⁶Ver apéndice B.

Como consecuencia del teorema de Frobenius, en un espacio-tiempo estático el campo de Killing ξ^a cumple con:

$$\xi_{[a} \nabla_b \xi_{c]} = 0. \quad (4.6)$$

Siguiendo las ideas planteadas en la demostración del teorema de rectificación de coordenadas ⁷, en un espacio-tiempo de cuatro dimensiones podemos dar coordenadas locales de la siguiente manera: si $\xi^a \neq 0$ en Σ , entonces por cada punto en Σ pasa una única curva integral de ξ^a , de modo que podemos considerar a estos puntos como condiciones iniciales del sistema $\frac{d\varphi_t}{dt} = \xi^a$. De este modo, damos coordenadas locales $(0, x^1, x^2, x^3)$ en Σ y a cada punto $x = \varphi_t(0, x^1, x^2, x^3)$ en una vecindad de Σ le asignamos las coordenadas (x^0, x^1, x^2, x^3) tales que $x^0 = t$. De este modo, φ_t mapea a la hoja Σ en otra hoja Σ_t que también es ortogonal a ξ^a .

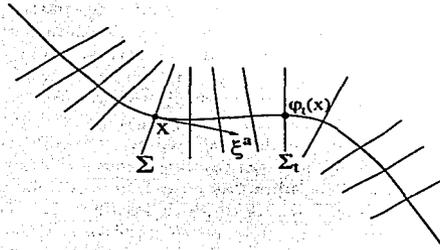


Figura 4.1: Curva temporal asociada a un campo de Killing ξ^a estático.

Analicemos ahora las restricciones que tenemos sobre la métrica, dadas las coordenadas anteriores. En un espacio-tiempo estacionario, como $\varphi_t(x)$ es un grupo de isometrías cuyo parámetro es precisamente la coordenada temporal, la métrica no depende del tiempo. Si además agregamos la condición de que el espacio-tiempo sea estático, esto nos lleva a que la expresión de la métrica no contenga términos cruzados que involucren al tiempo, pues ξ^a es ortogonal a Σ . Así, podemos escribir a la métrica de la siguiente manera:

$$ds^2 = -V^2 dt^2 + \sum_{\mu\nu} h_{\mu\nu}(x^1, x^2, x^3) dx^\mu dx^\nu, \quad (4.7)$$

donde $V^2 = V^2(x^1, x^2, x^3) = -\xi_a \xi^a$.

De la ecuación anterior, observamos que el difeomorfismo definido por $t \rightarrow -t$ (el mapeo que manda a cada punto en Σ_t a un punto con las mismas coordenadas espaciales en Σ_{-t}) es una isometría, pues el único elemento que involucra al tiempo aparece

⁷Ver capítulo 1, sección 1.3, teorema 1.22

elevado al cuadrado, por lo que no se ve afectado por un cambio de signo. Así, además de la simetría en las traslaciones temporales, $t \rightarrow t + cte$, que tienen los espacio-tiempos estacionarios, los espacio-tiempos estáticos son simétricos ante reflexiones temporales.

Por otro lado, se dice que un espacio-tiempo es *esféricamente simétrico* si su métrica se mantiene invariante ante rotaciones, es decir, las órbitas de un subgrupo del grupo de isometrías (isomorfo a $SO(3)$) son esferas de dimensión 2. En este caso, la métrica del espacio-tiempo induce una métrica en cada una de dichas esferas, la cual, debido a que existe una simetría rotacional, debe ser un múltiplo de la métrica en la esfera unitaria S^2 , por lo que el área A de cada esfera caracteriza a la métrica inducida en ella. Introduzcamos la función r , definida por:

$$r = \left(\frac{A}{4\pi} \right)^{\frac{1}{2}}. \quad (4.8)$$

Así, para las coordenadas angulares $(\phi$ y $\theta)$ de cada esfera, la métrica toma la forma

$$ds^2 = r^2 d\phi^2 + r^2 \sin^2 \phi d\theta^2. \quad (4.9)$$

En un espacio euclideo plano, tridimensional, r corresponde al valor del radio de la esfera ($A = 4\pi r^2$), sin embargo, en un espacio curvo no necesariamente sucede que r sea la distancia de una esfera a su centro, incluso no necesariamente se puede decir que la superficie tenga centro. De cualquier modo, nos referimos a r como la "coordenada radial" de la esfera.

Supongamos entonces un espacio-tiempo estático, esféricamente simétrico y cuyo campo de Killing estático, ξ^a , es único. Notemos que, si ξ^a varía sobre una esfera arbitraria, entonces la curva integral de ξ^a está sobre dicha superficie, así que podemos parametrizarla como una rotación, la cual es una isometría cuyo campo vectorial asociado es ortogonal a la foliación que tenemos como consecuencia de que el espacio-tiempo sea estático. Sin embargo, la unicidad de ξ^a implica que éste es el único campo vectorial para el que existe una foliación del espacio-tiempo con hojas espaciales ortogonales a él, por lo que ξ^a no debe variar sobre una esfera cualquiera, dadas las condiciones que hemos impuesto, es decir, ξ^a es invariante bajo isometrías rotacionales. De esto se sigue que la proyección de ξ^a sobre cada esfera es nula, es decir, ξ^a es ortogonal a las esferas que son órbitas de las isometrías rotacionales, por lo que deben estar completamente contenidas en las hojas Σ_t . Escogemos coordenadas en el espacio-tiempo del siguiente modo: en una hipersuperficie $\Sigma = \Sigma_{t=0}$, escogemos una esfera y damos coordenadas angulares (ϕ, θ) en ella. Como ya habíamos mencionado, la métrica en cada esfera es un múltiplo de la métrica en S^2 , así que "transportamos" las mismas coordenadas angulares a otras esferas de Σ por medio de geodésicas ortogonales a dichas esferas. Damos entonces coordenadas (r, ϕ, θ) en Σ , donde r especifica en qué esfera de Σ nos encontramos, siempre que $\nabla_a r \neq 0$. Finalmente, de acuerdo a lo que habíamos planteado al principio de esta sección, tomamos (t, r, ϕ, θ) como coordenadas del espacio-tiempo. Una vez hecho esto, la métrica de un espacio-tiempo estático, esféricamente simétrico

tiene la siguiente forma:

$$ds^2 = -f(r)dt^2 + h(r)dr^2 + r^2d\phi^2 + r^2\sin^2\phi d\theta^2. \quad (4.10)$$

Nos resta encontrar las funciones $f(r)$ y $h(r)$. Para hacer esto, podemos usar la ecuación de Einstein (4.4) en el vacío ($T_{ab} = 0 \Rightarrow R_{ab} = 0$), lo que nos remite a buscar una expresión del tensor de curvatura por medio del método de bases ortonormales ⁸.

Tomando en cuenta (4.10), introducimos la siguiente base ortonormal:

$$\begin{aligned} (e_0)_a &= f^{\frac{1}{2}}(dt)_a \\ (e_1)_a &= h^{\frac{1}{2}}(dr)_a \\ (e_2)_a &= r(d\phi)_a \\ (e_3)_a &= r\sin\phi(d\theta)_a. \end{aligned} \quad (4.11)$$

De acuerdo con (A.27), para conocer las 1-formas de conexión propias de esta base, necesitamos aplicar el operador ∂_a a los elementos de la base y tomar la parte antisimétrica:

$$\begin{aligned} \partial_a(e_0)_b &= f^{\frac{1}{2}}(dt)_a(dt)_b + \frac{1}{2}f^{-\frac{1}{2}}f'(dr)_a(dt)_b \Rightarrow \partial_{[a}(e_0)_{b]} = \frac{1}{2}f^{-\frac{1}{2}}f'(dr)_{[a}(dt)_{b]} \\ \partial_a(e_1)_b &= \frac{1}{2}h^{-\frac{1}{2}}h'(dr)_a(dr)_b \Rightarrow \partial_{[a}(e_1)_{b]} = 0 \\ \partial_a(e_2)_b &= (dr)_a(d\phi)_b + r(d\phi)_a(d\phi)_b \Rightarrow \partial_{[a}(e_2)_{b]} = (dr)_{[a}(d\phi)_{b]} \\ \partial_a(e_3)_b &= \sin\phi(dr)_a(d\theta)_b + r\cos\phi(d\phi)_a(d\theta)_b + r\sin\phi(d\theta)_a(d\theta)_b \\ &\Rightarrow \partial_{[a}(e_3)_{b]} = \sin\phi(dr)_{[a}(d\theta)_{b]} + r\cos\phi(d\phi)_{[a}(d\theta)_{b]}. \end{aligned} \quad (4.12)$$

