

2ej
15



Universidad Nacional Autónoma
de México

FACULTAD DE CIENCIAS

NOTAS DE GEOMETRIA DIFERENCIAL CON
APLICACIONES A LA FISICA

T E S I S

Que para obtener el título de
MATEMATICO

p r e s e n t a

CESAR HERNAN MENDIBURU SILVEIRA



México, D. F.

1987



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 está 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.

CONTENIDO

Introducción

I.- VARIEDADES

1. Variedades	pag. 1
2. Ejemplos de Variedades	5
3. Aplicaciones Diferenciables	10
4. Vectores Tangentes	12
5. La Diferencial	15
6. Curvas	16
7. Campos Vectoriales	18
8. Uno-Formas	21
9. Subvariedades	22
10. Inmersiones y Submersiones	25
11. Topología de Variedades	27
12. Algunas Variedades Especiales	31
13. Curvas Integrales	35

II.- TENSORES

14. Algebra Multilineal	pag. 41
15. Campos Tensoriales	42
16. Interpretaciones	43
17. Tensores en un Punto	44
18. Componentes Tensoriales	46
19. Contracción	48
20. Tensores Covariantes	49
21. Derivaciones Tensoriales	51
22. Formas Bilineales Simétricas	55
23. Productos Escalares	56

III.- FORMAS DIFERENCIALES

24. Formas Diferenciales	pag. 62
25. Derivada Exterior	65
26. Productos Interiores	68
27. Inverso del Lema de Poincaré	71
28. Cadenas Cúbicas	74
29. Integración en Espacios Euclidianos	80
30. Integración de p-Formas	83
31. Teorema de Stokes	87
32. Teorema de Frobenius	89
33. Dual del Teorema de Frobenius	92

IV.- GEOMETRIA SEMI-RIEMANNIANA

34. Variedades semi-Riemannianas	pag. 95
35. Isometrías	98
36. Conexión de Levi-Civita	100
37. Transporte Paralelo	105
38. Geodésicas	107
39. Aplicación Exponencial	111
40. Curvatura	114
41. Curvatura Seccional	118
42. Superficies semi-Riemannianas	121
43. Cambio de Tipo y Contracción Métrica	122
44. Campos de Referenciales	124
45. Algunos Operadores Diferenciales	125
46. Curvatura de Ricci y Curvatura Escalar	127
47. Isometrías Locales	130
48. Variedades con Conexión	133

V.- GEOMETRIA LORENTZIANA Y RELATIVIDAD

49. Carácter Causal	pag. 141
50. Conos de Tiempo y Orientación en el Tiempo	143
51. Relatividad Especial	146
52. Espaciotiempo Newtoniano	147
53. Espaciotiempo de Minkowski	150
54. Geometría Minkowskiana	152
55. Observadores	155
56. Algunos Efectos Relativistas	158
57. Impulso y Energía	162
58. Aplicaciones Biparamétricas y Campos de Jacobi	164
59. Relatividad General	166
60. Ecuación de Einstein	170

Bibliografía	172
--------------	-----

INTRODUCCION

El presente trabajo consiste en una recopilación de ciertos resultados de la geometría diferencial que, durante las últimas dos décadas, se han convertido en indispensables para el estudio de muchos temas de la física matemática.

El avance de la teoría general de la relatividad y su aplicación a las teorías cosmológicas, ha provocado que el "cálculo tensorial" clásico, que fue creado por Ricci y Levi-Civita a principios de este siglo para desarrollar las ideas geométricas de Riemann, y que fue adoptado por Einstein en su teoría de la gravitación, resulte ahora insuficiente. Actualmente, los físicos se familiarizan con éste cálculo tensorial a través de un curso operativo de relatividad general a nivel de licenciatura, pero al consultar la literatura reciente sobre el tema, se enfrentan a nociones tales como la de variedad diferenciable, forma diferenciable, derivada de Lie, diferencial exterior, etc., para las cuales, no están adecuadamente preparados.

El cálculo tensorial clásico y la geometría diferencial que le sirve de fundamento, se avocan sobre todo, al estudio de las propiedades locales de las variedades diferenciables utilizando fuertemente las coordenadas; la cosmología, en cambio, considera al universo como un todo mediante un tipo especial de variedad diferenciable con varias estructuras matemáticas adicionales. Ello ha obligado a los físicos a utilizar cada vez más la geometría diferencial moderna, que pone el énfasis en las cuestiones globales y en el enfoque invariante, es decir, independiente de las coordenadas. Es así que los trabajos recientes de físicos y matemáticos sobre relatividad y cosmología incorporan el formalismo de la geometría diferencial moderna.

Estas notas pretenden servir de introducción al estudio de la geometría diferencial para los estudiantes de física.

Por lo general, un curso introductorio de geometría diferencial comprende pocas nociones, como por ejemplo la noción de superficie diferenciable, pero en cambio, se las desarrolla con bastante profundidad. Esto es adecuado para los estudiantes de matemática pues, para ellos, lo importante es obtener experiencia y madurez matemática, más que obtener información utilizable. Por el contrario, para un físico es más importante obtener un conocimiento operativo, aunque no demasiado profundo, de una gran variedad de tópicos. Debido a ésto, los temas tratados en las notas son muy variados y muchos de ellos no son elementales, han sido escogidos sobre todo en razón de su utilidad en física. Antes que profundizar en estos temas, se pretende que el estudiante de física se familiarice con ellos lo suficiente como para estar en condiciones de entender la literatura reciente sobre relatividad y cosmología.

A pesar de estas características o, tal vez debido a ellas, creo que el presente trabajo puede ser útil también a los estudiantes de matemática, para quienes un estudio preliminar y

panorámico de esta rama tan importante y extensa de la matemática puede ser de bastante interés.

Los prerrequisitos para estudiar estas notas son modestos: un buen conocimiento del cálculo diferencial de varias variables, de los teoremas de existencia y unicidad de las ecuaciones diferenciales ordinarias y de los fundamentos del álgebra lineal, serán suficientes.

I. VARIETADES

1. VARIETADES

Es difícil imaginar un problema físico en el cual no esté involucrado algún tipo de "espacio continuo". Esto puede ser el "espacio físico tridimensional", el "espaciotiempo de 4 dimensiones", el "espacio fase" de un problema de mecánica clásica o de mecánica cuántica, el "espacio de los estados de equilibrio termodinámico", etc. Estos "espacios" tienen propiedades geométricas y topológicas diferentes, pero las propiedades que tienen en común están incorporadas en la definición de variedad, que es el sustituto matemático preciso para la palabra "espacio".

El concepto de variedad generaliza el concepto de una superficie o de una curva en \mathbb{R}^3 . Sin embargo, la definición será dada sin hacer referencia a un espacio "ambiente" en el cual estén contenidas muestras variedades. Por razones obvias éste enfoque es el más conveniente y natural cuando se utilizan las variedades en cosmología como modelos del universo. El concepto de variedad diferenciable generaliza el concepto de superficie diferenciable en \mathbb{R}^3 , esto es, de una superficie que tiene en cada punto un plano tangente que varía suavemente y aquí, también, los vectores tangentes a la variedad diferenciable serán definidos de manera intrínseca, es decir, sin hacer referencia al espacio ambiente \mathbb{R}^3 .

Una variedad topológica n-dimensional es un espacio topológico de Hausdorff tal que cada punto tiene una vecindad homeomorfa a un conjunto abierto de \mathbb{R}^n .

Desglosaremos ahora esta definición. Un espacio topológico es simplemente un conjunto de puntos provisto de cierta estructura, llamada topología, mediante la cual es posible establecer la idea de continuidad en el contexto más general y abstracto. Tenemos, por ejemplo, que \mathbb{R}^n es un espacio topológico pues en él se puede hablar de la continuidad de las funciones. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función definida en \mathbb{R}^n y con valores en \mathbb{R}^n . Sabemos, por la definición clásica de continuidad para f , que f es continua en el punto $x = (x_1, x_2, \dots, x_n)$ de \mathbb{R}^n si dado un número $\epsilon > 0$ existe otro $\delta > 0$ tal que $\|f(x) - f(y)\| < \epsilon$ siempre que $\|x - y\| < \delta$, donde $\|x - y\| = ((x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2)^{1/2}$. f es continua si satisface esta condición para cualquier punto $x \in \mathbb{R}^n$.

Un subconjunto N de \mathbb{R}^n se llama vecindad del punto $p \in \mathbb{R}^n$ si para algún número $r > 0$ el disco abierto con centro en p y radio r está completamente contenido en N , es decir, si $x \in \mathbb{R}^n$ es tal que $\|x - p\| < r$ entonces $x \in N$. Es fácil reestablecer la definición de continuidad que dimos arriba en la forma siguiente: f es continua si para cualquier $x \in \mathbb{R}^n$ y cualquier vecindad N de $f(x) \in \mathbb{R}^n$ se tiene que $f^{-1}(N) = \{y \in \mathbb{R}^n \mid f(y) \in N\}$ es una vecindad de $x \in \mathbb{R}^n$.

Esta noción de cada punto de \mathbb{R}^n teniendo una colección de vecindades por medio de las cuales es posible dar una buena definición de continuidad para funciones $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, constituye la estruc-

tura que hace de \mathbb{R}^n un espacio topológico, es su topología.

Nótese que al definir vecindades en un espacio euclidiano usamos muy fuertemente la distancia euclidiana $\|x-y\|$ entre puntos x, y de \mathbb{R}^n . Para definir un espacio topológico general nos gustaría retener el concepto de vecindad pero, al mismo tiempo, deshacernos de cualquier dependencia del concepto de distancia, que en espacios distintos de \mathbb{R}^n podría no estar a la mano. Para lograr éste objetivo, tomaremos las propiedades principales que tienen las vecindades de puntos de \mathbb{R}^n y haremos que esas mismas propiedades caractericen a las vecindades abstractas de un espacio topológico general. De éste modo, llegamos a la siguiente definición de un espacio topológico:

Un espacio topológico es un conjunto T tal que cada punto x de T tiene una familia no vacía de subconjuntos de T , llamados vecindades de x , y que tienen las siguientes propiedades:

- 1) x pertenece a cada una de sus vecindades.
- 2) La intersección de dos vecindades de x es una vecindad de x .
- 3) Si N es vecindad de x y si U es un subconjunto de T que contiene a N , entonces U también es vecindad de x .
- 4) Si N es una vecindad de x y si N° denota al conjunto $\{z \in N \mid \text{Nes vecindad de } z\}$, entonces N° es también una vecindad de x . El conjunto N° se llama interior de N .

La asignación de una colección de vecindades a cada punto $x \in T$ se conoce como una topología para T , pero es claro que puede haber muchas maneras de hacer ésta asignación.

No es difícil comprobar que las vecindades de los puntos de \mathbb{R}^n que definimos más arriba tienen las propiedades 1) a 4) y que, por lo tanto, definen una topología para \mathbb{R}^n , conocida como la topología usual de \mathbb{R}^n .

Podemos ahora decir, con precisión, lo que se entiende por una función continua y por un homeomorfismo. Sean T_1 y T_2 espacios topológicos, entonces una función $f: T_1 \rightarrow T_2$ es continua si para cada punto x de T_1 y para cada vecindad N de $f(x)$ en T_2 el conjunto $f^{-1}(N)$ es una vecindad de x en T_1 .

Intuitivamente, ésta definición nos dice que una función continua es tal que los puntos "próximos" en T_2 provienen de puntos de T_1 que son "próximos" entre sí, donde el concepto de "proximidad" depende solamente de las topologías respectivas de T_1 y T_2 .

Una función $h: T_1 \rightarrow T_2$ es llamada un homeomorfismo si es inyectiva (uno a uno), suprayectiva (sobre), continua y tiene inversa continua. Cuando tal función existe, los espacios T_1 y T_2 se llaman homeomorfos. El homeomorfismo h establece una correspondencia biunívoca entre los puntos de T_1 y de T_2 : a cada punto $x \in T_1$, corresponde un único $h(x) \in T_2$ y a cada $y \in T_2$ corresponde un único $h^{-1}(y) \in T_1$.

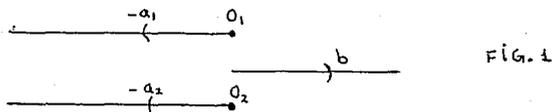
Puesto que para un homeomorfismo h tanto la función $h: T_1 \rightarrow T_2$ como la función $h^{-1}: T_2 \rightarrow T_1$ son continuas, de la definición de continuidad se sigue que a la vecindad N de x en T_1 le corresponde

la vecindad $f(N)$ de $f(x)$ en T_2 y que a la vecindad M de y en T_1 le corresponde la vecindad $f^{-1}(M)$ de $f^{-1}(y)$ en T_1 . Así que un homeomorfismo establece también una correspondencia biunívoca entre las topologías de T_1 y de T_2 ; cualquier propiedad topológica que tenga T_1 , es decir una propiedad que dependa de su topología, también la tendrá T_2 , es por esto que dos espacios topológicos homeomorfos también se llaman topológicamente equivalentes, pues comparten las mismas propiedades topológicas.

Un espacio topológico se llama de Hausdorff si cada par de puntos distintos poseen vecindades que no se intersectan. \mathbb{R}^n es obviamente un espacio de Hausdorff. Un subconjunto U de un espacio topológico se llama abierto si U es vecindad de cada uno de sus puntos. Así, por ejemplo, N° , el interior de una vecindad N , es un conjunto abierto de cualquier espacio topológico.

Con estas definiciones ya estamos en condiciones de entender y reinterpretar la definición de variedad que hemos dado más arriba. Diremos, entonces, que una variedad es un espacio localmente euclidiano, es decir, cada punto tiene una vecindad idéntica a \mathbb{R}^n , que además es un espacio de Hausdorff.

La propiedad de ser un espacio topológico de Hausdorff que imponemos a las variedades en su definición, es muy natural, sin embargo, ésta no se sigue en forma alguna de la propiedad de ser localmente euclidiano. Esto podría parecer un poco sorprendente, dado que una variedad es localmente idéntica a \mathbb{R}^n y \mathbb{R}^n es un espacio de Hausdorff. Para demostrar ésta afirmación basta presentar un ejemplo de espacio localmente euclidiano pero que no es de Hausdorff. Consideremos el conjunto M que consta de dos semirectas $(-\infty, 0]$ y de una semirecta abierta $(0, \infty)$. (Ver fig. 1).



Dado cualquier $x \in M$ con $x \neq 0_1$ y $x \neq 0_2$, definimos las vecindades de x como cualquier intervalo abierto que contenga a x ; pero si $x = 0_1$ ó $x = 0_2$, definimos la vecindad de x como cualquier subconjunto de M de la forma $(-a_1, 0] \cup (0, b)$ ó de la forma $(-a_2, 0] \cup (0, b)$, respectivamente. (Ver fig. 1). Es fácil ver que el conjunto de las vecindades de los puntos de M definidas de esa manera satisfacen los axiomas 1) a 4) en la definición de espacio topológico y por lo tanto, M es un espacio topológico. También es claro que M es localmente euclidiano ya que cada una de sus vecindades es homeomorfa a un intervalo abierto de la recta real. Sin embargo, M no es un espacio de Hausdorff pues cualquier vecindad de O_1 se intersecta con cualquier vecindad de O_2 .

Ahora consideraremos algunas definiciones que nos llevarán a la noción de variedad diferenciable.

Una carta (U, φ) de una variedad M es una pareja de objetos U y φ , en donde U es un subconjunto abierto de M y φ es un homeomorfismo $\varphi: U \rightarrow V$ de U sobre un subconjunto abierto V de \mathbb{R}^n .

La imagen $\varphi(x)$ del punto $x \in U \subset M$ pertenece a $V \subset \mathbb{R}^n$ y, por lo tanto, está representada por la n -ada de números reales (x^1, x^2, \dots, x^n) . Las componentes de ésta n -ada se llaman las coordenadas locales del punto $x \in M$ con respecto a la carta (U, φ) . Por ésto, una carta en M se llama también sistema local de coordenadas en M .

Un atlas de clase C^∞ sobre una variedad M es una familia de cartas $\{(U_\alpha, \varphi_\alpha)\}$ de la variedad M , tal que los dominios $\{U_\alpha\}$ cubren a M ($M = \bigcup U_\alpha$) y los homeomorfismos $\{\varphi_\alpha\}$ cumplen la siguiente condición de compatibilidad:

Las aplicaciones $\varphi_\beta \circ \varphi_\alpha^{-1}: \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta)$ son aplicaciones entre subconjuntos abiertos $\varphi_\alpha(U_\alpha \cap U_\beta)$ y $\varphi_\beta(U_\alpha \cap U_\beta)$ de \mathbb{R}^n y son aplicaciones diferenciables de clase C^∞ . En otras palabras, si (x^1, x^2, \dots, x^n) y (y^1, y^2, \dots, y^n) son las coordenadas locales del punto $x \in M$ con respecto a las cartas $(U_\alpha, \varphi_\alpha)$ y (U_β, φ_β) , respectivamente, la aplicación $\varphi_\beta \circ \varphi_\alpha^{-1}$ está definida en el dominio $\varphi_\alpha(U_\alpha \cap U_\beta)$ como un sistema de n funciones reales de n variables

$$y^j = f^j(x^1, \dots, x^n) \quad j=1, 2, \dots, n.$$

que poseen derivadas parciales continuas de todos los órdenes. (Ver fig. 2).

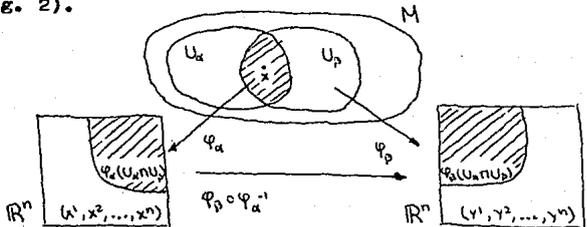


FIG. 2

Dos atlas de clase C^∞ , $\{(U_\alpha, \varphi_\alpha)\}$ y $\{(V_\beta, \psi_\beta)\}$, sobre la variedad M , se llaman equivalentes si y solo si el conjunto $\{(U_\alpha, \varphi_\alpha), (V_\beta, \psi_\beta)\}$, que es la unión de los dos, es también un atlas de clase C^∞ sobre M .

Una variedad (topológica) M junto con todos los atlas de clase C^∞ que son equivalentes a un cierto atlas dado de clase C^∞ , constituye una variedad con estructura diferenciable, también llamada variedad diferenciable o variedad suave. Así pues, una variedad diferenciable posee dos estructuras. Por un lado, tiene una topología localmente euclidiana y de Hausdorff y, por otro, una estructura diferenciable determinada por todos los atlas que son equivalentes a un cierto atlas dado.

En las definiciones anteriores vamos siempre a suponer que todas las cartas de un atlas $\{(U_\alpha, \varphi_\alpha)\}$ aplican subconjuntos abiertos U_α de M en subconjuntos abiertos $\varphi_\alpha(U_\alpha)$ de \mathbb{R}^n , en donde el número entero n permanece fijo. Se dice entonces que M es una variedad diferenciable n -dimensional. Se puede demostrar que si M es una variedad conexa, el número n considerado arriba permanece, efectivamente, fijo. De todas maneras, todas las variedades que consideraremos en estas notas serán conexas, de modo que tendrán una dimensión bien definida. Entenderemos por una variedad conexa, una variedad M que no puede ser considerada como la unión de dos partes distintas y separadas.

2 EJEMPLOS DE VARIEDADES

1) La esfera S^2 de \mathbb{R}^3 . Esta variedad diferenciable bidimensional se define como sigue: S^2 es el conjunto de los puntos (x^1, x^2, x^3) de \mathbb{R}^3 tales que $\sum (x^i)^2 = 1$, junto con todos los atlas equivalentes al atlas con 6 cartas $(D_{i\epsilon}, \varphi_{i\epsilon})$, en donde $D_{i\epsilon} = \{(x^1, x^2, x^3) \in S^2 \mid \epsilon x^i > 0\}$ $i=1,2,3$; $\epsilon = \pm 1$ y $\varphi_{i\epsilon}(x^1, x^2, x^3) = (x^j, x^k)$ donde i, j, k es una permutación circular de 1, 2, 3 si $\epsilon = 1$, y una permutación circular de 1, 3, 2 si $\epsilon = -1$. Los $D_{i\epsilon}$ corresponden a los 6 hemisferios de S^2 y, por lo tanto, cubren a S^2 . La aplicación $\varphi_{i\epsilon}$ es un homeomorfismo, pues simplemente es la proyección ortogonal del hemisferio $D_{i\epsilon}$ sobre el plano $x^i = 0$.

Para demostrar que las cartas sobre S^2 así definidas satisfacen la condición de compatibilidad en la definición de atlas, tomemos, por ejemplo $(i, \epsilon) = (3, +) = N$, $(i, \epsilon) = (3, -) = S$, $(i, \epsilon) = (2, +) = E$. Entonces

$$\varphi_N(x^1, x^2, x^3) = (x^1, x^2) \quad \text{y} \quad \varphi_E(x^1, x^2, x^3) = (x^3, x^1)$$

$$\varphi_S(x^3, x^1) = (x^1, (1 - (x^1)^2 - (x^3)^2)^{1/2}, x^3)$$

Por lo tanto, $(\varphi_N \circ \varphi_E^{-1})(x^3, x^1) = (x^1, (1 - (x^1)^2 - (x^3)^2)^{1/2})$ y esta aplicación es de clase C^∞ . De manera semejante se puede comprobar que las demás aplicaciones $\varphi_{i\epsilon} \circ \varphi_{j\epsilon'}^{-1}$ son de clase C^∞ .

También es claro que podemos construir un atlas semejante pero con $2n$ cartas para la esfera n -dimensional S^n contenida en \mathbb{R}^{n+1} .

2) Otro atlas sobre S^2 . Podemos construir otro atlas para S^2 equivalente al del ejemplo anterior usando las coordenadas de las proyecciones estereográficas. Sean P y Q los polos norte y sur, respectivamente; sean $W = S^2 - \{P\}$ y $V = S^2 - \{Q\}$, y sean g y h las proyecciones estereográficas desde los polos norte y sur sobre el plano $x^3 = 0$. (Ver fig. 3). Usando triángulos semejantes en la figura 3, vemos que

$$y^1 = g^1(x) = \frac{x^1}{1 - x^3}, \quad y^2 = g^2(x) = \frac{x^2}{1 - x^3}$$

determinan a $g: W \rightarrow \mathbb{R}^2$ y que

$$z^1 = h^1(x) = \frac{x^1}{1+x^3}$$

$$z^2 = h^2(x) = \frac{x^2}{1+x^3}$$

determinan a $h: V \rightarrow \mathbb{R}^2$.

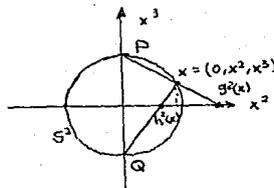
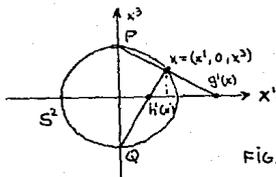


FIG. 3

Podemos verificar ahora que $\{(W, g), (V, h)\}$ es un atlas de clase C^∞ sobre S^2 con sólo dos cartas, ya que no es difícil convenirse de la veracidad de las siguientes afirmaciones:

- a) $S^2 = W \cup V$ y las aplicaciones h y g son claramente homeomorfismos.
 b) Sea $d = h \circ g^{-1}$, donde g^{-1} se entiende restringida a $g(W \cap V)$, entonces $d: \mathbb{R}^2 - \{0\} \rightarrow \mathbb{R}^2 - \{0\}$ es C^∞ y tiene inversa $d^{-1} = g \circ h^{-1}$, que es también C^∞ , puesto que tenemos

$$d \begin{cases} z^1 = y^1 / ((y^1)^2 + (y^2)^2) \\ z^2 = y^2 / ((y^1)^2 + (y^2)^2) \end{cases} \quad d^{-1} \begin{cases} y^1 = z^1 / ((z^1)^2 + (z^2)^2) \\ y^2 = z^2 / ((z^1)^2 + (z^2)^2) \end{cases}$$

- c) Este atlas es equivalente al del ejemplo 1) ya que es posible comprobar que el atlas formado por las 6 cartas del ejemplo 1) y las dos cartas de éste ejemplo, forman en conjunto, un atlas C^∞ sobre la esfera S^2 . La comprobación es fácil pero aburrida pues hay que estudiar todas las posibles intersecciones $D_i \cap W$ y $D_i \cap V$ y todos los correspondientes cambios de coordenadas definidos en éstas intersecciones.

La tremenda utilidad del concepto de variedad proviene de su generalidad. De hecho, abarca conjuntos que a primera vista no serían considerados como "espacios"; no obstante, por la definición de variedad, cualquier conjunto M que puede ser parametrizado local y continuamente es una variedad, cuya dimensión es el número de parámetros continuos e independientes que son necesarios para especificar sin ambigüedad cada uno de sus puntos. Así, consideremos el siguiente ejemplo:

3) El conjunto de todas las rotaciones de un cuerpo rígido en 3 dimensiones. Este conjunto es una variedad tridimensional ya que cada uno de sus puntos, es decir, cada rotación de un cuerpo rígido, puede ser especificada completamente por los tres ángulos de Euler, que en éste caso juegan el papel de parámetros.

4) El conjunto de todas las transformaciones de Lorentz entre dos sistemas de referencia inerciales con velocidad relativa v .

Este conjunto es una variedad tridimensional ya que como parámetros podemos tomar las tres componentes de la velocidad relativa $v=(v^1, v^2, v^3)$.

5) El espacio fase de un sistema mecánico de N partículas. Un punto de este espacio fase consiste de todas las posiciones y velocidades que en un momento dado adoptan todas las partículas del sistema. Se necesita un total de $6N$ números reales para especificar cada uno de los puntos del espacio fase. Es, por lo tanto, una variedad $6N$ -dimensional.

6) Un espacio vectorial real de n dimensiones. En este caso, tomemos cualquier base $\{e_1, e_2, \dots, e_n\}$ del espacio vectorial real de n dimensiones V . Entonces, cualquier vector $x \in V$ se puede expresar de manera única como una combinación lineal

$$x = a^1 e_1 + a^2 e_2 + \dots + a^n e_n \quad \text{con } a_i \in \mathbb{R}.$$

De esta manera queda establecida una biyección entre los elementos de V y los elementos de \mathbb{R}^n . Si $x = a^1 e_1 + \dots + a^n e_n$ y $y = b^1 e_1 + \dots + b^n e_n$ son dos elementos cualesquiera de V , definimos la distancia $d(x, y)$ entre ellos de acuerdo a la ecuación

$$d(x, y) = ((a^1 - b^1)^2 + \dots + (a^n - b^n)^2)^{1/2}$$

Con esta definición de distancia es posible definir vecindades de puntos en V y es obvio que la topología que se obtiene en V de esta manera será homeomorfa a la topología usual de \mathbb{R}^n .

En V tomaremos como atlas el que consiste de la única carta (V, I_V) en donde I_V es la función identidad sobre V , es decir, $I_V(x) = x$ para todo $x \in V$.

Como un caso particular de este ejemplo podemos considerar al propio espacio euclidiano \mathbb{R}^n , así pues, \mathbb{R}^n es una variedad diferenciable n -dimensional y en muchos sentidos puede ser considerada como la variedad más sencilla.

7) La topología de la esfera S^2 como un subespacio de \mathbb{R}^3 . Esta topología se define como sigue: N es vecindad de $x \in S^2$ si y solo si $N = N' \cap S^2$, donde N' es vecindad de x de acuerdo a la topología usual de \mathbb{R}^3 . Con esta definición es fácil demostrar que los conjuntos D_{x^+} , W y V de los ejemplos 1) y 2) son abiertos en la topología de S^2 como subespacio de \mathbb{R}^3 . En efecto, tenemos por ejemplo, que $D_{x^+} = D_N$, el hemisferio norte de S^2 , es la intersección de S^2 con $\{(x^1, x^2, x^3) \mid x^3 > 0\} = N'$. Pero en la topología usual de \mathbb{R}^3 , este conjunto N' es una vecindad de cada uno de sus propios puntos, entonces, por definición, $D_{x^+} \subset S^2$ es vecindad de cada uno de sus propios puntos y, por lo tanto, es un conjunto abierto según la topología de S^2 como subespacio de \mathbb{R}^3 .

Algunas variedades, que ejemplificaremos con el caso de la esfera, poseen la siguiente propiedad topológica: si $\{W_\alpha\}$ es una familia arbitraria de abiertos de la esfera con la propiedad de que $S^2 = \bigcup_\alpha W_\alpha$, entonces en la familia $\{W_\alpha\}$ existe una subfamilia finita

$W_{\alpha_1}, W_{\alpha_2}, \dots, W_{\alpha_n}$ que también cubre a la esfera, es decir, $S^2 = W_{\alpha_1} \cup \dots \cup W_{\alpha_n}$. Tal propiedad se llama compacidad.

Un subconjunto C de un espacio topológico T se llama cerrado si es el complemento de algún conjunto abierto A de T . Un subconjunto B de \mathbb{R}^n se llama acotado si B está contenido en el disco abierto con centro en el origen para algún radio $r > 0$.

El conocido teorema de Heine y Borel (Ver [S] en la bibliografía), afirma que un subconjunto C de \mathbb{R}^n es compacto si y solo si C es cerrado y acotado. Aplicando éste resultado a la esfera S^2 vemos que S^2 es compacta.

Por otro lado tenemos que el espacio \mathbb{R}^2 no es compacto. En efecto, consideremos la cubierta $\{B_r(n, m) \mid n, m \in \mathbb{Z}\}$ que consta de la familia de todas las bolas abiertas con centro en $(n, m) \in \mathbb{R}^2$ con n y m números enteros y radio $r=1$. Si quitamos de la familia $\{B_r(n, m)\}$ una sola bola, digamos $B_r(n', m')$, entonces el punto $(n', m') \in \mathbb{R}^2$ queda sin cubrir (recordemos que $B_r(n, m)$ no contiene a su frontera). Así pues, la familia $\{B_r(n, m)\}$ no puede contener a una subfamilia finita que cubra a \mathbb{R}^2 , luego, \mathbb{R}^2 no es compacto.

Ahora bien, nos preguntamos si es posible encontrar para la esfera S^2 un atlas que contenga una sola carta. La respuesta es negativa. Pues, en caso de existir, una carta $h: S^2 \rightarrow \mathbb{R}^2$ sería un homeomorfismo entre S^2 y \mathbb{R}^2 , y recordando nuestra discusión sobre los homeomorfismos, vemos que S^2 y \mathbb{R}^2 serían topológicamente equivalentes y tendrían que compartir todas sus propiedades topológicas. Pero hemos visto que S^2 es compacta, mientras que \mathbb{R}^2 no lo es. Por lo tanto S^2 no puede admitir un atlas con una sola carta.

Esta discusión muestra claramente la diferencia que hay entre las propiedades locales y las propiedades globales de las variedades. S^2 es localmente idéntica a \mathbb{R}^2 , de hecho, es localmente idéntica a cualquier variedad bidimensional. En éste nivel del desarrollo de la teoría de variedades la única propiedad determinante a nivel local es la dimensión. Sin embargo, a nivel global, S^2 y \mathbb{R}^2 son muy diferentes.

Para que no nos quedemos con la idea, errónea, de que cualquier cosa que parece un "espacio" es una variedad, consideremos el ejemplo siguiente:

9) El cono $C \subset \mathbb{R}^3$. Este espacio se define como el conjunto de los puntos de \mathbb{R}^3 tales que $(x^1)^2 - (x^2)^2 - (x^3)^2 = 0$ junto con la topología que C tiene como subespacio de \mathbb{R}^3 . En otras palabras, N es una vecindad en la topología del cono si y solo si $N = N' \cap C$, donde $N' \subset \mathbb{R}^3$ es una vecindad en la topología usual de \mathbb{R}^3 . Pues bien, éste espacio topológico de Hausdorff no es una variedad. (Ver fig. 4). El cono C no es un espacio localmente euclidiano. En efecto, consideremos cualquier vecindad del origen en \mathbb{R}^3 , como N_1 en la fig. 4. Esta vecindad intersecta al cono C en un conjunto N_2 que no es homeomorfo a un disco de \mathbb{R}^2 , más bien, N_2 es homeomorfo a 2 discos planos pegados en sus respectivos centros.

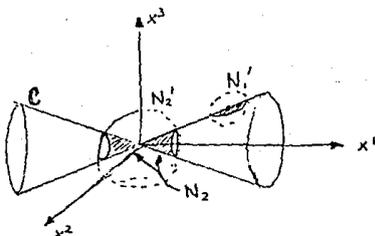


FIG. 4

N_2 no puede ser homeomorfa a un disco de \mathbb{R}^2 , pues en éste, la eliminación de un solo punto no provoca la aparición de dos partes disjuntas. En cambio, si en N_1 eliminamos el vértice del cono, N_2 se desconecta en dos partes. Así pues, N_2 y \mathbb{R}^2 (o un disco de \mathbb{R}^2) no pueden ser homeomorfos.

No hemos incluido en nuestra definición de variedad diferenciable un axioma que especifique la numerabilidad de las cartas de un atlas, pero en varias situaciones de la teoría de variedades éste axioma es necesario y es generalmente incluido.

Ahora consideraremos dos formas simples de obtener nuevas variedades a partir de variedades dadas.

10) Subvariedad abierta. Primero hagamos la observación de que si ξ es un sistema coordenado en la variedad M y si V es un conjunto abierto contenido en el dominio de ξ , entonces la restricción de ξ a V denotada por $\xi|_V$, es también un sistema coordenado en M . Sea ahora U un subconjunto abierto de una variedad diferenciable M . Sea A' el conjunto de todos los sistemas coordenados ξ en M tales que el dominio de ξ está contenido en U . Por la observación que hemos hecho al comenzar éste apartado, vemos que el conjunto A' es un atlas sobre U convirtiéndolo, por lo tanto, en una variedad diferenciable llamada subvariedad abierta de M . Los subconjuntos abiertos de una variedad M pueden ser siempre considerados como subvariedades abiertas de M .