Por (A.27), tomando en cuenta (A.7) y (A.9):

$$\begin{aligned} \frac{1}{2}f^{-\frac{1}{2}}f'(dr)_{[a}(dt)_{b]} &= f^{\frac{1}{2}}(dt)_{[a}w_{b]00} + h^{\frac{1}{2}}(dr)_{[a}w_{b]01} + r(d\phi)_{[a}w_{b]02} + r\sin\phi(d\theta)_{[a}w_{b]03} \\ &= h^{\frac{1}{2}}(dr)_{[a}w_{b]01} + r(d\phi)_{[a}w_{b]02} + r\sin\phi(d\theta)_{[a}w_{b]03}. \end{aligned} \quad (4.13)$$

$$\begin{aligned} 0 &= -f^{\frac{1}{2}}(dt)_{[a}w_{b]10} + h^{\frac{1}{2}}(dr)_{[a}w_{b]11} + r(d\phi)_{[a}w_{b]12} + r\sin\phi(d\theta)_{[a}w_{b]13} \\ &= f^{\frac{1}{2}}(dt)_{[a}w_{b]01} + r(d\phi)_{[a}w_{b]12} + r\sin\phi(d\theta)_{[a}w_{b]13}. \end{aligned} \quad (4.14)$$

$$\begin{aligned} (dr)_{[a}(d\phi)_{b]} &= -f^{\frac{1}{2}}(dt)_{[a}w_{b]20} + h^{\frac{1}{2}}(dr)_{[a}w_{b]21} + r(d\phi)_{[a}w_{b]22} + r\sin\phi(d\theta)_{[a}w_{b]23} \\ &= f^{\frac{1}{2}}(dt)_{[a}w_{b]02} - h^{\frac{1}{2}}(dr)_{[a}w_{b]12} + r\sin\phi(d\theta)_{[a}w_{b]23}. \end{aligned} \quad (4.15)$$

$$\begin{aligned} \sin\phi(dr)_{[a}(d\theta)_{b]} + r\cos\phi(d\phi)_{[a}(d\theta)_{b]} &= -f^{\frac{1}{2}}(dt)_{[a}w_{b]30} + h^{\frac{1}{2}}(dr)_{[a}w_{b]31} + r(d\phi)_{[a}w_{b]32} \\ &\quad + r\sin\phi(d\theta)_{[a}w_{b]33} \\ &= f^{\frac{1}{2}}(dt)_{[a}w_{b]03} - h^{\frac{1}{2}}(dr)_{[a}w_{b]13} - r(d\phi)_{[a}w_{b]23}. \end{aligned}$$

⁸Ver apéndice A.

(4.16)

Tomando en cuenta que $w_{a\mu\nu} = (e_\mu)^b \nabla_a (e_\nu)_b$, podemos pensar que la 1-forma de conexión $w_{a\mu\nu}$ nos da el valor del cambio de $(e_\nu)_b$ en la dirección de $(e_\mu)^b$. En este sentido, dado el planteamiento de la base en (4.11) y la forma de la métrica en (4.10), el cambio en los vectores angulares $(e_2)^a$ y $(e_3)^a$ se realiza dentro de una hoja Σ_{t_0} fija, por lo que no tiene componente en la dirección temporal $(e_0)^a$ y tiene sentido considerar:

$$w_{b02} = w_{b03} = 0, \quad (4.17)$$

de donde (4.13) toma la forma:

$$\frac{1}{2} f^{-\frac{1}{2}} f' (dr)_{[a} (dt)_{b]} = h^{\frac{1}{2}} (dr)_{[a} w_{b]01},$$

lo que nos lleva a proponer

$$w_{b01} = \frac{1}{2} \frac{f'}{(fh)^{\frac{1}{2}}} (dt)_b + \alpha_1 (dr)_b. \quad (4.18)$$

Sustituyendo (4.18) en (4.14):

$$0 = \alpha_1 f^{\frac{1}{2}} (dt)_{[a} (dr)_{b]} + r (d\phi)_{[a} w_{b]12} + r \sin \phi (d\theta)_{[a} w_{b]13},$$

por lo que $\alpha_1 = 0$. En el caso de w_{b12} y w_{b13} , sabemos por (4.11) que la expresión de los vectores angulares involucra a la coordenada radial, al igual que cierta relación entre las coordenadas angulares mismas, así que tomamos:

$$w_{b12} = \alpha_2 (d\phi)_b + \alpha_3 (d\theta)_b, \quad (4.19)$$

$$w_{b13} = \alpha_4 (d\theta)_b + \frac{\alpha_5}{\sin \phi} (d\phi)_b. \quad (4.20)$$

Haciendo la sustitución de $w_{b02} = 0$ y de w_{b12} en (4.15):

$$\begin{aligned} (dr)_{[a} (d\phi)_{b]} &= -\alpha_2 h^{\frac{1}{2}} (dr)_{[a} (d\phi)_{b]} - \alpha_3 h^{\frac{1}{2}} (dr)_{[a} (d\theta)_{b]} + r \sin \phi (d\theta)_{[a} w_{b]23} \\ \Rightarrow \alpha_2 &= -h^{-\frac{1}{2}} \text{ y} \\ w_{b23} &= \frac{h^{\frac{1}{2}}}{r \sin \phi} \alpha_3 (dr)_b + \alpha_5 (d\theta)_b. \end{aligned}$$

Por último, al sustituir en (4.16), llegamos a:

$$\alpha_3 = 0, \quad \alpha_4 = -h^{-\frac{1}{2}} \sin \phi, \quad \alpha_5 = -\cos \phi.$$

Así, las expresiones de las 1-formas de conexión son:

$$w_{b02} = w_{b03} = 0$$

$$\begin{aligned}
 w_{b01} &= \frac{1}{2} \frac{f'}{(fh)^{\frac{1}{2}}} (dt)_b \\
 w_{b12} &= -h^{-\frac{1}{2}} (d\phi)_b \\
 w_{b13} &= -h^{-\frac{1}{2}} \sin \phi (d\theta)_b \\
 w_{b23} &= -\cos \phi (d\theta)_b.
 \end{aligned} \tag{4.21}$$

De (A.13), recordando que para el tensor de Riemann se cumple que $R_{abcd} = -R_{abdc}$, tenemos que:

$$\begin{aligned}
 R_{ab01} &= -R_{ab10} = (\nabla_a w_{b01} - \nabla_b w_{a01}) - (w_{a00} w_{b01} - w_{b00} w_{a01}) \\
 &+ (w_{a10} w_{b11} - w_{b10} w_{a11}) + (w_{a20} w_{b21} - w_{b20} w_{a21}) + (w_{a30} w_{b31} - w_{b30} w_{a31})
 \end{aligned} \tag{4.22}$$

Tomando en cuenta (A.9), (4.17), la ortonormalidad de la base y (A.23):

$$\begin{aligned}
 R_{ab01} &= -R_{ab10} = (\nabla_a w_{b01} - \nabla_b w_{a01}) \\
 &= \nabla_{[a} w_{b]01} = \partial_{[a} w_{b]01} \\
 &= \left(\frac{d}{dr} f' (fh)^{-\frac{1}{2}} \right) (dr)_{[a} (dt)_{b]}.
 \end{aligned} \tag{4.23}$$

De igual modo se calcula el resto de las componentes Riemann ($R_{ab\mu\nu} = 0$, $\mu = 0, 1, 2, 3$):

$$\begin{aligned}
 R_{ab02} &= -R_{ab20} = f^{-\frac{1}{2}} h^{-1} f' (d\phi) [a (dt)_{b}], \\
 R_{ab03} &= -R_{ab30} = f^{-\frac{1}{2}} h^{-1} f' \sin \phi (d\theta) [a (dt)_{b}], \\
 R_{ab12} &= -R_{ab21} = h^{-\frac{3}{2}} h' (dr) [a (d\phi)_{b}], \\
 R_{ab13} &= -R_{ab31} = \sin \phi h^{-\frac{3}{2}} h' (dr) [a (d\theta)_{b}], \\
 R_{ab23} &= -R_{ab32} = 2(1 - h^{-1}) \sin \phi (d\phi) [a (d\theta)_{b}].
 \end{aligned} \tag{4.24}$$

Las componentes del tensor de Ricci se obtienen entonces usando (A.17), una vez que se han hecho las proyecciones de (4.24) en las direcciones correspondientes. Igualando los coeficientes de Ricci a cero, encontramos la ecuación de Einstein en el vacío, para un espacio-tiempo estático y esféricamente simétrico:

$$\begin{aligned}
 0 = R_{00} &= R_{010}^1 + R_{020}^2 + R_{030}^3 \\
 &= \frac{1}{2} (fh)^{-\frac{1}{2}} \left(\frac{d}{dr} f' (fh)^{-\frac{1}{2}} \right) + (rfh)^{-1} f',
 \end{aligned} \tag{4.25}$$

$$0 = R_{11} = -\frac{1}{2} (fh)^{-\frac{1}{2}} \left(\frac{d}{dr} f' (fh)^{-\frac{1}{2}} \right) + (rh^2)^{-1} h', \tag{4.26}$$

$$0 = R_{22} = -\frac{1}{2} (rfh)^{-1} f' + \frac{1}{2} (rh^2)^{-1} h' + r^{-2} (1 - h^{-1}). \tag{4.27}$$

Además, $R_{33} = R_{22}$ y las componentes fuera de la diagonal se anulan. Sumando (4.25) y (4.26), vemos que:

$$\begin{aligned} \frac{f'}{f} + \frac{h'}{h} &= 0 \\ \Rightarrow f &= Kh^{-1}, \end{aligned} \quad (4.28)$$

donde K es una constante. Haciendo un reescalamiento en la coordenada temporal, $t \rightarrow K^{\frac{1}{2}}t$, podemos tomar $K = 1$. Sustituyendo (4.28) en (4.27), llegamos a:

$$\begin{aligned} -f' + \frac{1-f}{r} &= 0 \\ \Rightarrow rf' + f &= 1 \\ \Rightarrow \frac{d}{dr}(rf) &= 1 = \frac{dr}{dr} \\ \Rightarrow \frac{d}{dr}(rf - r) &= 0 \\ \Rightarrow rf - r &= C \\ \Rightarrow f &= 1 + \frac{C}{r}, \end{aligned} \quad (4.29)$$

donde C es una constante. La expresión de $f = \frac{1}{h} = 1 + \frac{C}{r}$ resuelve las ecuaciones (4.25)-(4.27), por lo que hemos encontrado la solución general a la ecuación de Einstein en el vacío para un espacio-tiempo estático y esféricamente simétrico. Sustituyendo (4.29) en (4.10), tenemos una expresión precisa para la métrica, dada por:

$$ds^2 = -\left(1 + \frac{C}{r}\right) dt^2 + \left(1 + \frac{C}{r}\right)^{-1} dr^2 + r^2 d\phi^2 + r^2 \sin^2 \phi d\theta^2, \quad (4.30)$$

la cual se conoce como la métrica de Schwarzschild. Notemos que, si $r \rightarrow \infty$, las componentes de la métrica tienden a ser las de la métrica de Minkowsky. Esto nos permite interpretar a la métrica de Schwarzschild como aquella asociada al exterior del campo gravitacional generado por un cuerpo aislado. El significado de la constante C se obtiene por medio de una comparación entre el comportamiento de una carga de prueba cuando el campo gravitacional es débil ($r \rightarrow \infty$) y el de una carga de prueba en un campo newtoniano el resultado de esta comparación es que la carga de prueba con parámetro C en la solución de Schwarzschild se comporta como una carga de prueba en el potencial newtoniano con masa $M = \frac{-C}{2} \Rightarrow C = -2M$. Escribimos entonces la métrica en su forma final:

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\phi^2 + r^2 \sin^2 \phi d\theta^2. \quad (4.31)$$

De esta expresión podemos observar que la solución de Schwarzschild tiene dos singularidades: en $r = 2M$ y en $r = 0$. La primera de ellas resulta ser una singularidad debida a las coordenadas que tomamos, es decir, que para $r = 2M$, la hipótesis de que $\nabla_a r \neq 0$ deja de cumplirse. En este caso, el problema se resuelve tomando un

cambio de coordenadas, en concreto, se toman las llamadas *coordenadas de Eddington Finkelstein* para la geometría de Schwarzschild, (v, r, ϕ, θ) , donde

$$v = t + r + 2M \ln(r - 2M). \quad (4.32)$$

En tales coordenadas la métrica de Schwarzschild toma la forma:

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dv^2 + 2dvdr + r^2 d\phi^2 + r^2 \sin^2 \phi d\theta^2. \quad (4.33)$$

Por otro lado, cuando $r = 0$, tenemos una singularidad real relacionada con la formación de un hoyo negro; una vez que se tiene esto en mente, el radio $r = 2M$ cobra un nuevo significado ya que este valor de r asigna una superficie nula conocida como el *horizonte del hoyo negro* que corresponde a aquélla a partir de la cual tanto los rayos de luz como las partículas quedan atrapadas en el hoyo negro.

4.4 Mecánica de Hoyos Negros.

En términos físicos, un hoyo negro es una región del espacio-tiempo donde la gravedad es tan fuerte que nada, ni siquiera la luz, puede escapar y a su frontera se le conoce como horizonte. Como ya habíamos mencionado, un cuerpo suficientemente masivo tiende a sufrir un colapso gravitacional, produciendo una singularidad en el espacio-tiempo que está asociada a un hoyo negro como producto final del colapso⁹.

La teoría de los hoyos negros en relatividad general clásica¹⁰ se desarrolló a finales de la década de los 60 y principios de los 70, obteniéndose diversos resultados notables. Uno de ellos, probado por Penrose, se refiere a la noción que de que la energía de un hoyo negro no necesariamente se pierde si éste tiene momento angular. La idea de Penrose es inyectar una partícula tangente al horizonte de un hoyo negro de manera que ésta se parte en dos nuevas partículas, una de las cuales cae en el hoyo, mientras que la otra escapa con una energía mayor a la que tenía la partícula original. El efecto neto es una extracción de la energía rotacional del hoyo, junto con una correspondiente disminución en su masa y momento angular.

Más tarde, Demetrios Christodoulou determinó que la eficiencia del proceso descrito por Penrose está limitada por una cantidad que no puede decrecer y que está en función de la masa y del momento angular; poco después Stephen Hawking mostró que la cantidad que no puede decrecer en la interacción de un hoyo negro con su exterior es precisamente el área del horizonte. Los resultados anteriores tienen una fuerte similitud con la segunda ley de la termodinámica, la cual establece que en cualquier proceso físico el total de la entropía de la materia en el universo no puede decrecer. Tomando en

⁹A la idea de que una singularidad en el espacio-tiempo siempre deba aparecer como un hoyo negro se conoce como la *conjetura de la censura cósmica*, que "prohíbe" que aparezcan singularidades "desnudas", "vistiéndolas" con un hoyo negro y su respectivo horizonte.

¹⁰Es decir, no cuántica.

cuenta esta similitud, se desarrollaron leyes de la mecánica de hoyos negros análogas a cada una de las leyes de la termodinámica en las que la masa del hoyo negro M , la gravedad superficial κ en su horizonte y el área del horizonte A desempeñan papeles equivalentes a los de la energía E , la temperatura T y la entropía S de un sistema termodinámico.

Inicialmente se pensó que la fuerte similitud entre las leyes de la termodinámica y las leyes de la mecánica de hoyos negros eran de carácter puramente formal, puesto que de manera clásica, por la definición misma de los hoyos negros, resultaba contradictorio asociarles cantidades termodinámicas. Sin embargo, utilizando mecánica cuántica, Hawking demostró que las fluctuaciones del vacío permiten afirmar que un hoyo negro emite energía independientemente de que tenga momento angular o no y por lo tanto es posible asignarle una temperatura.

Ciertamente, la deducción las leyes de la mecánica de hoyos negros tienen un carácter más geométrico que físico y se demuestran a partir de la suposición de diversas simetrías en el espacio-tiempo. Por lo general, se considera un hoyo negro estacionario (el caso estático es aún más sencillo) y a partir de las propiedades de campos de Killing, se definen determinadas cantidades a las que después se les asocia un significado físico preciso. Así, la función κ puede definirse en términos de un campo de Killing ξ^a y del operador de derivada covariante compatible con la métrica:

$$\xi^a \nabla_a \xi^b = \kappa \xi^b. \quad (4.34)$$

Mediante distintas manipulaciones, es posible encontrar una expresión de κ en la que es fácil interpretarla como la gravedad superficial. Una vez hecho esto, haciendo uso de diversas propiedades, se demuestra la *ley cero de la mecánica de hoyos negros*, que establece que κ es constante en el horizonte de un hoyo negro estacionario, del mismo modo que la ley cero de la termodinámica dice que la temperatura es uniforme en un sistema en equilibrio térmico.

De igual manera, se encuentra que la masa de un hoyo negro está dada en función del tensor de energía momento, el área del horizonte del hoyo negro y el momento angular del mismo. A partir de esto se puede deducir una forma diferencial de la masa que involucra a los dos términos anteriores y que constituye la *primera ley de la mecánica de hoyos negros*, análoga a la primera ley de la termodinámica (ver tabla 4.1).

En 1993 Robert Wald realizó una nueva deducción de la primera ley de la mecánica de hoyos negros que ilustra el empleo de los conceptos revisados a lo largo de este trabajo.