11) Producto de variedades. Si M y N son variedades, sean $\xi = (x^1, \dots, x^m): U \rightarrow \mathbb{R}^m$ y $\eta = (y^1, \dots, y^n): V \rightarrow \mathbb{R}^n$ sistemas de coordenadas en M y N , respectivamente. Definamos la función producto $\xi \times \eta (p, q) = (x^1(p), \dots, x^m(p), y^1(q), \dots, y^n(q))$, $\xi \times \eta: M \times N \rightarrow \mathbb{R}^{m+n}$. Evidentemente $\xi \times \eta$ es un sistema coordenado con dominio $U \times V \subset M \times N = \{(p, q) \mid p \in M, q \in N\}$. El espacio $M \times N$ es de Hausdorff y es fácil ver que cualesquiera dos de tales sistemas coordenados producto en $M \times N$ cumplen la condición de compatibilidad. $M \times N$ con el atlas que contiene a todos éstos sistemas coordenados producto, es una variedad, llamada la variedad producto de M y N .

La dimensión de $M \times N$ es $\dim M + \dim N$. Esta construcción se puede extender de manera obvia al producto de cualquier número finito de variedades.

3 APLICACIONES DIFERENCIABLES

Así como un espacio topológico posee una estructura (su topología) por medio de la cual es posible hablar de la continuidad de las funciones que aplican unos espacios topológicos en otros, así, una variedad, gracias a su estructura diferenciable, permite generalizar a las aplicaciones entre variedades los conceptos del cálculo diferencial.

Consideremos primero el caso especial de una función f con valores reales y definida sobre una variedad M . Si $\xi:U \rightarrow \mathbb{R}^m$ es un sistema coordenado en M , entonces la función compuesta $f \circ \xi^{-1}: \xi(U) \rightarrow \mathbb{R}$ definida sobre la imagen $\xi(U)$ de ξ en \mathbb{R}^m se llama la expresión coordenada de f en términos de ξ . De hecho, tenemos

$$f = (f \circ \xi^{-1})(x^1, x^2, \dots, x^m) \text{ sobre } U.$$

Es natural entonces definir una función $f: M \rightarrow \mathbb{R}$ como diferenciable ó suave si para todo sistema coordenado ξ en M la expresión coordenada $f \circ \xi^{-1}$ es diferenciable en el sentido usual del cálculo elemental de varias variables. Denotaremos por $\mathcal{F}(M)$ al conjunto de todas las funciones suaves con valores reales definidas sobre la variedad M . Si f y g son funciones suaves sobre M , también lo serán las funciones $f+g$ y fg . Las reglas algebraicas usuales son válidas para éstas dos operaciones de manera que el conjunto $\mathcal{F}(M)$ es un anillo conmutativo. Sin embargo, los inversos multiplicativos no existen en general, pero si $f \in \mathcal{F}(M)$ es tal que nunca se anula sobre M , entonces $1/f \in \mathcal{F}(M)$.

La noción de diferenciable se puede extender de las funciones con valores reales a las aplicaciones arbitrarias entre variedades usando la misma idea: que las expresiones coordenadas sean diferenciables en el sentido usual del cálculo.

Llegamos así a la siguiente definición:

Sean M^m y N^n variedades de dimensiones m y n , respectivamente. Una aplicación $\phi: M \rightarrow N$ es diferenciable o suave si para cualesquiera sistemas coordenados ξ en M y η en N la expresión coordenada $\eta \circ \phi \circ \xi^{-1}$ es diferenciable, siendo una aplicación definida en un subconjunto abierto de \mathbb{R}^m que toma valores en \mathbb{R}^n .

Explícitamente, si $U \subset M$ y $V \subset N$ son los dominios de ξ y de η , entonces para todo $p \in \phi^{-1}(V) \cap U$ las coordenadas $y^1(\phi p), \dots, y^n(\phi p)$ asignadas por η a $\phi(p)$, dependen suavemente de las coordenadas $x^1(p), \dots, x^m(p)$ asignadas por ξ a p .

Hagamos ahora algunas observaciones.

1) Es suficiente verificar la condición de suavidad para aquellos sistemas coordenados que cubren a M y a N ya que la condición de compatibilidad de la definición de variedad, automáticamente comprueba la diferenciable de todas las expresiones coordenadas de ϕ , con respecto a todas las demás cartas en las estructuras diferenciables de M y N .

2) Una función diferenciable $\psi: U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ es también diferenciable en el sentido generalizado de la definición anterior, ya que

ψ es su propia expresión coordinada con respecto a los sistemas coordinados I_U e $I_{\mathbb{R}^n}$ que cubren a $U \subset \mathbb{R}^m$ y a \mathbb{R}^n , en donde I_U e $I_{\mathbb{R}^n}$ son las identidades sobre U y \mathbb{R}^n , respectivamente.

- 3) La aplicación identidad I_M sobre una variedad M es suave; cualquier composición de aplicaciones suaves es una aplicación suave.
- 4) Los sistemas coordinados ξ y las funciones coordinadas x^i son ejemplos de aplicaciones suaves sobre el dominio de ξ .
- 5) La diferenciable o suavidad es una propiedad local. En efecto, definimos $\phi: M \rightarrow N$ suave en $p \in M$ si la restricción de ϕ a una vecindad V de p , denotada por $\phi|_V$, es una aplicación suave. Entonces tenemos que ϕ es suave si y solo si ϕ es suave en todo punto $p \in M$.
- 6) Las aplicaciones suaves son continuas.

La siguiente consecuencia de la observación 5) será usada con frecuencia más adelante. Para cada índice $\alpha \in A$ sea U_α un conjunto abierto en una variedad M y sea $\phi_\alpha: U_\alpha \rightarrow N$ una aplicación suave. Si para todo par de índices $\alpha, \beta \in A$, $\phi_\alpha = \phi_\beta$ sobre $U_\alpha \cap U_\beta$, entonces éstas aplicaciones se combinan para dar una sola aplicación $\phi: \bigcup U_\alpha \rightarrow N$ tal que $\phi|_{U_\alpha} = \phi_\alpha$ para toda $\alpha \in A$. Puesto que la suavidad es una propiedad local, ϕ es suave.

Un difeomorfismo $\phi: M \rightarrow N$ es una aplicación suave que tiene una aplicación inversa que también es suave.

Las aplicaciones identidad sobre variedades, las composiciones de difeomorfismos y la aplicación inversa de un difeomorfismo dado, son, a su vez, difeomorfismos. Si existe un difeomorfismo ϕ de M a N , entonces M y N se dicen difeomorfas bajo ϕ . Por ejemplo, cualquier intervalo abierto (a, b) en \mathbb{R} es difeomorfo a $(-1, 1)$ bajo una transformación lineal adecuada, y $(-1, 1)$ es difeomorfo a \mathbb{R} bajo $\phi(t) = t/(1-t^2)$.

La topología diferencial o teoría de variedades puede definirse como el estudio de aquellas propiedades que permanecen invariantes bajo difeomorfismos, del mismo modo que la topología puede ser considerada como el estudio de aquellas propiedades de los espacios topológicos que permanecen invariantes bajo homeomorfismos. Así pues, variedades difeomorfas se consideran idénticas desde el punto de vista de la topología diferencial.

Si ϕ es una función biyectiva de un conjunto C sobre una variedad M , entonces evidentemente hay sólo una manera de convertir a C en una variedad (esto es, hay una única topología y una única estructura diferenciable sobre C) de modo que ϕ sea un difeomorfismo.

Puesto que las aplicaciones suaves son continuas, un difeomorfismo es en particular un homeomorfismo. Sin embargo, un homeomorfismo suave no tiene porqué ser un difeomorfismo, ya que su inversa no tiene porqué ser diferenciable. El ejemplo clásico de ésta situación es la aplicación $t \rightarrow t^3$ sobre la recta real.

Todo sistema coordenado ξ es un difeomorfismo de su dominio U a $\xi(U) \subset \mathbb{R}^n$. Inversamente, todo difeomorfismo ϕ de un conjunto abierto $V \subset M$ a $\phi(V) \subset \mathbb{R}^n$ es un sistema coordenado de M . De hecho, para cualquier sistema coordenado ξ las aplicaciones $\phi \circ \xi^{-1}$ y $\xi \circ \phi^{-1}$ son suaves y, por la condición de compatibilidad, se sigue que ϕ pertenece a la estructura diferenciable de M . Así, si componemos un sistema coordenado con un difeomorfismo, obtenemos de nuevo un sistema coordenado. En particular, si $p \in M$ siempre existe un sistema coordenado tal que $\xi(p) = 0$.

El soporte de una función $f \in \mathcal{F}(M)$, denotado $\text{supp } f$, se define como el más pequeño conjunto cerrado de M que contiene al conjunto $\{p \in M \mid f(p) \neq 0\}$. Entonces, $M - \text{supp } f$ es el más grande subconjunto abierto de M sobre el cual f es idénticamente nula.

El anillo conmutativo $\mathcal{F}(M)$ de todas las funciones real-valoradas, suaves y definidas sobre M , contiene una gran riqueza de funciones. Para tener una idea de esto, demostraremos que dada cualquier vecindad U de un punto p en M existe una función $f \in \mathcal{F}(M)$, llamada un pronotorio en p , y que tiene las siguientes propiedades:

- 1) $0 \leq f \leq 1$ sobre M .
- 2) $f = 1$ sobre alguna vecindad de p .
- 3) $\text{supp } f \subset U$.

Demostración. Comencemos considerando la función suave que es igual a e^{-1/t^2} si $t > 0$, y es igual a cero si $t < 0$. Usando el cálculo elemental es fácil construir a partir de esta función, para cualquier $\epsilon > 0$, una función suave h sobre \mathbb{R} tal que $h(t) = 1$ para $t \leq \epsilon$, $h(t) = 0$ para $t \geq 2\epsilon$, y $0 \leq h \leq 1$.

Ahora sea $\xi: V \rightarrow \mathbb{R}^n$ un sistema coordenado en M con $\xi(p) = 0$ y $V \subset U$; si $\epsilon > 0$ es suficientemente pequeño, entonces $\xi(V)$ contiene a la vecindad $\{p \in \mathbb{R}^n \mid \|p\|^2 < 3\epsilon\}$ de 0 en \mathbb{R}^n . Sea $g = \sum (x^i)^2$ sobre V . Para h como antes, definimos $f = h \circ g$ sobre V e igual a cero sobre $M - V$. Entonces la función f es suave y tiene las propiedades requeridas.

4 VECTORES TANGENTES

En el cálculo diferencial sobre \mathbb{R}^n nos encontramos con el concepto de derivada de una función en la dirección de un vector dado. Para definir los vectores tangentes a una variedad de manera intrínseca, es decir, sin hacer referencia a un espacio exterior a nuestra variedad, invertiremos la situación anterior. Así, en lugar de tener dados los vectores y, por medio de ellos, definir la derivada direccional de las funciones, definiremos los vectores tangentes a una variedad por medio de las propiedades abstractas que debe tener una derivada direccional de funciones definidas sobre la variedad. Hacemos entonces la siguiente definición.

Sea p un punto de una variedad M . Un vector tangente a M en p

es una función $v: \mathcal{F}(M) \rightarrow \mathbb{R}$ que es

1) \mathbb{R} -lineal: $v(af+bg) = av(f) + bv(g)$, y

2) Leibniziana: $v(fg) = v(f)g(p) + f(p)v(g)$ para todo $a, b \in \mathbb{R}$ y $f, g \in \mathcal{F}(M)$.

En cada punto $p \in M$ sea $T_p(M)$ el conjunto de todos los vectores tangentes a M en p . Por medio de las definiciones usuales de adición de funciones y de multiplicación de funciones por escalares (números) reales, $T_p(M)$ se convierte en un espacio vectorial sobre los números reales. En forma explícita, definimos

$$(v+w)(f) = v(f) + w(f),$$

$$(av)(f) = av(f) \text{ para toda } f \in \mathcal{F}(M), a \in \mathbb{R},$$

y $T_p(M)$ se llama el espacio tangente a M en p .

Para definir la derivación parcial sobre una variedad, la idea es considerar a la función f como definida sobre el espacio euclídeo \mathbb{R}^n , mediante una de sus expresiones coordenadas y tomar la derivada parcial usual.

Sea $\xi = (x^1, \dots, x^n)$ un sistema coordenado de M en p . Si $f \in \mathcal{F}(M)$, sea

$$\frac{\partial f}{\partial x^i}(p) = \frac{\partial (f \circ \xi^{-1})}{\partial u^i}(p) \quad (1 \leq i \leq n),$$

donde u^1, \dots, u^n son las funciones coordenadas naturales de \mathbb{R}^n , las cuales son tales que $u^i(x) = x^i$ para todo $x \in \mathbb{R}^n$.

Un cálculo sencillo demuestra que la función

$$\partial_i|_p = \frac{\partial}{\partial x^i} \Big|_p: \mathcal{F}(M) \rightarrow \mathbb{R}$$

que envía cada $f \in \mathcal{F}(M)$ a $(\partial f / \partial x^i)(p)$ es un vector tangente a M en p . Podemos visualizar a $\partial_i|_p$ como una flecha en p tangente a la curva coordenada x^i que pasa por p .

Lema 4.1 Sea $v \in T_p(M)$. (1) Si $f, g \in \mathcal{F}(M)$ son iguales sobre una vecindad de p , entonces $v(f) = v(g)$. (2) Si $h \in \mathcal{F}(M)$ es constante sobre una vecindad de p , entonces $v(h) = 0$.

Demostración. Por la linealidad de v es suficiente demostrar que si $f=0$ sobre una vecindad U de p , entonces $v(f)=0$. Sea g una función protuberancia en p con soporte en U ; entonces $fg=0$ sobre toda M . Pero $v(0) = v(0+0) = v(0) + v(0)$ implica que $v(0)=0$. Así pues,

$$0 = v(fg) = v(f)g(p) + f(p)v(g) = v(f),$$

ya que $f(p)=0$ y $g(p)=1$. Esto demuestra la afirmación (1).

Para demostrar la segunda afirmación, vemos que por (1) podemos suponer que h es constante $=c$ para toda M . Si 1 es la función constante con valor 1, entonces

$$v(1) = v(1 \cdot 1) = v(1)1 + 1v(1) = 2v(1).$$

De donde, $v(1)=0$ y $v(h)=v(c \cdot 1)=cv(1)=0$. ##

El lema anterior es una forma de expresar el hecho de que los vectores tangentes son objetos locales. Otra forma de expresarlo es ésta: Si U es un conjunto abierto en M entonces, siendo una subvariedad abierta, U tiene un espacio tangente $T_p(U)$ en $p \in U$. Si $v \in T_p(U)$ definimos $\forall(f)=v(f|U)$ para toda $f \in \mathcal{F}(M)$. Evidentemente $\forall \in T_p(M)$, y la función $v \rightarrow \forall$ es un isomorfismo lineal. A partir de ahora ignoraremos éste isomorfismo y escribiremos $T_p(U)=T_p(M)$.

El siguiente resultado, llamado teorema de la base, es el enlace fundamental entre las coordenadas y los vectores tangentes.

Teorema 4.2 Si $\xi=(x^1, \dots, x^n)$ es un sistema coordenado de M en p , entonces sus vectores coordenados $\partial_1|_p, \dots, \partial_n|_p$ forman una base para el espacio tangente $T_p(M)$; y

$$v = \sum_{i=1}^n v(x^i) \partial_i|_p \quad \text{para todo } v \in T_p(M).$$

Demostración. Por los comentarios anteriores, podemos limitarnos a trabajar en la vecindad coordenada que es el dominio U de ξ . Como $v(c)=0$, no hay pérdida de generalidad si suponemos que $\xi(p) \in \mathbb{R}^n$. Considerando un dominio más pequeño que U si es necesario, podemos suponer que $\xi(U) = \{q \in \mathbb{R}^n \mid \|q\| < \epsilon\}$ para algún ϵ .

Si g es una función suave sobre $\xi(U)$ entonces para cada $1 \leq i \leq n$ definimos

$$g_i(q) = \int_0^1 \frac{\partial g}{\partial u^i}(tq) dt \quad \text{para todo } q \in \xi(U).$$

Se sigue utilizando el teorema fundamental del cálculo que

$$g = g(0) + \sum g_i u^i \quad \text{sobre } \xi(U).$$

Así si $f \in \mathcal{F}(M)$, haciendo $g = f \circ \xi^{-1}$ se obtiene

$$f = f(p) + \sum f_i x^i \quad \text{sobre } U.$$

Aplicando $\partial/\partial x^i$ se obtiene $f_i(p) = (\partial f / \partial x^i)(p)$. De manera que si aplicamos el vector tangente v a ésta fórmula tenemos

$$v(f) = 0 + \sum v(f_i) x^i(p) + \sum f_i(p) v(x^i) = \sum v(x^i) \frac{\partial f}{\partial x^i}(p).$$

Dado que esto es cierto para toda $f \in \mathcal{F}(M)$, los vectores tangentes v y $\sum v(x^i) \partial_i|_p$ son iguales. Con esto queda demostrado que los vectores coordenados $\partial_1|_p, \dots, \partial_n|_p$ generan al espacio $T_p(M)$.

Falta demostrar que los vectores coordenados son linealmente independientes. Pero si $\sum a_i \partial_i|_p = 0$, entonces aplicando ésta combinación lineal a x^j obtenemos

$$0 = \sum a_i \frac{\partial x^j}{\partial x^i}(p) = \sum a_i \delta_{ij} = a_j. \quad \#\#$$

En particular vemos que la dimensión del espacio vectorial $T_p(M)$ es la misma que la dimensión de la variedad M .

5 LA DIFERENCIAL

La idea básica del cálculo diferencial es aproximar los objetos suaves o diferenciables por objetos lineales. En la sección anterior vimos como una variedad diferenciable puede ser aproximada en la vecindad de cada uno de sus puntos $p \in M$ por el espacio lineal $T_p(M)$. Ahora aproximaremos a una aplicación diferenciable $\phi: M \rightarrow N$ en la vecindad de cada punto $p \in M$ por medio de una transformación lineal entre espacios tangentes. Hagamos notar primero que si $v \in T_p(M)$ entonces la función $v_\phi: \mathcal{F}(N) \rightarrow \mathbb{R}$, que envía cada $g \in \mathcal{F}(N)$ a $v(g \circ \phi)$, es un vector tangente a N en $\phi(p)$. Debemos verificar que v_ϕ satisface las propiedades 1) y 2) de la definición de vector tangente (pag. 13). En efecto, para probar la propiedad Leibniziana, por ejemplo, sean $f, g \in \mathcal{F}(N)$; entonces

$$\begin{aligned} v_\phi(fg) &= v(fg \circ \phi) = v((f \circ \phi)(g \circ \phi)) = v(f \circ \phi)g(\phi p) + f(\phi p)v(g \circ \phi) = \\ &= v_\phi(f)g(\phi p) + f(\phi p)v_\phi(g). \end{aligned}$$

Así pues, hacemos la siguiente definición. Sea $\phi: M \rightarrow N$ una aplicación suave. Para cada $p \in M$ la función

$$d\phi_p: T_p(M) \rightarrow T_{\phi(p)}(N)$$

que envía v a v_ϕ (definida arriba) se llama la diferencial de ϕ en p . (Ver fig. 5).

Entonces, $d\phi_p$ está caracterizada por la ecuación

$$d\phi_p(v)(g) = v(g \circ \phi)$$

para todo $v \in T_p(M)$ y toda $g \in \mathcal{F}(N)$. De aquí es obvio que la diferencial de ϕ en p es una transformación lineal. Para toda $g \in \mathcal{F}(N)$ y cualquier $\phi: M \rightarrow N$ suave, la composición $g \circ \phi$ pertenece a $\mathcal{F}(M)$ y se llama la retracción de g mediante ϕ , denotada por $\phi^*(g)$. A veces, la diferencial de ϕ en p se denota por ϕ_{*p} . Con estas notaciones la diferencial de ϕ en p queda definida por la ecuación

$$\phi_{*p}(v)(g) = v(\phi^*(g)) \quad \text{para todo } v \in T_p(M) \text{ y toda } g \in \mathcal{F}(N).$$

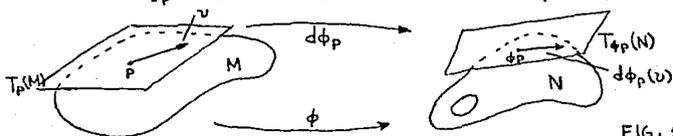


FIG. 5

Lema 5.1 Sea $\phi: M \rightarrow N$ una aplicación suave. Si ξ es un sistema coordenado de M en p , y si η es un sistema coordenado de N en $\phi(p)$, entonces

$$d\phi_p \left(\frac{\partial}{\partial x^i} \Big|_p \right) = \sum_{j=1}^n \frac{\partial (y^j \circ \phi)}{\partial x^j} (p) \frac{\partial}{\partial y^j} \quad (1 \leq i \leq m).$$

Demostración. Denotemos por $w \in T_{p,p}(N)$ al lado izquierdo de esta ecuación. Por el teorema de la base (teorema 4.2), $w = \sum w(y^i) \partial/\partial y^i$. Pero por la definición de diferencial,

$$w(y^i) = d\phi_p \left(\frac{\partial}{\partial x^j} \right) (y^i) = \frac{\partial (y^i \circ \phi)}{\partial x^j} (p). \quad \#\#$$

La matriz de la aplicación lineal $d\phi_p$ con respecto a éstas bases coordenadas es

$$\left[\frac{\partial (y^i \circ \phi)}{\partial x^j} (p) \right]_{1 \leq i \leq n, 1 \leq j \leq m}$$

y se llama la matriz jacobiana de ϕ en p con respecto a ξ y η .

La clásica regla de la cadena para la matriz jacobiana de una composición de aplicaciones se sigue inmediatamente de la fórmula que demostraremos en el siguiente lema.

Lema 5.2 Si $\phi: M \rightarrow N$ y $\psi: N \rightarrow P$ son aplicaciones suaves, entonces para cada $p \in M$,

$$d(\psi \circ \phi)_p = d\psi_p \circ d\phi_p.$$

Demostración. Si $v \in T_p(M)$ y $g \in T_p(P)$, entonces

$$d(\psi \circ \phi)(v)(g) = v(g \circ \psi \circ \phi) = d\phi(v)(g \circ \psi) = (d\psi d\phi(v))(g). \quad \#\#$$

De aquí en adelante y cuando no haya lugar a confusión, omitiremos el subíndice p de $d\phi_p$.

En términos de la teoría de variedades el teorema de la función inversa puede ser reformulado en la siguiente forma.

Teorema 5.3 Sea $\phi: M \rightarrow N$ una aplicación suave. La diferencial $d\phi_p$ en un punto $p \in M$ es un isomorfismo lineal si y solo si existe una vecindad V de p en M tal que $\phi|_V$ es un difeomorfismo de V sobre una vecindad $\phi(V)$ de $\phi(p)$ en N .

Para demostrar ésto es suficiente aplicar el teorema clásico de la función inversa (consultar [Sp], pag. 32) a la expresión coordenada de ϕ en una vecindad de p .

Debido a éste resultado una aplicación suave $\phi: M \rightarrow N$ tal que toda $d\phi_p$ es un isomorfismo lineal, se llama un difeomorfismo local. Si ϕ es también inyectiva y suprayectiva, entonces es un difeomorfismo.

6 CURVAS

Una curva en una variedad M es una aplicación diferenciable $\alpha: I \rightarrow M$, donde I es un intervalo abierto de la recta real \mathbb{R} . (Se permite la posibilidad de que I sea una semirecta infinita o toda la recta \mathbb{R} .) Siendo una subvariedad abierta de \mathbb{R} , I tiene un sistema coordenado compuesto por la función coordenada natural

u de \mathbb{R} . En cada $t \in \mathbb{R}$ podemos visualizar el vector coordenado $(d/du)(t) \in T_t(\mathbb{R})$ como un vector unitario basado en t y con la dirección del eje u positivo.

Sea $\alpha: I \rightarrow M$ una curva. El vector velocidad de α en $t \in I$ está definido por

$$\alpha'(t) = d\alpha \left(\left. \frac{d}{du} \right|_t \right) \in T_{\alpha(t)}(M).$$

Intuitivamente, $\alpha'(t)$ es la razón de cambio vectorial de α en $t \in I$. Ennumeremos algunas de sus propiedades básicas:

1) Derivada direccional. Por la definición de $d\alpha$, el vector tangente $\alpha'(t)$ aplicado a la función $f \in \mathcal{F}(M)$ nos da

$$\alpha'(t)f = \frac{d(f \circ \alpha)}{du}(t).$$

De éste modo, si α es cualquier curva con, digamos, $\alpha'(0) = v$ entonces, $v(f) = (d(f \circ \alpha)/dt)(0)$.

2) Expresión coordinada. Sea $\xi = (x^1, \dots, x^n)$ un sistema coordinado de M en un punto $\alpha(t)$ de α . Por el teorema de la base y por la observación 1), tenemos

$$\alpha'(t) = \sum_{i=1}^n \frac{d(x^i \circ \alpha)}{du}(t) \delta_i \Big|_{\alpha(t)}$$

3) Reparametrización. Si $\alpha: I \rightarrow M$ es una curva y $h: J \rightarrow I$ es una función suave sobre un intervalo abierto J , entonces $\beta = \alpha \circ h: J \rightarrow M$ es una curva llamada una reparametrización de α . Además,

$$\beta'(s) = (dh/du)(s) \alpha'(h(s)) \quad \text{para todo } s \in J.$$

Esto se sigue de la regla de la cadena (lema 5.2), al igual que la siguiente observación.

4) Efecto de una aplicación. Si $\alpha: I \rightarrow M$ es una curva en M , entonces la aplicación suave $\phi: M \rightarrow N$ transforma la curva α en la curva $\phi \circ \alpha: I \rightarrow N$ en N . La diferencial de ϕ preserva velocidades, es decir,

$$\alpha \phi(\alpha'(t)) = (\phi \circ \alpha)'(t) \quad \text{para todo } t \in I.$$

Esta es en general una manera más eficiente de obtener información acerca de $d\phi$ que los cálculos con coordenadas como los del lema 5.1.

Una curva α es regular si $\alpha'(t) \neq 0$ para todo t . Si $[a, b]$ es un intervalo cerrado en \mathbb{R} , entonces un segmento de curva $\alpha: [a, b] \rightarrow M$ es una aplicación que tiene una extensión diferenciable a un intervalo abierto que contiene a $[a, b]$. De éste modo α' está bien definida incluso en los puntos extremos a y b .

Una aplicación $\beta: [a, b] \rightarrow M$ es un segmento de curva suave por pedazos si existe una partición $a = t_0 < t_1 < \dots < t_{k-1} = b$ del intervalo $[a, b]$ tal que cada restricción $\beta|_{[t_i, t_{i+1}]}$ es un segmento de curva. Según ésto, β puede tener dos vectores velocidad en los puntos de ruptura t_i . Para un intervalo abierto I , $\beta: I \rightarrow M$ se llama suave por pedazos si para cualesquiera $a < b$ en I la restricción $\beta|_{[a, b]}$

es suave por pedazos en el sentido antes definido. (Así queda asegurado que los puntos de ruptura no pueden acumularse alrededor de un punto de I .)

Para ninguno de los tipos de curvas que hemos definido, la reparametrización $t \rightarrow x(t+c)$, con c constante, cambia en algo sus propiedades. Por eso, para simplificar la notación vamos a suponer en adelante, sin mencionarlo explícitamente, que el dominio de una curva contiene al número 0.

7 CAMPOS VECTORIALES

Un campo vectorial V sobre una variedad M es una función que asigna a cada punto p de M un vector tangente V_p de $T_p(M)$. Intuitivamente, V es una colección de flechas, una para cada punto de M y que pueden servir para representar fuerzas, velocidades, aceleraciones, etc. Si V es un campo vectorial sobre M y si $f \in \mathcal{F}(M)$, entonces Vf denota a la función de valores reales sobre M dada por

$$(Vf)(p) = V_p(f) \quad \text{para todo } p \in M.$$

Entonces, V se llama suave o diferenciable si Vf es suave para toda $f \in \mathcal{F}(M)$.

Los campos vectoriales sobre M pueden sumarse o ser multiplicados por funciones $f \in \mathcal{F}(M)$, de acuerdo a las siguientes reglas obvias:

$$\begin{aligned} (fV)_p &= f(p)V_p, \\ (V+W)_p &= V_p + W_p \end{aligned} \quad \text{para todo } p \in M.$$

Si V y W son suaves, también lo son los campos vectoriales $V+W$ y fV . Estas dos operaciones convierten al conjunto $\mathcal{V}(M)$ de todos los campos vectoriales suaves sobre la variedad M en un módulo sobre el anillo $\mathcal{F}(M)$. (La definición de módulo sobre un anillo conmutativo con unidad es formalmente la misma que la de espacio vectorial sobre un campo, con la diferencia de que los escalares provienen de un anillo en lugar de un campo.)

Si $\xi = (x^1, \dots, x^n)$ es un sistema coordenado en UCM, entonces para cada $1 \leq i \leq n$ el campo vectorial ∂_i sobre U que envía cada p a $\partial_i|_p$ se llama el i -ésimo campo vectorial coordenado de ξ . Estos campos vectoriales son suaves, ya que $\partial_i(f) = df/\partial x^i$. Se sigue inmediatamente del teorema de la base que para cualquier campo vectorial suave V

$$V = \sum V(x^i) \partial_i \quad \text{sobre UCM.}$$

Para variedades de dimensión pequeña la notación con índices puede ser difícil de interpretar. Por eso sobre el plano euclidiano, por ejemplo, frecuentemente denotaremos por x, y a las funciones coordenadas naturales como en el cálculo elemental; en éste caso, los campos vectoriales coordenados se denotan por ∂_x, ∂_y .

Una derivación en $\mathcal{F}(M)$ es una función $\mathcal{D}: \mathcal{F}(M) \rightarrow \mathcal{F}(M)$ que satisface las siguientes propiedades:

- 1) \mathbb{R} -lineal: $\mathcal{D}(af+bg) = a\mathcal{D}(f) + b\mathcal{D}(g)$, $(a, b \in \mathbb{R})$.
- 2) Leibniziana: $\mathcal{D}(fg) = \mathcal{D}(f)g + f\mathcal{D}(g)$.

La definición de vector tangente demuestra que para un campo vectorial $V \in \mathcal{K}(M)$ la función $f \rightarrow Vf$ es una derivación en $\mathcal{F}(M)$. Inversamente, toda derivación \mathcal{D} en $\mathcal{F}(M)$ proviene de un campo vectorial. De hecho, para cada $p \in M$ definamos $V_p: \mathcal{F}(M) \rightarrow \mathbb{R}$ mediante la ecuación $V_p(f) = \mathcal{D}(f)(p)$. Entonces, las propiedades 1) y 2) en la definición de derivación implican que V_p es un vector tangente a M en p ; así que V es un campo vectorial bien definido sobre M . Además $Vf = \mathcal{D}(f) \in \mathcal{F}(M)$ para toda $f \in \mathcal{F}(M)$, por lo que V es suave y determina a la derivación dada \mathcal{D} .

Así pues, siempre que sea conveniente, consideraremos que los campos vectoriales sobre M son derivaciones en $\mathcal{F}(M)$. Esta interpretación conduce a una operación crucial sobre los campos. Si $V, W \in \mathcal{K}(M)$ sea $[V, W] = VW - WV$. Esta es una función de $\mathcal{F}(M)$ a $\mathcal{F}(M)$ que envía cada f a $V(Wf) - W(Vf)$. Un cálculo directo muestra que $[V, W]$ es una derivación en $\mathcal{F}(M)$, es decir, un campo vectorial suave sobre M . $[V, W]$ se llama paréntesis (de Lie) de V y W .

En términos de la definición original de los campos vectoriales, $[V, W]$ asigna a cada $p \in M$ el vector tangente $[V, W]_p$ tal que

$$[V, W]_p(f) = V_p(Wf) - W_p(Vf).$$

Lema 7.1 El paréntesis de Lie tiene las siguientes propiedades:

- 1) \mathbb{R} -bilinealidad: $[aV+bW, X] = a[V, X] + b[W, X]$,
 $[X, aV+bW] = a[X, V] + b[X, W]$.
- 2) Antisimetría: $[W, V] = -[V, W]$.
- 3) Identidad de Jacobi: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

Demostración. Estas identidades son puramente formales y se cumplen para cualesquiera derivaciones sobre cualquier álgebra lineal: 1) es inmediata de la linealidad de las derivaciones, y 2) es obvia. Para probar 3) basta con sustituir la definición del paréntesis de Lie, entonces los doce términos resultantes se cancelan por parejas. ##

La operación de paréntesis sobre $\mathcal{K}(M)$ aún cuando es \mathbb{R} -bilineal no es $\mathcal{F}(M)$ -bilineal. De hecho,

$$[fV, gW] = fg[V, W] + f(Vg)W - g(Wf)V.$$

Para probar esto, hay que verificar que ambos miembros de esta ecuación tienen el mismo efecto sobre cualquier función $h \in \mathcal{F}(M)$.

En dos casos especiales el paréntesis es siempre nulo. Para cualquier $V \in \mathcal{K}(M)$, $[V, V] = 0$. De hecho, dada la bilinealidad, esto es equivalente a la antisimetría. Para cualesquiera dos campos vectoriales coordenados del mismo sistema coordenado, $[\partial_i, \partial_j] = 0$,

ya que ésto equivale a la identidad $\partial^i f / \partial x^i \partial x^k = \partial^2 f / \partial x^k \partial x^i$, que es válida para funciones f diferenciables de clase C^2 por lo menos.

Ejemplo. Sean x, y las funciones coordenadas naturales de \mathbb{R}^2 , y consideremos los campos vectoriales $V = y\partial_y$ y $W = x\partial_y$. Entonces $[V, W] = -W$, ya que para toda $f \in \mathcal{F}(M)$,

$$\begin{aligned} [V, W]f &= [y\partial_y, x\partial_y]f = y\partial_y(x\partial_y f) - x\partial_y(y\partial_y f) \\ &= yx \frac{\partial^2 f}{\partial y^2} - x \frac{\partial f}{\partial y} - xy \frac{\partial^2 f}{\partial y^2} = -x \partial_y f = -Wf. \end{aligned}$$

Cálculos de éste tipo pueden abreviarse omitiendo la función f y también las segundas derivadas ya que, como sabemos, deben cancelarse.

La diferencial de una aplicación $\phi: M \rightarrow N$ manda vectores tangentes individuales de M a N , pero en general la diferencial no proporciona una manera de mandar campos vectoriales de M a N (o de N a M). Esta dificultad puede ser solucionada en cierta medida mediante la siguiente definición.

Sea $\phi: M \rightarrow N$ una aplicación suave. Los campos vectoriales X sobre M y Y sobre N están ϕ -relacionados, lo que se denotará por $X \sim_\phi Y$, si

$$d\phi(X_p) = Y_{\phi p} \quad \text{para todo } p \in M.$$

Lema 7.2 Los campos vectoriales $X \in \mathcal{X}(M)$ y $Y \in \mathcal{X}(N)$ están ϕ -relacionados si y solo si $X(g \circ \phi) = Yg \circ \phi$ para toda $g \in \mathcal{F}(N)$.

Demostración. Las siguientes afirmaciones son equivalentes:

$$\begin{array}{ll} X(g \circ \phi) = Yg \circ \phi, & g \in \mathcal{F}(N), \\ X(g \circ \phi)(p) = (Yg \circ \phi)(p), & p \in M, \\ X_p(g \circ \phi) = (Yg)(\phi p), & g \in \mathcal{F}(N), \\ (d\phi X_p)g = Y_{\phi p}(g), & p \in M, \\ d\phi(X_p) = Y_{\phi p}. & g \in \mathcal{F}(N), \end{array} \quad \#\#$$

Usando éste criterio repetidamente, se prueba fácilmente que los paréntesis de Lie no cambian cuando los campos están ϕ -relacionados, concretamente, tenemos el siguiente lema:

Lema 7.3 Si $X_1 \sim_\phi Y_1$ y $X_2 \sim_\phi Y_2$, entonces $[X_1, X_2] \sim_\phi [Y_1, Y_2]$.

Cuando la aplicación $\phi: M \rightarrow N$ es un difeomorfismo la dificultad de la que hablábamos no se presenta, ya que para cada $X \in \mathcal{X}(M)$ existe un único campo vectorial $d\phi(X) \in \mathcal{X}(N)$ que está ϕ -relacionado a X . De hecho, la única posibilidad es definir $(d\phi X)_q = d\phi(X_p)$ para todo $q = \phi(p) \in N$. Entonces $d\phi(X)$ es suave, ya que si $g \in \mathcal{F}(N)$, ésta definición y la definición de diferencial nos conducen a la fórmula $(d\phi X)_g = X(g \circ \phi) \circ \phi^{-1} \in \mathcal{F}(N)$. Llamaremos a $d\phi(X)$ el campo vec-

torial transferido de X mediante .

8 UNO-FORMAS

Las uno-formas sobre una variedad suave M son los objetos duales a los campos vectoriales. En un punto p de M el espacio dual $T_p^*(M)$ del espacio tangente $T_p(M)$ es llamado el espacio cotangente de M en p . Los elementos de $T_p^*(M)$, a veces llamados covectores, son aplicaciones lineales de $T_p(M)$ a \mathbb{R} .

Una uno-forma θ sobre una variedad M es una función que asigna a cada punto p un elemento θ_p de el espacio cotangente $T_p^*(M)$.

Así θ asigna un número a todo vector tangente y es lineal sobre los vectores tangentes en cada punto.

Si θ es una uno-forma sobre M y X es un campo vectorial sobre M , denotaremos por θX la función de valores reales sobre M cuyo valor en cada punto p es el valor de θ_p sobre X_p . Una uno-forma es suave si θX es suave para todo $X \in \mathcal{X}(M)$.

Sea $\mathcal{F}(M)$ el conjunto de todas las uno-formas suaves sobre M . Dos uno-formas se suman, y una uno-forma se multiplica por una función de valores reales, en la misma manera que las correspondientes operaciones para campos vectoriales. Explícitamente,

$$(\theta + \omega)_p = \theta_p + \omega_p, \quad (f\theta)_p = f(p)\theta_p$$

para todo $p \in M$. Así pues, $\mathcal{F}(M)$ se convierte en un módulo sobre $\mathcal{F}(M)$.

Existe una operación notable que convierte funciones en uno-formas. La diferencial de $f \in \mathcal{F}(M)$ es la uno-forma df tal que $(df)(v) = v(f)$ para todo vector tangente v a M .

Claramente df es una uno-forma ya que en cada punto p , la función $(df)_p : T_p(M) \rightarrow \mathbb{R}$ es lineal, y si $V \in \mathcal{X}(M)$ la función $(df)(V) = V(f)$ es suave.

Si x^1, \dots, x^n es un sistema coordinado sobre $U \subset M$, tomando la diferencial de cada una de las funciones coordenadas $x^i \in \mathcal{F}(U)$, obtenemos las uno-formas coordenadas dx^1, \dots, dx^n sobre U . En cada punto de U éstas proporcionan una base a la base de los campos vectoriales coordenados $\partial_1, \dots, \partial_n$, ya que $dx^i(\partial_k) = \partial x^i / \partial x^k = \delta_{ik}$. Se sigue que para cualquier uno-forma θ ,

$$\theta = \sum_{i=1}^n \theta(\partial_i) \cdot dx^i \quad \text{sobre } U.$$

Esta fórmula es la correspondiente a la del teorema de la base para campos vectoriales. Para verificarla, basta con aplicarla a los campos vectoriales coordenados. Como un caso particular de esta fórmula, si $f \in \mathcal{F}(M)$, y como $df(\partial_i) = \partial f / \partial x^i$, obtenemos

$$df = \sum \frac{\partial f}{\partial x^i} dx^i \quad \text{sobre } U.$$

Lema 8.1 La diferencial de una función $f \in \mathcal{F}(M)$ tiene las siguientes propiedades:

- 1) $d: \mathcal{F}(M) \rightarrow \mathcal{F}^*(M)$ es \mathbb{R} -lineal.
- 2) Regla de Leibniz: Si $f, g \in \mathcal{F}(M)$, entonces $d(fg) = gdf + fdg$.
- 3) Si $f \in \mathcal{F}(M)$ y $h \in \mathcal{F}(\mathbb{R})$, entonces $d(h(f)) = h'(f)df$.

Demostración. Estas propiedades se siguen directamente de la definición. 3) es un caso especial del lema 5.2 (pag. 16). ##

9 SUBVARIETADES

A grandes rasgos, una subvariedad P de una variedad M es un subconjunto de M que tiene una estructura diferenciable determinada por la estructura diferenciable de M . En particular, requeriremos que P sea un subespacio topológico de M , es decir, que tenga la topología inducida. La topología inducida de P en M se define de la manera siguiente:

Un subconjunto V de P es abierto en la topología inducida si y solo si existe un subconjunto abierto V' en la topología de M tal que $V = V' \cap P$.

Una variedad P es una subvariedad de una variedad M si:

- 1) P es un subespacio topológico de M .
- 2) La aplicación inclusión $j: P \hookrightarrow M$ que manda cada $p \in P$ a $p = j(p) \in M$, es suave y en cada $p \in P$ su diferencial dj es inyectiva.

Si P es una subvariedad de M y $\phi: M \rightarrow N$ es una aplicación suave, entonces también es suave la restricción $\phi|_P$ de ϕ a P , ya que P es simplemente $\phi \circ j$. En particular, si $f \in \mathcal{F}(M)$, entonces $f|_P$ pertenece a $\mathcal{F}(P)$. Debido a que la inclusión j es una aplicación bastante obvia y a que cada $dj_p: T_p(P) \rightarrow T_p(M)$ es inyectiva, es costumbre ignorar dj y considerar que el espacio tangente $T_p(P)$ es un subespacio vectorial de $T_p(M)$.

Las subvariedades abiertas (que definimos en la sección 2, num. 10) son ejemplos triviales de subvariedades. La n -esfera S^n (ver sección 2, num. 1)) es una subvariedad de \mathbb{R}^{n+1} . Los sistemas coordenados pueden servir para producir ejemplos de subvariedades. Por ejemplo, el plano $z=1$ en \mathbb{R}^3 es una subvariedad, que es difeomorfa a \mathbb{R}^2 con respecto al difeomorfismo $(x, y, 1) \rightarrow (x, y)$. Más en general, si $\xi: U \rightarrow \mathbb{R}^n$ es un sistema coordenado en una variedad M , entonces manteniendo fijas cualesquiera $n-m$ de las funciones coordenadas de ξ , se obtiene una subvariedad m -dimensional llamada una rebanada coordenada de ξ . Ahora demostraremos que toda subvariedad puede ser construida 'pegando' tales rebanadas coordenadas.

Para precisar éstas nociones, hagamos primero una definición. Sea P un subconjunto de M . Un sistema coordenado $\xi: U \rightarrow \mathbb{R}^n$ en M está adaptado a P si $U \cap P$ es una rebanada coordenada de ξ en U .

Por medio de un cambio en los nombres de las coordenadas, podemos suponer que son las últimas $n-m$ funciones coordenadas de ξ las que se mantienen constantes. Entonces, sea $\xi_P: U \cap P \rightarrow \mathbb{R}^m$ la restricción de $\xi = (x^1, \dots, x^m)$ a $U \cap P$.

Cuando P es una subvariedad de M , se sigue del teorema de la función inversa (teorema 5.3), que ξ_P es un difeomorfismo sobre su imagen (abierto) en \mathbb{R}^m , y por lo tanto, un sistema coordenado en P .

Proposición 9.1 Si P^m es una subvariedad de M^n , entonces en cada punto de P hay un sistema coordenado de M adaptado a P .

Demostración. Sea $\xi = (x^1, \dots, x^n)$ un sistema coordenado para M en $p \in M$ y sea (y^1, \dots, y^m) un sistema coordenado para P en p . Puesto que la diferencial $d\xi_P$ es inyectiva, su matriz jacobiana

$$\left(\frac{\partial x^i}{\partial y^j}(p) \right)_{1 \leq i \leq n, 1 \leq j \leq m}$$

tiene rango m . Así, reetiquetando los x^i si es necesario, podemos suponer que los primeros m renglones de esta matriz son linealmente independientes. Entonces, la restricción a P de las funciones coordenadas x^1, \dots, x^m forman un sistema coordenado para P en una vecindad W de p . Si f^k es la expresión coordenada de $x^k|_P$ con respecto a (x^1, \dots, x^m) , entonces

$$x^k = f^k(x^1, \dots, x^m) \text{ sobre } W.$$

Se sigue que las funciones

$$z^k = x^k - f^k(x^1, \dots, x^m)$$

están bien definidas sobre una vecindad de p en M .

Si $\zeta = (x^1, \dots, x^m, z^{m+1}, \dots, z^n)$ entonces en p la matriz jacobiana de ζ con respecto a $\xi = (x^1, \dots, x^n)$ tiene la siguiente forma

$$\begin{pmatrix} I_m & 0 \\ A & I_{n-m} \end{pmatrix},$$

en donde I_m es la matriz identidad de orden m . Evidentemente esta matriz es invertible (su determinante es $\neq 0$), por lo cual es un sistema coordenado sobre una vecindad U de p en M . Como P es un subespacio topológico de M , podemos escoger U de manera que $U \cap P \subset W$. Como $A = (x^1, \dots, x^m)(U \cap P)$ es un subconjunto abierto de \mathbb{R}^m , tomando una U más pequeña si es necesario, podemos también suponer que $(x^1, \dots, x^m)(U) \subset A$.

Ahora bien, sobre $U \cap P$ las funciones z^k son todas cero, por lo que $U \cap P$ está contenido en la rebanada \sum de ζ en U dada por

$$z^{m+1} = 0, \dots, z^n = 0.$$

Inversamente, se verifica fácilmente que $\sum \subset U \cap P$. Por lo tanto, $U \cap P = \sum$. ##

El siguiente resultado, de apariencia trivial, depende del hecho de que las subvariedades tienen la topología inducida.

Corolario 9.2 Sea P^m una subvariedad de M^n . Si $\phi: N \rightarrow M$ es una aplicación suave tal que $\phi(N) \subset P$, entonces la aplicación inducida $\bar{\phi}: N \rightarrow P$ es suave también. (Si j es la inclusión de P en M , $\bar{\phi}$ es tal que $\phi = j \circ \bar{\phi}$.)

Demostración. Si $q \in N$ sea x^1, \dots, x^n un sistema coordinado sobre una vecindad $U \subset M$ de $\phi(q)$ que está adaptada a P . Siendo suave, ϕ es continua, y dado que P es un subespacio topológico de M , $\bar{\phi}$ es continua. Así que hay una vecindad V de q en N tal que $\bar{\phi}(V)$ está contenido en la vecindad $U \cap P$ de $\bar{\phi}(q)$ en P . Ahora bien, $x^1|_P, \dots, x^m|_P$ es un sistema coordinado sobre $U \cap P$. Además,

$$(x^i|_P) \circ \bar{\phi} = x^i \circ \phi = x^i \circ \phi$$

Estas funciones son suaves debido a que ϕ y las x^i son suaves. Se sigue que $\bar{\phi}$ es suave debido a que sus componentes $(x^i|_P) \circ \bar{\phi}$ son suaves. ##

Corolario 9.3 Un subconjunto P de una variedad suave M puede convertirse en una subvariedad de M de una sola manera.

Demostración. Por definición debe darse a P la topología inducida. Supongamos que es posible asignar dos estructuras diferenciales al espacio P con lo cual se obtienen dos subvariedades P_1 y P_2 de M . Las aplicaciones inclusión de P_1 y P_2 en M son suaves, entonces, por el corolario anterior, las aplicaciones identidad $P_1 \rightarrow P_1$ y $P_2 \rightarrow P_1$ son suaves. Por lo tanto, éstas identidades son difeomorfismos inversos y P_1 y P_2 son difeomorfas. Se sigue que P_1 y P_2 son idénticas como variedades. ##

Debido a esto, tiene sentido decir que un subconjunto P de una variedad M es (o no es) una subvariedad de M . Un criterio básico para hacer ésta determinación está dado por la siguiente proposición.

Proposición 9.4 Un subconjunto P de una variedad M es una subvariedad m -dimensional si (y solo si) en cada punto p de P hay un sistema coordinado de M adaptado a P compuesto por rebanadas de dimensión m .

Demostración. Asignemos a P la topología inducida. Sea $\xi: U \rightarrow \mathbb{R}^n$ un sistema coordinado de M en $p \in P$ tal que $U \cap P$ es la rebanada $x^j = x^j(p)$ para $m+1 \leq j \leq n$. Podemos suponer que $\xi(p) = 0$. Entonces, como ξ es un homeomorfismo, la aplicación $\xi_p = (x^1, \dots, x^m)|_P$ es un homeomorfismo de $U \cap P$ al subconjunto abierto $\xi(U) \cap \mathbb{R}^m$ de \mathbb{R}^m .

Afirmamos que todos éstos sistemas coordinados topológicos ξ_p

forman un atlas para P . Ciertamente ellos cubren a P y cualesquiera dos de ellos cumplen la condición de compatibilidad, ya que para $l \leq i \leq m$,

$$u^i \circ \xi_p \circ \eta_p^{-1} = u^i \circ (\xi \circ \eta^{-1}) \circ \mathbb{R}^m,$$

donde \mathbb{R}^m se considera como el hiperplano m -dimensional que consiste de las primeras m coordenadas de \mathbb{R}^n .

Sólo falta demostrar que éste atlas determina una estructura diferenciable que hace de P una subvariedad de M . Para un sistema coordinado como el ξ dado más arriba, las funciones $x^i|_P = x^i \circ j$ para $1 \leq i \leq m$ son suaves, ya que ellas constituyen el sistema coordinado ξ_p . Por lo tanto, la aplicación inclusión $j: P \hookrightarrow M$ es suave pues son suaves sus componentes $x^i \circ j$. Evidentemente la matriz jacobiana de ξ con respecto a ξ_p contiene una matriz identidad de orden m , por lo tanto, dj es siempre inyectiva. ##

Un campo vectorial X sobre M es tangente a una subvariedad P de M si $X_p \in T_p(P)$ para todo p en P . (Recuerdese que para una subvariedad P , $T_p(P)$ se considera como un subespacio de $T(M)$.)

Proposición 9.5 Sea P una subvariedad de M . (1) Si $X \in \mathfrak{X}(M)$ es tangente a P , entonces su restricción $X|_P$ a P es un campo vectorial suave sobre P . (2) Más aún, si $Y \in \mathfrak{X}(M)$ es también tangente a P , entonces el paréntesis de Lie $[X, Y]$ es tangente a P y

$$[X, Y]|_P = [X|_P, Y|_P].$$

La prueba es directa, usando un sistema coordinado adaptado a P para (1) y aplicando el lema 7.3 (pag. 20) para probar (2).

10 INMERSIONES Y SUBMERSIONES

Esta sección trata con dos tipos especiales de aplicaciones diferenciables.

Lema 10.1 Sea $\phi: M^m \rightarrow N^n$ una aplicación suave, y sea p un punto de M . Entonces, las siguientes afirmaciones son equivalentes:

- (1) La diferencial $d\phi_p$ es inyectiva.
- (2) La matriz jacobiana de $d\phi_p$ tiene rango m con respecto a una (y entonces para toda) elección de sistemas coordinados.
- (3) Si y^1, \dots, y^n es un sistema coordinado de N en $\phi(p)$, entonces existen enteros $1 \leq i_1 < \dots < i_m < n$ tales que las funciones $y^{i_1} \circ \phi, \dots, y^{i_m} \circ \phi$ forman un sistema coordinado sobre una vecindad de p en M .

La demostración de éste lema es una ligera generalización de la primera parte de la demostración de la proposición 9.1 (pag. 23).

Una inmersión $\phi: M^m \rightarrow N^n$ es una aplicación suave tal que $d\phi_p$ es inyectiva para todo $p \in M$.

Por ejemplo, toda curva regular (α' nunca es cero) es una inmersión.

Un encajamiento de una variedad P en una variedad M es una inmersión $\phi: P \rightarrow M$ tal que es inyectiva y la aplicación inducida $P \rightarrow \phi(P)$ es un homeomorfismo sobre el subespacio $\phi(P)$ de M. El ejemplo estándar es la aplicación $(a_1, \dots, a_m) \rightarrow (a_1, \dots, a_m, 0, \dots, 0)$ de \mathbb{R}^m en \mathbb{R}^n . El lema 10.1, num. (3), dice que localmente toda inmersión es un encajamiento. Explícitamente, la restricción de una inmersión a un conjunto abierto suficientemente pequeño es un encajamiento.

Las subvariedades y los encajamientos están relacionados: Si P es una subvariedad de M, entonces la inclusión $j: P \hookrightarrow M$ es un encajamiento. Inversamente, si $\phi: P \rightarrow M$ es un encajamiento, su imagen $\phi(P)$ es una variedad difeomorfa a P; éste difeomorfismo está dado por la aplicación inducida $\phi: P \rightarrow \phi(P)$. Entonces, $\phi(P)$ es un subespacio de M, y la inclusión $j: \phi(P) \hookrightarrow M$ es igual a $\phi \circ \phi^{-1}$, la cual por la regla de la cadena es una inmersión. Así $\phi(P)$ es una subvariedad de M.

El término "subvariedad" se aplica algunas veces a un objeto más general: Sea P una variedad que es simplemente un subconjunto de una variedad M. Si la inclusión $j: P \hookrightarrow M$ es una inmersión, diremos que P es una subvariedad inmersa de M. Según ésta definición, las subvariedades son subvariedades inmersas, pero no inversamente, ya que éstas últimas no necesitan tener la topología inducida.

El lema 10.1 tiene un resultado asociado que puede ser considerado como su dual:

Lema 10.2 Sea $\psi: M^m \rightarrow N^n$ una aplicación suave, y sea $p \in M$. Las siguientes afirmaciones son equivalentes:

- (1) La diferencial $d\psi_p$ es suprayectiva.
- (2) La matriz jacobiana de $d\psi_p$ tiene rango n con respecto a una (y entonces para toda) elección de sistemas coordenados.
- (3) Si y^1, \dots, y^n es un sistema coordenado para N en $\psi(p)$, existe un sistema coordenado para M en p de la forma

$$(y^1 \circ \psi, \dots, y^n \circ \psi, x^{n+1}, \dots, x^m).$$

Demostración. Es fácil ver que dada una elección de coordenadas, (2) implica (1) y que para toda elección de coordenadas (1) implica (2). Para unas coordenadas como las de (3), (2) es válida. Supongamos que (2) es verdadera para las coordenadas y^1, \dots, y^n en $\psi(p)$ y x^1, \dots, x^m en p, donde necesariamente $m > n$. En p la matriz jacobiana $(\partial(y^i \circ \psi) / \partial x^j)$ de orden $n \times m$ tiene, por hipótesis, n columnas linealmente independientes. Reetiquetando si es necesario, podemos suponer que éstas son las primeras n columnas. Pero entonces, $\xi = (y^1 \circ \psi, \dots, y^n \circ \psi, x^{n+1}, \dots, x^m)$ es un sistema coordenado en p, ya que el determinante de la matriz jacobiana de ξ en p es distinto de cero. ##

Un punto $q \in N$ se llama un valor regular de una aplicación suave $\psi: M \rightarrow N$ si $d\psi_p$ es suprayectiva para todo $p \in \psi^{-1}(q)$.

Corolario 10.3 Si $q \in \psi(M)$ es un valor regular de una aplicación suave $\psi: M \rightarrow N$, entonces $\psi^{-1}(q)$ es una subvariedad de M y $\dim M = \dim N + \dim \psi^{-1}(q)$.

Demostración. Utilizaremos la caracterización de las subvariedades mediante rebanadas coordenadas que demostramos en la proposición 9.4 (pag. 24). Para $p \in \psi^{-1}(q)$, la parte (3) del lema anterior demuestra, reordenando las funciones coordenadas, que

$$(x^{n+1}, \dots, x^m, y^1 \circ \psi, \dots, y^n \circ \psi)$$

es un sistema coordinado sobre una vecindad U de p en M . Ahora bien, la rebanada coordinada de U que pasa por p es precisamente $U \cap \psi^{-1}(q)$. Según la proposición 9.4, esto significa que $\psi^{-1}(q)$ es una subvariedad de M . ##

Una hipersuperficie en una variedad M es una subvariedad P de M cuya codimensión = $\dim M - \dim P$ es igual a 1. Aplicando el corolario anterior en el caso en que $N = \mathbb{R}$ obtenemos el siguiente método para construir hipersuperficies.

Corolario 10.4 Sea $c \in \mathbb{R}$ un valor de la función $f \in \mathcal{F}(M)$. Si en cada punto del conjunto $f^{-1}(c) = \{p \in M \mid f(p) = c\}$ la diferencial df_p es diferente de cero, entonces $f^{-1}(c)$ es una subvariedad de M de codimensión 1 llamada una hipersuperficie de nivel de f .

Por ejemplo, sobre \mathbb{R}^{n+1} sea $f = \sum (u^i)^2$, en donde u^i son las coordenadas naturales de \mathbb{R}^{n+1} . Entonces $df = \sum 2u^i du^i$, y tanto f como df son cero solamente en el origen. De éste modo, la n -esfera $S^n(r) = f^{-1}(r^2)$ de radio $r > 0$ es una hipersuperficie de nivel de f en \mathbb{R}^{n+1} .

Una submersión $\psi: M \rightarrow B$ es una aplicación suave y suprayectiva tal que $d\psi_p$ es suprayectiva para todo $p \in M$.

De acuerdo con ésta definición, todo valor de una submersión es un valor regular de manera que M se descompone en las subvariedades $\psi^{-1}(q)$ para todo $q \in B$. Para $m \geq n$, la proyección $\mathbb{R}^m \rightarrow \mathbb{R}^n$ que envía (t_1, \dots, t_m) a (t_1, \dots, t_n) es claramente una submersión. El lema 10.2, parte (3), implica que localmente toda submersión es de ésta forma.

11 TOPOLOGIA DE VARIEDADES

Las variedades son espacios localmente euclidianos; es decir, cada punto de una variedad M tiene una vecindad homeomorfa (me-

diante un sistema coordinado) a un subconjunto abierto de un espacio euclidiano. Debido a esto las variedades deben compartir todas las propiedades locales de los espacios euclidianos. Entre estas propiedades locales destacan la conexidad local y la compacidad local. La primera significa que cada punto de una variedad M tiene una vecindad conexa y la segunda afirma que cada punto tiene una vecindad compacta. Un espacio localmente euclidiano no tiene porque ser de Hausdorff, tal como vimos en la sección 1 (pag. 3). Es por esto que en la definición de variedad incluimos el axioma de Hausdorff.

A. Conexidad

Como para cualquier espacio topológico, una variedad es conexa si no puede ser expresada como la unión de dos conjuntos abiertos no vacíos con intersección vacía. En una variedad los sistemas coordinados demuestran que cada punto puede ser conectado a todo punto suficientemente cercano mediante un segmento de curva suave. Se sigue que una variedad es conexa si y solo si cualesquiera dos de sus puntos pueden ser unidos mediante un segmento de curva suave.

Un espacio topológico arbitrario es la unión disjunta de sus componentes conexas que son sus subespacios conexos más grandes. Como una variedad es localmente euclidiana, sus componentes conexas son subconjuntos abiertos y, por lo tanto, subvariedades abiertas.

B. Segunda Numerabilidad

Un espacio topológico S se llama segundo numerable si su topología tiene una base numerable B , esto es, una colección numerable de conjuntos abiertos tal que todo conjunto abierto es la unión de una subcolección de B . Por ejemplo, \mathbb{R}^n tiene una base numerable consistente de la colección de todos los conjuntos $\{p \in \mathbb{R}^n \mid a_i < p_i < b_i\}$ donde a_i y b_i son números racionales. Es fácil verificar que (1) cualquier subespacio de un espacio segundo numerable es segundo numerable, (2) el producto cartesiano de espacios segundo numerables es segundo numerable, y (3) un espacio que es la unión numerable de subespacios abiertos segundo numerables es segundo numerable.

Una cubierta (abierta) C de un espacio topológico S es una colección de subconjuntos (abiertos) de S cuya unión es S . Un espacio segundo numerable S tiene la propiedad de Lindelöf: Toda cubierta abierta de S tiene una subcubierta numerable (es decir, una subcolección de la cubierta que todavía cubre a S). En particular, una variedad segundo numerable tiene a lo más un número numerable de componentes conexas.

Existen ejemplos de variedades conexas que no son segundo numerables pero se trata de meras curiosidades. Por eso, de ahora en adelante, supondremos que las variedades con las que tratamos

son segundo numerables.

C. Particiones de la Unidad

Una colección L de subconjuntos de un espacio S se llama localmente finita si cada punto de S tiene una vecindad que intersecta a sólo un número finito de miembros de L . Sea $\{f_\alpha | \alpha \in A\}$ una colección de funciones suaves de valores reales sobre una variedad M tal que $\{\text{supp } f_\alpha | \alpha \in A\}$ es localmente finita. Entonces la suma $\sum_\alpha f_\alpha$ es una función suave bien definida sobre M , ya que sobre alguna vecindad de cada punto, todas excepto un número finito de las funciones f_α son cero idénticamente.

Una partición suave de la unidad sobre una variedad M es una colección $\{f_\alpha | \alpha \in A\}$ de funciones $f_\alpha \in \mathcal{F}(M)$ tales que

- (1) $0 \leq f_\alpha \leq 1$ para toda $\alpha \in A$.
- (2) $\{\text{supp } f_\alpha | \alpha \in A\}$ es localmente finita.
- (3) $\sum_\alpha f_\alpha = 1$.

Se dice que la partición está subordinada a una cubierta abierta C de M si cada conjunto $\text{supp } f_\alpha$ está contenido en algún elemento de C .

Las particiones de la unidad son una herramienta indispensable para ensamblar objetos matemáticos localmente definidos y convertirlos en objetos definidos globalmente sobre una variedad o, también, para descomponer un objeto global en una suma de objetos locales.

Proposición 11.1 Si M es una variedad segundo numerable, entonces dada cualquier cubierta abierta C de M , existe una partición suave de la unidad subordinada a C .

Demostración. Se puede demostrar (ver [H y Y], pag. 79) que cualquier cubierta de una variedad segundo numerable tiene una subcubierta localmente finita; ésta propiedad se llama paracompacidad. En otras palabras, toda variedad conexa y segundo numerable es paracompacta.

Sea $C' = \{U_\alpha | \alpha \in A\}$ la subcubierta localmente finita de C que existe por la paracompacidad de M . Para cada $\alpha \in A$ y $p_\alpha \in U_\alpha$ sea $h_\alpha \in \mathcal{F}(M)$ la función promotorio en p_α , cuya definición y existencia probamos en la página 12. Entonces la familia $\{f_\alpha \in \mathcal{F}(M) | \alpha \in A\}$ es la partición suave de la unidad subordinada a C que estábamos buscando, en donde para cada $\alpha \in A$ se define

$$f_\alpha = \frac{h_\alpha}{\sum_\beta h_\beta} \quad \#\#$$

Es debido a la utilidad de las particiones de la unidad y a la proposición anterior, que supondremos en adelante que nuestras variedades son segundo numerables.

D. Orientabilidad

Esta es una propiedad topológica muy sutil que tiene varias caracterizaciones diferentes; para las variedades la siguiente es la más simple.

Una variedad M es orientable si existe una colección O de sistemas coordenados de M cuyos dominios cubren a M y tal que para cada $\xi, \eta \in O$, $\xi = (x^1, \dots, x^n)$ y $\eta = (y^1, \dots, y^n)$, la función determinante jacobiano $J(\xi, \eta) = \det(\partial y^i / \partial x^j)$ es positiva. O se llama un atlas orientado para M .

Es evidente que \mathbb{R}^m es orientable, ya que puede ser cubierto por un solo sistema coordenado. La esfera S^2 es orientable, ya que es fácil comprobar que los atlas dados en los ejemplos 1) y 2) de la sección 2, son atlas orientados para S^2 . La botella de Klein y la banda de Möbius son ejemplos de variedades no orientables.

Sean ahora M y N variedades y $\phi: M \rightarrow N$ un difeomorfismo. Es inmediato verificar que M es orientable si y sólo si N es orientable. Si, además de ésto, M y N son conexas y están orientadas, es decir, si se han elegido atlas orientados para M y N , el difeomorfismo ϕ induce una orientación para N que puede o no coincidir con la orientación inicial de N . En el primer caso se dice que ϕ preserva la orientación y en el segundo, que ϕ revierte la orientación.

Proposición 11.2 Si la variedad M puede ser cubierta por dos vecindades coordenadas V_1 y V_2 de manera que la intersección $V_1 \cap V_2$ es conexa, entonces M es orientable.

Demostración. Como para los sistemas coordenados $\xi = (x^1, \dots, x^n): V_1 \rightarrow \mathbb{R}^n$ y $\eta = (y^1, \dots, y^n): V_2 \rightarrow \mathbb{R}^n$ el determinante jacobiano $J(\xi, \eta)$ es distinto de cero y es una función continua, no puede cambiar de signo en el conjunto conexo $V_1 \cap V_2$. Si el signo de $J(\xi, \eta)$ fuera negativo en un punto, cambiando el signo de una de las coordenadas se puede lograr que $J(\xi, \eta)$ sea positiva. ##

Usando ésta proposición, se sigue inmediatamente del ejemplo 2) de la sección 2, que la esfera S^2 es orientable.

E. Otra Definición de Variedad

La definición de variedad que hemos dado comienza presuponiendo un espacio topológico, pero en realidad el atlas de una variedad determina su topología. Es importante tener ésto en cuenta, pues así tenemos un método más eficiente de construir variedades.

Proposición 11.3 Sea Σ un conjunto sin estructura, y para cada $\alpha \in A$ sea ξ_α una función biyectiva de un subconjunto U_α de Σ a un conjunto abierto $\xi_\alpha(U_\alpha)$ en \mathbb{R}^n . Supongamos que
(1) Los dominios $\{U_\alpha | \alpha \in A\}$ cubren a Σ .

(2) Para todo $\alpha, \beta \in A$ la función $\xi_\beta \circ \xi_\alpha^{-1}$ es diferenciable sobre el conjunto abierto $\xi_\alpha(U_\alpha \cap U_\beta)$ en \mathbb{R}^n .

(3) Si $p \neq q$ en \mathcal{Z} , entonces p y q pertenecen a un solo U_α o bien existen $\alpha, \beta \in A$ tales que $p \in U_\alpha$, $q \in U_\beta$, con U_α y U_β disjuntos.

Entonces existe una única topología de Hausdorff y una única estructura diferencial sobre \mathcal{Z} tal que cada ξ_α es un sistema coor- denado de la variedad resultante. Además, si un número numerable de los U_α cubren a \mathcal{Z} la variedad es segundo numerable.

Demostración. Como cada ξ_α debe convertirse en un homeomorfismo la única posibilidad que tenemos es definir un subconjunto V de \mathcal{Z} como abierto si y sólo si $\xi_\alpha(V \cap U_\alpha)$ es abierto en $\xi_\alpha(U_\alpha)$ y, por lo tanto, en \mathbb{R}^n para toda $\alpha \in A$. Es fácil verificar que tales conjuntos abiertos constituyen una topología sobre \mathcal{Z} . Nótese que por (2) los dominios U_α son abiertos.

Afirmamos que cada ξ_α es un homeomorfismo. Si V es abierto en U_α y por lo tanto abierto en \mathcal{Z} , entonces por definición $\xi_\alpha(V)$ es abierto en $\xi_\alpha(U_\alpha)$. Inversamente, si W es un conjunto abierto en $\xi_\alpha(U_\alpha)$ debemos demostrar que $\xi_\alpha^{-1}(W)$ es abierto. Si $\beta \in A$ entonces aplicando (2) tanto a α, β como a β, α vemos que $\xi_\beta \circ \xi_\alpha^{-1}$ es un homeomorfismo de $\xi_\alpha(U_\alpha \cap U_\beta)$ a $\xi_\beta(U_\alpha \cap U_\beta)$. Pero entonces la expresión en el lado derecho de la fórmula

$$\xi_\beta(\xi_\alpha^{-1}(W) \cap U_\beta) = (\xi_\beta \circ \xi_\alpha^{-1})(W \cap \xi_\alpha(U_\alpha \cap U_\beta))$$

muestra que éste conjunto es abierto. Consecuentemente, $\xi_\alpha^{-1}(W)$ es abierto en \mathcal{Z} .