La idea es utilizar una formulación lagrangiana de la Teoría General de la Relatividad en la que se estudia a un hoyo negro con perturbaciones cercanas a un estado estacionario. Tomando un campo de Killing ξ^a normal al horizonte Σ y suponiendo invarianza del lagrangiano L ante isomorfismos (en particular L es invariante ante la acción del grupo monoparamétrico asociado a ξ^a), se puede hacer uso del teorema de

LEYES	CONTEXTO	
	Termodinámica	Hoyos Negros
Cero	T es constante en todo un cuerpo que se encuentra en equilibrio térmico	κ es constante sobre el horizonte de un hoyo negro estacionario
Primera	$dE = TdS +$ términos de trabajo	$dM = \frac{1}{8\pi} \kappa dA +$ términos de momento angular
Segunda	$dS \geq 0$ en cualquier proceso	$dA \geq 0$ en cualquier proceso

Tabla 4.1: Comparación entre las leyes de la mecánica de hoyos negros y la termodinámica

Noether¹¹ y encontrar entonces una cantidad conservada Q , también llamada *carga de Noether*. Haciendo uso de la simetría estacionaria, de que ξ^a se anula en Σ y de la relación entre ξ^a y la gravedad superficial κ dada por (4.34), se encuentra la siguiente expresión de la primera ley de la mecánica de los hoyos negros:

$$\frac{\kappa}{2\pi} \delta A = \delta \epsilon - \Omega_H^{(\mu)} \delta J_{(\mu)}, \quad (4.35)$$

donde

$$\delta A = 2\pi \int_{\Sigma} \tilde{Q}; \quad (4.36)$$

y $\delta \epsilon$ y $\Omega_H^{(\mu)} \delta J_{(\mu)}$ son términos asociados a la energía y el momento angular generalizados, respectivamente.

Esta deducción tiene la ventaja de dar una expresión precisa del área del horizonte del hoyo negro en términos de la carga de Noether y, más aún, la carga de Noether está dada como una cantidad geométrica intrínseca al hoyo negro, independiente de los campos que aparecen en el lagrangiano.

¹¹Ver sección 3.3

Apéndice A

El Método de Bases Ortonormales para Calcular la Curvatura

Para resolver la ecuación de Einstein es necesario conocer la expresión del tensor de curvatura en términos de la métrica del espacio-tiempo. Muchas veces es conveniente calcular la curvatura por medio de una base ortonormal ya que este método no sólo puede resultar menos laborioso que otros, sino que incluso permite encontrar el tensor de Riemann aunque no se tenga una expresión precisa de la métrica y/o su operador de derivada covariante asociado, además de que suele requerir cálculos que son relativamente sencillos.

Consideremos el espacio-tiempo de dimensión n y tomemos entonces una base ortonormal de campos vectoriales suaves $\{(e_\mu)^a\}$,

$$(e_\mu)^a (e_\nu)_a = \eta_{\mu\nu}, \quad (\text{A.1})$$

donde

$$\eta_{\mu\nu} = \begin{cases} 0 & \text{si } \mu \neq \nu \\ \pm 1 & \text{si } \mu = \nu \end{cases},$$

$\mu, \nu = 1, \dots, n$. Si consideramos a $\eta_{\mu\nu}$ como las entradas de una matriz de $n \times n$, entonces la inversa de esta matriz es ella misma, por lo que $\eta^{\mu\nu} = \eta_{\mu\nu}$. Los índices griegos indican qué vector de la base estamos tomando y los índices latinos indican el tipo de tensor, de acuerdo con la convención que estamos siguiendo¹. En el caso $n = 4$ y a veces también en contextos más generales, a la base $\{(e_\mu)^a\}$ se le conoce como *tétrada* y al método que estamos desarrollando se le llama *el método de tétradas*.

Notemos que:

$$\sum_{\mu, \nu} \eta^{\mu\nu} \eta_{\sigma\nu} (e_\mu)^a = (e_\sigma)^a = \delta_b^a (e_\sigma)^b,$$

donde δ_b^a puede verse como el mapeo identidad entre vectores. Además, por la ortogo-

¹ Ver Wald [10], capítulo 2.

100 El Método de Bases Ortonormales para Calcular la Curvatura

nalidad, tenemos $(e_\nu)_b(e_\sigma)^b = \eta_{\sigma\nu}$, de donde

$$\begin{aligned} & \sum_{\mu,\nu} \eta^{\mu\nu} (e_\mu)^a (e_\nu)_b (e_\sigma)^b = \delta_b^a (e_\sigma)^b \\ \Rightarrow & \left(\sum_{\mu,\nu} \eta^{\mu\nu} (e_\mu)^a (e_\nu)_b \right) (e_\sigma)^b = \delta_b^a (e_\sigma)^b \\ \Rightarrow & \sum_{\mu,\nu} \eta^{\mu\nu} (e_\mu)^a (e_\nu)_b = \delta_b^a. \end{aligned} \quad (\text{A.2})$$

El método de tétradas, se basa en considerar una base ortonormal y, a partir de esto, buscar expresiones equivalentes a los siguientes tres puntos claves en la relación entre la curvatura y la métrica, en términos de la base:

i) Tomamos un operador de derivada covariante ∇_a compatible con la métrica, es decir,

$$\nabla_a g_{bc} = 0. \quad (\text{A.3})$$

ii) Como ∇_a es un operador de derivada covariante (independientemente de que sea o no compatible con la métrica), cumple con cinco propiedades, a saber: es un operador lineal, que sigue la regla de Leibniz, que conmuta con las contracciones, que es consistente con la noción de tangencia de los campos vectoriales, es decir, que $\forall f \in \mathcal{F}$ y $\forall v^a$ campo vectorial,

$$v(f) = v^a \nabla_a f. \quad (\text{A.4})$$

Además, ∇_a es libre de torsión, lo que quiere decir que $\forall f \in \mathcal{F}$,

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f. \quad (\text{A.5})$$

Justamente nos interesa expresar esta última condición por medio de la base ortonormal.

iii) El tensor de Riemann está relacionado con ∇_a por la ecuación

$$\nabla_a \nabla_b w_c - \nabla_b \nabla_a w_c = R_{abc}^d w_d, \quad (\text{A.6})$$

donde w_d es un vector dual.

Definimos las *1-formas de conexión*, $w_{a\mu\nu}$, como:

$$w_{a\mu\nu} = (e_\mu)^b \nabla_a (e_\nu)_b. \quad (\text{A.7})$$

Las componentes $w_{\lambda\mu\nu}$ de $w_{a\mu\nu}$ se conocen como los *coeficientes de rotación de Ricci* y están dados por la proyección de $w_{a\mu\nu}$ sobre $(e_\lambda)^a$:

$$w_{\lambda\mu\nu} = (e_\lambda)^a (e_\mu)^b \nabla_a (e_\nu)_b. \quad (\text{A.8})$$

Mediante la regla de Leibniz y la ortonormalidad de la base $\{(e_\mu)^a\}$, notemos que:

$$\nabla_a (e_\mu)^b (e_\nu)_b = \nabla_a \eta_{\mu\nu} = 0$$

$$\begin{aligned} \Leftrightarrow & (e_\mu)^b \nabla_a (e_\nu)_b + (e_\nu)_b \nabla_a (e_\mu)^b = 0 \\ \Leftrightarrow & w_{a\mu\nu} = -(e_\nu)_b \nabla_a (e_\mu)^b = -(e_\nu)_b \nabla_a g^{bc} (e_\mu)_c = -(e_\nu)_b g^{bc} \nabla_a (e_\mu)_c - (e_\nu)_b (e_\mu)_c \nabla_a g^{bc} \\ \Leftrightarrow & w_{a\mu\nu} + (e_\nu)^c \nabla_a (e_\mu)_c = w_{a\mu\nu} + w_{a\nu\mu} = -(e_\nu)_b (e_\mu)_c \nabla_a g^{bc}. \end{aligned}$$

Por lo tanto, la condición

$$w_{a\mu\nu} = -w_{a\nu\mu} \quad (\text{A.9})$$

es equivalente a la condición $\nabla_a g^{bc} = \nabla_a g_{bc} = 0$. De este modo hemos encontrado una manera de expresar el hecho de que el operador ∇_a es compatible con la métrica (el punto (i)) en términos de la base ortonormal.