Así pues, $\{\xi_\alpha \mid \alpha \in A\}$ es un atlas sobre el espacio topológico \mathcal{Z} , el cual por la condición (3) es de Hausdorff. Como cada U_α es homeomorfo a un conjunto abierto de \mathbb{R}^n , cada U_α es segundo numerable, y como $\mathcal{Z} = \bigcup_{\alpha \in A} U_\alpha$, se sigue que \mathcal{Z} es segundo numerable en el caso de que la familia $\{U_\alpha \mid \alpha \in A\}$ sea numerable. ##

12 ALGUNAS VARIETADES ESPECIALES

A. Variedades Producto

Consideraremos ahora cómo el cálculo sobre una variedad producto (ver ejemplo 11), pag. 9) se deriva de aquellos sobre las variedades M y N separadamente. Esta descomposición es análoga a la manera en que el cálculo sobre el plano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ se deriva del cálculo sobre la recta real.

Usando sistemas coordenados producto sobre $M \times N$ es fácil verificar que

(a) Las proyecciones

$$\pi: M \times N \rightarrow M \quad \text{que manda } (p, q) \text{ a } p,$$

$$\sigma: M \times N \rightarrow N \quad \text{que manda } (p, q) \text{ a } q$$

son aplicaciones suaves, de hecho submersiones.

(b) Una aplicación $\phi: P \rightarrow M \times N$ es suave si y solo si tanto $\pi \circ \phi$ como

son suaves.

(c) Para cada $(p, q) \in M \times N$ los subconjuntos

$$M \times q = \{(r, q) \in M \times N \mid r \in M\},$$

$$p \times N = \{(p, r) \in M \times N \mid r \in N\}$$

son subvariedades de $M \times N$.

(d) Para cada (p, q)

$$\pi|_{M \times q} \text{ es un difeomorfismo de } M \times q \text{ a } M;$$

$$\sigma|_{p \times N} \text{ es un difeomorfismo de } p \times N \text{ a } N.$$

Por (b) los espacios tangentes

$$T_{(p,q)}M = T_{(p,q)}(M \times q) \quad \text{y} \quad T_{(p,q)}N = T_{(p,q)}(p \times N)$$

son subespacios del espacio tangente a $M \times N$ en (p, q) .

Lema 12.1 $T_{(p,q)}(M \times N)$ es la suma directa de sus subespacios $T_{(p,q)}M$ y $T_{(p,q)}N$; es decir, cada elemento de $T_{(p,q)}(M \times N)$ tiene una expresión única de la forma

$$x + v, \quad \text{donde } x \in T_{(p,q)}M \text{ y } v \in T_{(p,q)}N.$$

Demostración. Puesto que $\pi|_{p \times N}$ es constante, $d\pi$ en (p, q) envía todo el espacio $T_{(p,q)}N$ a 0. Pero por (c), $d\pi|_{T_{(p,q)}M}$ es un isomorfismo lineal. Por lo tanto, $T_{(p,q)}M \cap T_{(p,q)}N = 0$. El resultado se sigue ahora del álgebra lineal, ya que por (c) la suma de las dimensiones de éstos dos subespacios es $\dim(M \times N)$. ##

Para poder relacionar el cálculo en la variedad producto $M \times N$ con los cálculos en las variedades factor M y N , la noción fundamental es la de levantamiento.

Si $f \in \mathcal{X}(M)$ el levantamiento de f a $M \times N$ es $\tilde{f} = f \circ \pi \in \mathcal{X}(M \times N)$.

Si $x \in T_p(M)$ y $q \in N$ entonces el levantamiento \tilde{x} de x a (p, q) es el único vector en $T_{(p,q)}(M)$ tal que $d\pi(\tilde{x}) = x$.

Si $X \in \mathcal{X}(M)$ el levantamiento de X a $M \times N$ es el campo vectorial \tilde{X} cuyo valor en cada (p, q) es el levantamiento de X_p a (p, q) . Usando sistemas coordenados producto se puede demostrar que \tilde{X} es suave. Así pues, tenemos que el levantamiento de $X \in \mathcal{X}(M)$ a $M \times N$ es el único elemento de $\mathcal{X}(M \times N)$ que está π -relacionado a X y está σ -relacionado al campo vectorial cero sobre N . El conjunto de todos éstos levantamientos horizontales \tilde{X} se denota por $\mathcal{L}(M)$.

Las funciones, los vectores tangentes y los campos vectoriales sobre N pueden ser levantados a $M \times N$ de manera semejante usando ésta vez la proyección σ . Nótese que $\mathcal{L}(M)$ y simétricamente el conjunto de los levantamientos verticales $\mathcal{L}(N)$, son subespacios vectoriales de $\mathcal{X}(M \times N)$ pero que (excepto en casos triviales) ninguno de los dos puede ser considerado como un submódulo de $\mathcal{X}(M \times N)$ sobre $\mathcal{F}(M \times N)$.

Por ejemplo, sobre \mathbb{R}^2 el campo vectorial coordinado $\partial_x = \partial/\partial x$ es el levantamiento horizontal del campo vectorial d/dx sobre \mathbb{R} , pero $y\partial_x$ no es un levantamiento.

Corolario 12.2 (1) Si $\tilde{X}, \tilde{Y} \in \mathcal{L}(M)$ entonces $[\tilde{X}, \tilde{Y}] = [X, Y] \sim \epsilon \mathcal{L}(M)$, y similarmente para $\mathcal{L}(N)$.
 (2) Si $\tilde{X} \in \mathcal{L}(M)$ y $\tilde{V} \in \mathcal{L}(N)$, entonces $[\tilde{X}, \tilde{V}] = 0$.

Demostración. Ambas afirmaciones se siguen del lema 7.3, pag. 20. En el caso de la afirmación (2), por ejemplo, $[\tilde{X}, \tilde{V}]$ está $\tilde{\alpha}$ -relacionado a $[X, 0] = 0$ y está σ -relacionado a $[0, V] = 0$. El resultado se sigue entonces del lema 12.1. ##

Para establecer éstos resultados hemos usado una notación bastante complicada. En la práctica, sin embargo, el resultado del lema 12.1 se escribe simplemente como $T_{p,q}(M \times N) = T_p(M) \times T_p(N)$, y la tilde (\sim) se omite de los levantamientos.

B. Espacios Vectoriales como Variedades

Sea V un espacio vectorial n -dimensional sobre los números reales. Si ξ y η son isomorfismos lineales de V a \mathbb{R}^n , entonces $\xi \circ \eta^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un isomorfismo lineal y por lo tanto un difeomorfismo. Debido a ésto es muy fácil aplicar en éste caso la proposición 11.3, para demostrar que existe una sola manera de convertir al espacio V en una variedad de modo que todo isomorfismo lineal $\xi: V \rightarrow \mathbb{R}^n$ sea un sistema coordinado.

La siguiente notación conveniente será usada con frecuencia.

Si $p, v \in V$, sea $v_p \in T_p(V)$ la velocidad inicial $\alpha'(0)$ de la curva $\alpha(t) = p + tv$.

Podemos imaginar a v_p como una flecha que va de p a $p + v$.

Lema 12.3 Si x^1, \dots, x^n es un sistema lineal de coordenadas sobre V , entonces

$$v_p = \sum x^i(v) \partial_i|_p.$$

Demostración. Dado que las coordenadas x^i son lineales,
 $x^i(\alpha(t)) = x^i(p) + tx^i(v)$.

Entonces, por la fórmula de la velocidad,

$$v_p = \alpha'(0) = \sum \frac{d(x^i \circ \alpha)}{dt}(0) \partial_i|_p = \sum x^i(v) \partial_i|_p. \quad ##$$

Así pues, v_p es el vector tangente en p con las mismas componentes coordenadas que $v \in V$. Se sigue inmediatamente que (1) para $p \in V$ fijo, la función $v \mapsto v_p$ es un isomorfismo lineal, de modo que V y $T_p(V)$ son isomorfos.

(2) Para $p, q \in V$, la función $v_p \mapsto v_q$ es un isomorfismo lineal, de modo que $T_p(V)$ y $T_q(V)$ son isomorfos.

Como en el caso familiar de $V = \mathbb{R}^3$ estos isomorfismos canónicos permiten intercambiar libremente los conceptos de v como un punto

de V , de v como una flecha que va de 0 a v , y de v como una flecha que va de p a $p + v$.

El campo vectorial de posición $Pf: \mathcal{K}(V)$ asigna a cada $p \in V$ el vector tangente $p_p \in T_p(V)$ (intuitivamente se trata de un duplicado, que comienza en p , de la flecha que va de 0 a p). En términos de un sistema lineal de coordenadas (x^1, \dots, x^n) , tenemos que

$$p = \sum x^i \partial_i.$$

C. El Haz Tangente

Para una variedad M , sea TM el conjunto $\bigcup \{T_p(M) \mid p \in M\}$ de todos los vectores tangentes a M . Hay aquí un detalle técnico: Debemos reemplazar para cada $p \in M$ el vector $0 \in T_p(M)$ por 0_p , ya que los espacios tangentes son espacios distintos sin elementos comunes. Así pues, para cada $v \in TM$ existe un único $T_p(M)$ que contiene a v , y la proyección $\pi: TM \rightarrow M$ manda v a p . Entonces $\pi^{-1}(p) = T_p(M)$.

Hay una manera natural de convertir a TM en una variedad, llamada el haz tangente de M . Sea ξ un sistema coordenado sobre $U \subset M$. Si v es un vector tangente a M en p de U , entonces v está determinado de manera única por las coordenadas de p y por las componentes de v con respecto a la base $\partial_1, \dots, \partial_n$ de $T_p(M)$. Para formalizar estas ideas, sea \hat{x}^i la función real valuada sobre $\pi^{-1}(U) \subset TM$ dada por $\hat{x}^i(v) = v(x^i)$. Entonces definimos $\tilde{\xi}: \pi^{-1}(U) \rightarrow \mathbb{R}^{2n}$ por

$$\tilde{\xi} = (x^1 \circ \pi, \dots, x^n \circ \pi, \hat{x}^1, \dots, \hat{x}^n).$$

Ahora aplicaremos la proposición 11.3 para convertir a TM en una variedad para la cual las funciones $\tilde{\xi}$ serán los sistemas coordenados de un atlas.

Por el teorema de la base tenemos que si $v \in \pi^{-1}(U)$, entonces $v = \sum \hat{x}^i(v) \partial_i|_{\pi(v)}$. Así vemos que $\tilde{\xi}$ es una función biyectiva de $\pi^{-1}(U)$ sobre el conjunto abierto $\xi(U) \times \mathbb{R}^n$ de \mathbb{R}^{2n} .

Veamos ahora que cualesquiera dos de tales funciones $\tilde{\xi}$ y $\tilde{\eta}$ cumplen la condición de compatibilidad (condición (2) de la proposición 11.3). Si $(a, b) \in \eta(U \cap V) \times \mathbb{R}^n$, entonces para $1 \leq i \leq n$

$$x^i \tilde{\xi}^{-1}(a, b) = x^i \tilde{\eta}^{-1}(a, b) = x^i \eta^{-1}(a).$$

Puesto que $\partial/\partial y^i = \sum_k (\partial x^k / \partial y^i) \partial / \partial x^k$, también tenemos

$$y^{n+i} \tilde{\xi}^{-1}(a, b) = \hat{x}^i \tilde{\eta}^{-1}(a, b) = \sum_k b^k \frac{\partial x^k}{\partial y^k}(\eta^{-1} a).$$

Por lo tanto, $\tilde{\xi} \circ \tilde{\eta}^{-1}$ es suave.

Es fácil verificar las otras dos condiciones de la proposición 11.3, por lo cual concluimos que TM es una variedad diferenciable de Hausdorff y segundo numerable cuya dimensión es el doble de la de M .

Un campo vectorial $X \in \mathcal{K}(M)$ se puede considerar como una sección suave de TM , es decir, una función suave $X: M \rightarrow TM$ tal que $\pi \circ X = I_M$, en donde I_M es la función identidad sobre M . Esto sugiere una generalización que es útil.

Un campo vectorial Z sobre una aplicación suave $\phi: P \rightarrow M$ es una función $Z: P \rightarrow TM$ tal que $\pi \circ Z = \phi$, donde π es la proyección $TM \rightarrow M$.

Según esto, Z asigna a cada punto $p \in P$ un vector tangente a M en $\phi(p)$. (Por ejemplo, la velocidad α' es un campo vectorial sobre la curva α de M .) Z es suave considerado como la aplicación $Z: P \rightarrow TM$ si y sólo si $f \in \mathcal{F}(M)$ implica que $Zf \in \mathcal{F}(P)$, donde $(Zf)(p) = Z(p)f$ para todo $p \in P$. El conjunto $\mathcal{K}(\phi)$ de todos los campos vectoriales suaves sobre la aplicación $\phi: P \rightarrow M$ es, naturalmente, un módulo sobre $\mathcal{F}(P)$.

13 CURVAS INTEGRALES

Veremos en esta sección que un campo vectorial sobre una variedad puede ser interpretado como una ecuación diferencial y caracterizaremos geoméricamente las soluciones de dichas ecuaciones.

Una curva $\alpha: I \rightarrow M$ es una curva integral de $V \in \mathcal{K}(M)$ si $\alpha' = V_\alpha$; es decir, $\alpha'(t) = V_{\alpha(t)}$ para todo $t \in I$.

Así, en cada punto la curva α tiene la velocidad determinada por el campo V . Si la ecuación $\alpha' = V_\alpha$ se expresa en términos de un sistema coordenado $\xi = (x^1, \dots, x^n)$, obtenemos un sistema de ecuaciones diferenciales ordinarias y de primer orden

$$\frac{d(x^i \circ \alpha)}{dt} = F^i(x^1 \circ \alpha, \dots, x^n \circ \alpha) \quad (1 \leq i \leq n),$$

donde F^i es la expresión coordenada para Vx^i .

Puesto que el parámetro t no aparece en el lado derecho, podemos considerar a V como determinando la velocidad de un fluido en M con un régimen de flujo estacionario. Esta visualización puede ser útil para comprender las definiciones que haremos más adelante. El teorema fundamental de existencia y unicidad para las soluciones de tales sistemas de ecuaciones diferenciales (ver [I y V], capítulo 2) tiene la siguiente consecuencia en términos de la teoría de variedades.

Proposición 13.1 Si $V \in \mathcal{K}(M)$ entonces para cada $p \in M$ existen un intervalo I que contiene a 0 y una única curva integral $\alpha: I \rightarrow M$ de V tal que $\alpha(0) = p$.

Nótese que si α es una curva integral de V , entonces $t \rightarrow \alpha(t+c)$ también es una curva integral de V , para toda constante $c \in \mathbb{R}$.

Corolario 13.2 Si $\alpha, \beta: I \rightarrow M$ son curvas integrales de V tales que $\alpha(a) = \beta(a)$ para alguna $a \in I$, entonces $\alpha = \beta$.

Demostración. Por la continuidad de α y β sabemos que el conjunto

en donde α y β coinciden, $A = \{t \in I \mid \alpha(t) = \beta(t)\}$, es un conjunto cerrado. Si A es también abierto, puesto que por hipótesis A es no vacío, debemos tener que $A = I$; de otra manera, llegaríamos a la contradicción de que el intervalo I , que es conexo por definición, se descompone como $I = A \cup (I - A)$, en donde tanto A como $(I - A)$ son abiertos no vacíos. Para demostrar que A es abierto, fijemos un $t \in A$. Entonces $s \rightarrow \alpha(t + s)$ y $s \rightarrow \beta(t + s)$ son curvas integrales de V que coinciden en $s = 0$. Por lo tanto, la proposición 13.1 implica que estas curvas integrales coinciden en una vecindad de $s = 0$. Esto demuestra que todo $t \in A$ está en una vecindad totalmente contenida en A , por lo que A es abierto. ##

Consideremos la colección de todas las curvas integrales $\alpha: I_\alpha \rightarrow M$ de V que arrancan de $p \in M$, es decir, tales que $\alpha(0) = p$. Para cualesquiera dos de tales curvas, el corolario 13.2 muestra que $\alpha = \beta$ sobre $I_\alpha \cap I_\beta$. Así, según el comentario que hicimos en la pag. 11, después de la observación 6), vemos que todas estas curvas definen una sola curva integral $\alpha_p: I_p \rightarrow M$ donde $I_p = \bigcup I_\alpha$. Llamaremos a α_p la curva integral maximal de V que comienza en p . Este dominio maximal I_p no tiene porqué coincidir con toda la recta \mathbb{R} .

Ejemplo. Sobre el plano \mathbb{R}^2 sea $V = x\delta_x - y\delta_y$. Entonces $\alpha(t) = (x(t), y(t))$ es una curva integral de V si y sólo si $dx/dt = x$, $dy/dt = -y$. Por lo tanto, $x(t) = Ae^t$ y $y(t) = Be^{-t}$. Así, la curva integral maximal que comienza en $p = (p_1, p_2)$ es

$$\alpha_p(t) = (p_1 e^t, p_2 e^{-t}) \quad \text{para todo } t \in \mathbb{R}.$$

Estas curvas parametrizan a las hipérbolas $xy = \text{constante}$.

El siguiente refinamiento del corolario 13.2 demuestra que si dos curvas integrales maximales de V se encuentran en un punto, entonces ellas sólo difieren en la parametrización.

Lema 13.3 Para $V \in \mathcal{L}(M)$ sea $q = \alpha_p(s)$. Entonces $s + I_q = I_p$, y $\alpha_p(s+t) = \alpha_q(t)$ para todo $t \in I_q$.

Demostración. Sea $\beta(u) = \alpha_q(u - s)$; entonces β es una curva integral de V definida sobre $s + I_q = \{s+t \mid t \in I_q\}$. Puesto que $\beta(s) = q = \alpha_p(s)$, el corolario 13.2 implica que $\beta = \alpha_p$ sobre $(s + I_q) \cap I_p$. Así, β y α_p se combinan para dar una sola curva integral sobre $(s + I_q) \cup I_p$. Como I_p es maximal, $s + I_q \subset I_p$. De donde $\alpha_p(s+t) = \beta(s+t) = \alpha_q(t)$ para toda $t \in I_q$. Pero como I_q es maximal, $I_q \supset -s + I_p$. De donde, $s + I_q = I_p$. ##

Una curva no constante $\gamma: \mathbb{R} \rightarrow M$ se llama periódica si existe un número $c > 0$ tal que $\gamma(t+c) = \gamma(t)$ para todo t . Puesto que la curva γ es no constante, al más pequeño de tales números $c > 0$, se le conoce como el período de γ . Si γ es inyectiva sobre un intervalo $[a, a+c)$,

entonces ψ se llama simplemente periódica. Según esto, una figura en forma de numeral ocho es periódica pero no simplemente periódica.

Se sigue del lema anterior que una curva integral maximal sólo puede ser inyectiva, simplemente periódica o constante.

Un campo vectorial se llama completo si cada una de sus curvas integrales está definida sobre toda la recta real. La siguiente definición nos proporciona un medio para ensamblar en una sola aplicación a todas las curvas integrales de un campo vectorial dado V , considerando primero el caso en que V sea completo.

El flujo de un campo vectorial completo V sobre M es la aplicación $\psi: M \times \mathbb{R} \rightarrow M$ dada por

$$\psi(p, t) = \alpha_p(t),$$

donde α_p es la curva integral maximal que comienza en p .

Si p se mantiene fijo, entonces la función $t \rightarrow \psi(p, t)$ es simplemente la curva integral α_p . Por otra parte, si t se mantiene fijo, entonces la función $p \rightarrow \psi(p, t)$ denotada por $\psi_t: M \rightarrow M$, hace "fluir" cada punto $p \in M$ por la curva integral maximal α_p durante t unidades de tiempo. Llamaremos a ψ_t la t -ésima etapa del flujo ψ . A veces llamaremos a $\{\psi_t \mid t \in \mathbb{R}\}$ el flujo de V .

Lema 13.4 Si ψ es el flujo de un campo vectorial completo, entonces:

- (1) ψ_0 es la aplicación identidad I_M sobre M .
- (2) $\psi_s \circ \psi_t = \psi_{s+t}$ para todo $s, t \in \mathbb{R}$ (así, las etapas de ψ conmutan).
- (3) Cada etapa es un difeomorfismo, con $\psi_{-t}^{-1} = \psi_t$.

Demostración. (1) Esto es obvio, ya que $\psi_0(p) = \alpha_p(0) = p$. (2) Esto se sigue inmediatamente del lema 13.3. La afirmación (3) se sigue de las anteriores, ya que $\psi_t \circ \psi_{-t} = I_M = \psi_{-t} \circ \psi_t$. ###

Si el campo vectorial V no es completo, entonces para cada punto p de M tenemos al menos un flujo local $\psi: U \times I \rightarrow M$ definido por la misma ecuación que antes, pero ahora U es una vecindad de p en M e I es un intervalo alrededor del 0 en \mathbb{R} . De acuerdo con la teoría de las ecuaciones diferenciales (ver [I y V], pags. 48 y 49), si U e I son suficientemente pequeños, entonces ψ es suave. Para estos flujos locales son válidas propiedades análogas a las del lema 13.4:

- (1) ψ_0 es la aplicación identidad sobre U .
- (2) $\psi_{s+t} = \psi_s \circ \psi_t$ siempre que s, t y $s+t$ estén en I .
- (3) Para todo $t \in I$, $\psi_t: U \rightarrow \psi_t(U)$ es un difeomorfismo.

Si V no es necesariamente completo entonces su flujo $\psi(p, t) = \alpha_p(t)$ posee un dominio maximal

$$D = \{(p, t) \in M \times \mathbb{R} \mid t \in I_p\}.$$

Se puede demostrar (ver por ejemplo L, teorema 5, capítulo IV) que el dominio \mathcal{D} es un subconjunto abierto de $M \times \mathbb{R}$ y que el flujo $\psi: \mathcal{D} \rightarrow M$ que se obtiene al ensamblar todos los flujos locales suaves es también suave.

Considerando el comportamiento de las curvas $\alpha: I \rightarrow M$ en las cercanías de los extremos del intervalo I , es suficiente concentrarse en el extremo derecho y suponer que $I = [0, B]$; se hace la suposición análoga para tratar con el extremo izquierdo. (Por convención, $b < \infty$ pero $B < \infty$.)

Una curva suave por pedazos $\alpha: [0, B] \rightarrow M$ se llama extensible si tiene una extensión continua $\bar{\alpha}: [0, B] \rightarrow M$. En este caso, $q = \bar{\alpha}(B)$ se llama un punto extremo de α .

En forma equivalente, existe un punto $q \in M$ tal que para toda sucesión $\{s_i\}$ en $[0, B)$ que tiende a B , la sucesión $\{\alpha(s_i)\}$ converge a q .

Una curva extensible en general no tiene una extensión suave por pedazos, sin embargo esto suele ocurrir en casos especiales que son importantes.

Lema 13.5 Sea $\alpha: [0, b] \rightarrow M$, $b < \infty$, una curva integral de $V \in \mathcal{F}(M)$. Las siguientes afirmaciones son equivalentes:

- (1) α no es maximal; es decir, α es extensible como curva integral de V a una curva integral más grande definida sobre el intervalo $[0, b + \epsilon)$.
- (2) α es extensible.
- (3) α está contenida en un subconjunto compacto de M .
- (4) Existe una sucesión $\{s_i\}$ que tiende a b tal que $\{\alpha(s_i)\}$ converge.

Demostración. Claramente tenemos que (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4). Para probar que (4) \Rightarrow (1), sea U una vecindad de $\lim \alpha(s_i)$ tal que esté definido un flujo de V sobre $U \times (-\delta, \delta)$. Existe un número natural n tal que $b - \delta < s_n$ y $\alpha(s_n) \in U$. La curva integral de V que comienza en $\alpha(s_n)$ está definida sobre $[0, \delta)$. Por lo tanto, aplicando el lema 13.3 podemos extender α más allá de b . ##

Finalmente, consideremos dos aplicaciones del concepto de flujo.

Lema 13.6 Si V es un campo vectorial y p es un punto tal que $V_p \neq 0$, entonces existe un sistema coordenado x^1, \dots, x^n en p tal que $V = \partial/\partial x^1$ sobre el dominio del sistema coordenado.

Demostración. Sea $\psi: U \times I \rightarrow M$ un flujo local para V , donde U es una vecindad de p sobre la cual V nunca se anula. Sea S una hipersuperficie que contiene a p en U tal que $V_p \notin T_p(S)$, y consideremos la restricción $\Psi: S \times I \rightarrow M$ de ψ .

La aplicación $\psi|(S \times 0)$ es trivialmente un difeomorfismo a S , y

$d\psi_{(p,0)}$ simplemente identifica $T_{(p,0)}(S)$ con $T_p(S)$. Puesto que $d\psi(\partial/\partial t|_{(p,0)}) = V_p \notin T_p(S)$, se sigue que $d\psi_{(p,0)}$ es un isomorfismo. Así, por el teorema de la función inversa, ψ es un difeomorfismo de una vecindad $Q \times J$ de $(p,0)$ sobre una vecindad W de p en M .

Podemos suponer que y^2, \dots, y^n es un sistema coordenado para S sobre Q . Sea $y^1 = t$ la coordenada natural sobre $J \subset \mathbb{R}$. Transfiriendo el sistema coordenado producto y^1, \dots, y^n a W mediante el difeomorfismo $\psi|_{(Q \times J)}$ obtenemos el sistema coordenado buscado. De hecho, $\partial/\partial x^1 = d\psi(\partial/\partial t) = V$, ya que $\psi(p, t) = \alpha_p(t)$. ##

El paréntesis de Lie de un par de campos vectoriales se puede describir utilizandò flujos. Recordemos primero que una función suave $F: I \rightarrow V$ con valores en un espacio vectorial V de dimensión finita tiene la derivada $F': I \rightarrow V$ definida como

$$F'(s) = \lim_{t \rightarrow 0} \frac{1}{t} [F(s+t) - F(s)] = \sum \frac{df^i}{dt}(s) e_i,$$

donde $F = \sum f^i e_i$ con respecto a la base e_1, \dots, e_n de V .

Ahora tomemos la derivada de un campo vectorial W con respecto a un campo vectorial V , de acuerdo con las siguientes consideraciones. Si $p \in M$ sea α_p la curva integral de V que comienza en p . Usemos el flujo de V para mover el valor de W en $\alpha_p(s)$ y traerlo de regreso a $T_p(M)$. Entonces tomemos la derivada vectorial que hemos recordado arriba en $s = 0$. Afirmamos que el resultado de estas operaciones es el paréntesis de Lie $[V, W]_p$.

Proposición 13.7 Si $V, W \in \mathcal{X}(M)$, sea ψ un flujo local de V cerca de $p \in M$. Entonces

$$[V, W]_p = \lim_{t \rightarrow 0} \frac{1}{t} [d\psi_t(W_{\psi_t p}) - W_p].$$

Demostración. Escribiendo $F_p(t) = d\psi_t(W_{\psi_t p})$, vemos que el lado derecho de la ecuación de arriba es $F'_p(0)$.

Caso 1. $V_p \neq 0$. Escojamos un sistema coordenado x^1, \dots, x^n tal como en el lema anterior, de modo que $V = \partial_1$. Así, el flujo de V solamente hace cambiar la coordenada x^1 de los puntos q en una vecindad de p :

$$x^1(\psi_t q) = x^1(q) + t, \quad x^j(\psi_t q) = x^j(q) \quad \text{para } 2 \leq j \leq n.$$

se sigue que $d\psi_t(\partial_i) = \partial_i$ para todo i y todo t . Así si $W = \sum W^i \partial_i$,

$$F_p(t) = \sum W^i(\psi_t p) \partial_i|_p.$$

Omitiendo el subíndice p podemos calcular

$$\begin{aligned} F'_p(0) &= \sum \frac{d}{dt} (W^i \circ \alpha_p)(0) \partial_i = \sum V_p(W^i) \partial_i = \\ &= \sum \frac{\partial W^i}{\partial x^1}(0) \partial_i = [\partial_1, W]_p = [V, W]. \end{aligned}$$

Caso 2. $V = 0$ sobre una vecindad de p . En este caso, $[V, W]_p = 0$.

Las curvas integrales que comienzan en puntos de la vecindad en cuestión son constantes y, por lo tanto, $\psi_t =$ identidad para todo t . Así pues, F_p es constante y tenemos que $F_p'(0) = 0$.

Caso 3. $V_p = 0$, pero p es el límite de una sucesión $\{p_i\}$ tal que $V_{p_i} \neq 0$ para toda i . Las expresiones en términos de coordenadas tanto de $F_p'(0)$ como de $[V, W]_p$ demuestran que ambos dependen continuamente de p y, por lo tanto, el resultado se sigue del caso 1. ##

+++++

II. TENSORES

La noción de campo tensorial sobre una variedad generaliza las nociones de función con valores reales, de campo vectorial, y de uno-forma, y así proporciona los mecanismos matemáticos para describir objetos más complicados sobre una variedad. Los tensores aparecen en muchas situaciones que son de interés para la física, pero la propiedad que los caracteriza es siempre la multilinealidad. La definición que usaremos pone el énfasis en esta propiedad de los tensores y tiene la ventaja de convertirse fácilmente en la descripción clásica de los tensores mediante las coordenadas. En las últimas secciones de esta parte, consideraremos el concepto general de producto interior en un espacio vectorial de dimensión finita, el cual juega un papel básico en la geometría semi-Riemanniana.

14 ALGEBRA MULTILINEAL

Las siguientes definiciones son lo bastante generales como para servirnos en los dos casos que serán de más interés para nosotros: el módulo $\mathcal{M}(M)$ sobre el anillo $\mathcal{F}(M)$, y el espacio vectorial $T_p(M)$ sobre el campo \mathbb{R} .

Sean V_1, \dots, V_s módulos sobre el anillo K . Entonces $V_1 \times \dots \times V_s$ será el conjunto de todas las s -adas (v_1, \dots, v_s) con $v_i \in V_i$. Como es usual, definiremos la suma de dos s -adas (v_1, \dots, v_s) y (w_1, \dots, w_s) de $V_1 \times \dots \times V_s$ como la s -ada $(v_1 + w_1, \dots, v_s + w_s)$, y definiremos el producto de la s -ada (v_1, \dots, v_s) por el elemento $k \in K$ como la s -ada (kv_1, \dots, kv_s) . Con estas operaciones el conjunto $V_1 \times \dots \times V_s$ se convierte en un módulo sobre el anillo K , llamado el producto directo de los módulos V_1, \dots, V_s (también se llama la suma directa si la notación \times se reemplaza por \oplus). Si W es también un módulo sobre K , una función

$$A: V_1 \times \dots \times V_s \rightarrow W$$

se llama K -multilineal si A es K -lineal en cada una de sus variables; es decir, si para todo $1 \leq i \leq s$ y $v_j \in V_j$ ($j \neq i$), la función

$$v \rightarrow A(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_s)$$

es K -lineal.

Si V es un módulo sobre K , sea V^* el conjunto de todas las funciones K -lineales definidas en V y con valores en K . Las definiciones usuales de adición de funciones y multiplicación de funciones por elementos de K convierten a V^* en un módulo sobre K , llamado el módulo dual de V .

Si $V_i = V$ para $1 \leq i \leq s$, la notación $V_1 \times \dots \times V_s$ se abrevia como V^s .

Para enteros $r \leq 0$ y $s \leq 0$, no ambos cero, una función K -multilineal $A: (V^*)^r \times V^s \rightarrow K$ se llama un tensor de tipo (r, s) sobre V . (Se entiende que $A: V^s \rightarrow K$ si $r=0$, y que $A: (V^*)^r \rightarrow K$ si $s=0$.)

El conjunto $\mathcal{T}_{rs}^r(V)$ de todos los tensores de tipo (r, s) sobre V es

un módulo sobre K , si nuevamente definimos la suma de tensores del mismo tipo y la multiplicación de tensores por elementos de K , de la misma manera que las correspondientes definiciones para funciones. Un tensor del tipo $(0,0)$ sobre V es simplemente un elemento de K .

15 CAMPOS TENSORIALES

Un campo tensorial A sobre una variedad M es un tensor sobre el $\mathcal{F}(M)$ -módulo $\mathcal{K}(M)$, tal como lo hemos definido arriba. Así si A es del tipo (r,s) , A es una función $\mathcal{F}(M)$ -multilineal

$$A: \mathcal{K}^*(M)^r \times \mathcal{K}(M)^s \rightarrow \mathcal{F}(M).$$

Podemos decir que A es una máquina multilineal que cuando se le alimenta con r uno-formas $\theta^1, \dots, \theta^r$ y con s campos vectoriales X_1, \dots, X_s produce una función de valores reales

$$f = A(\theta^1, \dots, \theta^r, X_1, \dots, X_s) \in \mathcal{F}(M).$$

Se dice que θ^i ocupa la i -ésima posición contravariante, y que X_j ocupa la j -ésima posición covariante en el campo tensorial A .

El conjunto $\mathcal{F}_0^r(M)$ de todos los campos tensoriales sobre M del tipo (r,s) es entonces un módulo sobre $\mathcal{F}(M)$. En el caso excepcional $r=s=0$, un campo tensorial sobre M de tipo $(0,0)$ es simplemente una función $f \in \mathcal{F}(M)$; es decir, $\mathcal{F}_0^0(M) = \mathcal{F}(M)$.

Para demostrar que una función dada $A: \mathcal{K}^*(M)^r \times \mathcal{K}(M)^s \rightarrow \mathcal{F}(M)$ es un campo tensorial debemos demostrar que es $\mathcal{F}(M)$ -lineal en cada variable. La aditividad en cada variable es casi siempre obvia, de modo que la cuestión crucial es verificar que las funciones $f \in \mathcal{K}(M)$ se pueden factorizar fuera de cada variable:

$$A(\theta^1, \dots, \theta^r, X_1, \dots, fX_i, \dots, X_s) = fA(\theta^1, \dots, \theta^r, X_1, \dots, X_i, \dots, X_s).$$

Consideremos los siguientes ejemplos. (1) La función evaluación $E: \mathcal{K}^*(M) \times \mathcal{K}(M) \rightarrow \mathcal{F}(M)$ dada por $E(\theta, X) = \theta X$. Claramente E es $\mathcal{F}(M)$ -lineal en cada variable y es, por lo tanto, un campo tensorial sobre M de tipo $(1,1)$.

(2) Fijemos una uno-forma $\omega \neq 0$ y definamos $F: \mathcal{K}(M) \times \mathcal{K}(M) \rightarrow \mathcal{F}(M)$ por $F(X, Y) = X(\omega Y)$ para todo X, Y . Entonces F es $\mathcal{F}(M)$ -lineal en la variable X , pero sólo es aditiva en la variable Y . De hecho,

$$F(X, fY) = X(\omega(fY)) = X(f\omega Y) = (Xf)\omega Y + fF(X, Y).$$

Así pues, F no es un campo tensorial.

En estos ejemplos hacemos notar que el tensor (1) sólo involucra operaciones algebraicas, mientras que en (2) aparece la diferenciación.

Si bien sólo es posible sumar tensores del mismo tipo, cualesquiera dos tensores se pueden multiplicar como sigue: Si $A \in \mathcal{F}_s^r(M)$ y $B \in \mathcal{F}_t^u(M)$, definimos

$$A \otimes B: \mathcal{K}^*(M)^{r+r'} \times \mathcal{K}^*(M)^{s+s'} \longrightarrow \mathcal{F}(M)$$

por

$$\begin{aligned} (A \otimes B)(\theta^1, \dots, \theta^{r+r'}, X_1, \dots, X_{s+s'}) &= \\ &= A(\theta^1, \dots, \theta^r, X_1, \dots, X_s) B(\theta^{r+1}, \dots, \theta^{r+r'}, X_{s+1}, \dots, X_{s+s'}). \end{aligned}$$

Entonces $A \otimes B$ es un tensor de tipo $(r+r', s+s')$, llamado el producto tensorial de A y B . Si $r'=s'=0$, de modo que B es una función $f \in \mathcal{F}(M)$, definimos

$$A \otimes f = f \otimes A = fA.$$

También si A es del tipo $(0,0)$, el producto tensorial se reduce a la multiplicación ordinaria de elementos de $\mathcal{F}(M)$.

Evidentemente el producto tensorial es $\mathcal{F}(M)$ -bilineal, es decir,

$$(fA + gA') \otimes B = fA \otimes B + gA' \otimes B,$$

con una identidad similar para B . Más aún, es inmediato de la definición que el producto tensorial es asociativo; así $A \otimes B \otimes C$ está bien definido para tensores de cualquier tipo. Sin embargo, el producto tensorial no es conmutativo en general. Por ejemplo, sobre una vecindad coordinada tenemos

$$\begin{aligned} (dx^1 \otimes dx^2)(\partial_1, \partial_2) &= dx^1(\partial_1) dx^2(\partial_2) = 1, \\ (dx^2 \otimes dx^1)(\partial_1, \partial_2) &= dx^2(\partial_1) dx^1(\partial_2) = 0, \end{aligned}$$

por lo que $dx^1 \otimes dx^2 \neq dx^2 \otimes dx^1$. Por otra parte, las funciones conmutan con el producto tensorial:

$$f(A \otimes B) = fA \otimes B = A \otimes fB.$$

16 INTERPRETACIONES

Si ω es una uno-forma suave, es decir, un elemento de $\mathcal{K}^*(M)$, entonces la función $X \rightarrow \omega(X)$ es una función $\mathcal{F}(M)$ -lineal de $\mathcal{K}(M) \rightarrow \mathcal{F}(M)$ o sea que se trata de un campo tensorial de tipo $(0,1)$. Inversamente, todo campo tensorial del tipo $(0,1)$ surge de esta manera a partir de una única uno-forma. Así pues, queda claro que los módulos sobre $\mathcal{F}(M)$, $\mathcal{X}_1^0(M)$ y $\mathcal{K}^*(M)$ son isomorfos. Escribiremos esto simplemente como

$$\mathcal{X}_1^0(M) = \mathcal{K}^*(M).$$

Hay otras dos interpretaciones menos obvias que usaremos con frecuencia en lo que sigue.

(1) Si V es un campo vectorial suave sobre M , definimos

$$V(\theta) = \theta(V) \quad \text{para toda } \theta \in \mathcal{K}^*(M).$$

Esta función $V: \mathcal{K}^*(M) \rightarrow \mathcal{F}(M)$ es $\mathcal{F}(M)$ -lineal, por lo tanto, define un campo tensorial del tipo $(1,0)$. Inversamente, todo campo tensorial de tipo $(1,0)$ sobre M surge de esta manera a partir de un único campo

vectorial. Para ver esto, recordemos del álgebra lineal (ver [H], pag. 188) que si W es un espacio vectorial de dimensión finita y $\phi \in (W^*)^*$, entonces existe un único $v \in W$ tal que $\alpha(v) = \phi(\alpha)$, $\alpha \in W^*$. La aplicación $\phi \rightarrow v$ es un isomorfismo lineal entre $(W^*)^* = W^{**}$ y W ; gracias a este isomorfismo, podemos hacer la identificación $W = W^{**}$, para espacios vectoriales de dimensión finita. Ahora bien, si $Z \in \mathcal{X}'_0(M)$, entonces $Z_p \in T_p(M)^{**} = T_p(M)$, por lo que la correspondencia $p \rightarrow Z_p$ nos define el campo vectorial sobre M que estábamos buscando.

Así pues, podemos hacer la identificación

$$\mathcal{X}'_0(M) = \mathcal{X}(M).$$

(2) Si $A: \mathcal{X}(M)^s \rightarrow \mathcal{X}(M)$ es $\mathcal{X}(M)$ -multilineal, definamos $\bar{A}: \mathcal{X}^*(M) \times \mathcal{X}(M)^s \rightarrow \mathcal{X}(M)$ por

$$\bar{A}(\theta, X_1, \dots, X_s) = \theta(A(X_1, \dots, X_s)), \quad \text{para toda } \theta \text{ y } X_i.$$

Evidentemente \bar{A} es $\mathcal{X}(M)$ -multilineal y, por lo tanto, es un campo tensorial de tipo $(1, s)$. En lo sucesivo, consideraremos que la propia A es un campo tensorial, utilizando la fórmula anterior cuando sea necesario.

Los tensores de tipo $(0, s)$ se conocen como tensores covariantes, mientras que los tensores de tipo $(r, 0)$ con $r \geq 1$ se llaman tensores contravariantes. Por ejemplo, las funciones con valores reales y las uno-formas son tensores covariantes; los campos vectoriales, en cambio, son tensores contravariantes. Un tensor de tipo (r, s) se llama mixto si $r \neq 0$ y $s \neq 0$. Nótese que por la definición del producto tensorial, se sigue que si A es covariante y B es contravariante entonces, $A \otimes B = B \otimes A$.

17 TENSORES EN UN PUNTO

El objetivo de esta sección es demostrar que al igual que para los campos vectoriales y las uno-formas, cualquier campo tensorial A sobre M puede ser considerado como un campo sobre M ; es decir, una aplicación que asigna un valor A_p a cada punto p de M . El hecho fundamental es que cuando A se evalúa sobre uno-formas y campos vectoriales para dar una función con valores reales

$$A(\theta^1, \dots, \theta^r, X_1, \dots, X_s),$$

el valor de esta función en un punto $p \in M$ no depende de como sean las uno-formas $\theta^1, \dots, \theta^r$ y los campos vectoriales X_1, \dots, X_s , ni siquiera depende de los valores que tomen en una vecindad de p , sino que únicamente depende de sus valores $\theta^i|_p, \dots, \theta^r|_p$ y $X_1|_p, \dots, X_s|_p$ en el propio punto p . Formalmente, tenemos el siguiente resultado:

Proposición 17.1 Sea $p \in M$ y $A \in \mathcal{X}^r(M)$. Sean $\bar{\theta}^1, \dots, \bar{\theta}^r$ y $\theta^1, \dots, \theta^r$ uno-formas tales que $\theta^i|_p = \bar{\theta}^i|_p$ ($1 \leq i \leq r$); sean $\bar{X}_1, \dots, \bar{X}_s$ y X_1, \dots, X_s campos vectoriales tales que $\bar{X}_j|_p = X_j|_p$ ($1 \leq j \leq s$). Entonces

$$A(\bar{\theta}^1, \dots, \bar{\theta}^r, \bar{X}_1, \dots, \bar{X}_s)(p) = A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)(p).$$

La demostración de esta proposición será muy fácil una vez que quede establecido el siguiente resultado.

Lema 17.2 Si cualquiera de las uno-formas $\theta^1, \dots, \theta^r$ o de los campos vectoriales X_1, \dots, X_s es cero en p , entonces

$$A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)(p) = 0.$$

Demostración. Supongamos que, por ejemplo, $X_s|_p = 0$. Sea x^1, \dots, x^n un sistema coordinado sobre una vecindad U de p . Entonces $X_s = \sum X^i \partial_i$ sobre U , donde $X^i = X_s x^i \in \mathcal{F}(U)$ son las componentes de X_s en estas coordenadas. Sea f una función de promontorio suave en p con soporte contenido en U (ver pag. 12). Con este promontorio hemos logrado que fX^i sea una función suave definida sobre toda la variedad M , es decir, $fX^i \in \mathcal{F}(M)$; similarmente, $f\partial_i \in \mathcal{K}(M)$. Por lo tanto,

$$\begin{aligned} f^2 A(\theta^1, \dots, X_s) &= A(\theta^1, \dots, f^2 X_s) = A(\theta^1, \dots, \sum fX^i f\partial_i) \\ &= \sum_{i=1}^n fX^i A(\theta^1, \dots, f\partial_i). \end{aligned}$$

Puesto que $X_s|_p = 0$, cada $X^i(p) = 0$; también $f(p) = 1$. Por lo tanto, al evaluar la fórmula anterior en p , obtenemos

$$A(\theta^1, \dots, X_s)|_p = 0. \quad \#\#$$

Demostración de la proposición 17.1 Por claridad supongamos que $r=1$ y $s=2$. Consideremos la siguiente identidad telescópica, que se puede extender sin dificultad para cualesquiera r y s :

$$A(\bar{\theta}, \bar{X}, \bar{Y}) - A(\theta, X, Y) = A(\bar{\theta} - \theta, \bar{X}, \bar{Y}) + A(\theta, \bar{X} - X, \bar{Y}) + A(\theta, X, \bar{Y} - Y).$$

Por hipótesis, $\bar{\theta} - \theta$, $\bar{X} - X$, y $\bar{Y} - Y$ se anulan en el punto p . Entonces, por el lema anterior, $A(\bar{\theta}, \bar{X}, \bar{Y}) = A(\theta, X, Y)$ en el punto p . $\#\#$

Se sigue inmediatamente de la proposición 17.1 que un campo tensorial $A \in \mathcal{T}_s^r(M)$ tiene un valor A_p en cada punto p de M , a saber, la función

$$A_p: (T_p M^*)^r \times (T_p M)^s \rightarrow \mathbb{R}$$

definida como sigue. Si $\alpha^1, \dots, \alpha^r \in T_p(M)^*$ y $x_1, \dots, x_s \in T_p(M)$, sea

$$A_p(\alpha^1, \dots, \alpha^r, x_1, \dots, x_s) = A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)(p),$$

donde $\theta^1, \dots, \theta^r$ son cualesquiera uno-formas sobre M tales que $\theta^i|_p = \alpha^i$ ($1 \leq i \leq r$) y X_1, \dots, X_s son cualesquiera campos vectoriales sobre M tales que $X_j|_p = x_j$ ($1 \leq j \leq s$).

Es fácil verificar que la función A_p es \mathbb{R} -multilineal; entonces, por definición, A_p es un tensor de tipo (r, s) sobre $T_p(M)$. Podemos así considerar a $A \in \mathcal{T}_s^r(M)$ como un campo que asigna suavemente a cada punto $p \in M$ un tensor A_p de tipo (r, s) . Tal como un campo vectorial es una sección suave del haz tangente TM , un campo tensorial

puede ser considerado como una sección suave del haz tensorial de tipo (r,s) sobre la variedad M ; este último se denota por TM_s^r y se puede obtener a partir de TM reemplazando cada espacio $T_p(M)$ por el espacio lineal $T_p(M)_s^r$ de todos los tensores de tipo (r,s) sobre $T_p(M)$. En la próxima sección quedará claro que el espacio $T_p(M)_s^r$ es de dimensión n^{r+s} en donde $n = \dim M$. Así pues, mientras que TM es una variedad de dimensión $n + n$, TM_s^r será una variedad de dimensión $n + n^{r+s}$.

Por otra parte, una sección suave del haz tensorial, digamos por ejemplo, $p \rightarrow B_p \in T_p(M)_2^1$ surge del único tensor $B \in \mathcal{T}_2^1(M)$ dado por

$$B(\theta, X, Y)(p) = B_p(\theta_p, X_p, Y_p)$$

para todos los puntos p , uno-formas θ , y campos vectoriales X, Y . (Que la sección sea suave quiere decir que los valores de B pertenecen a $\mathcal{T}(M)$.)

En particular, la interpretación en términos de secciones suaves muestra que si $A \in \mathcal{T}_s^r(M)$ y U es un subconjunto abierto de M , entonces la restricción $A|_U$ de A a U es un campo tensorial bien definido sobre U .

18 COMPONENTES TENSORIALES

Las fórmulas con coordenadas $X = \sum X(x^i) \partial_i$ para un campo vectorial y $\theta = \sum \theta(\partial_i) dx^i$ para una uno-forma se pueden extender al caso de tensores arbitrarios de cualquier tipo.

Sea $\xi = (x^1, \dots, x^n)$ un sistema coordinado sobre $U \subset M$. Si $A \in \mathcal{T}_s^r(M)$ las componentes de A con respecto al sistema coordinado ξ son las funciones con valores reales definidas por

$$A_{j_1 \dots j_s}^{i_1 \dots i_r} = A(dx^{i_1}, \dots, dx^{i_r}, \partial_{j_1}, \dots, \partial_{j_s}) \text{ sobre } U,$$

donde todos los índices corren de 1 a $n = \dim M$.

Evidentemente para un tensor de tipo $(0,1)$, es decir, una uno-forma, estas componentes coinciden con las de la fórmula $\theta = \sum \theta(\partial_i) dx^i$. Para ver que la correspondiente coincidencia se sostiene para un campo vectorial X , debemos usar su interpretación como un campo tensorial de tipo $(1,0)$. Por la definición de arriba, la i -ésima componente de X con respecto a ξ es $X(dx^i)$, la cual se interpreta como $dx^i(X) = Xx^i$.

Similarmente, cuando un campo tensorial de tipo $(1,s)$ está dado en la forma $A: \mathcal{K}(M)^s \rightarrow \mathcal{K}(M)$, sus componentes están determinadas directamente por la ecuación

$$A(\partial_{i_1}, \dots, \partial_{i_s}) = \sum_j A_{i_1 \dots i_s}^j \partial_j,$$

ya que para su interpretación $\bar{A} \in \mathcal{T}_s^1(M)$ tenemos

$$\bar{A}(dx^j, \partial_{i_1}, \dots, \partial_{i_s}) = dx^j(A(\partial_{i_1}, \dots, \partial_{i_s})) = \sum_k A_{i_1 \dots i_s}^k dx^j(\partial_k) = A_{i_1 \dots i_s}^j.$$

La evaluación de un campo tensorial sobre uno-formas y campos vectoriales puede ser descrita en términos de coordenadas escribiéndolo todo en componentes. Por ejemplo, supongamos que A es un campo tensorial de tipo $(1,2)$. Escribamos $\theta = \sum \theta_k dx^k$ para una uno-forma arbitraria y $X = \sum X^i \partial_i$, $Y = \sum Y^j \partial_j$ para campos vectoriales arbitrarios. La $\mathcal{F}(M)$ -multilinealidad de A implica que

$$A(\theta, X, Y) = \sum_{i,j,k} A(dx^k, \partial_i, \partial_j) \theta_k X^i Y^j = \sum_{i,j,k} A_{ij,k}^i \theta_k X^i Y^j.$$

Para un sistema fijo de coordenadas, Las componentes de una suma de tensores son las sumas de las componentes respectivas de los sumandos. (Recuérdese que sólo está permitido sumar tensores del mismo tipo.) Las componentes de un producto tensorial están dadas por

$$(A \otimes B)_{j_1, \dots, j_{s+1}}^{i_1, \dots, i_{r+r'}} = A_{j_1, \dots, j_s}^{i_1, \dots, i_r} \cdot B_{j_{s+1}, \dots, j_{s+r'}}^{i_{r+1}, \dots, i_{r+r'}},$$

donde, como siempre, todos los índices corren de 1 a $n = \dim M$. Para verificar esta fórmula, supongamos que A es de tipo $(1,2)$ y B es de tipo $(1,1)$. Entonces $A \otimes B$ es un tensor de tipo $(2,3)$ con componentes:

$$\begin{aligned} (A \otimes B)_{ijp}^{kq} &= (A \otimes B)(dx^k, dx^q, \partial_i, \partial_j, \partial_p) \\ &= A(dx^k, \partial_i, \partial_j) B(dx^q, \partial_p) = A_{ij}^k B_p^q. \end{aligned}$$

Sea ξ un sistema coordenado sobre UCM. Entonces tal como para un campo vectorial o para una uno-forma, cualquier tensor tiene una única expresión sobre U en términos de componentes relativas al sistema ξ . Supongamos por ejemplo que $r=1$ y $s=2$. Entonces $\partial_k \otimes dx^i \otimes dx^j$ es un campo tensorial de tipo $(1,2)$ sobre U para todo $1 \leq i, j, k \leq n$. Afirmamos que si A es cualquier tensor de tipo $(1,2)$, entonces

$$A = \sum A_{ij}^k \partial_k \otimes dx^i \otimes dx^j \quad \text{sobre } U,$$

donde cada índice es un índice mudo de suma que corre de 1 a n . Puesto que ambos miembros de esta ecuación son $\mathcal{F}(U)$ -multilineales, es suficiente verificar que ambos toman el mismo valor sobre $dx^m, \partial_p, \partial_q$ para todo $1 \leq m, p, q \leq n$. Pero esto se sigue inmediatamente de

$$(\partial_k \otimes dx^i \otimes dx^j)(dx^m, \partial_p, \partial_q) = dx^m(\partial_k) dx^i(\partial_p) dx^j(\partial_q) = \delta_k^m \delta_i^p \delta_j^q,$$

donde para mantener el balance entre subíndices y superíndices usamos las siguientes definiciones de las deltas de Kronecker

$$\delta_{ij} = \delta_i^j = \delta^{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

En general tenemos el siguiente resultado.

Lema 18.1 Sea x^1, \dots, x^n un sistema coordenado sobre UCM. Si A es un campo tensorial de tipo (r,s) , entonces sobre U tenemos,

$$A = \sum A_{j_1, \dots, j_s}^{i_1, \dots, i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s},$$

donde se suma con respecto a cada índice de 1 a n .

19 CONTRACCION

Existe una operación notable llamada contracción que convierte tensores de tipo (r, s) en tensores de tipo $(r-1, s-1)$. La definición general de esta operación surge del siguiente caso especial.

Lema 19.1 Existe una única función $\mathcal{F}(M)$ -lineal $C: \mathcal{L}_1^1(M) \rightarrow \mathcal{L}_0^0(M) = \mathcal{F}(M)$, llamada la contracción (1,1), tal que $C(X \otimes \theta) = \theta X$ para todo $X \in \mathcal{L}_1^1(M)$ y toda $\theta \in \mathcal{L}_0^1(M)$.

Demostración. Dado que C debe ser $\mathcal{F}(M)$ -lineal, C debe ser una operación puntual. Sobre una vecindad coordinada U , un campo tensorial A de tipo $(1,1)$ se puede escribir como $\sum A_j^i \delta_i \otimes dx^j$. Puesto que $C(\delta_i \otimes dx^j)$ debe ser igual a $dx^j(\delta_i) = \delta_j^i$, no nos queda otra posibilidad que definir

$$C(A) = \sum A_i^i = \sum A(dx^i, \delta_i).$$

Entonces C tiene las propiedades requeridas sobre UCM. Para ver que la C que hemos definido sobre U está bien definida sobre toda la variedad M , debemos demostrar que esta definición es independiente de la elección de sistema coordinado. Pero tenemos que

$$\begin{aligned} A(dy^m, \frac{\partial}{\partial y^m}) &= A(\sum \frac{\partial y^m}{\partial x^i} dx^i, \sum \frac{\partial x^j}{\partial y^m} \frac{\partial}{\partial x^j}) \\ &= \sum_{i,j,m} \frac{\partial y^m \partial x^i}{\partial x^i \partial y^m} A(dx^i, \frac{\partial}{\partial x^j}) = \sum_{i,j} \delta_j^i A(dx^i, \frac{\partial}{\partial x^j}) \\ &= \sum_i A(dx^i, \frac{\partial}{\partial x^i}). \quad \#\# \end{aligned}$$

Es evidente que la contracción $(1,1)$ está relacionada con el concepto de traza del álgebra matricial.

Para extender la contracción $(1,1)$ a tensores de tipo general, la idea es especificar en estos tensores una posición covariante y una posición contravariante, y aplicar la función C en este caso.

Supongamos que $A \in \mathcal{L}_s^r(M)$, que $1 \leq i \leq r$ y que $1 \leq j \leq s$. Fijemos las unoformas $\theta^1, \dots, \theta^{r-1}$ y los campos vectoriales X_1, \dots, X_{s-1} . Entonces la función

$$(\theta, X) \rightarrow A(\theta^1, \dots, \underbrace{\theta^j}_{j\text{-ÉSIMA POSICIÓN}}, \dots, \theta^{r-1}, \underbrace{X_1, \dots, X_j, \dots, X_{s-1}}_{i\text{-ÉSIMA POSICIÓN})$$

es un campo tensorial de tipo $(1,1)$ que puede denotarse como

$$A(\theta^1, \dots, \theta^j, \dots, \theta^{r-1}, X_1, \dots, X_j, \dots, X_{s-1}).$$

Aplicando la contracción $(1,1)$ a este tensor obtenemos una función con valores reales que se denota por

$$(C_j^i A)(\theta^1, \dots, \theta^{r-1}, X_1, \dots, X_{s-1}).$$

Es evidente que $C_j^i A$ es una función $\mathcal{F}(M)$ -multilineal en sus argumen-

tos. Por lo tanto, $C_j^i A$ es un tensor de tipo $(r-1, s-1)$ llamado la contracción de A con respecto a la i-ésima posición contravariante y la j-ésima posición covariante.

Por ejemplo, si A es un tensor de tipo $(2, 3)$, entonces $C_j^i(A)$ es un tensor de tipo $(1, 2)$ dado por

$$(C_j^i A)(\theta, X, Y) = C_j^i(A(\cdot, \cdot, X, Y, \cdot))$$

Con respecto a un sistema coordenado las componentes de $C_j^i A$ son

$$\begin{aligned} (C_j^i A)_{ij}^k &= (C_j^i A)(dx^k, \partial_i, \partial_j) = C_j^i(A(\cdot, dx^k, \partial_i, \partial_j, \cdot)) \\ &= \sum_m A(dx^m, dx^k, \partial_i, \partial_j, \partial_m) = \sum_m A_{ijm}^m{}^k, \end{aligned}$$

donde hemos usando la expresión en términos de coordenadas para la función C que encontramos en el lema 19.1.

En general se tiene el siguiente resultado.

Corolario 19.2 Sean $1 \leq i \leq r$ y $1 \leq j \leq s$. Si con respecto a un sistema coordenado el campo tensorial $A \in \mathcal{X}_s^r(M)$ tiene las componentes $A_{j_1, \dots, j_s}^{i_1, \dots, i_r}$, entonces $C_j^i A$ tiene las componentes

$$\sum_m A_{j_1, \dots, j_s}^{i_1, \dots, i_r, m} \quad \begin{array}{l} \text{--- } i\text{-ésimo índice} \\ \text{--- } j\text{-ésimo índice} \end{array}$$

20 TENSORES COVARIANTES

Mediante una aplicación suave entre la variedad M y la variedad N, es posible "jalar" todo campo tensorial covariante definido sobre N y considerarlo como si fuera un campo tensorial del mismo tipo pero definido sobre M. En la siguiente definición se usa la interpretación de los campos tensoriales como secciones suaves del haz tensorial.

Sea $\phi: M \rightarrow N$ una aplicación suave. Si $A \in \mathcal{X}_s^0(N)$ con $s \geq 1$ es un campo tensorial covariante sobre N, definimos

$$(\phi^* A)(v_1, \dots, v_s) = A(d\phi v_1, \dots, d\phi v_s)$$

para todo $v_i \in T_p(M)$ y todo $p \in M$. Entonces $\phi^*(A)$ se llama la retracción de A por ϕ .

En cada punto p de M, $\phi^*(A)$ da una función \mathbb{R} -multilineal de $T_p(M)$ a \mathbb{R} , es decir, un tensor sobre $T_p(M)$ de tipo $(0, s)$. Un cómputo simple usando coordenadas, demuestra que $\phi^*(A)$ es un campo tensorial covariante suave definido sobre M. En efecto, sea $\phi: M^m \rightarrow N^n$ una aplicación suave que aplica la vecindad coordenada $U \subset M$ de $\xi = (x^1, \dots, x^m)$ en la vecindad coordenada $V \subset N$ de $\eta = (y^1, \dots, y^n)$. Si A es un campo tensorial covariante sobre N que es suave y si, para simplificar suponemos que es de tipo $(0, 2)$, vemos que sobre U es válida la siguiente ecuación

$$(\phi^* A) \left(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b} \right) = \sum_{a, b=1}^n \frac{\partial(y^a \circ \phi)}{\partial x^a} \frac{\partial(y^b \circ \phi)}{\partial x^b} A \left(\frac{\partial}{\partial y^a}, \frac{\partial}{\partial y^b} \right) \circ \phi.$$

Esta ecuación se sigue inmediatamente de la definición de $\phi^*(A)$ y de la ecuación

$$d\phi\left(\frac{\partial}{\partial x^i}\right) = \sum_{\alpha=1}^n \frac{\partial(y^\alpha \circ \phi)}{\partial x^\alpha} \frac{\partial}{\partial y^\alpha} \quad (1 \leq i < m),$$

que es la expresión en coordenadas de la diferencial $d\phi$. Ahora vemos que en el lado izquierdo de esa ecuación tenemos una función suave sobre M , es decir, un elemento de $\mathcal{F}(M)$, como queríamos demostrar. (Para ver esto es suficiente evaluar ambos miembros de la ecuación en un punto $p \in U$.)

En el caso especial de un tensor de tipo $(0,0)$, $f \in \mathcal{F}(N)$, la retracción a M se define como $\phi^*(f) = f \circ \phi \in \mathcal{F}(M)$. Nótese que $\phi^*(df) = d(\phi^*f)$. En efecto, para todo $p \in M$ y para todo $v \in T_p(M)$ tenemos

$$\phi^*(df)(v) = df(d\phi v) = (df \circ d\phi)(v) = d(f \circ \phi)(v) = d(\phi^*f)(v).$$

Las siguientes propiedades de la operación de retracción son muy fáciles de verificar a partir de las definiciones.

Lema 20.1 (1) Si $\phi: M \rightarrow N$ es una aplicación suave, entonces $\phi^*: \mathcal{F}_s^0(N) \rightarrow \mathcal{F}_s^0(M)$ es una función \mathbb{R} -lineal para cada $s \geq 0$, y además

$$\phi^*(A \otimes B) = \phi^*(A) \otimes \phi^*(B)$$

para tensores arbitrarios A y B de tipos $(0,s)$ y $(0,t)$, respectivamente.

(2) Si $\psi: N \rightarrow P$ es también una aplicación suave, entonces

$$(\psi \circ \phi)^* = \phi^* \circ \psi^*: \mathcal{F}_s^0(P) \rightarrow \mathcal{F}_s^0(M)$$

para toda $s \geq 0$.

La propiedad de retracción es exclusiva de los tensores covariantes, ya que, en general, un tensor de tipo (r,s) con $r \geq 1$ no puede ser movido ni de M a N ni de N a M mediante una aplicación arbitraria $\phi: M \rightarrow N$.

Sea A un tensor covariante o contravariante (pero no mixto) de tipo 2, por lo menos. A se llama simétrico si al transponer cualesquiera dos de sus argumentos el valor de A no cambia. A se llama antisimétrico o alternante si al intercambiar cualesquiera dos de sus argumentos sólo se produce un cambio de signo en el valor de A . Las funciones, las uno-formas y los campos vectoriales se consideran, por convención, tanto simétricos como antisimétricos. Una s-forma (diferencial) es un campo tensorial antisimétrico covariante de tipo $(0,s)$. Para mayores detalles acerca del cálculo con s -formas, consultar la parte III de estas notas.

Lema 20.2 Sea ω una n -forma sobre una variedad de dimensión n . Si $V_i = \sum_{j=1}^n A_{ij} W_j$ para $1 \leq i \leq n$, entonces

$$\omega(V_1, \dots, V_n) = (\det A) \omega(W_1, \dots, W_n).$$

Demostración. Esta demostración se hace mediante un argumento combinatorio algo complicado, sin embargo, podemos convencernos de la validez de este resultado, considerando el caso $n = 2$. Supongamos entonces que $V_1 = A_{11} W_1 + A_{12} W_2$ y que $V_2 = A_{21} W_1 + A_{22} W_2$. Puesto que ω es lineal en su primera variable, su valor en los campos vectoriales V_1 y V_2 es

$$A_{11} \omega(W_1, A_{21} W_1 + A_{22} W_2) + A_{12} \omega(W_2, A_{21} W_1 + A_{22} W_2).$$

Al usar la linealidad de ω en su segunda variable, obtenemos

$$A_{11} A_{21} \omega(W_1, W_1) + A_{11} A_{22} \omega(W_1, W_2) + A_{12} A_{21} \omega(W_2, W_1) + A_{12} A_{22} \omega(W_2, W_2).$$

Entonces, por la antisimetría de ω tenemos que $\omega(W_1, W_2) = -\omega(W_2, W_1)$ de donde $\omega(W_1, W_1) = \omega(W_2, W_2) = 0$ y

$$\omega(V_1, V_2) = (A_{11} A_{22} - A_{12} A_{21}) \omega(W_1, W_2) = (\det A) \omega(W_1, W_2). \quad \#\#$$

Corolario 20.3 Sean M y N variedades de dimensión n , ω una n -forma sobre N y sea $\phi: M \rightarrow N$ una aplicación suave que manda la vecindad $UC M$ de $\xi = (x^1, \dots, x^n)$ en la vecindad coordenada $VC N$ de $\eta = (y^1, \dots, y^n)$. Entonces,

$$\phi^* \omega \left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right) = \det \left(\frac{\partial (y^i \circ \phi)}{\partial x^j} \right) \omega \left(\frac{\partial}{\partial y^1}, \dots, \frac{\partial}{\partial y^n} \right),$$

es la expresión coordenada para la retracción de ω por ϕ .

21 DERIVACIONES TENSORIALES

En las secciones anteriores hemos considerado principalmente el álgebra tensorial sobre variedades; consideraremos ahora algo del cálculo diferencial de tensores sobre variedades.

Una derivación tensorial D sobre una variedad suave M es un conjunto de funciones \mathbb{R} -lineales

$$D = D_s^r: \mathcal{F}_s^r(M) \rightarrow \mathcal{F}_s^r(M) \quad (r \geq 0, s \geq 0)$$

tales que para cualesquiera tensores A y B :

$$(1) D(A \otimes B) = DA \otimes B + A \otimes DB,$$

$$(2) D(CA) = C(DA) \text{ para cualquier contracción } C.$$

Así pues, D es \mathbb{R} -lineal, preserva el tipo tensorial, obedece la regla de Leibnitz para la derivada de un producto, y conmuta con todas las contracciones. Para una función $f \in \mathcal{F}(M)$ recuérdese que $fA = f \otimes A$, por lo tanto, $D(fA) = (Df)A + fDA$.

En el caso especial $r = 0 = s$, D_s^r es una derivación sobre $\mathcal{F}_0^0(M) = \mathcal{F}(M)$ así que, como vimos en la parte I, existe un único campo vectorial $V_f \in \mathcal{X}(M)$ tal que

$$Df = V_f \text{ para toda } f \in \mathcal{F}(M).$$

Puesto que las derivaciones tensoriales no son $\mathcal{F}(M)$ -lineales, el valor de DA en un punto $p \in M$ no se puede encontrar, por lo general, sólo a partir del valor A_p de A en p . Sin embargo, el valor de DA en p sí puede encontrarse a partir de los valores de A sobre una vecindad arbitrariamente pequeña del punto p . Este carácter local de las derivaciones tensoriales puede expresarse como sigue.

Proposición 21.1 Si D es una derivación sobre M y U es un subconjunto abierto de M , entonces existe una única derivación tensorial D_U sobre U tal que

$$D_U(A|U) = (DA)|U \quad \text{para todos los tensores } A \text{ sobre } M.$$

(D_U se llama la restricción de D a U y, de ahora en adelante, omitiremos el subíndice U cuando no haya lugar a confusión.)

Demostración. Sólo indicaremos los puntos más importantes de la demostración. Sea $B \in \mathcal{X}_s^r(U)$. Si $p \in U$, sea f una función de promontorio en p con soporte en U . Así pues, tenemos $fB \in \mathcal{X}_s^r(M)$. Definamos

$$(D_U B)_p = D(fB)_p.$$

Entonces, debemos demostrar que: (1) Esta definición es independiente de la elección de la función de promontorio. (2) $D_U B$ es un campo tensorial suave sobre U . (3) D_U es una derivación tensorial sobre U . (4) D_U tiene la propiedad de restricción que estamos buscando. (5) D_U es única. ##

La fórmula Leibniziana $D(A \otimes B) = DA \otimes B + A \otimes DB$ puede ser reformulada como sigue:

Proposición 21.2 (Regla del producto). Sea D una derivación tensorial sobre M . Si $A \in \mathcal{X}_s^r(M)$ entonces

$$\begin{aligned} D[A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)] &= (DA)(\theta^1, \dots, \theta^r, X_1, \dots, X_s) \\ &+ \sum_{i=1}^r A(\theta^1, \dots, D\theta^i, \dots, \theta^r, X_1, \dots, X_s) \\ &+ \sum_{j=1}^s A(\theta^1, \dots, \theta^r, X_1, \dots, DX_j, \dots, X_s). \end{aligned}$$

(La colocación de los paréntesis es importante en esta fórmula: en el lado izquierdo D se aplica a una función, en el lado derecho D se aplica al tensor A , a las uno-formas θ^i y a los campos vectoriales X_j .)