Por otro lado, las componentes del tensor de Riemann en términos de $\{(e_\mu)^a\}$ se obtienen aplicando R_{abc}^d a $(e_\nu)_d$ y proyectando sobre los vectores de la base:

$$\begin{aligned} R_{\rho\sigma\mu\nu} &= (e_\rho)^a (e_\sigma)^b (e_\mu)^c R_{abc}^d (e_\nu)_d \\ &= (e_\rho)^a (e_\sigma)^b (e_\mu)^c (\nabla_a \nabla_b - \nabla_b \nabla_a) (e_\nu)_c. \end{aligned} \quad (\text{A.10})$$

Usando la regla de Leibniz, tenemos que

$$\begin{aligned} \nabla_a \{(e_\mu)^c \nabla_b (e_\nu)_c\} &= (e_\mu)^c \nabla_a \nabla_b (e_\nu)_c + [\nabla_a (e_\mu)^c][\nabla_b (e_\nu)_c] \\ \Rightarrow (e_\mu)^c \nabla_a \nabla_b (e_\nu)_c &= \nabla_a \{(e_\mu)^c \nabla_b (e_\nu)_c\} - [\nabla_a (e_\mu)^c][\nabla_b (e_\nu)_c]. \end{aligned} \quad (\text{A.11})$$

Además, por (A.2),

$$\begin{aligned} [\nabla_a (e_\mu)^c][\nabla_b (e_\nu)_c] &= [\nabla_a (e_\mu)^f] \delta_f^c [\nabla_b (e_\nu)_c] \\ &= \sum_{\alpha, \beta} \eta^{\alpha\beta} [\nabla_a (e_\mu)^f] (e_\alpha)^c (e_\beta)_f [\nabla_b (e_\nu)_c] \\ &= \sum_{\alpha, \beta} \eta^{\alpha\beta} w_{a\beta\mu} w_{b\alpha\nu}. \end{aligned} \quad (\text{A.12})$$

Haciendo un desarrollo similar para $(e_\mu)^c \nabla_b \nabla_a (e_\nu)_c$ y sustituyendo en (A.10) llegamos a:

$$R_{\rho\sigma\mu\nu} = (e_\rho)^a (e_\sigma)^b \{ \nabla_a w_{b\mu\nu} - \nabla_b w_{a\mu\nu} - \sum_{\alpha, \beta} \eta^{\alpha\beta} (w_{a\beta\mu} w_{b\alpha\nu} - w_{b\beta\mu} w_{a\alpha\nu}) \}. \quad (\text{A.13})$$

Ahora bien, tomando en cuenta la regla de Leibniz y la ortonormalidad, observamos que:

$$\begin{aligned} \nabla_a (e_\sigma)^b w_{b\mu\nu} &= (e_\sigma)^b \nabla_a w_{b\mu\nu} + [\nabla_a (e_\sigma)^b] w_{b\mu\nu} \\ \Rightarrow (e_\sigma)^b \nabla_a w_{b\mu\nu} &= \nabla_a (e_\sigma)^b w_{b\mu\nu} - [\nabla_a (e_\sigma)^b] w_{b\mu\nu} = \nabla_a w_{\sigma\mu\nu} - [\nabla_a (e_\sigma)^b] w_{b\mu\nu} \\ \Rightarrow (e_\rho)^a (e_\sigma)^b \nabla_a w_{b\mu\nu} &= (e_\rho)^a \nabla_a w_{\sigma\mu\nu} - (e_\rho)^a [\nabla_a (e_\sigma)^b] w_{b\mu\nu} \\ &= (e_\rho)^a \nabla_a w_{\sigma\mu\nu} - (e_\rho)^a [\nabla_a (e_\sigma)^f] \delta_f^b w_{b\mu\nu} \\ &= (e_\rho)^a \nabla_a w_{\sigma\mu\nu} - \sum_{\alpha, \beta} \eta^{\alpha\beta} (e_\rho)^a [\nabla_a (e_\sigma)^f] (e_\alpha)^b (e_\beta)_f w_{b\mu\nu} \end{aligned}$$

102 El Método de Bases Ortonormales para Calcular la Curvatura

$$\begin{aligned}
 &= (e_\rho)^a \nabla_a w_{\sigma\mu\nu} - \sum_{\alpha,\beta} \eta^{\alpha\beta} (e_\rho)^a (e_\alpha)^b w_{a\beta\sigma} w_{b\mu\nu} \\
 &= (e_\rho)^a \nabla_a w_{\sigma\mu\nu} - \sum_{\alpha,\beta} \eta^{\alpha\beta} w_{\rho\beta\sigma} w_{\alpha\mu\nu}. \tag{A.14}
 \end{aligned}$$

Procediendo de igual manera con $(e_\rho)^a (e_\sigma)^b \nabla_b w_{a\mu\nu}$, sustituyendo en (A.13) y haciendo el resto de las proyecciones, obtenemos la expresión del tensor de Riemann en términos de los coeficientes de rotación de Ricci:

$$\begin{aligned}
 R_{\rho\sigma\mu\nu} &= (e_\rho)^a \nabla_a w_{\sigma\mu\nu} - (e_\sigma)^a \nabla_a w_{\rho\mu\nu} \\
 &\quad - \sum_{\alpha,\beta} \eta^{\alpha\beta} \{ w_{\rho\beta\mu} w_{\sigma\alpha\nu} - w_{\sigma\beta\mu} w_{\rho\alpha\nu} + w_{\rho\beta\sigma} w_{\alpha\mu\nu} - w_{\sigma\beta\sigma} w_{\alpha\mu\nu} \}. \tag{A.15}
 \end{aligned}$$

Por (A.4), para cualesquiera dos operadores de derivada covariante ∇_a y $\tilde{\nabla}_a$, se cumple:

$$\begin{aligned}
 v^a \nabla_a f &= v(f) = v^a \tilde{\nabla}_a f \\
 \Rightarrow \nabla_a f &= \tilde{\nabla}_a f, \tag{A.16}
 \end{aligned}$$

es decir, los operadores covariantes coinciden cuando tenemos campos escalares. Dado que $w_{\sigma\mu\nu}$ y $w_{\rho\mu\nu}$ son escalares, podemos sustituir el operador de derivada covariante por una derivada ordinaria ∂_a , lo que nos permite calcular el tensor de curvatura por medio de (A.15). Las componentes del tensor de Ricci se calculan entonces usando la fórmula:

$$R_{\rho\mu} = \sum_{\sigma,\nu} \eta^{\sigma\nu} R_{\rho\sigma\mu\nu}. \tag{A.17}$$

La ecuación (A.15) nos permite expresar el punto (iii). Lo que nos resta ahora es buscar una condición equivalente al hecho de que ∇_a es libre de torsión.

Tomemos un vector dual w_b y consideremos la diferencia $\tilde{\nabla}_a(f w_b) - \nabla_a(f w_b)$. Por la regla de Leibniz y (A.16), tenemos:

$$\begin{aligned}
 \tilde{\nabla}_a(f w_b) - \nabla_a(f w_b) &= w_b \tilde{\nabla}_a f + f \tilde{\nabla}_a w_b - w_b \nabla_a f - f \nabla_a w_b \\
 &= -f \{ \tilde{\nabla}_a w_b - \nabla_a w_b \}. \tag{A.18}
 \end{aligned}$$

Ahora consideremos un vector dual w'_b cuyo valor en un punto dado x_0 coincide con el valor de w_b . Usando una versión del lema 1.15 para vectores duales, como $w_b|_{x_0} = w'_b|_{x_0}$, sabemos que existen funciones suaves $f_{(\alpha)}$ que se anulan en x_0 y campos vectoriales duales suaves $u_b^{(\alpha)}$, tales que:

$$w'_b - w_b = \sum_{\alpha=1}^n f_{(\alpha)} u_b^{(\alpha)}.$$

Entonces, por (A.18), tenemos que en el punto x_0 se cumple:

$$\begin{aligned}\bar{\nabla}_a(w'_b - w_b) - \nabla_a(w'_b - w_b) &= \sum_{\alpha=1}^n \bar{\nabla}_a(f_{(\alpha)} u_b^{(\alpha)}) - \nabla_a(f_{(\alpha)} u_b^{(\alpha)}) \\ &= \sum_{\alpha=1}^n f_{(\alpha)} \{ \bar{\nabla}_a u_b^{(\alpha)} - \nabla_a U_b^{(\alpha)} \} \\ &= 0,\end{aligned}\tag{A.19}$$

pues $f_{(\alpha)} = 0$ en x_0 . Por lo tanto,

$$(\bar{\nabla}_a - \nabla_a)w'_b = (\bar{\nabla}_a - \nabla_a)w_b,\tag{A.20}$$

así que la diferencia $(\bar{\nabla}_a - \nabla_a)w_b$ sólo depende del valor de w_b en x_0 . Como consecuencia de esto, $(\bar{\nabla}_a - \nabla_a)$ define un mapeo lineal que manda vectores duales en x_0 a tensores de tipo $(0, 2)$ en x_0 , es decir, $(\bar{\nabla}_a - \nabla_a)$ puede ser visto como un tensor de tipo $(1, 2)$ en x_0 , que se denota como C_{ab}^c , y

$$\nabla_a w_b = \bar{\nabla}_a w_b - C_{ab}^c w_c.\tag{A.21}$$

En el caso en que el operador $\bar{\nabla}_a$ es igual a la derivada ordinaria ∂_a ($\bar{\nabla}_a = \partial_a$), C_{ab}^c se denota como Γ_{ab}^c y se llama símbolo de Christóffel. Tomemos $w_b = \nabla_b f = \bar{\nabla}_b f$, por (A.21) y el hecho de que ∇_a y $\bar{\nabla}_a$ son libres de torsión, vemos que:

$$\begin{aligned}C_{ab}^c \nabla_c f &= \bar{\nabla}_a \bar{\nabla}_b f - \nabla_a \nabla_b f = \bar{\nabla}_b \bar{\nabla}_a f - \nabla_b \nabla_a f = C_{ba}^c \nabla_c f \\ &\Rightarrow C_{ab}^c = C_{ba}^c.\end{aligned}\tag{A.22}$$