Demostración. Para simplificar la notación, supongamos que $r=s=1$. Afirmamos que

$$A(\theta, X) = \bar{C}(A \otimes \theta \otimes X),$$

donde \bar{C} es la composición de dos contracciones. De hecho, con respecto a un sistema coordenado, $A \otimes \theta \otimes X$ tiene componentes $A^i_j \theta_k X^l$,

mientras que $A(\theta, X) = \sum A_i^j \theta_i X^j$. Así pues, tenemos

$$\begin{aligned} D(A(\theta, X)) &= D\bar{C}(A \otimes \theta \otimes X) = \bar{C}D(A \otimes \theta \otimes X) \\ &= \bar{C}(DA \otimes \theta \otimes X) + \bar{C}(A \otimes D\theta \otimes X) + \bar{C}(A \otimes \theta \otimes DX) \\ &= (DA)(\theta, X) + A(D\theta, X) + A(\theta, DX). \quad \#\# \end{aligned}$$

Para un campo tensorial expresado como una función $\mathcal{F}(M)$ -multilineal $A: \mathcal{K}(M)^s \rightarrow \mathcal{K}(M)$, la derivación tensorial obedece la misma regla formal del producto, a saber,

$$D(A(X_1, \dots, X_s)) = (DA)(X_1, \dots, X_s) + \sum_{i=1}^s A(X_1, \dots, DX_i, \dots, X_s).$$

En ambas versiones de la regla del producto, se despeja por lo general al término que contiene a DA . De este modo se obtiene una fórmula para la derivación D de un tensor arbitrario en términos de la derivación D aplicada únicamente a funciones, campos vectoriales y uno-formas. Pero, para uno-formas, la fórmula da

$$(D\theta)(X) = D(\theta X) - \theta(DX).$$

Así pues, una derivación tensorial D queda completamente determinada si se conoce cual es su efecto sobre funciones y campos vectoriales.

Corolario 21.3 Si las derivaciones tensoriales D_1 y D_2 concuerdan al aplicarse a funciones de $\mathcal{F}(M)$ y campos de $\mathcal{K}(M)$, entonces $D_1 = D_2$.

Más aún, se puede contruir una derivación tensorial especificando su comportamiento sobre funciones de $\mathcal{F}(M)$ y campos de $\mathcal{K}(M)$.

Teorema 21.4 Dado un campo vectorial $V \in \mathcal{K}(M)$ y dada una función \mathbb{R} -lineal $\delta: \mathcal{K}(M) \rightarrow \mathcal{K}(M)$ tal que

$$\delta(fX) = VfX + f\delta(X) \quad \text{para toda } f \in \mathcal{F}(M) \text{ y todo } X \in \mathcal{K}(M),$$

existe una única derivación tensorial D sobre M tal que $D_0^0 = V: \mathcal{F}(M) \rightarrow \mathcal{F}(M)$ y $D_0^1 = \delta$.

Demostración. D_0^0 y D_0^1 están dadas. La fórmula que escribimos antes del corolario 21.3 muestra que D debe estar definida sobre uno-formas θ por

$$(D\theta)(X) = V(\theta X) - \theta(\delta X) \quad \text{para todo } X \in \mathcal{K}(M).$$

Usando la fórmula dada para δ es fácil verificar que $D\theta$ es $\mathcal{F}(M)$ -lineal y, por lo tanto, una uno-forma. También se ve que

$$D = D_0^0: \mathcal{K}^*(M) \rightarrow \mathcal{K}^*(M) \text{ es } \mathbb{R}\text{-lineal.}$$

Por la regla del producto del lema 21.2, D aplicada a un campo tensorial A de tipo (r, s) con $r + s \geq 2$ debe estar definida por

$$\begin{aligned}
 (DA)(\theta^1, \dots, \theta^r, X_1, \dots, X_s) &= V(A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)) \\
 &\quad - \sum_{i=1}^r A(\theta^1, \dots, D\theta^i, \dots, \theta^r, X_1, \dots, X_s) \\
 &\quad - \sum_{j=1}^s A(\theta^1, \dots, \theta^r, X_1, \dots, \delta X_j, \dots, X_s).
 \end{aligned}$$

(En el lado derecho D aplicada a una uno-forma se define como antes.)
 Nuevamente es fácil verificar que DA es $\mathcal{F}(M)$ -multilineal y es, por tanto, un campo tensorial de tipo (r, s) ; también, se ve que $D: \mathcal{F}_1^1(M) \rightarrow \mathcal{F}_1^1(M)$ es \mathbb{R} -lineal. Además, un cómputo directo muestra que $D(A \otimes B) = DA \otimes B + A \otimes DB$. (Para hacer esta verificación se pueden tomar A y B de tipo $(1, 1)$.)

Para demostrar que D conmuta con las contracciones, consideremos primero el caso de $C: \mathcal{F}_1^1(M) \rightarrow \mathcal{F}(M)$. Que $DC = CD$ para productos tensoriales de la forma $\theta \otimes X$ se sigue inmediatamente de la definición de D para uno-formas. Por lo tanto tenemos que $DC = CD$ para sumas de términos de la forma $\theta \otimes X$. Puesto que D es de carácter local y C es de carácter puntual, es suficiente probar que $DC = CD$ sobre vecindades coordinadas. Pero el lema 18.1 (pag. 47), demuestra que sobre una vecindad coordinada todo campo tensorial se escribe como una suma de esa forma.

La extensión de este resultado a contracciones arbitrarias es un fácil ejercicio en el uso de paréntesis. Tomando $A \in \mathcal{F}_2^1(M)$, por ejemplo, tenemos

$$\begin{aligned}
 (DC_1^1 A)(X) &= D((C_1^1 A)(X)) - (C_1^1 A)(DX) \\
 &= D(C\{A(\cdot, X, \cdot)\}) - C\{A(\cdot, DX, \cdot)\} \\
 &= C\{D(A(\cdot, X, \cdot)) - A(\cdot, DX, \cdot)\} \\
 &= C\{(DA)(\cdot, X, \cdot)\} = (C_1^1 DA)(X).
 \end{aligned}$$

Por lo tanto, $DC_1^1 A = C_1^1 DA$.

##

Como una aplicación de este teorema, podemos hacer la siguiente definición.

Si $V \in \mathcal{X}(M)$ la derivación tensorial L_V tal que

$$\begin{aligned}
 L_V(f) &= Vf && \text{para toda } f \in \mathcal{F}(M), \\
 L_V(X) &= [V, X] && \text{para todo } X \in \mathcal{X}(M)
 \end{aligned}$$

se llama la derivada de Lie con respecto a V .

Esta definición es válida puesto que L_V aplicada a campos vectoriales satisface la correspondiente hipótesis que hicimos para δ en el teorema:

$$L_V(fX) = [V, fX] = VfX + f[V, X] = VfX + fL_V X.$$

22 FORMAS BILINEALES SIMÉTRICAS

En la geometría semi-Riemanniana que estudiaremos más adelante, juega un papel fundamental un cierto campo tensorial de tipo $(0,2)$ y que es simétrico. Este campo asocia con cada espacio vectorial tangente una función bilineal simétrica. Debido a su importancia, estudiaremos a estas funciones definidas sobre un espacio vectorial real de dimensión finita V . Una forma bilineal simétrica sobre V es una función \mathbb{R} -bilineal $b: V \times V \rightarrow \mathbb{R}$, tal que $b(v,w) = b(w,v)$, para todo $v, w \in V$.

Una forma bilineal simétrica sobre V se llama

- (1) definida positiva [negativa] si $v \neq 0$ implica que $b(v,v) > 0$ [< 0],
- (2) semidefinida positiva [negativa] si $b(v,v) \geq 0$ [≤ 0] para todo $v \in V$,
- (3) no degenerada si $b(v,w) = 0$ para todo $w \in V$ implica que $v = 0$.

También b se llama definida [semidefinida] si se cumple la condición (1) [(2)]. Si b es definida es claro que también es semidefinida y no degenerada; la afirmación inversa también es cierta.

Si b es una forma bilineal simétrica sobre V , entonces para cualquier subespacio W de V la restricción $b|_{(W \times W)}$, denotada simplemente por $b|_W$, es nuevamente bilineal y simétrica. Si b es [semi-]definida, también lo será $b|_W$.

El índice ν de una forma bilineal simétrica b sobre V es el más grande entero que es la dimensión de un subespacio $W \subset V$ sobre el cual $b|_W$ es definida negativa.

Según esto, $0 \leq \nu \leq \dim V$, y $\nu = 0$ si y sólo si b es semidefinida positiva.

La función $q: V \rightarrow \mathbb{R}$ dada por $q(v) = b(v,v)$ es la forma cuadrática asociada con b . Con frecuencia, es más fácil tratar con q que con b , y al hacerlo así no hay una pérdida de información, ya que b se puede reconstruir a partir de q por medio de la identidad de polarización

$$b(v,w) = \frac{1}{2} \{q(v+w) - q(v) - q(w)\}.$$

Si e_1, \dots, e_n es una base para V , entonces la matriz de $n \times n$ $(b_{ij}) = b(e_i, e_j)$ se llama la matriz de b con respecto a la base e_1, \dots, e_n . Puesto que b es simétrica, esta matriz es simétrica. Claramente, el conocimiento de esta matriz es suficiente para determinar a b , ya que

$$b(\sum v_i e_i, \sum w_j e_j) = \sum b_{ij} v_i w_j.$$

Lema 22.1 Una forma bilineal simétrica es no degenerada si y sólo si su matriz con respecto a una (y , por lo tanto, cualquier) base es invertible (no singular).

Demostración. Sea e_1, \dots, e_n una base para V . Si $v \in V$, entonces $b(v,w) =$

= 0 si y sólo si $b(v, e_i) = 0$ para $i = 1, \dots, n$. Puesto que (b_{ij}) es simétrica,

$$b(v, e_i) = b(\sum v_j e_j, e_i) = \sum b_{ij} v_j.$$

Así pues, b es degenerada si y sólo si existen números v_1, \dots, v_n no todos nulos tales que $\sum b_{ij} v_j = 0$ para $i = 1, \dots, n$. Pero esto es equivalente a la dependencia lineal de las columnas de la matriz (b_{ij}) , es decir, a que (b_{ij}) sea singular. ##

23 PRODUCTOS ESCALARES

Un producto escalar g sobre un espacio vectorial V es una forma bilineal simétrica no degenerada sobre V .

Un producto interno es un producto escalar definido positivo. El ejemplo canónico de producto interno es el producto punto sobre \mathbb{R}^n , que está dado por $v \cdot w = \sum v_i w_i$. Muchas de las propiedades de los productos interiores también son poseídas por los productos escalares, sin embargo, aparecen algunos fenómenos nuevos cuando g es indefinido.

Mediante un cambio de signo en la definición del producto punto sobre \mathbb{R}^2 , se obtiene el ejemplo más simple de un producto escalar indefinido.

Ejemplo. Definamos $g: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ por

$$g(v, w) = v_1 w_1 - v_2 w_2.$$

Es obvio que g es simétrico y bilineal. Tomando $w = (1, 0)$ y después $w = (0, 1)$, se comprueba que g es no degenerado. Así pues, g es un producto escalar sobre \mathbb{R}^2 . La forma cuadrática asociada con g es

$$q(v) = v_1^2 - v_2^2,$$

y vemos que g es indefinido.

En el resto de esta sección, V denotará un espacio con producto escalar, es decir, un espacio vectorial real de dimensión finita provisto con un producto escalar g sobre V . Un vector $v \in V$ se llama nulo si $g(v, v) = q(v) = 0$ pero $v \neq 0$. Evidentemente los vectores nulos existen si y sólo si g es indefinido. En el ejemplo anterior, los vectores nulos son todos aquellos que están situados sobre las dos rectas que forman ángulos de 45° con los ejes coordenados, con la única excepción del origen (0 no es un vector nulo). Para $c \neq 0$ los conjuntos $q = c$ y $q = -c$ son hipérbolas cuyas asíntotas son las líneas nulas (ver fig. 6).

Los vectores $v, w \in V$ se llaman ortogonales, lo cual se denota por $v \perp w$, si $g(v, w) = 0$. Los subconjuntos A y B de V se llaman ortogonales, y escribimos $A \perp B$, si $v \perp w$ para todo $v \in A$ y todo $w \in B$.

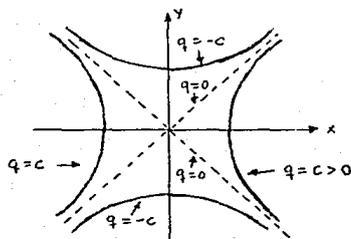


FIG. 6

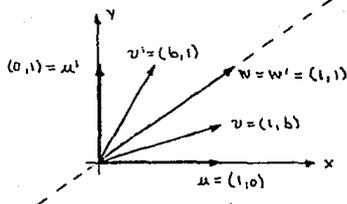


FIG. 7

Cuando el producto escalar g es indefinido, ya no podemos imaginar a los vectores ortogonales como formando ángulos rectos entre sí. Volviendo al ejemplo del producto escalar más simple \mathbb{R}^2 , en la figura 7 hemos dibujado tres pares de vectores $z \perp z'$; el caso $z = w$ y $z' = w'$ ilustra el hecho de que un vector nulo es un vector distinto de cero que es ortogonal a sí mismo.

Si W es un subespacio de V , sea

$$W^\perp = \{v \in V \mid v \perp W\}.$$

W^\perp es un subespacio de V , llamado W perp. No es posible llamar a W^\perp el complemento ortogonal de W en V , ya que en general $W + W^\perp$ no es siempre igual a todo el espacio V . Por ejemplo, si en la figura 7 W es el subespacio de \mathbb{R}^2 generado por el vector $(1,1)$, entonces $W^\perp = W$. No obstante, la operación perp sí tiene algunas propiedades más familiares:

Lema 2.3.1 Si W es un subespacio de un espacio con producto escalar V , entonces

- (1) $\dim W + \dim W^\perp = n = \dim V$,
- (2) $(W^\perp)^\perp = W$.

Demostración. (1) Sea e_1, \dots, e_n una base para V adaptada a W , es decir, tal que e_1, \dots, e_k es una base para W . Ahora bien, $v \in W^\perp$ si y sólo si $g(v, e_i) = 0$ para $i = 1, \dots, k$, lo cual se escribe en términos de coordenadas como

$$\sum_{j=1}^n \varepsilon_{ij} v_j = 0 \quad (i = 1, \dots, k).$$

Este es un sistema de k ecuaciones lineales algebraicas con n incógnitas y, de acuerdo con el lema 22.1, los renglones de la matriz de coeficientes son linealmente independientes, es decir, ésta tiene rango igual a k . Según el álgebra lineal, esto quiere decir que la dimensión del espacio de soluciones de este sistema es $n - k$. Pero por construcción las soluciones (v_1, \dots, v_n) del sistema dan exactamente los vectores $v = \sum v_j e_j$ de W^\perp .

(2) Puesto que $v \in (W^\perp)^\perp$ equivale a $v \perp W^\perp$, y todos los vectores de W tienen esta última propiedad, tenemos que $W \subset (W^\perp)^\perp$. Por (1), estos dos subespacios tienen la misma dimensión, por lo tanto, ellos deben ser iguales. ##

Nótese que la no degeneración de g sobre todo el espacio V es equivalente a que $V^\perp = 0$.

Un subespacio W de V se llama no degenerado si $g|_W$ es no degenerada. Cuando V es un espacio con producto interior, todo subespacio W es nuevamente un espacio con el producto interior $g|_W$ y es, por lo tanto, un subespacio no degenerado. Sin embargo, cuando g es indefinido existirán siempre subespacios degenerados; por ejemplo, cualquier vector nulo genera un subespacio degenerado. Así pues, un subespacio de un espacio con producto escalar no siempre es, a su vez, un espacio con producto escalar. Esta dificultad está relacionada con las peculiaridades de la operación perp que consideramos más arriba.

Lema 23.2 Un subespacio W de V es no degenerado si y sólo si V es la suma directa de W y W^\perp .

Demostración. Recordemos (ver [H], pag. 183) la siguiente identidad de la teoría de espacios vectoriales de dimensión finita:

$$\dim(W + W^\perp) + \dim(W \cap W^\perp) = \dim W + \dim W^\perp.$$

Según el lema 23.1, el lado derecho de esta ecuación es $n = \dim V$. Por lo tanto, $W + W^\perp = V$ si y sólo si $W \cap W^\perp = 0$. Así, cualquiera de estas dos condiciones es equivalente a $V = W \oplus W^\perp$. Pero $W \cap W^\perp = \{w \in W \mid w \perp W\}$, por lo que el hecho de que este subespacio sea cero equivale a la no degeneración del subespacio W . ##

Puesto que $(W^\perp)^\perp = W$, se sigue que W es no degenerado si y sólo si W^\perp es no degenerado.

Como $q(v) = g(v, v)$ puede ser negativo, la norma de un vector v , denotada por $|v|$, se define por $|g(v, v)|^{1/2}$. Un vector unitario u es un vector de norma = 1, es decir, $g(u, u) = \pm 1$. Como es usual, un conjunto de vectores mutuamente ortogonales y unitarios se llama ortonormal, y para $n = \dim V$, cualquier conjunto de n vectores ortonormales es necesariamente una base para V .

Lema 23.3 Todo espacio con producto escalar $V \neq 0$ tiene una base ortonormal.

Demostración. Puesto que g es no degenerada, existe un vector $v \in V$ tal que $g(v, v) \neq 0$. De donde, $v/|v|$ es un vector unitario. Así, por inducción es suficiente demostrar que cualquier conjunto ortonormal e_1, \dots, e_k con $k < n$ puede ser agrandado con un elemento más. Por el lema 22.1 sabemos que estos vectores generan un subespacio no degenerado W de dimensión k . Sólo resta encontrar un vector unitario en $W^\perp \neq 0$. Pero como hicimos notar antes, W^\perp es también no degenerado, de modo que el argumento anterior demuestra que W^\perp contiene un vector unitario. ##

La matriz de g con respecto a una base ortonormal e_1, \dots, e_n para V es diagonal; de hecho,

$$g(e_i, e_j) = \delta_{ij} \epsilon_j, \quad \text{donde } \epsilon_j = g(e_j, e_j) = \pm 1.$$

Siempre que sea conveniente, ordenaremos los vectores de una base ortonormal de manera tal que los signos negativos, si los hay, aparezcan en los primeros lugares de la llamada signatura $(\epsilon_1, \dots, \epsilon_n)$.

Tomando en cuenta estos signos, todavía se puede realizar una expansión ortonormal.

Lema 23.4 Sea e_1, \dots, e_n una base ortonormal para V , con $\epsilon_i = g(e_i, e_i)$. Entonces cada $v \in V$ tiene una única expresión de la forma

$$v = \sum \epsilon_i g(v, e_i) e_i.$$

Demostración. Es fácil verificar que el vector

$$v - \sum \epsilon_i g(v, e_i) e_i$$

es ortogonal a todos los e_j ; así, por la no degeneración de g , este vector debe ser cero. ##

La proyección ortogonal π de V sobre un subespacio W es la transformación lineal que manda W^\perp a 0 y deja fijo a cada vector de W . Una base e_1, \dots, e_k para W puede siempre ser agrandada hasta obtener una base para V ; así

$$\pi(v) = \sum_{j=1}^k \epsilon_j g(v, e_j) e_j.$$

Es costumbre referirse al índice v del producto escalar g de V como el índice de V , y se escribe $v = \text{ind } V$.

Lema 23.5 Para cualquier base ortonormal e_1, \dots, e_n de V el número de signos negativos en la signatura $(\epsilon_1, \dots, \epsilon_n)$ es el índice v de V .

Demostración. Supongamos que los primeros m signos ϵ_i son negativos. El resultado es trivial si g es definido, así, supongamos que $0 < m < n$. Es evidente que g es definido negativo sobre el subespacio S generado por e_1, \dots, e_m ; así que $v \geq m$.

Para demostrar la desigualdad inversa, sea W un subespacio arbi-

trario sobre el cual g es definido negativo, y definamos $\pi: W \rightarrow S$ por

$$\pi(w) = - \sum_{i \leq m} g(w, e_i) e_i.$$

Evidentemente π es lineal. Así que es suficiente demostrar que π es inyectiva, ya que entonces se sigue $\dim W \leq \dim S = m$, de donde $v \leq m$. Si $\pi(w) = 0$, entonces mediante una expansión ortonormal, escribimos

$$w = \sum_{j > m} g(w, e_j) e_j.$$

Pero dado $w \in W$,

$$0 \geq g(w, w) = \sum_{j > m} g(w, e_j)^2.$$

Por lo tanto, $g(w, e_j) = 0$ para $j > m$, de donde $w = 0$. ##

Se sigue que para un subespacio no degenerado W de V

$$\text{ind } V = \text{ind } W + \text{ind } W^\perp,$$

ya que por el lema 23.4, podemos encontrar una base ortonormal para V adaptada a la suma directa $V = W + W^\perp$.

Sean V y \bar{V} espacios con los productos escalares g y \bar{g} , respectivamente. Una transformación lineal $T: V \rightarrow \bar{V}$ preserva los productos escalares si

$$\bar{g}(Tv, Tw) = g(v, w) \quad \text{para todo } v, w \in V.$$

En este caso, T necesariamente es inyectiva, ya que si $Tv = 0$ entonces $g(v, w) = 0$ para todo w , de donde $v = 0$.

Nótese que T preserva productos escalares si y sólo si T preserva las correspondientes formas cuadráticas, es decir

$$\bar{q}(Tv) = q(v) \quad \text{para todo } v \in V.$$

Una de estas implicaciones es obvia; la otra se sigue mediante la identidad de polarización.

Un isomorfismo lineal $T: V \rightarrow W$ que preserva los productos escalares se llama una isometría lineal. Por las observaciones anteriores, tenemos que una transformación lineal $T: V \rightarrow W$ es una isometría lineal si y sólo si $\dim V = \dim W$ y T preserva los productos escalares (o bien, las formas cuadráticas correspondientes).

Lema 23.6 Los espacios con producto escalar V y W tienen la misma dimensión e índice si y sólo si existe una isometría lineal de V a W .

Demostración. Suponiendo que los invariantes \dim e ind son los mismos, elijamos bases ortonormales e_1, \dots, e_n para V y e'_1, \dots, e'_n para W . Por el lema 23.5, podemos suponer que $g_V(e_i, e_j) = g_W(e'_i, e'_j)$ para toda i . Sea T una transformación lineal tal que $Te_i = e'_i$ para toda i . Entonces $g_W(Te_i, Te_j) = g_W(e'_i, e'_j)$ para toda i, j . Por la linealidad de T y por la bilinealidad de los productos escalares g_V y g_W , se sigue de

las igualdades anteriores que T es una isometría lineal.

Inversamente, si $T:V \rightarrow W$ es una isometría lineal, entonces T transforma una base ortonormal para V en una base ortonormal para W . Por lo tanto, $\dim V = \dim W$ y, por el lema 23.5, $\text{ind } V = \text{ind } W$. ##

+++++

III. FORMAS DIFERENCIALES

De todos los tipos de campos tensoriales, los que son covariantes y antisimétricos (ver sección 20), es decir, las formas diferenciales, son los que se encuentran con más frecuencia y los que tienen las más variadas aplicaciones. La teoría electromagnética (ecuaciones de Maxwell) puede formularse de manera muy elegante y concisa utilizando las formas diferenciales, además, esta formulación no cambia al pasar al espacio-tiempo de la relatividad. Las formas diferenciales han sido usadas por de Rham para expresar una muy profunda relación entre la estructura topológica de una variedad y ciertos aspectos del análisis vectorial sobre la variedad. En el trabajo del famoso geómetra francés E. Cartan, él usa casi exclusivamente las formas diferenciales para formular y desarrollar sus resultados sobre los sistemas diferenciales y la geometría riemanniana. Las generalizaciones de los teoremas de Stokes y de la divergencia a dimensiones superiores y a espacios más generales (variedades), es muy complicada a menos que se use sistemáticamente el cálculo de formas diferenciales. Son este cálculo y su aplicación a la teoría de integración en variedades, además del estudio del método dual de las distribuciones vectoriales, los temas principales de esta tercera parte de las notas.

24 FORMAS DIFERENCIALES

Una p-forma (diferencial) sobre una variedad M es un campo tensorial antisimétrico de tipo $(0, p)$ sobre la variedad M (ver sección 20). En otras palabras, una p -forma ω es un campo tensorial de $\mathcal{F}_p^0(M)$ con la propiedad de que

$$\omega(X_1, \dots, X_i, \dots, X_j, \dots, X_p) = -\omega(X_1, \dots, X_j, \dots, X_i, \dots, X_p),$$

para todos los pares de índices i y j y para todos los $X_k \in \mathcal{X}(M)$.

Así pues, una cero-forma es una función $f \in \mathcal{F}(M)$ y una uno-forma es un elemento $\theta \in \mathcal{F}_1^0(M) = \mathcal{X}^*(M)$ (ver sección 8).

Una s -forma se puede construir a partir de un tensor general de tipo $(0, s)$, aplicando a éste último un operador especial A , llamado operador alternante. Si denotamos por $\Lambda^s(M)$ al conjunto de todas las s -formas sobre M , el operador alternante $A: \mathcal{F}_s^0(M) \rightarrow \Lambda^s(M)$ se define por

$$AT(X_1, \dots, X_s) = \frac{1}{s!} \sum_{(j_1, \dots, j_s)} \text{sgn}(j_1, \dots, j_s) T(X_{j_1}, \dots, X_{j_s}), \text{ para todo } T \in \mathcal{F}_s^0(M),$$

donde la suma se extiende sobre todas las $s!$ permutaciones de los s enteros $(1, \dots, s)$ y $\text{sgn}(j_1, \dots, j_s) = \pm 1$, dependiendo de que (j_1, \dots, j_s) sea una permutación par o impar de $(1, \dots, s)$; la ecuación anterior se supone válida para todos los $(X_1, \dots, X_s) \in \mathcal{X}(M)^s$.

Está claro que el operador alternante $A: \mathcal{F}_s^0(M) \rightarrow \Lambda^s(M)$ es lineal y que si $T \in \mathcal{F}_s^0(M)$ ya es un tensor antisimétrico, entonces el efecto de A sobre T es simplemente el de reproducir a T : $AT = T$.

Si $\xi = (x^1, \dots, x^n)$ es un sistema coordinado para una vecindad U de una variedad M de dimensión n , sabemos que dx^1, \dots, dx^n es una base lo-

cal para las uno-formas sobre M ; es decir, si θ es una uno-forma sobre M , entonces podemos escribir

$$\theta = \sum \theta_i dx^i \quad \text{sobre } U.$$

Si $T \in \mathcal{L}_s^0(M)$, tenemos en estas condiciones que

$$T = \sum T_{j_1, \dots, j_s} dx^{j_1} \otimes \dots \otimes dx^{j_s} \quad \text{sobre } U,$$

en donde $T_{j_1, \dots, j_s} = T(\delta_{j_1}, \dots, \delta_{j_s})$ (con los s índices corriendo de 1 a n), son las n componentes del tensor T con respecto al sistema coordinado ξ .

Para obtener una base local para las s -formas sobre M , debemos antisimetrizar a los tensores básicos $dx^{j_1} \otimes \dots \otimes dx^{j_s}$ por medio del operador alternante, de este modo obtenemos:

$$dx^{j_1} \wedge \dots \wedge dx^{j_s} = A(dx^{j_1} \otimes \dots \otimes dx^{j_s}), \quad j = 1, \dots, n; k = 1, \dots, s.$$

En el lado izquierdo de esta ecuación tenemos la s -forma básica $dx^{j_1} \wedge \dots \wedge dx^{j_s}$ con respecto al sistema ξ , y se dice que se ha obtenido tomando el producto exterior de las uno-formas $dx^{j_1}, \dots, dx^{j_s}$.

Podría pensarse que como en la ecuación anterior tenemos s índices j_1, \dots, j_s , cada uno de los cuales corre de 1 a n , se obtienen entonces n^s s -formas básicas $dx^{j_1} \wedge \dots \wedge dx^{j_s}$. Sin embargo, esto no es cierto, ya que por la definición de $dx^{j_1} \wedge \dots \wedge dx^{j_s}$, si cambiamos de lugar cualquier par de diferenciales, hay un cambio de signo:

$$\begin{aligned} dx^{j_1} \wedge \dots \wedge dx^{j_k} \wedge \dots \wedge dx^{j_l} \wedge \dots \wedge dx^{j_s} = \\ = -dx^{j_1} \wedge \dots \wedge dx^{j_l} \wedge \dots \wedge dx^{j_k} \wedge \dots \wedge dx^{j_s} \end{aligned}$$

De esta observación se sigue que sólo cuando todos los índices j_1, \dots, j_s son diferentes, la s -forma $dx^{j_1} \wedge \dots \wedge dx^{j_s}$ será diferente de cero. Por lo tanto, hay tantas s -formas básicas sobre una variedad de dimensión n , como haya sucesiones finitas $j_1 < j_2 < \dots < j_s$, donde cada j_k corre de 1 a n . En otras palabras, si $s \leq n$ hay un total de

$$\binom{n}{s} = \frac{n!}{s!(n-s)!}$$

s -formas básicas $dx^{j_1} \wedge \dots \wedge dx^{j_s}$ sobre una variedad de dimensión n , y si $s > n$ la única s -forma básica es la s -forma cero.

Así, por ejemplo, el conjunto $\{dx^i \wedge dx^j \mid i < j\}$ que contiene unas $n_1/2_1(n-2)_1$ dos-formas, es una base local para las dos-formas sobre M . También, la n -forma básica $dx^1 \wedge \dots \wedge dx^n$ constituye por sí sola una base local para las n -formas sobre M .

Podemos resumir los comentarios anteriores en el siguiente teorema.

Teorema 24.1 Sea $\xi = (x^1, \dots, x^n)$ un sistema coordinado sobre el subconjunto U de la variedad n -dimensional M . Toda s -forma $\omega \in \mathcal{L}_s^0(M)$ tiene una expresión local en coordenadas

$$\omega = \sum_{j_1 < \dots < j_s} \omega_{j_1, \dots, j_s} dx^{j_1} \wedge \dots \wedge dx^{j_s} \quad \text{sobre } U,$$

en donde la suma se toma sobre las $n_1/s_1(n-s)_1$ sucesiones finitas

j_1, \dots, j_s , con cada $j_k = 1, \dots, n$; y $\omega_{j_1, \dots, j_s} = \omega(\partial_{j_1}, \dots, \partial_{j_s})$ son las componentes de ω con respecto al sistema coordinado ξ .

Puesto que en este capítulo trataremos exclusivamente con productos exteriores de formas diferenciales, podemos simplificar ligeramente nuestra notación. Así, escribiremos $dx^i dx^j$ en lugar de $dx^i \wedge dx^j$.

Por ejemplo, si $\omega \in \Lambda^2(M)$, escribiremos

$$\omega = \sum \omega_{ij} dx^i dx^j = \omega_{12} dx^1 dx^2 + \omega_{13} dx^1 dx^3 + \omega_{23} dx^2 dx^3.$$

Definiremos las componentes de una p-forma como sus componentes con respecto a la base $\{dx^{i_1} \dots dx^{i_p}\}$ de p-formas y no con respecto a la base $\{dx^{i_1} \otimes \dots \otimes dx^{i_p}\}$ de tensores covariantes, como lo habíamos hecho en el capítulo II. Así, ahora decimos que las componentes de la dos-forma que consideramos arriba son ω_{12} , ω_{13} , y ω_{23} , mientras que en el capítulo II, y puesto que

$$dx^i dx^j = \frac{1}{2}(dx^i \otimes dx^j - dx^j \otimes dx^i),$$

habríamos dicho que las componentes de ω son $b_{11} = 0$, $b_{12} = \frac{1}{2}\omega_{12}$, $b_{13} = \frac{1}{2}\omega_{13}$, $b_{21} = -b_{12}$, etc. También en el ejemplo anterior es fácil comprobar que

$$dx^1 dx^2 (\partial_1, \partial_2) = \frac{1}{2}, \quad dx^1 dx^2 (\partial_1, \partial_3) = 0, \quad dx^1 dx^2 (\partial_2, \partial_3) = 0,$$

$$dx^1 dx^3 (\partial_1, \partial_2) = 0, \quad dx^1 dx^3 (\partial_1, \partial_3) = \frac{1}{2}, \quad dx^1 dx^3 (\partial_2, \partial_3) = 0,$$

$$dx^2 dx^3 (\partial_1, \partial_2) = 0, \quad dx^2 dx^3 (\partial_1, \partial_3) = 0, \quad dx^2 dx^3 (\partial_2, \partial_3) = \frac{1}{2}.$$

De aquí debe quedar claro que la regla general para evaluar las p-formas básicas sobre los campos vectoriales básicos es

$$dx^{i_1} \dots dx^{i_p} (\partial_{j_1}, \dots, \partial_{j_p}) = \frac{1}{p!} \delta_{j_1}^{i_1} \dots \delta_{j_p}^{i_p},$$

donde i_1, \dots, i_p y j_1, \dots, j_p son conjuntos crecientes de índices.

También es claro que si θ_{i_1, \dots, i_p} son las componentes de una p-forma θ , entonces

$$\theta(\partial_{j_1}, \dots, \partial_{j_p}) = \frac{1}{p!} \theta_{j_1, \dots, j_p}.$$

Dadas cualquier p-forma ω y cualquier q-forma η , definimos su producto exterior por la regla

$$\omega \wedge \eta = A(\omega \otimes \eta).$$

Así pues, el producto exterior de una p-forma y de una q-forma es una (p+q)-forma, la cual será nula si $p+q > n$.

El producto exterior es claramente asociativo y distributivo, pero no es conmutativo en general. De hecho, para un sistema coordinado x^1, \dots, x^n y, por la definición, tenemos

$$\omega \wedge \eta = (\sum \omega_{j_1, \dots, j_p} dx^{j_1} \dots dx^{j_p}) \wedge (\sum \eta_{k_1, \dots, k_q} dx^{k_1} \dots dx^{k_q}),$$

donde (j_1, \dots, j_p) y (k_1, \dots, k_q) son sucesiones estrictamente crecien-

tes. De aquí tenemos que

$$\begin{aligned}\omega \wedge \eta &= (-1)^{pq} (\sum_{k_1, \dots, k_q} \eta_{k_1, \dots, k_q} dx^{k_1} \dots dx^{k_q}) \wedge (\sum_{j_1, \dots, j_p} \omega_{j_1, \dots, j_p} dx^{j_1} \dots dx^{j_p}) \\ &= (-1) \eta \wedge \omega,\end{aligned}$$

ya que cada uno de los elementos básicos $dx^{k_1}, \dots, dx^{k_q}$ debe sufrir p intercambios de posición antes de que $\omega \wedge \eta$ pueda ser llevado a la forma de $\eta \wedge \omega$. La propiedad

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$$

del producto exterior se llama anticonmutatividad.