Por lo que C_{ab}^b es un tensor simétrico ($C_{[ab]}^b = 0$). De esto se sigue que, para cualquier vector dual w_b ,

$$\nabla_{[a} w_{b]} = \partial_{[a} w_{b]}.\tag{A.23}$$

Observemos la siguiente consecuencia de la ortonormalidad:

$$\begin{aligned}(e_\rho)^a w_{a\lambda\sigma} - (e_\lambda)^b w_{b\rho\sigma} &= (e_\lambda)^b w_{b\rho\sigma} - (e_\rho)^a w_{a\sigma\lambda} \\ &= (e_\lambda)^b \left(\sum_{\mu,\nu} \eta^{\mu\nu} \eta_{\rho\mu} w_{b\sigma\nu} \right) - (e_\rho)^a \left(\sum_{\mu,\nu} \eta^{\mu\nu} \eta_{\lambda\mu} w_{a\sigma\nu} \right) \\ &= \sum_{\mu,\nu} \eta^{\mu\nu} \eta_{\rho\mu} (e_\lambda)^b w_{b\sigma\nu} - \sum_{\mu,\nu} \eta^{\mu\nu} \eta_{\lambda\mu} (e_\rho)^a w_{a\sigma\nu} \\ &= \sum_{\mu,\nu} \eta^{\mu\nu} \{ \eta_{\rho\mu} (e_\lambda)^b w_{b\sigma\nu} - \eta_{\lambda\mu} (e_\rho)^a w_{a\sigma\nu} \} \\ &= \sum_{\mu,\nu} \eta^{\mu\nu} \{ (e_\rho)^a (e_\mu)_a (e_\lambda)^b w_{b\sigma\nu} - (e_\lambda)^b (e_\mu)_b (e_\rho)^a w_{a\sigma\nu} \} \\ &= \sum_{\mu,\nu} \eta^{\mu\nu} (e_\rho)^a (e_\lambda)^b \{ (e_\mu)_a w_{b\sigma\nu} - (e_\mu)_b w_{a\sigma\nu} \}\end{aligned}$$

104 El Método de Bases Ortonormales para Calcular la Curvatura

$$= (e_\rho)^a (e_\lambda)^b \sum_{\mu, \nu} \eta^{\mu\nu} \{ (e_\mu)_a w_{b\sigma\nu} - (e_\mu)_b w_{a\sigma\nu} \}. \quad (\text{A.24})$$

Por definición de las 1-formas de conexión, $(e_\rho)^a w_{a\lambda\sigma} - (e_\lambda)^b w_{b\rho\sigma} = (e_\rho)^a (e_\lambda)^b \{ \nabla_a (e_\sigma)_b - \nabla_b (e_\sigma)_a \}$ de modo que, sustituyendo en (A.24), tenemos:

$$(e_\rho)^a (e_\lambda)^b \{ \nabla_a (e_\sigma)_b - \nabla_b (e_\sigma)_a \} = (e_\rho)^a (e_\lambda)^b \sum_{\mu, \nu} \eta^{\mu\nu} \{ (e_\mu)_a w_{b\sigma\nu} - (e_\mu)_b w_{a\sigma\nu} \}.$$

Por lo tanto,

$$\nabla_a (e_\sigma)_b - \nabla_b (e_\sigma)_a = \sum_{\mu, \nu} \eta^{\mu\nu} \{ (e_\mu)_a w_{b\sigma\nu} - (e_\mu)_b w_{a\sigma\nu} \}. \quad (\text{A.25})$$

Usando la notación para tensores antisimétricos, podemos escribir esta última ecuación como:

$$\nabla_{[a} (e_\sigma)_{b]} = \sum_{\mu, \nu} \eta^{\mu\nu} (e_\mu)_{[a} w_{b]\sigma\nu} \quad (\text{A.26})$$

y por (A.23), llegamos a:

$$\partial_{[a} (e_\sigma)_{b]} = \sum_{\mu, \nu} \eta^{\mu\nu} (e_\mu)_{[a} w_{b]\sigma\nu}, \quad (\text{A.27})$$

es decir, (A.27) es consecuencia de (A.5). Recíprocamente, si para la base $\{(e_\mu)^a\}$ exigimos la condición de que se cumpla la ecuación (A.27), entonces, por (A.26), se tiene:

$$\Rightarrow \nabla_a (e_\sigma)_b - \nabla_b (e_\sigma)_a = \partial_{[a} (e_\sigma)_{b]} = \partial_a (e_\sigma)_b - \partial_b (e_\sigma)_a,$$

pero, por (A.21), $\nabla_a (e_\sigma)_b = \partial_a (e_\sigma)_b - \Gamma_{ab}^c (e_\sigma)_c$, de donde

$$\Rightarrow \begin{cases} \Gamma_{ba}^c (e_\sigma)_c - \Gamma_{ab}^c (e_\sigma)_c = 0 \\ \Gamma_{ba}^c (e_\sigma)_c = \Gamma_{ab}^c (e_\sigma)_c. \end{cases}$$

En particular, si tomamos $w_c = \nabla_c f$, entonces

$$\begin{aligned} \Rightarrow \partial_a \nabla_b f - \nabla_a \nabla_b f &= \Gamma_{ba}^c \nabla_c f \\ \Rightarrow \partial_a \nabla_b f - \nabla_a \nabla_b f &= \partial_b \nabla_a f - \nabla_b \nabla_a f \\ \Rightarrow \nabla_a \nabla_b f &= \nabla_b \nabla_a f. \end{aligned}$$

Por lo tanto, para una base ortonormal $\{(e_\mu)^\alpha\}$, la condición dada por la ecuación (A.27) es equivalente a que ∇_α sea libre de torsión. Esto significa que hemos encontrado una expresión equivalente al punto (ii), lo cual completa el desarrollo del método de bases ortonormales.

Para calcular la curvatura por el método de tétradas, se define una base ortonormal $\{(e_\mu)^\alpha\}$ y se imponen como condiciones las ecuaciones (A.9), (A.15) y (A.27). Por medio de (A.27) se determinan las 1-formas de conexión y se sustituyen en (A.15), lo que da el tensor de curvatura. Notemos que este método no requiere que conozcamos la expresión precisa de la métrica o del operador de derivada covariante asociado a ella, de hecho, el método mismo sirve para encontrar una expresión de la métrica, dadas ciertas condiciones.

Apéndice B

Teorema de Frobenius (versión dual) y Campos de Killing

B.1 Formas Diferenciales y el Teorema de Frobenius

Definición B.1 Sea M una variedad de dimensión m . Una p -forma diferencial es un tensor de tipo $(0, p)$ completamente antisimétrico, de manera que $w_{a_1 \dots a_p}$ es una p -forma si:

$$w_{a_1 \dots a_p} = w_{[a_1 \dots a_p]}. \tag{B.1}$$

Para el caso $p = 0$, una 0-forma es simplemente un tensor de tipo $(0, 0)$, es decir, una función diferenciable $f : M \rightarrow \mathbb{R} \in \mathcal{F}(M)$. Si $p = 1$, una 1-forma es sencillamente un vector dual, o covector, w_a ($w : V \rightarrow \mathbb{R}$). Denotamos el espacio vectorial de p -formas en un punto x por Λ_x^p y a la colección de campos de p -formas por Λ^p . Si tomamos el producto tensorial de una p -forma $w_{a_1 \dots a_p}$ y una q -forma $u_{b_1 \dots b_q}$, obtenemos un tensor de tipo $(0, p+q)$, el cual no necesariamente es completamente antisimétrico, así que no necesariamente es una $(p+q)$ -forma. Sin embargo, podemos tomar la parte antisimétrica de $w_{a_1 \dots a_p} u_{b_1 \dots b_q}$ y definir el siguiente mapeo:

$$\begin{aligned} \Lambda : \Lambda_x^p \times \Lambda_x^q &\rightarrow \Lambda_x^{p+q} \\ (w_{a_1 \dots a_p}, u_{b_1 \dots b_q}) &\rightarrow (w \wedge u)_{a_1 \dots a_p b_1 \dots b_q} = \binom{p+q}{p!q!} w_{[a_1 \dots a_p} u_{b_1 \dots b_q]}. \end{aligned}$$

Definimos el espacio vectorial de todas las formas diferenciales en x como la suma directa de Λ_x^p , $p = 0, \dots, m$.

Si tenemos un operador de derivada covariante ∇_a , podemos definir el siguiente mapeo suave:

$$w_{a_1 \dots a_p} \rightarrow (p+1) \nabla_{[b} w_{a_1 \dots a_p]}. \tag{B.2}$$

Tomemos ahora al operador $\tilde{\nabla}_a$ y definamos un mapeo de manera similar:

$$w_{a_1 \dots a_p} \rightarrow (p+1) \tilde{\nabla}_{[b} w_{a_1 \dots a_p]}$$

Para ver qué tanto difieren estos dos mapeos, hay que tomar la siguiente diferencia:

$$\nabla_{[b} w_{a_1 \dots a_p]} - \tilde{\nabla}_{[b} w_{a_1 \dots a_p]} = \sum_{j=1}^p C_{[b a_j}^d w_{a_1 \dots [d] \dots a_p]} = 0, \quad (\text{B.3})$$

pues C_{ab}^c es un tensor simétrico en a y b ¹. Por lo tanto, el mapeo (B.2) no depende del operador de derivada covariante, es decir, está bien definido sin que sea necesario tomar un operador preferencial. Se suele denotar al operador ∇ como d y se puede utilizar al gradiente ∂_x para calcular el mapeo anterior.

Sea M una variedad de dimensión de dimensión m . Frecuentemente nos encontramos con el siguiente problema: en cada punto $x \in M$ tenemos un subespacio $\mathcal{D}_x \subset T_x M$, $\dim \mathcal{D}_x = n < m$. El subespacio debe variar suavemente junto con x en el sentido de que, para cada $x \in M$, podemos encontrar una vecindad abierta U de x tal que, en U , \mathcal{D}_x es generado por n campos vectoriales suaves evaluados en x . Denotamos como \mathcal{D} a la colección de subespacios \mathcal{D}_x . Lo que queremos saber es bajo qué condiciones podemos encontrar subvariedades integrales de \mathcal{D} , es decir, cuándo podemos encontrar, para cada x , una subvariedad encajada N tal que $x \in N$ y $T_x N$ coincide con \mathcal{D}_x . Un caso especial de este problema se tiene cuando hay una métrica en M y se desea saber si un campo vectorial ξ^a es ortogonal a una familia de hipersuperficies, o bien, cuándo tenemos subespacios \mathcal{D} $(m+1)$ -dimensionales integrables que sean ortogonales a ξ^a .

De acuerdo a lo que vimos en el capítulo 1, la solución al problema general está dada por el teorema de Frobenius², el cual también tiene una formulación dual en términos de 1-formas diferenciales. Dado $\mathcal{D}_x \subset T_x M$, podemos considerar las 1-formas $w \in (T_x M)^*$ que satisfacen:

$$w_a v^a = 0, \quad (\text{B.4})$$

para todo $v^a \in \mathcal{D}_x$. Estas 1-formas w generan un subespacio de dimensión $m-n$, $T_x^* \subset (T_x M)^*$, del espacio dual al espacio tangente en x . Recíprocamente, un subespacio T_x^* $(m-n)$ -dimensional de $(T_x M)^*$ define un subespacio $\mathcal{D}_x \subset T_x M$ de dimensión n . Entonces, podemos reformular el problema anterior en términos de $(T_x^*)^*$: nos preguntamos ahora bajo qué condiciones la colección T^* de subespacios $(m-n)$ dimensionales de 1-formas tiene la propiedad de que la distribución \mathcal{D} de espacios vectoriales \mathcal{D}_x asociados a cada T_x^* tenga subvariedades integrales.

De acuerdo con el teorema de Frobenius, las subvariedades integrales existen si y sólo si, para todo $w_a \in T^*$ y para todos $v^a, u^a \in \mathcal{D}$ ($w_a v^a = w_a u^a = 0$), se tiene:

$$w_a [v, u]^a = 0 \quad (\text{B.5})$$

¹Ver apéndice A.

²Ver capítulo 1, sección 1.4, teorema 1.32

Notemos que, por (B.4) y la regla de leibniz, se cumple:

$$\begin{aligned} & \nabla_b w_a v^a = 0 \\ \Rightarrow & w_a \nabla_b v^a + v^a \nabla_b w_a = 0 \\ \Rightarrow & w_a \nabla_b v^a = -v^a \nabla_b w_a \end{aligned} \quad (\text{B.6})$$

y, analogamente, $w_a \nabla_b v^a = -v^a \nabla_b w_a$. Por (A.4), tenemos que $w_a [v, w]^a = w_a (v^b \nabla_b w^a - w^b \nabla_b v^a)$, entonces, para saber qué tipo de condición impone el teorema de Frobenius sobre w_a , sustituimos en (B.5):

$$\begin{aligned} 0 &= w_a (v^b \nabla_b w^a - w^b \nabla_b v^a) \\ &= -v^a v^b \nabla_b w_a + v^a w^b \nabla_b w_a \\ &= 2v^{[a} w^{b]} \nabla_b w_a \end{aligned} \quad (\text{B.7})$$

La ecuación anterior se cumple para v^a y w^a en el subespacio anulado por T^* si y sólo si $\nabla_{[b} w_a]$ puede expresarse en términos de elementos de T^* , es decir:

$$\nabla_{[b} w_a] = \sum_{\alpha=1}^{m-1} \mu_{[a}^{\alpha} v_{b]}^{\alpha}, \quad (\text{B.8})$$

donde cada v_{α}^a es una 1-forma arbitraria y cada $\mu_{\alpha}^a \in T^*$. Así, podemos reformular el teorema de Frobenius en términos de formas diferenciales como sigue:

Teorema B.2 (Frobenius, versión dual).

Sea T^* una colección suave de subespacios de 1-formas de dimensión $(m - n)$. La distribución \mathcal{D} de subespacios vectoriales m -dimensionales \mathcal{D}_x , contenidos en el espacio tangente a cada x , admite una subvariedad integral si y sólo si, para cada $w_a \in T^*$ se cumple la ecuación (B.8) con $\mu_{\alpha}^a \in T^*$. Esto también puede escribirse como

$$dw = \sum_{\alpha} \mu^{\alpha} \wedge v^{\alpha}. \quad (\text{B.9})$$

La formulación dual del teorema de Frobenius da un criterio útil para saber cuándo un campo vectorial ξ^a es ortogonal a una hipersuperficie. Tomamos T^* el subespacio de una dimensión generado por $\xi_a = g_{ab} \xi^b$; por el teorema de Frobenius ξ^a es ortogonal a una hipersuperficie (de dimensión $m - 1$) si y sólo si $\nabla_{[a} \xi_{b]} = \xi_{[a} v_{b]}$ ($\mu_a = \xi_a$ ya que T^* tiene dimensión uno). Observemos que:

$$\begin{aligned} & \nabla_{[a} \xi_{b]} = \xi_{[a} v_{b]} \\ \Leftrightarrow & \nabla_a \xi_b - \nabla_b \xi_a = \xi_a v_b - \xi_b v_a \\ \Leftrightarrow & \xi_c \nabla_a \xi_b - \xi_c \nabla_b \xi_a = \xi_c \xi_a v_b - \xi_c \xi_b v_a \\ \Leftrightarrow & \xi_c \nabla_a \xi_b - \xi_c \nabla_b \xi_a + \xi_b \nabla_c \xi_a - \xi_b \nabla_a \xi_c + \xi_a \nabla_b \xi_c - \xi_a \nabla_c \xi_b \\ & = \xi_c \xi_a v_b - \xi_c \xi_b v_a + \xi_b \xi_c v_a - \xi_b \xi_a v_c + \xi_a \xi_b v_c - \xi_a \xi_c v_b \\ \Leftrightarrow & \xi_{[a} \nabla_b \xi_{c]} = 0. \end{aligned}$$

por lo tanto, ξ^a es ortogonal a una superficie $(m - 1)$ -dimensional si y sólo si

$$\xi_{[a} \nabla_b \xi_{c]} = 0. \quad (\text{B.10})$$

B.2 Mapeo Diferencial para Tensores y Campos de Killing

Según el capítulo 1, para una función entre variedades $F : M \rightarrow N$, el mapeo diferencial "manda" vectores tangentes en p a vectores tangentes en $F(p)$.

$$\begin{aligned} F_* : T_p M &\rightarrow T_{F(p)} N \\ v &\rightarrow F_*(v) = v(f \circ F) \quad \forall f \in \mathcal{F}(N). \end{aligned} \quad (B.11)$$

De manera similar, podemos emplear la función F para "atraer" vectores duales en $F(p)$ a vectores duales en p . Definimos siguiente mapeo:

$$\begin{aligned} F_* : (T_{F(p)} N)^* &\rightarrow (T_p M)^* \\ w_\alpha &\rightarrow (F_* w)_\alpha v^\alpha = w_\alpha (F^* v)^\alpha \quad \forall v^\alpha \in T_p^* M. \end{aligned} \quad (B.12)$$

Podemos extender la acción de F_* como un mapeo que "atrae" tensores de tipo $(0, l)$ en $F(p)$ a tensores de tipo $(0, l)$ en p :

$$(F_* T)_{b_1 \dots b_l} (v_1)^{b_1} \dots (v_l)^{b_l} = T_{b_1 \dots b_l} (F^* v_1)^{b_1} \dots (F^* v_l)^{b_l}. \quad (B.13)$$