25 DERIVADA EXTERIOR

La derivada exterior de una p -forma θ es una $(p+1)$ -forma que denotaremos por $d\theta$. Ya en la sección 5 habíamos definido $d\theta$ para el caso $p = 0$. Hay varias maneras de dar la definición de la derivada exterior, cada una de las cuales nos da información importante acerca de las propiedades del operador $d: \Lambda^p(M) \rightarrow \Lambda^{p+1}(M)$.

A. Definición Coordenada

En términos de un sistema coordenado, el operador d simplemente opera sobre las funciones componentes. Definimos entonces,

$$d\theta = \sum_{j_1 < \dots < j_p} (d\theta_{j_1, \dots, j_p}) \wedge dx^{j_1} \dots dx^{j_p}. \quad (25.1)$$

No es inmediatamente claro que esta sea una buena definición, ya que el lado derecho de esta ecuación podría depender de la elección del sistema de coordenadas x^1, \dots, x^n . Sin embargo, se comprueba fácilmente que esta fórmula satisface los axiomas para d que daremos más adelante. Puesto que dichos axiomas serán independientes del sistema coordenado y determinan completamente a d , se seguirá que la ecuación (25.1) es independiente de la elección de coordenadas.

En el caso particular de que M sea \mathbb{R}^3 con las coordenadas cartesianas x, y, z la fórmula adquiere una forma muy parecida a los operadores diferenciales clásicos del cálculo elemental grad , rot y div :

$$df = f_x dx + f_y dy + f_z dz,$$

$$\begin{aligned}d(fdx + gdy + hdz) &= df \wedge dx + dg \wedge dy + dh \wedge dz \\ &= (f_x dx + f_y dy + f_z dz) \wedge dx \\ &\quad + (g_x dx + g_y dy + g_z dz) \wedge dy \\ &\quad + (h_x dx + h_y dy + h_z dz) \wedge dz \\ &= (h_y - g_z) dy dz + (f_z - h_x) dz dx + (g_x - f_y) dx dy, \\ d(fdy dz + gdz dx + hdx dy) &= df \wedge dy dz + dg \wedge dz dx + dh \wedge dx dy \\ &= (f_x + g_y + h_z) dx dy dz.\end{aligned}$$

(En estas fórmulas hemos indicado las derivadas parciales mediante

el subíndice adecuado). Las discrepancias entre estas fórmulas y las fórmulas usuales de los operadores grad (gradiente), rot (rotacional) y div (divergencia) pueden hacerse desaparecer si consideramos el producto interno usual de \mathbb{R}^3 , para el cual dx, dy, dz , es una base ortonormal en cada punto de \mathbb{R}^3 . Esto nos da un isomorfismo entre vectores contravariantes y vectores covariantes $a\partial_x + b\partial_y + c\partial_z \longleftrightarrow adx + bdy + cdz$; nosotros ignoraremos aquí este isomorfismo y trataremos sólo con los vectores covariantes.

Ahora bien, en la variedad tridimensional \mathbb{R}^3 , tenemos los espacios $\Lambda^0(\mathbb{R}^3) = \mathcal{F}(\mathbb{R}^3)$, $\Lambda^1(\mathbb{R}^3)$, $\Lambda^2(\mathbb{R}^3)$, $\Lambda^3(\mathbb{R}^3)$ de cero-formas, uno-formas, dos-formas y tres-formas sobre \mathbb{R}^3 , respectivamente (los espacios $\Lambda^s(\mathbb{R}^3)$ con $s > 3$, son todos ellos iguales a cero). Las formas básicas para estos espacios, con respecto a las coordenadas cartesianas x, y, z , son:

1 para $\Lambda^0(\mathbb{R}^3)$ que es un espacio de dimensión 1;
 dx, dy, dz para $\Lambda^1(\mathbb{R}^3)$ que es un espacio de dimensión 3;
 $dydz, dzdx, dx dy$ para $\Lambda^2(\mathbb{R}^3)$ que es un espacio de dimensión 3;
 $dx dy dz$ para $\Lambda^3(\mathbb{R}^3)$ que es un espacio de dimensión 1.

Debido a las dimensiones de estos espacios, podemos esperar encontrar unos isomorfismos lineales entre los espacios $\Lambda^1(\mathbb{R}^3)$ y $\Lambda^2(\mathbb{R}^3)$, así como entre los espacios $\Lambda^0(\mathbb{R}^3)$ y $\Lambda^3(\mathbb{R}^3)$. Definimos, entonces, un operador $\mathcal{F}(\mathbb{R}^3)$ -lineal, llamado operador estrella de Hodge

$$*: \Lambda^s(\mathbb{R}^3) \longrightarrow \Lambda^{3-s}(\mathbb{R}^3) \quad s = 0, 1, 2, 3.$$

especificando su efecto sobre las formas básicas como sigue:

$$*(dx) = dydz, *(dy) = dzdx, *(dz) = dx dy, \text{ para el caso } s = 1;$$

$*(dx dy dz) = 1$, para el caso $s = 3$; y $** = \text{identidad}$, para los demás casos.

Usando el operador estrella, ahora tenemos

$$*d(fdx + gdy + hdz) = (h_y - g_z)dx + (f_z - h_x)dy + (g_x - f_y)dz,$$

$$*d*(fdx + gdy + hdz) = f_x + g_y + h_z.$$

Por lo tanto, ahora vemos que la versión covariante de los operadores rot y div está dada por los operadores $\text{rot} = *d$ y $\text{div} = *d*$, ambos operando sobre uno-formas. La versión covariante del grad es simplemente $\text{grad} = d$, operando sobre cero-formas.

B. Definición Axiomática

Existen unas cuantas propiedades del operador $d: \Lambda^s(M) \longrightarrow \Lambda^{s+1}(M)$ que son suficientes para determinar completamente a d y que, por lo tanto, tomaremos como axiomas para d :

(1) Si f es una cero-forma, entonces df coincide con la definición que habíamos dado en la parte I de las notas; es decir, $df(X) = Xf$, para todo campo vectorial $X \in \mathcal{X}(M)$.

(2) Se tiene una regla, análoga a la regla de Leibniz, para calcular el efecto del operador d sobre un producto exterior de formas:

si $\theta \in \wedge^p(M)$ y $\tau \in \wedge^q(M)$, entonces

$$d(\theta \wedge \tau) = d\theta \wedge \tau + (-1)^p \theta \wedge d\tau.$$

Debido a esta propiedad y a la linealidad que daremos en el axioma (4), se dice que d es una antiderivación.

(3) Cuando d se aplica dos veces el resultado siempre es cero: $d(d\theta) = 0$ para toda $\theta \in \wedge^p(M)$; escribimos entonces $d^2 = 0$.

(4) El operador d es \mathbb{R} -lineal: si θ y τ son p -formas y si a y b son números reales, entonces

$$d(a\theta + b\tau) = ad\theta + bd\tau.$$

Teorema 25.1 La definición en términos de coordenadas (ecuación (25.1)) y la definición axiomática del operador derivada exterior son equivalentes.

Demostración. La definición coordenada (25.1) es una consecuencia fácil de los axiomas. De hecho, de los axiomas (2) y (3) tenemos

$$d(dx^{i_1} \dots dx^{i_p}) = (d^2 x^{i_1}) \wedge dx^{i_2} \dots dx^{i_p} - dx^{i_1} \wedge (d^2 x^{i_2}) \wedge dx^{i_3} \dots dx^{i_p} + \dots + (-1)^{p-1} dx^{i_1} \dots dx^{i_{p-1}} \wedge d^2 x^{i_p} = 0.$$

Así tenemos

$$\begin{aligned} d(f dx^{i_1} \dots dx^{i_p}) &= df \wedge dx^{i_1} \dots dx^{i_p} + f d(dx^{i_1} \dots dx^{i_p}) = \\ &= df \wedge dx^{i_1} \dots dx^{i_p}, \end{aligned}$$

de donde, debido a la linealidad dada en el axioma (4) y a que toda p -forma se escribe como una suma de términos de la forma $f dx^{i_1} \dots dx^{i_p}$, obtenemos la fórmula (25.1).

Probar la proposición inversa, es decir, que la fórmula (25.1) satisface los axiomas, es un poco más difícil. Sin embargo, probar que satisface los axiomas (1) y (4) es trivial. Para probar que (25.1) satisface el axioma (2) necesitamos recordar la regla de Leibniz para funciones: $d(fg) = (df)g + fdg$. Las componentes de $\theta \wedge \tau$ son sumas de productos de las funciones componentes de θ y τ . Aplicando la fórmula (25.1) a cada término de la suma $\theta \wedge \tau$ y usando la regla de Leibniz, tenemos

$$\begin{aligned} d(\theta_{i_1 \dots i_p} \tau_{j_1 \dots j_q} dx^{i_1} \dots dx^{i_p} dx^{j_1} \dots dx^{j_q}) &= \\ &= d(\theta_{i_1 \dots i_p} \tau_{j_1 \dots j_q}) \wedge dx^{i_1} \dots dx^{i_p} dx^{j_1} \dots dx^{j_q} = \\ &= d\theta_{i_1 \dots i_p} \wedge dx^{i_1} \dots dx^{i_p} \tau_{j_1 \dots j_q} dx^{j_1} \dots dx^{j_q} + \\ &\quad + \theta_{i_1 \dots i_p} (-1)^p dx^{i_1} \dots dx^{i_p} \wedge d\tau_{j_1 \dots j_q} \wedge dx^{j_1} \dots dx^{j_q} \end{aligned}$$

De esta fórmula y de la linealidad (axioma (4)), se obtiene el axioma (2). El axioma (3) se sigue del hecho de que las derivadas parciales de segundo orden de una función no dependen del orden en el que sean calculadas. Esto, junto con la antisimetría de los produc-

tos exteriores nos lleva al resultado buscado. En efecto, para un término típico $fdx^{i_1} \dots dx^{i_p}$ de una p -forma tenemos

$$\begin{aligned} d^2(fdx^{i_1} \dots dx^{i_p}) &= d(df \wedge dx^{i_1} \dots dx^{i_p}) = d\left\{ \sum_k f_{x^k} dx^k \wedge dx^{i_1} \dots dx^{i_p} \right\} = \\ &= \sum_{k, i_1, \dots, i_p} f_{x^k x^{i_1}} dx^k dx^{i_1} \dots dx^{i_p}. \end{aligned}$$

Pero como $f_{x^k x^{i_1}} = f_{x^{i_1} x^k}$ y $dx^k dx^{i_1} \dots dx^{i_p} = -dx^{i_1} dx^k dx^{i_2} \dots dx^{i_p}$, tenemos que

$$\sum_{k, i_1, \dots, i_p} f_{x^k x^{i_1}} dx^k dx^{i_1} \dots dx^{i_p} = - \sum_{k, i_1, \dots, i_p} f_{x^k x^{i_1}} dx^{i_1} dx^k dx^{i_2} \dots dx^{i_p},$$

de donde $d^2(fdx^{i_1} \dots dx^{i_p}) = 0$. ##

El axioma (3) se conoce como lema de Poincaré, aunque hay cierta confusión en la literatura, pues en algunos libros la proposición inversa, "si $d\theta = 0$, entonces existe una forma τ tal que $\theta = d\tau$ ", se llama lema de Poincaré. Esta proposición inversa es verdadera sólo localmente y será demostrada en la sección 27.

C. Definición Invariante

Hay una fórmula invariante para definir a d , es decir, una fórmula que da $d\theta$ directamente como una forma diferencial, sin hacer referencia a un sistema coordinado. Esta fórmula involucra al paréntesis de Lie y resulta ilustrativo compararla con la fórmula del producto que demostramos en la proposición 21.2.

Sólo daremos la fórmula para los casos en que el grado de la forma diferencial es pequeño, pues estos casos son los más útiles.

Si f es una cero-forma: $df(X) = Xf$.

Si θ es una uno-forma: $d\theta(X, Y) = \frac{1}{2}\{X\theta(Y) - Y\theta(X) - \theta[X, Y]\}$.

Si θ es una dos-forma: $d\theta(X, Y, Z) = \frac{1}{3}\{X\theta(Y, Z) + Y\theta(Z, X) + Z\theta(X, Y) - \theta([X, Y], Z) - \theta([Y, Z], X) - \theta([Z, X], Y)\}$.

(Los extraños factores $\frac{1}{2}, \frac{1}{3}, \dots$ pueden eliminarse usando otra definición de los productos exteriores. Esta definición alternativa, que no altera en nada las propiedades esenciales de los productos exteriores, aparece frecuentemente en la literatura. Esta sólo difiere de la definición que nosotros hemos dado por la aparición de un factor entero de la forma $(p+q)!/p!q!$.)

26 PRODUCTOS INTERIORES

El producto interior por X es un operador $i(X): \Lambda^p(M) \rightarrow \Lambda^{p-1}(M)$ definido para cada campo vectorial $X \in \mathfrak{X}(M)$. Esencialmente, $i(X)\theta$ fija la primera variable de la p -forma θ en X , pero deja libres a las restantes $p-1$ variables. En fórmulas tenemos

$$[i(X)\theta](X_1, \dots, X_{p-1}) = p\theta(X, X_1, \dots, X_{p-1}),$$

en donde $\theta \in \wedge^p(M)$ y $X_1, \dots, X_{p-1} \in \mathfrak{X}(M)$.

Para cero-formas, definimos $i(X)\theta = 0$.

Es claro que $i(X): \wedge^p(M) \rightarrow \wedge^{p-1}(M)$ es un operador lineal.

Ejemplo. Calculemos $i(\partial_i)$ aplicado a las p -formas básicas $dx^{i_1} \dots dx^{i_p}$, donde $i_1 < \dots < i_p$.

(a) Si $i_1 \neq 1$ tenemos para $j_1 < \dots < j_{p-1}$,

$$i(\partial_{i_1})(dx^{i_1} \dots dx^{i_p}) = p dx^{i_1} \dots dx^{i_p} (\partial_{i_1}, \partial_{j_1}, \dots, \partial_{j_{p-1}}) = 0 \quad (\text{ver pag. 64}).$$

Así, $i(\partial_{i_1})(dx^{i_1} \dots dx^{i_p}) = 0$, ya que todas sus componentes son cero.

(b) Si $i_1 = 1$ tenemos que $p dx^1 dx^{i_2} \dots dx^{i_p} (\partial_1, \partial_{j_1}, \dots, \partial_{j_{p-1}})$ es cero si

$(i_2, \dots, i_p) \neq (j_1, \dots, j_{p-1})$ y es $p/p_j = 1/(p-1)$; si $(i_2, \dots, i_p) = (j_1, \dots, j_{p-1})$. Puesto que estos son los mismos valores que $dx^{i_2} \dots dx^{i_p}$ tiene sobre $(\partial_{j_1}, \dots, \partial_{j_{p-1}})$ se sigue que

$$i(\partial_1)(dx^1 dx^{i_2} \dots dx^{i_p}) = dx^{i_2} \dots dx^{i_p}.$$

La acción de $i(\partial_i)$ sobre todas las otras formas puede ahora obtenerse usando la linealidad de $i(\partial_i)$.

Proposición 26.1 El operador $i(X): \wedge^p(M) \rightarrow \wedge^{p-1}(M)$ es una antiderivación, es decir, $i(X)$ es una transformación lineal y si θ es una p -forma y τ es una q -forma, $i(X)$ satisface la regla del producto:

$$i(X)(\theta \wedge \tau) = i(X)\theta \wedge \tau + (-1)^p \theta \wedge i(X)\tau.$$

Demostración. El operador $i(X)$ es puramente algebraico, así que el valor de $i(X)\theta$ en un punto depende solamente de los valores de X y θ en dicho punto. En particular, si X es cero en un punto, entonces ambos miembros en la fórmula de la regla del producto son cero en ese punto. Por lo tanto, necesitamos sólo considerar el caso donde $X \neq 0$, y podemos también escoger un sistema coordenado tal que $X = \partial_{i_1}$.

Si escribimos θ y τ en términos de sus expresiones coordenadas y desarrollamos ambos miembros de la fórmula del producto usando la ley distributiva para el producto exterior y la linealidad de $i(\partial_{i_1})$, resulta claro que sólo necesitamos probar la fórmula para el caso en donde θ y τ son las formas coordenadas $dx^{i_2} \dots dx^{i_p}$ by $dx^{j_1} \dots dx^{j_q}$, respectivamente. Así pues, nos quedan cuatro subcasos dependiendo de que i_1 y j_1 sean iguales a 1 o no. Estos subcasos pueden ser abordados usando el ejemplo anterior. ##

La composición de estos operadores producto interior es antisimétrica, es decir, $i(X)i(Y) = -i(Y)i(X)$:

$$\begin{aligned} i(X)i(Y)\theta(\dots) &= p i(Y)\theta(X, \dots) = p(p-1)\theta(Y, X, \dots) = \\ &= -p(p-1)\theta(X, Y, \dots) = -i(Y)i(X)\theta(\dots). \end{aligned}$$

Como hay una gran analogía formal entre derivaciones tensoriales y antiderivaciones de formas, no debe sorprendernos el hecho de que

una antiderivación esté completamente determinada si se conoce cual es su efecto sobre cero-formas y uno-formas. En efecto, sólo tenemos que expresar una p-forma arbitraria en términos de cero-formas y de uno-formas y aplicar repetidamente la regla del producto. Este es precisamente el caso de las antiderivaciones d e $i(X)$.

Debido a que en la definición invariante del operador d que dimos en la sección anterior aparece la derivada de Lie en la forma de paréntesis de Lie, no debe sorprendernos tampoco que exista una relación entre la derivada de Lie para formas y las antiderivaciones d e $i(X)$.

Teorema 26.2 Sobre formas diferenciales, la derivada de Lie con respecto a X , L_X , está dada por la ecuación operacional

$$L_X = i(X)d + di(X).$$

(Nótese que en cada miembro de esta ecuación tenemos un operador lineal de $\wedge^p(M) \rightarrow \wedge^p(M)$.)

Demostración. Hemos visto que L_X es una derivación tensorial; es decir, que es un operador lineal que preserve el tipo tensorial y que satisface la regla del producto de Leibniz (ver sección 21). Demostraremos que $i(X)d + di(X)$ es también una derivación:

$$\begin{aligned} [i(X)d + di(X)](\theta \wedge \tau) &= i(X)(d\theta \wedge \tau + (-1)^p \theta \wedge d\tau) + \\ &\quad + d(i(X)\theta \wedge \tau + (-1)^p \theta \wedge i(X)\tau) \\ &= i(X)d\theta \wedge \tau + (-1)^{p+1} d\theta \wedge i(X)\tau \\ &\quad + (-1)^p i(X)\theta \wedge d\tau + (-1)^{2p} \theta \wedge i(X)d\tau \\ &\quad + di(X)\theta \wedge \tau + (-1)^{p-1} i(X)\theta \wedge d\tau \\ &\quad + (-1)^p d\theta \wedge i(X)\tau + (-1)^{2p} \theta \wedge di(X)\tau \\ &= (i(X)d + di(X))\theta \wedge \tau + \theta \wedge (i(X)d + di(X))\tau. \end{aligned}$$

Así, si $L_X = i(X)d + di(X)$ coinciden sobre cero-formas y uno-formas, entonces coinciden sobre todas las p-formas.

Sobre cero-formas tenemos $L_X f = Xf$, mientras que $i(X)df + di(X)f = i(X)df + d0 = df(X) = Xf$.

Sobre una uno-forma df tenemos que $L_X df = d(Xf)$. Para demostrar esto, denotemos por $\langle Y, df \rangle$ a la función que resulta al evaluar la uno-forma df sobre el campo vectorial Y . Entonces tenemos $L_X \langle Y, df \rangle = X \langle Y, df \rangle = X \langle Yf \rangle$. Pero, por otra parte,

$$\begin{aligned} L_X \langle Y, df \rangle &= \langle L_X Y, df \rangle + \langle Y, L_X df \rangle = \langle [X, Y], df \rangle + \langle Y, L_X df \rangle = \\ &= XYf - YXf + \langle Y, L_X df \rangle, \end{aligned}$$

de donde $\langle Y, L_X df \rangle = YXf = \langle Y, d(Xf) \rangle$. Como Y es arbitrario, hemos demostrado que $L_X df = d(Xf)$.

Por otra parte,

$$[i(X)d + di(X)]df = i(X)d^2f + di(X)df = 0 + d(Xf).$$

##

Corolario 26.3 Los operadores d y L_X conmutan al ser aplicados a formas diferenciales; es decir, para toda p -forma θ , $dL_X\theta = L_Xd\theta$.

Demostración. Usando la fórmula para L_X que acabamos de demostrar en el teorema anterior, y el hecho de que $d^2 = 0$, tenemos que

$$dL_X\theta = d(i(X)d\theta + di(X)\theta) = di(X)d\theta,$$

mientras que

$$L_Xd\theta = (i(X)d + di(X))d\theta = di(X)d\theta. \quad \#\#$$

Cuando la fórmula del teorema se escribe como

$$(p+1)d\theta(X, \dots) = [L_X\theta - d(i(X)\theta)](\dots),$$

obtenemos un medio para determinar a d sobre p -formas a partir de la derivada de Lie y d sobre $(p-1)$ -formas. Esto sugiere que cuando deseamos probar alguna propiedad del operador d y tenemos la propiedad correspondiente para derivadas de Lie, podemos utilizar la inducción matemática con respecto al grado (tipo tensorial) de las formas involucradas.

27 INVERSO DEL LEMA DE POINCARÉ

Consideremos los siguientes resultados que aparecen con frecuencia en los textos de física teórica, los cuales se dan sin mayor discusión y que pertenecen al análisis vectorial en \mathbb{R}^3 .

- (a) Si $\text{grad } f = 0$, entonces f es constante.
- (b) Si $\text{rot } X = 0$, entonces existe una función f tal que $X = \text{grad } f$.
- (c) Si $\text{div } X = 0$, entonces existe un campo vectorial Y tal que $\text{rot } Y = X$.
- (d) Para toda función f existe un campo vectorial X tal que $\text{div } X = f$.

Cada una de estas afirmaciones es defectuosa, aún cuando (d) sólo requiere una modesta suposición de diferenciabilidad. Sin embargo, tales suposiciones de diferenciabilidad no pueden servir para reparar a las afirmaciones (a), (b) y (c), puesto que su mayor defecto radica en que no especifican ciertas suposiciones topológicas acerca de los dominios de definición de f y X . En (a) debe suponerse que el dominio de f es conexo; en (b) que el dominio de X es simplemente conexo, es decir, que toda curva cerrada simple en el dominio de X es la curva frontera de una superficie de extensión finita contenida en el dominio de X ; en (c) debe suponerse que toda superficie compacta en el dominio de X es la frontera de una región acotada contenida en el dominio de X .

El propósito de esta sección es unificar y generalizar las versiones correctas de las afirmaciones (a), (b), (c) y (d). Esto se logra mediante la traducción de estos resultados en afirmaciones acerca de p -formas sobre variedades. Las condiciones topológicas sobre los do-

minios se reemplazan por una sólo condición más fuerte que implica todos los casos especiales: Suponemos que el dominio contiene a un cubo coordenado $U = \{m \in M \mid a^i < x^i(m) < b^i\}$ de algún sistema coordenado x^1, \dots, x^n de M .

Una p -forma τ se llama cerrada si $d\tau = 0$. Si $p > 0$, decimos que τ es exacta si existe una $(p-1)$ -forma θ tal que $\tau = d\theta$; una p -forma es exacta si es constante. Si para todo punto m en el dominio de τ hay una vecindad U de m tal que $\tau|_U$, la restricción de τ a U , es exacta, entonces decimos que τ es localmente exacta.

Es obvio que la exactitud implica la exactitud local. El axioma (3) para el operador d (sección 25), es decir, el lema de Poincaré, muestra que la exactitud local implica la cerradura. De hecho, si $\tau|_U = d\theta|_U$, entonces $d\tau|_U = d(d\theta|_U) = d^2\theta|_U = 0$, y puesto que esto es verdad para alguna U alrededor de todo punto, $d\tau = 0$. Demostraremos ahora el inverso local del lema de Poincaré:

Teorema 27.1 Si τ es una p -forma cerrada, $p = 1, \dots, d = \dim M$, entonces para toda vecindad cúbica de coordenadas $U = \{m \in M \mid a^i < x^i(m) < b^i\}$ contenida en el dominio de τ , existe una $(p-1)$ -forma θ definida sobre U tal que $d\theta = \tau|_U$. Una p -forma cerrada definida sobre U es constante sobre U . En particular, toda p -forma cerrada es localmente exacta.

Demostración. Para una p -forma τ , $d\tau = 0$ significa que en U , $\partial_i \tau = 0$, $i = 1, \dots, d$. Se sigue entonces que τ es constante sobre cualquier curva en U , y dado que U es conexa, τ es constante en U .

Sin pérdida de generalidad podemos suponer que el origen $0 \in \mathbb{R}^d$ corresponde a algún punto de U bajo el sistema coordenado x^1, \dots, x^d ; es decir, $a^i < 0 < b^i$, $i = 1, \dots, d$.

Para concluir la demostración construiremos lo que se conoce como una homotopía algebraica H de d sobre formas definidas en U . Esto significa que $H: \wedge^p(M) \rightarrow \wedge^{p-1}(M)$ es una transformación lineal tal que para toda p -forma τ se tiene

$$Hd\tau + dH\tau = \tau;$$

es decir, $Hd + dH$ es la transformación identidad sobre formas. Habiendo construido dicha homotopía algebraica H es trivial resolver el problema de encontrar θ , ya que $d\tau = 0$ implica que $dH\tau = \tau$; así pues, podemos hacer $\theta = H\tau$.

Definamos H para formas del tipo $\alpha = f(x^1, \dots, x^d) dx^{i_1} \dots dx^{i_p}$, donde $i_1 < \dots < i_p$, por

$$H\alpha = \left[\int_0^1 f(0, \dots, 0, tx^{i_1}, x^{i_2}, \dots, x^d) dt \right] x^{i_1} dx^{i_2} \dots dx^{i_p}.$$

Si α es una p -forma, definimos $H\alpha = 0$. Extenderemos H a las p -formas arbitrarias por medio de la linealidad de H .

Se requiere un cómputo algo largo para verificar que $Hd\alpha + dH\alpha = \alpha$. Tomar la derivada exterior de $H\alpha$ involucra derivadas parciales de una

integral con respecto a parámetros en el integrando. Un teorema muy conocido del cálculo avanzado justifica la posibilidad de pasar los operadores de derivación parcial dentro de los signos de integral. Aparte de este resultado, necesitamos observar que

$$x^i \partial_i f(0, \dots, 0, tx^i, x^{i+1}, \dots, x^d) = \frac{d}{dt} f(0, \dots, 0, tx^i, x^{i+1}, \dots, x^d)$$

y que

$$\begin{aligned} tx^i \partial_i f(0, \dots, 0, tx^i, x^{i+1}, \dots, x^d) + f(0, \dots, 0, tx^i, x^{i+1}, \dots, x^d) &= \\ &= \frac{d}{dt} [tf(0, \dots, 0, tx^i, x^{i+1}, \dots, x^d)]. \end{aligned}$$

Completar los detalles de la demostración es un fácil ejercicio. ##

Ejemplos. Demostremos como funciona H en los casos (b), (c) y (d) que consideramos al principio de esta sección.

(b) Supongamos que $\tau = f dx + g dy + h dz$, $d\tau = 0$, y τ está definida sobre una región cúbica de \mathbb{R}^3 . A partir de $d\tau = 0$ tenemos que $f_y = g_x$, $f_z = h_x$, y $g_z = h_y$, donde los subíndices indican derivadas parciales. Entonces $\theta = H\tau$ es una cero-forma dada por

$$\theta(x, y, z) = x \int_0^1 f(tx, y, z) dt + y \int_0^1 g(0, ty, z) dt + z \int_0^1 h(0, 0, tz) dt.$$

De aquí obtenemos

$$\begin{aligned} d\theta &= \left(\int_0^1 [f(tx, y, z) + xtf_x(tx, y, z)] dt \right) dx + \\ &+ \left(\int_0^1 [xf_y(tx, y, z) + g(0, ty, z) + ytg_y(0, ty, z)] dt \right) dy + \\ &+ \left(\int_0^1 [xf_z(tx, y, z) + yg_z(0, ty, z) + h(0, 0, tz) + zth_z(0, 0, tz)] dt \right) dz \\ &= (f(x, y, z) - 0) dx + ([g(x, y, z) - g(0, y, z)] + [g(0, y, z) - 0]) dy \\ &+ ([h(x, y, z) - h(0, y, z)] + [h(0, y, z) - h(0, 0, z)] + [h(0, 0, z) - 0]) dz \\ &= f dx + g dy + h dz = \tau. \end{aligned}$$

(c) Si $\tau = f dx dy + g dx dz + h dy dz$, entonces $d\tau = 0$ implica que $f_z + g_y + h_x = 0$. Para $\theta = H\tau$ tenemos

$$\theta = \left(\int_0^1 f(tx, y, z) dt \right) x dy + \left(\int_0^1 g(tx, y, z) dt \right) x dz + \left(\int_0^1 h(0, ty, z) dt \right) y dz.$$

La verificación de que $d\theta = \tau$ se realiza con la misma técnica que en el ejemplo (b).

(d) Si $\tau = f dx dy dz$ y

$$\theta = \left(\int_0^1 f(tx, y, z) dt \right) x dy dz,$$

entonces es obvio que $d\theta = \tau$.

Debe quedar claro que la solución θ cuya existencia garantiza el teorema 27.1 no es única, excepto cuando τ es una uno-forma. De hecho, si α es una $(p-2)$ -forma arbitraria, entonces $d(\theta + d\alpha) = d\theta + d^2\alpha = \tau$.

28 CADENAS CUBICAS

Los objetos sobre los cuales integraremos a las p -formas serán un poco más generales que las subvariedades orientadas de dimensión p . Nosotros integraremos sobre p -cubos suaves y sobre sumas formales de tales p -cubos (p -cadenas). Desde luego, el dominio de integración que surge en un problema de física rara vez está dado como una cadena, así que al aplicar esta teoría de integración debemos desarrollar la habilidad de representar los dominios que se encuentran comúnmente en forma de cadenas, es decir, debemos parametrizar los dominios. En las aplicaciones de la matemática pocas veces se parametrizan los dominios explícitamente, más bien se usa el hecho de que una amplia clase de dominios son parametrizables.

Un p -cubo ($p > 0$) en \mathbb{R}^p es una vecindad cúbica cerrada con respecto a las coordenadas cartesianas:

$$U = \{(u^1, \dots, u^p) \in \mathbb{R}^p \mid b^i \leq u^i \leq b^i + c^i, i=1, \dots, p\},$$

donde los b^i y c^i son constantes dadas con $c^i > 0$, $i=1, \dots, p$. No permitiremos que las cotas de las u^i tomen valores infinitos, entonces U es cerrado y acotado, por lo tanto, compacto.

Un p -cubo suave α en una variedad M es una aplicación suave $\alpha: U \rightarrow M$ donde U es un p -cubo en \mathbb{R}^p .

Un p -cubo orientado es una pareja (α, ω) , en donde α es un p -cubo y ω es una orientación de \mathbb{R}^p . De acuerdo con la definición que dimos en la sección 11, ω es un atlas de cartas sobre \mathbb{R}^p relacionadas entre sí por medio de determinantes jacobianos positivos. Sin embargo, para nuestros propósitos en esta sección es mejor expresar la orientación por medio de p -formas. Si las coordenadas x^1, \dots, x^p y y^1, \dots, y^p están relacionadas por el determinante jacobiano $J = \det(\partial x^i / \partial y^j)$, entonces $dx^1 \dots dx^p = J dy^1 \dots dy^p$ (ver corolario 20.3, pag. 51). Así, siempre podemos decir cuando dos sistemas coordenados están consistentemente orientados comparando sus "elementos coordenados de volumen" $dx^1 \dots dx^p$ y $dy^1 \dots dy^p$. Puesto que uno de los dos sistemas globales de coordenadas cartesianas u^1, \dots, u^p o $-u^1, \dots, u^p$ debe estar consistentemente orientado con ω , nosotros identificaremos ω con uno de los elementos de volumen $du^1 \dots du^p$ o $d(-u^1) du^2 \dots du^p$.

Si $-\omega$ es la orientación opuesta a ω , entonces decimos que $(\alpha, -\omega)$ es el negativo de (α, ω) .

Completamos nuestras definiciones incluyendo el caso $p = 0$ y definimos un cero-cubo en M como un punto $m \in M$. Un cero-cubo orientado en M será un punto de M junto con uno de los números $+1$ o -1 , es decir, $(m, +1)$ o $(m, -1)$; estos son negativos uno del otro.

Si U , como antes, es el dominio de un p -cubo α , definimos las $(p-1)$ -caras de α como los $(p-1)$ -cubos $\alpha_{i\varepsilon}$, $i=1, \dots, p$, $\varepsilon=0, 1$, definidos por

$$\alpha_{i\varepsilon}(v^1, \dots, v^{p-1}) = \alpha(v^1, \dots, v^{i-1}, b^i + \varepsilon c^i, v^i, \dots, v^{p-1}),$$

donde $b^j \leq v^j \leq b^j + c^j$ para $j=1, \dots, i-1$ y $b^{j+1} \leq v^j \leq b^{j+1} + c^{j+1}$ para $j=i, \dots, p-1$.

Así pues, vemos que α tiene $2p$ de estas $(p-1)$ -caras. Las s-caras de α se definen recursivamente como las s -caras de las $(s+1)$ -caras de α , $s=0, \dots, p-2$. Estas se designan por $\alpha_{i_1, i_2, \dots, i_h, \epsilon_h}$, en lugar de la notación más complicada $(\dots(\alpha_{i_1, \epsilon_1})_{i_2, \epsilon_2} \dots)_{i_h, \epsilon_h}$, donde $h = p-s$. En particular, los vértices o cero-caras de α son los puntos

$$(b^1 + \epsilon, c^1, \dots, b^p + \epsilon p c^p),$$

así que un p -cubo tiene 2^p vértices.