Análogamente, podemos extender la acción de F^* para "mandar" tensores de tipo $(k, 0)$ en p a tensores de tipo $(k, 0)$ en $F(p)$:

$$(F^* T)^{a_1 \dots a_k} (w_1)_{a_1} \dots (w_k)_{a_k} = T^{a_1 \dots a_k} (F_* w_1)_{a_1} \dots (F_* w_k)_{a_k}. \quad (B.14)$$

Si F es un difeomorfismo, podemos usar F^{-1} para extender la definición de F a tensores de todo tipo, tomando en cuenta que $(F^{-1})^*$ es un mapeo de $T_{F(p)} N$ a $T_p M$. Si $T_{a_1 \dots a_k}^{b_1 \dots b_k}$ es un tensor de tipo (k, l) en p , definimos el tensor $(F^* T)_{a_1 \dots a_k}^{b_1 \dots b_k}$ en $F(p)$ como:

$$(F^* T)_{a_1 \dots a_k}^{b_1 \dots b_k} (w_1)_{a_1} \dots (w_k)_{a_k} (v_1)^{b_1} \dots (v_l)^{b_l} = T_{b_1 \dots b_l}^{a_1 \dots a_k} (F_* w_1)_{a_1} \dots ((F^{-1})^* v_l)^{b_l}. \quad (B.15)$$

Ahora bien, dado $F : M \rightarrow M$ un difeomorfismo y T un campo tensorial en M , podemos comparar T y $F^* T$; si $F^* T = T$, entonces, pese a que hemos "movido" a T por medio de F , el tensor "permanece igual", es decir, F es una *transformación de simetría del campo tensorial* T . En el caso de la métrica g_{ab} , una transformación de simetría se llama una isimetría ³.

Así como podemos extender la acción del mapeo diferencial a los tensores, también podemos hacer lo mismo con la derivada de Lie. Según la sección 1.3, $\forall f \in \mathcal{F}$, $L_v(f) = v(f)$, donde v es un campo vectorial con un grupo monoparamétrico asociado, φ_t . Para un tensor de tipo (k, l) , su derivada de Lie queda definida como:

$$L_v T_{b_1 \dots b_l}^{a_1 \dots a_k} = \lim_{t \rightarrow 0} \left\{ \frac{\varphi_t^* T_{b_1 \dots b_l}^{a_1 \dots a_k} - T_{b_1 \dots b_l}^{a_1 \dots a_k}}{t} \right\}. \quad (B.16)$$

³Notemos que esto es una generalización de la idea de isimetría planteada en el capítulo 2

De acuerdo con las ideas desarrolladas en la sección 2.1, φ_t es una transformación de simetría de $T_{b_1 \dots b_l}^{a_1 \dots a_k}$ si y sólo si su derivada de Lie es cero.

Si tomamos un sistema coordenado adaptado a v^a , las componentes de la derivada de Lie de un tensor $T_{b_1 \dots b_l}^{a_1 \dots a_k}$ en dicho sistema son:

$$L_v T_{b_1 \dots b_l}^{a_1 \dots a_k} = \frac{\partial T_{b_1 \dots b_l}^{a_1 \dots a_k}}{\partial x^1} \tag{B.17}$$

para un campo vectorial w^a se tiene entonces:

$$L_v w^\mu = \frac{\partial w^\mu}{\partial x^1} \tag{B.18}$$

además, como $v^a = \left(\frac{\partial}{\partial x^1}\right)^a$ y $w^a = \sum_\mu w^\mu \left(\frac{\partial}{\partial x^1}\right)^a$, el conmutador de v y w está dado por

$$[v, w]^\mu = \sum_\nu v^\nu \frac{\partial w^\mu}{\partial x^\nu} - w^\nu \frac{\partial v^\mu}{\partial x^\nu} = \frac{\partial w^\mu}{\partial x^1} \tag{B.19}$$

Entonces, las componentes de $L_v w^a$ y $[v, w]^a$ son iguales en un sistema coordenado adaptado a v . Sin embargo, dado que ambos objetos están definidos de manera independiente a un sistema coordenado, obtenemos que:

$$L_v w^a = [v, w]^a \tag{B.20}$$

En base a (B.20) y retomando (A.4), podemos construir una expresión de la derivada de Lie de un campo tensorial $T_{b_1 \dots b_l}^{a_1 \dots a_k}$ en términos de un operador de derivada covariante:

$$L_v T_{b_1 \dots b_l}^{a_1 \dots a_k} = v^c \nabla_c T_{b_1 \dots b_l}^{a_1 \dots a_k} - \sum_{i=1}^k T_{b_1 \dots b_l}^{a_1 \dots a_{i-1} a_{i+1} \dots a_k} \nabla_c v^{a_i} + \sum_{j=1}^l T_{b_1 \dots b_{j-1} b_{j+1} \dots b_l}^{a_1 \dots a_k} \nabla_{b_j} v^c \tag{B.21}$$

En el caso de g_{ab} , si ∇_a es compatible con la métrica, tenemos:

$$L_v g_{ab} = v^c \nabla_c g_{ab} + g_{cb} \nabla_a v^c + g_{ac} \nabla_b v^c = \nabla_a v_b + \nabla_b v_a \tag{B.22}$$

Definición B.3 Sea $\varphi_t : M \rightarrow M$ un grupo monoparamétrico de isometrías, $\varphi_t^* g_{ab} = g_{ab}$. El campo vectorial ξ^a que genera a φ_t se llama *campo vectorial de Killing*

Para que φ_t sea un grupo de isometrías, la derivada de Lie de la métrica en la dirección del campo de Killing ξ^a debe ser nula:

$$\begin{aligned} L_\xi g_{ab} &= 0 \\ \Leftrightarrow \nabla_a \xi_b + \nabla_b \xi_a &= 0. \end{aligned} \tag{B.23}$$

⁴Ver teorema 1.22

112 Teorema de Frobenius (versión dual) y Campos de Killing

A la ecuación anterior se le conoce como ecuación de Killing, donde ∇_a es el operador de derivada covariante asociado a g_{ab} .

Proposición B.4

Sea ξ^a un campo de Killing y γ una geodésica con tangente u^a , entonces $\xi_a u^a$ es constante a lo largo de γ .

Prueba.

Tenemos que

$$u^b \nabla_b (\xi_a u^a) = u^b u^a \nabla_b \xi_a + \xi_a u^b \nabla_b u^a, \quad (\text{B.24})$$

por regla de Leibniz. El segundo término se anula por estar a lo largo de γ . En cuanto al primero, notemos que, por (B.23)

$$\begin{aligned} u^b u^a \nabla_a \xi_b + u^b u^a \nabla_b \xi_a &= 0 \\ \Rightarrow 2u^b u^a \nabla_b \xi_a &= 0, \end{aligned} \quad (\text{B.25})$$

por lo que el primer término también se anula y

$$u^b \nabla_b (\xi_a u^a) = 0. \quad (\text{B.26})$$

□

Dado que en relatividad general las geodésicas de tipo temporal representan los movimientos de caída libre de partículas en el espacio-tiempo y las geodésicas nulas representan las trayectorias de los rayos de luz, la proposición anterior tiene un significado similar al teorema de Noether en el sentido de que un grupo monoparamétrico de simetrías (en este caso la simetría está asociada con la métrica) da origen a una ley de conservación para partículas y rayos de luz.

Bibliografía

- [1] Arnold, Vladimir I. *Mathematical Methods of Classical Mechanics*. Springer Verlag, 1984.
- [2] Dubrovin, B. A., Fomenko A.T., Novikov, S. P. *Modern Geometry-Methods and Applications*. Springer Verlag, 1992.
- [3] D'Eath, P. D. *Black Holes Gravitational Interactions*. Clarendon Press, 1996.
- [4] Hawking, S. W. *300 Years of Gravitation*. Cambridge University Press, 1987.
- [5] Kawakubo, K. *The Theory of Transformations Groups*. Oxford University Press, 1991.
- [6] Landau, L. D., Lifshitz, E. M. *Mechanics*. Pergamon Press, 1976.
- [7] Olver, P. J. *Applications of Lie Groups to Differential Equations*. Springer Verlag, 1993.
- [8] O'Neill, B. *Semi-Riemannian Geometry*. Academic Press, 1983.
- [9] Sharpe, R. W. *Differential Geometry: Cartan's generalization of Klein's Erlangen Program*. Springer Verlag, 1996.
- [10] Wald, R. M. *General Relativity*. The University of Chicago Press, 1984
- [11] Wald, R. M. *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. The University of Chicago Press, 1994.
- [12] Wald, R. M. *Black Hole entropy is the Noether charge*. Physical Review , 1993.
- [13] Warner, F. W. *Foundation of Differentiable Manifolds and Lie Groups*. Scott Foresman, 1971.