Para definir las $(p-1)$ -caras de un p -cubo orientado (α, ω) debemos dar a $\alpha_{i, \epsilon}$ una orientación $\omega_{i, \epsilon}$. Puesto que queremos que $(-\omega)_{i, \epsilon} = -(\omega_{i, \epsilon})$ tomamos esta propiedad como parte de la definición y restringimos nuestra atención a $\omega = du^1 \dots du^p$. Entonces hacemos

$$\omega_{i, \epsilon} = (2\epsilon - 1)(-1)^{i-1} dv^1 \dots dv^{p-1}.$$

Sobre la cara de U para la cual $u^i = b^i + \epsilon c^i$, el vector coordenado correspondiente a la coordenada $(2\epsilon - 1)u^i$ está dirigido hacia afuera con respecto al interior de U . El sistema de coordenadas que consta de $(2\epsilon - 1)u^i$ y $(2\epsilon - 1)(-1)^{i-1}v^1, v^2, \dots, v^{p-1}$, será un sistema con una orientación consistente con ω , ya que el signo $(-1)^{i-1}$ sirve para compensar el corrimiento en la posición de u^i . Así, hemos elegido la orientación en las caras de la frontera de U de acuerdo con la "convención de la normal dirigida hacia afuera" (ver fig. 8).

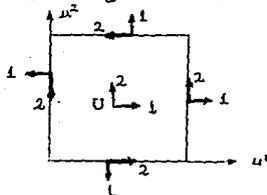
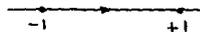


Fig. 8



Las cero-caras de un uno-cubo orientado (α, du^1) se definen como $(\alpha(b^1), -1)$ y $(\alpha(b^1 + c^1), +1)$.

Es interesante e importante darse cuenta de que no es posible definir consistentemente una orientación para las $(p-2)$ -caras de un p -cubo orientado. De hecho, tenemos

Proposición 28.1 Sea α un p -cubo ($p > 1$), sea ω una orientación de α , y sean $1 \leq i < j \leq p$. Entonces los $(p-2)$ -cubos $\alpha_{i, \delta(i-1), \epsilon}$ y $\alpha_{j, \epsilon, i, \delta}$ son la misma $(p-2)$ -cara de α , y las orientaciones determinadas en ellas por ω a través de $\alpha_{i, \delta}$ y $\alpha_{j, \epsilon}$ son negativas entre sí, es decir, $\omega_{i, \delta(j-1), \epsilon} = -\omega_{j, \epsilon, i, \delta}$.

Demostración. Es evidente que $\alpha_{i, \delta(i-1), \epsilon}$ y $\alpha_{j, \epsilon, i, \delta}$ se obtienen ambos a partir de α dando a u^i y u^j los valores $b^i + \delta c^i$ y $b^j + \epsilon c^j$, respectivamente. Los signos asociados a la primera de las restantes coordenadas,

los cuales determinan las orientaciones $\omega_i \delta^{(i-1)\epsilon}$ y $\omega_j \epsilon_i \delta$ en el caso $\omega = du^1 \dots du^p$, son $(2\delta-1)(-1)^{i-1}(2\epsilon-1)(-1)^j$ y $(2\epsilon-1)(-1)^{i-1}(2\delta-1)(-1)^{i-1}$, respectivamente. Estos signos son claramente negativos entre sí. ##

Decimos que los p -cubos α y β son equivalentes si existe un difeomorfismo φ entre conjuntos abiertos que contienen a los dominios de α y β tal que aplica las s -caras de U (= dominio de α) sobre las s -caras de V (= dominio de β), $s=0, \dots, p$, y tal que el diagrama

$$\begin{array}{ccc} \bar{U} & \xrightarrow{\alpha} & M \\ \varphi \downarrow & & \nearrow \beta \\ \bar{V} & & \end{array} \quad \begin{array}{l} U \subset \bar{U} \subset \mathbb{R}^p \\ V \subset \bar{V} \subset \mathbb{R}^p \end{array}$$

conmuta; es decir, $\beta \circ \varphi = \alpha$. Es claro que la equivalencia de p -cubos es una relación de equivalencia. También es obvio que los p -cubos equivalentes tienen la misma imagen. Se pueden obtener ejemplos simples de p -cubos equivalentes haciendo que φ sea una translación, sea una multiplicación de algunas de las coordenadas por números reales distintos de cero, sea una permutación de las coordenadas o que sea una combinación de todas estas operaciones.

Si α es equivalente a β por medio del difeomorfismo φ y ω es una orientación para α , entonces puesto que φ es, en particular, un sistema coordenado para una vecindad de \mathbb{R}^p , debe estar consistentemente orientado con (u^1, \dots, u^p) o con $(-u^1, \dots, u^p)$. Dependiendo de cual sea el caso, definimos a (α, ω) como equivalente a (β, ω) o a $(\beta, -\omega)$. Nuevamente esta es una relación de equivalencia entre p -cubos orientados.

Proposición 28.2 Si (α, ω) es equivalente a $(\beta, \delta\omega)$, $\delta = \pm 1$, entonces las $(p-1)$ -caras orientadas de (α, ω) son equivalentes a las $(p-1)$ -caras de $(\beta, \delta\omega)$, cuando estas últimas se toman en algún orden especial.

(Esto se sigue inmediatamente de las definiciones.)

Una p -cadena es una suma formal finita $\sum r_i C_i$ de p -cubos orientados C_i con números reales r_i como coeficientes. Una p -cadena $\sum r_i C_i$ es equivalente a una p -cadena $\sum s_j D_j$ si para todo p -cubo orientado (α, ω) tenemos

$$\begin{aligned} \sum \{r_i | C_i \text{ es equivalente a } (\alpha, \omega)\} \\ - \sum \{r_i | C_i \text{ es equivalente a } (\alpha, -\omega)\} \\ = \sum \{s_j | D_j \text{ es equivalente a } (\alpha, \omega)\} \\ - \sum \{s_j | D_j \text{ es equivalente a } (\alpha, -\omega)\}. \end{aligned}$$

Decimos que una p -cadena $\sum t_i E_i$ es irreducible si para toda i y j ($i \neq j$), E_i no es equivalente al negativo de E_j , o a E_j , o al negativo de E_j . Se sigue que toda p -cadena es equivalente a una p -cadena irreducible. Para cada p permitimos la posibilidad de sumas formales vacías (sin términos), en este caso hablamos de la p -cadena nula, denotada por 0.

Para una cero-cadena (no confundir con la cadena 0), combinamos los signos que definen la orientación con los coeficientes y simplemente la escribimos como una suma $\sum r_i m_i$, en donde $m_i \in M$ para toda i .

Las cadenas se pueden sumar y pueden ser multiplicadas por números reales de manera obvia, y estas operaciones son compatibles con la relación de equivalencia entre cadenas. Es claro que para cada p , el conjunto de las p -cadenas es un espacio vectorial sobre los números reales.

Decimos que x es un punto regular de α si $x \in U^\circ$, el interior del dominio cúbico U de α , y si $d\alpha$ es biyectiva en x . Denotaremos por U° al conjunto de todos los puntos regulares de α .

El significado geométrico de las p -cadenas proviene de la posibilidad de representar por medio de ellas cierta región de una subvariedad orientada de dimensión p . Haremos esto sólo para el caso de p -cadenas irreducibles con coeficientes $t_i = 1$. Decimos que un p -cubo orientado (α, ω) parametriza una región S de la subvariedad N si α es inyectiva sobre U° , S es la imagen de U bajo α ($S = \alpha(U)$), y siempre que $d\alpha$ sea biyectiva en x y v_1, \dots, v_p es una base para $T_x(\mathbb{R}^p)$ que es consistente con la orientación ω , entonces la base $dx(v_1), \dots, dx(v_p)$ de $T_{\alpha x}(N)$ es consistente con la orientación de N . [Una base de vectores tangentes v_i en x es consistente con ω si existe un sistema coordenado consistentemente orientado con ω tal que $\partial_i|_x = v_i$. Si un p -cubo parametriza una región, la misma región es parametrizada por cualquier otro p -cubo equivalente al primero.]

Una p -cadena irreducible (α_i, ω_i) parametriza una región S de N si

- Cada (α_i, ω_i) parametriza una región S_i de N .
- S es la unión de las S_i .
- Para todo $i \neq j$, $\alpha_i(U_i^\circ)$ y $\alpha_j(U_j^\circ)$ son disjuntos, donde U_i° es el conjunto de puntos regulares del dominio U_i de α_i .

Nótese que no hemos requerido que los p -cubos se encuentren unos a otros sobre sus caras de alguna manera regular, aún cuando exigir que los p -cubos se encuentren de una manera semejante es razonable cuando consideramos la parametrización de variedades con frontera, (ver más adelante).

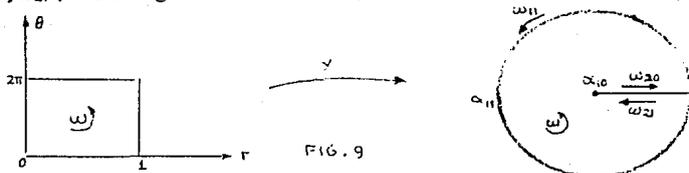
Una p -cadena parametriza una región S , cuando existe una p -cadena irreducible que parametriza la misma región S y que es equivalente a la p -cadena dada.

Ejemplo. Las coordenadas polares sobre el plano parametrizan al disco unitario cerrado por medio de un dos-cubo $\alpha(r, \theta) = (r \cos \theta, r \sin \theta)$ definido sobre el rectángulo $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$ (ver fig. 9). Los puntos regulares son los puntos interiores. Las cuatro caras están dadas por:

$\alpha_{1,0} = (0 \cos \theta, 0 \sin \theta) = (0, 0)$, de modo que $\alpha_{1,0}$ es un uno-cubo constante.

$\alpha_{11}\theta = (\cos\theta, \text{sen}\theta)$, de modo que α_{11} parametriza al círculo unitario.

$\alpha_{20}r = \alpha_{21}r = (r, 0)$, de modo que α_{20} y α_{21} son equivalentes y parametrizan un segmento unitario del eje x. Nótese que para cualquiera de las dos orientaciones ω de α , las caras orientadas $(\alpha_{20}, \omega_{20})$ y $(\alpha_{21}, \omega_{21})$ son negativas entre sí.



La frontera de un p-cubo orientado (α, ω) es la $(p-1)$ -cadena que consiste de la suma de todas las $(p-1)$ -caras orientadas de α :

$$\sum_{i=1, \dots, p} (\alpha_{i\varepsilon}, \omega_{i\varepsilon}).$$

La frontera de (α, ω) se denota por $\partial(\alpha, \omega)$. La frontera de una p-cadena $\sum r_i C_i$ es $\partial \sum r_i C_i = \sum r_i \partial C_i$. Así pues, ∂ es un operador lineal entre el espacio vectorial de las p-cadenas y el espacio vectorial de las $(p-1)$ -cadenas, llamado operador frontera. Este operador también se comporta bien con respecto a la equivalencia; es decir, si la p-cadena C es equivalente a la p-cadena D, entonces ∂C es equivalente a ∂D .

Proposición 28.3 Para cualquier p-cadena C, $p > 1$, $\partial^2 C = \partial \partial C = 0$.

Demostración. Para un p-cubo orientado (α, ω) , $\partial(\alpha, \omega) = 0$ se sigue inmediatamente de la proposición 28.1, ya que las $(p-2)$ -caras se cancelan por parejas pues tienen orientaciones contrarias. Entonces tenemos

$$\partial \partial C = \partial \partial \sum r_i C_i = \partial \sum r_i \partial C_i = \sum r_i \partial \partial C = 0. \quad \#\#$$

Para un uno-cubo (α, ω) , la frontera consiste esencialmente del punto final menos el punto inicial. Así, la suma de los coeficientes de $\partial(\alpha, \omega)$ es cero. En general, definiremos la suma de los coeficientes de una cero-cadena C como el índice de Kronecker de C, y lo denotaremos por $IC = I(\sum r_i m_i) = \sum r_i$. Se sigue que para cualquier uno-cadena D, $I\partial D = 0$. El inverso de este resultado no es cierto en general, pero la condición para que se verifique es de naturaleza topológica y nos da una idea de la relación que existe entre el álgebra de cadenas y la topología.

Proposición 28.4 Una variedad M es conexa si y sólo si toda cero-cadena C en M tal que $IC = 0$ es la frontera de una uno-cadena D en M: $\partial D = C$.

Demostración. Supongamos que M es conexa y que C es una cero-cadena en M tal que $IC = 0$. Entonces $C = \sum r_i m_i$, donde $\sum r_i = 0$. Tomemos un punto $m_0 \in M$. Para cada punto m_i podemos encontrar una curva suave α_i cuyos puntos extremos sean m_0 y m_i debido a que M es conexa. Podemos suponer que α_i está parametrizada de 0 a 1, de modo que α_i es un uno-cubo definido sobre $[0,1]$. Entonces la uno-cadena D sobre M , dada por $D = \sum r_i (\alpha_i, du)$ tiene frontera

$$\begin{aligned} D &= \sum r_i \partial(\alpha_i, du) = \sum r_i (\alpha_i(1) - \alpha_i(0)) \\ &= \sum r_i m_i - (\sum r_i) m_0 = \sum r_i m_i = C. \end{aligned}$$

Inversamente, si M no es conexa, entonces para cada componente conexa M_τ de M definimos el índice parcial de Kronecker I_τ sobre cero-cadenas por $I_\tau(\sum r_i m_i) =$ suma de aquellos r_i tales que $m_i \in M_\tau$. Puesto que un uno-cubo está completamente contenido en M_τ o está completamente fuera de M_τ , todavía tenemos que $I_\tau \partial D = 0$ para toda uno-cadena D . Sean ahora M_0 y M_1 componentes conexas diferentes de M y tomemos $m_0 \in M_0$ y $m_1 \in M_1$. Entonces $m_1 - m_0$ es una cero-cadena C tal que $IC = 0$, pero $I_\tau(m_1 - m_0) = 1$, de modo que $m_1 - m_0$ no puede ser una frontera. ##

Una clase bastante amplia de regiones en variedades pueden ser parametrizadas por medio de cadenas. A continuación daremos, sin demostración, un par de teoremas acerca de este tema que más bien pertenecen a la topología algebraica.

Teorema 28.5 Sea M una variedad compacta y orientada de dimensión d . Entonces existe una d -cadena C en M que parametriza a la propia M y para la cual ∂C es equivalente a 0.

Otra clase importante de espacios parametrizables por medio de cadenas son las variedades compactas orientables y con frontera. Un subconjunto $N = N \cup B$ de una variedad M es una subvariedad p -dimensional con frontera B si

- N es una subvariedad abierta de una subvariedad p -dimensional N^+ de M .
- B es una subvariedad $(p-1)$ -dimensional de N^+ .
- B es la frontera topológica de N con respecto a la topología de N^+ ; es decir, todo abierto de N^+ que contenga a un punto de B , debe contener puntos de N y puntos de $N^+ - N$.
- En cada punto $b \in B$ hay coordenadas x^1, \dots, x^p , sobre una vecindad U de b en N^+ tales que $B \cap U = \{n \mid x^i(n) = 0\}$ y $N \cap U = \{n \mid x^i(n) < 0\}$.

Si N está orientada, entonces B tiene la correspondiente orientación inducida, ésta es tal que siempre que las coordenadas x^i en (d) son consistentes con la orientación de N , entonces las coordenadas x^1, \dots, x^p , restringidas a B , son consistentes con la orientación de B .

Teorema 28.6 Si N es una variedad compacta y orientada con frontera B , entonces existe una cadena C que parametriza a N y tal que ∂C parametriza a B con la orientación inducida.

Para cerrar esta sección consideraremos brevemente la cuestión de los elementos de volumen. Si M es una variedad orientada de dimensión d , definimos un elemento de volumen sobre M como una d -forma Ω sobre M , que nunca es cero y que es consistente con la orientación de M , en el siguiente sentido: Para todo sistema coordenado x^1, \dots, x^d sobre M que es consistente con la orientación de M , la expresión coordenada de Ω es $\Omega = f dx^1 \dots dx^d$, donde f es una función suave y positiva sobre M . Para cualquier función g suave y positiva sobre todo M , $g\Omega$ es un elemento de volumen sobre M si Ω lo es, e inversamente, cualesquiera dos elementos de volumen son múltiplos positivos uno del otro. De hecho, cualquier d -forma se obtiene de Ω multiplicando a esta última por una función suave.

Para demostrar que siempre es posible definir un elemento de volumen sobre una variedad orientada, se puede utilizar el dispositivo de las particiones de la unidad para ensamblar los diferentes elementos coordenados locales de volumen $dx^1 \dots dx^d$ y obtener un elemento de volumen global sobre M .

Si Ω es un elemento de volumen sobre M y (α, ω) es un d -cubo que parametriza una región de M , entonces la consistencia de la orientación dada por α en sus puntos regulares implica el hecho de que $\alpha^*\Omega = f\omega$, donde $f \geq 0$ y $f > 0$ en los puntos regulares.

Ejemplo. Sea $M = \mathbb{R}^3$ y sea N la superficie cilíndrica cerrada:

$$N = \{(x, y, z) \mid x^2 + y^2 = 1, 0 \leq z \leq 1\}.$$

Entonces N es una variedad bidimensional con frontera. La frontera de N consiste de dos círculos. Como conjunto N en la definición de variedad con frontera podemos tomar al cilindro infinito

$$\{(x, y, z) \mid x^2 + y^2 = 1\}.$$

N puede ser parametrizada por un sólo dos-cubo dado por

$$\alpha(u, v) = (\cos u, \sin u, v), \quad 0 \leq u \leq 2\pi, \quad 0 \leq v \leq 1.$$

29 INTEGRACION EN ESPACIOS EUCLIDIANOS

Revisaremos en esta sección la teoría de integración que puede encontrarse en cualquier libro de cálculo avanzado (ver, por ejemplo, [Sp], en la bibliografía).

La medida (estandar) de un p -cubo en \mathbb{R}^p dado por

$U = \{(u^1, \dots, u^p) \mid a^i \leq u^i \leq b^i\}$, en donde u^i son coordenadas cartesianas, es $\mu_p U = (b^1 - a^1)(b^2 - a^2) \dots (b^p - a^p)$. Así pues, la función μ_p

asigna un número real a cada uno de estos conjuntos cúbicos de \mathbb{R}^p . La integral de Riemann de una función con valores reales definida sobre U , si existe, está definida por

$$\int_U f d\mu_p = \lim_{\mu_p U_j \rightarrow 0} \sum_{j=1}^N f(x_j) \mu_p U_j,$$

donde U ha sido partido en N p -cubos más pequeños U_j y un punto x_j ha sido escogido en cada U_j . Por la existencia del límite anterior queremos decir que debe ser posible hacer que la suma se aproxime tanto como lo deseemos a un valor límite al hacer que todos los U_j sean suficientemente pequeños, sin importar la elección de los x_j . Se prueba que la integral de Riemann de f existe cuando f es continua.

Esta definición es muy natural desde el punto de vista de la matemática aplicada, en donde se considera que generaliza la situación para una función constante f . Por ejemplo, si la densidad de una sustancia es constante, la masa se obtiene multiplicando el volumen (medida) por la densidad. Para una densidad variable f , que por lo general se supone continua, es muy natural considerar que la masa está aproximadamente dada por la suma de productos $f(x_j) \mu_p U_j$, donde los U_j son pequeños cubos sobre los cuales f tiene prácticamente el valor constante $f(x_j)$, $x_j \in U_j$. Así, la integral de Riemann $\int_U f d\mu_p$ es una definición razonable de la masa presente en el cubo U . En la mayoría de las aplicaciones físicas de la integración se pueden encontrar definiciones de este tipo para muchas cantidades físicas.

Sin embargo, tales límites de sumas son difíciles de evaluar (aún cuando, gracias a las computadoras, se pueden obtener valores aproximados). Por esta razón, las integrales de Riemann se evalúan por medio de la relación que guardan con otros objetos más sencillos como son las integrales simples iteradas. La justificación de este método de evaluación se da en el siguiente teorema.

Teorema 29.1 (Fubini) Si f es una función de valores reales definida sobre U , entonces las integrales definidas

$$f_p(u^1, \dots, u^{p-1}) = \int_{a^p}^{b^p} f(u^1, \dots, u^{p-1}, u^p) du^p$$

son funciones continuas de los parámetros u^1, \dots, u^{p-1} , y la integral de Riemann de f está dada por

$$\int_U f d\mu_p = \int_{U_{p-1}} f_p d\mu_{p-1},$$

donde μ_{p-1} es la medida del $(p-1)$ -cubo $U_{p-1} = \{(u^1, \dots, u^{p-1}) \mid a^i \leq u^i \leq b^i, i = 1, \dots, p-1\}$. Se sigue por iteración que

$$\int_U f d\mu_p = \int_{a^1}^{b^1} \left(\dots \int_{a^{p-1}}^{b^{p-1}} \left(\int_{a^p}^{b^p} f(u^1, \dots, u^p) du^p \right) du^{p-1} \dots \right) du^1.$$

Desde luego, ésto sólo reduce el problema a uno más simple de tipo similar, la evaluación de integrales definidas simples, las cuales se definen también como límites de sumas de Riemann. La evaluación de las integrales definidas simples se hace casi siempre por medio del

teorema fundamental del cálculo elemental, el cual las relaciona con el proceso de encontrar las antiderivadas.

Aunque es conveniente para ciertos propósitos restringir el dominio de las funciones a cubos rectilíneos, en la mayoría de las aplicaciones es difícil encontrar tales dominios de integración. Para definir la integral de una función sobre un dominio acotado más general D , lo que hacemos es incluir a D en un cubo rectilíneo U y tomar

$$\int_D f d\mu_p = \int_U \Phi_D f d\mu_p,$$

donde Φ_D es la función característica de D , definida por

$$\Phi_D x = \begin{cases} 1 & \text{si } x \in D \\ 0 & \text{si } x \notin D. \end{cases}$$

Nuevamente, esta definición no es muy conveniente para propósitos de evaluación, así que generalmente se reducen las integrales sobre D a integrales sobre cubos por medio de una aplicación biyectiva suave entre un cubo y el dominio D . En esta situación se aplica el Teorema de Cambio de Variables para la Integral de Riemann.

Teorema 29.2 Si E y D son regiones en \mathbb{R}^p , $\varphi: E \rightarrow D$ es una aplicación biyectiva suave, e $I = \int_D f d\mu_p$ existe, entonces $\int_E (f \circ \varphi) |J_\varphi| d\mu_p$ existe y es igual a I , donde J_φ es el determinante jacobiano de φ ; si φ está dada por las ecuaciones $u^i = F^i(v^1, \dots, v^p)$, $i = 1, \dots, p$, donde u^i son coordenadas cartesianas sobre D y las v^i son coordenadas cartesianas sobre E , entonces

$$J_\varphi(v^1, \dots, v^p) = \det(\partial F^i(v^1, \dots, v^p) / \partial v^j).$$

Nótese que si φ preserva la orientación en puntos en donde $d\varphi$ es invertible, entonces $J_\varphi \geq 0$ y podemos omitir el símbolo de valor absoluto. Nótese también que en los puntos en los que $d\varphi$ es singular, $J_\varphi = 0$, así que el teorema puede fortalecerse un poco requiriendo solamente que la aplicación sea biyectiva sobre el conjunto regular de φ , es decir, donde $d\varphi$ es invertible (no singular).

Ilustraremos este teorema de cambio de variables en el caso $p = 2$ demostrando como puede ser usado para obtener una forma útil del teorema de Fubini que aparece con frecuencia. Supongamos que

$$D = \{(x, y) \mid a \leq x \leq b, \text{ y para cada } x, h(x) \leq y \leq k(x)\},$$

donde h y k son funciones suaves tales que $h(x) \leq k(x)$ para $a \leq x \leq b$. Transformaremos el rectángulo $E = \{(u, v) \mid a \leq u \leq b, 0 \leq v \leq 1\}$ en la región D por medio de la aplicación suave $\varphi: E \rightarrow D$ dada por las ecuaciones $x = u$, $y = vk(u) + (1-v)h(u)$. Entonces, $J_\varphi = 1(k(u) - h(u)) - 0(\dots) = k(u) - h(u) \geq 0$. Por el teorema de cambio de variables seguido del teorema de Fubini, tenemos

$$\begin{aligned} \int_D f(x, y) d\mu_2 &= \int_E f(u, vk(u) + (1-v)h(u)) (k(u) - h(u)) d\mu_2 \\ &= \int_a^b \left(\int_0^1 f(u, vk(u) + (1-v)h(u)) (k(u) - h(u)) dv \right) du. \end{aligned}$$

Ahora, podemos aplicar el teorema de cambio de variable a la integral del paréntesis interior. Manteniendo fija la variable u y haciendo $y = vk(u) + (1-v)h(u)$ tenemos $dy = (k(u) - h(u))dv$, $y = h(u)$ cuando $v = 0$, y $y = k(u)$ cuando $v = 1$, así que la integral interior se convierte en $\int_{h(u)}^{k(u)} f(u, y) dy$. Ahora, cambiando la variable muda x por la variable u , obtenemos la otra forma del teorema de Fubini que estábamos buscando:

$$\int_D f(x, y) d\mu_2 = \int_a^b \left(\int_{h(x)}^{k(x)} f(x, y) dy \right) dx.$$

30 INTEGRACION DE P-FORMAS

Cuando consideramos la integración de p -formas sobre p -cadenas, aparece la nueva cuestión de la orientación. Esto aparece en forma natural en las aplicaciones. Por ejemplo, el trabajo mecánico que se realiza al recorrer una curva bajo la influencia de un campo de fuerza depende, en signo, de la dirección a lo largo de la cual se recorra la curva. Sería natural, desde el punto de vista físico, formular la definición usando límites de sumas. Sin embargo, la dificultad para manejar tales límites de sumas, se ve aumentada en este caso por la necesidad de integrar sobre objetos curvos. Daremos la definición en términos de las integrales de Riemann, para las cuales, como vimos en la sección anterior, el problema de evaluación se resuelve por medio del teorema de Fubini.

Sea θ una p -forma definida sobre una región de una variedad M que contiene a la imagen de un p -cubo orientado (α, ω) , donde $\alpha: U \rightarrow M$. Entonces usaremos la retracción de θ a U por medio de α para obtener una p -forma $\alpha^*\theta$ definida sobre U . En el corolario 20.3 (pag. 51), dimos una expresión para $\alpha^*\theta$ la cual se puede obtener fácilmente al sustituir las fórmulas coordenadas para α en la expresión coordenada para θ . Puesto que hemos definido la orientación ω por medio de la p -forma $\sum du \dots du$, ω es por sí sola una base para las p -formas sobre \mathbb{R}^p . Así, podemos escribir $\alpha^*\theta = f\omega$, donde $f \in \mathcal{F}(U)$. Si definimos un producto interno $\langle \cdot, \cdot \rangle_p$ para el espacio de las p -formas sobre \mathbb{R}^p con la propiedad de que ω sea unitaria, es decir, $\langle \omega, \omega \rangle_p = 1$, entonces $f = \langle \alpha^*\theta, \omega \rangle_p$. La definición de la integral de la p -forma θ sobre el p -cubo (α, ω) es entonces

$$\int_{(\alpha, \omega)} \theta = \int_U \langle \alpha^*\theta, \omega \rangle_p d\mu_p.$$

La integral de una cero-forma θ sobre un cero-cubo $m \in M$ se define como el valor θ_m de θ en m . La integral de una p -forma sobre una p -cadena $\sum r_i C_i$ se define de la manera más obvia en términos de las integrales sobre p -cubos:

$$\int_{\sum r_i C_i} \theta = \sum r_i \int_{C_i} \theta.$$

Ejemplos.

(a) El círculo $S^1 = \{(x,y) | x^2 + y^2 = 1\}$ en \mathbb{R}^2 , con la orientación en sentido contrario a la agujas del reloj, está parametrizado por (α, du) , en donde α está definido sobre $[0, 2\pi]$ por $\alpha(u) = (\cos u, \sin u)$. Las ecuaciones coordenadas para α son $x = \cos u$, $y = \sin u$. Si tenemos una dos-forma $\theta = (xdy - ydx)/(x^2 + y^2)$ entonces

$$\begin{aligned}\alpha^*\theta &= (\cos u \, d(\sin u) - \sin u \, d(\cos u))/(\cos^2 u + \sin^2 u) \\ &= (\cos^2 u \, du + \sin^2 u \, du)/(\cos^2 u + \sin^2 u) = du.\end{aligned}$$

Ahora tenemos que $\langle du, du \rangle = 1$, de donde

$$\int_{(\alpha, du)} \theta = \int_{[0, 2\pi]} 1 \, d\mu_1 = \int_0^{2\pi} du = 2\pi.$$

(b) La esfera S^2 está parametrizada por el dos-cubo α definido sobre $U = [0, \pi] \times [0, 2\pi]$ y dado por las ecuaciones

$(x, y, z) = (\cos u, \sin u \cos v, \sin u \sin v) = \alpha(u, v)$. Definimos la orientación positiva de S^2 como aquella para la cual las coordenadas y, z , restringidas a S^2 , forman un sistema consistentemente orientado en una vecindad del polo norte $(1, 0, 0)$. Esta definición está de acuerdo con la convención de la normal dirigida hacia afuera, en el sentido de que ∂_x apunta hacia afuera de la bola de \mathbb{R}^3 delimitada por S^2 y las coordenadas x, y, z definen la que se considera la orientación positiva de \mathbb{R}^3 . Así pues, $(\alpha, dudv)$ es el dos-cubo orientado que parametriza a S^2 con esta orientación positiva. Sea $\tau = (xdydz + ydzdx + zdxdy)/r^3$ una dos-forma definida sobre $\mathbb{R}^3 - \{0\}$, donde $r^3 = (x^2 + y^2 + z^2)^{3/2}$. Calculemos ahora $\alpha^*\tau$:

$$\alpha^* dx = -\sin u \, du,$$

$$\alpha^* dy = \cos u \cos v \, du - \sin u \, \sin v \, dv,$$

$$\alpha^* dz = \cos u \sin v \, du + \sin u \cos v \, dv,$$

$$\alpha^* r^3 = 1,$$

$$\alpha^*(dxdy) = \alpha^* dx \wedge \alpha^* dy = \sin^2 u \, \sin v \, dudv,$$

$$\begin{aligned}\alpha^*(dydz) &= \cos u \, \sin u \, \cos^2 v \, dudv - \sin u \, \cos u \, \sin^2 v \, dvdu = \\ &= \cos u \, \sin u \, dudv,\end{aligned}$$

$$\alpha^*(dzdx) = \sin^2 u \, \cos v \, dudv,$$

$$\begin{aligned}\alpha^*\tau &= (\cos^2 u \, \sin u + \sin^3 u \, \cos^2 v + \sin^3 u \, \sin^2 v) \, dudv = \\ &= \sin u \, dudv.\end{aligned}$$

La integral de superficie de τ sobre $(\alpha, dudv)$ que parametriza a S^2 se puede evaluar ahora fácilmente:

$$\int_{(\alpha, dudv)} \tau = \int_0^\pi \int_0^{2\pi} \sin u \, dv \, du = 4\pi.$$

Para asegurar que las integrales que hemos definido tienen sentido geométrico es necesario establecer dos resultados acerca de la independen

dencia de la parametrización. El primero es que las integrales de una p-forma θ sobre p-cubos equivalentes coinciden. El segundo es que la integral de θ sobre la parametrización de un "subconjunto orientado" (y en particular, sobre una subvariedad orientada con frontera) es independiente de la parametrización. El primer resultado nos permite ignorar la distinción, en esta teoría de integración, entre p-cadenas equivalentes. El segundo nos permite definir la integral de una p-forma sobre un subconjunto orientado de una variedad.

Teorema 30.1 Si (α, ω) y $(\beta, \delta\omega)$, $\delta = \pm 1$, son p-cubos equivalentes, entonces para cualquier p-forma θ definida sobre la imagen de α y β ,

$$\int_{(\alpha, \omega)} \theta = \int_{(\beta, \delta\omega)} \theta.$$

Demostración. Puesto que (α, ω) y $(\beta, \delta\omega)$ son equivalentes, existe un difeomorfismo $\varphi: U \rightarrow V$ tal que $\beta \circ \varphi = \alpha$, donde $\alpha: U \rightarrow M$ y $\beta: V \rightarrow M$. Se sigue inmediatamente del lema 20.1 (pag. 50) que $\varphi^* \circ \beta^* = \alpha^*$. Así tenemos que si $\alpha^* \theta = f\omega$ y $\beta^* \theta = g\delta\omega$, entonces $f\omega = \varphi^*(\beta^* \theta) = \varphi^*(g\delta\omega) = (g \circ \varphi) \varphi^*(\delta\omega)$. Sin embargo, φ debe llevar las coordenadas sobre V que están consistentemente orientadas con $\delta\omega$ en las coordenadas sobre U que están consistentemente orientadas con ω . Esto significa que el signo del jacobiano de φ es el mismo que el de δ .

Supongamos que las ecuaciones de φ son $v^i = F^i(u^1, \dots, u^p)$, $i = 1, \dots, p$, donde las u^i son coordenadas cartesianas sobre U y v^i son coordenadas cartesianas sobre V . Entonces si, por ejemplo, $\omega = dv^1 \dots dv^p$ (igualdad de p-formas), tenemos (ver corolario 20.3, pag. 51),

$$\varphi^*(\delta\omega) = \delta \varphi^*(dv^1 \dots dv^p) = \delta \det(\partial F^j / \partial u^i) du^1 \dots du^p = \delta J_\varphi \omega.$$

Obtenemos el mismo resultado si $\omega = -du^1 \dots du^p$, puesto que en este caso sólo tenemos un signo menos en todas las cantidades intermedias. Así, $f\omega = (g \circ \varphi) \varphi^*(\delta\omega) = (g \circ \varphi) |J_\varphi|$, ya que los signos de J_φ y δ son iguales. Por lo tanto, $f = (g \circ \varphi) |J_\varphi|$. Ahora, por el teorema de cambio de variables tenemos,

$$\int_{(\beta, \delta\omega)} \theta = \int_V g d\mu_p = \int_U (g \circ \varphi) |J_\varphi| d\mu_p = \int_U f d\mu_p = \int_{(\alpha, \omega)} \theta. \quad \#\#$$

Corolario 30.2 Si C y D son p-cadenas equivalentes, entonces

$$\int_C \theta = \int_D \theta.$$

Demostración. Además de la igualdad obvia de las sumas de integrales sobre p-cubos equivalentes, necesitamos una trivialidad adicional:

$$\int_{(\alpha, -\omega)} \theta = - \int_{(\alpha, \omega)} \theta. \quad \#\#$$

Teorema 30.3 Si las p-cadenas C y D parametrizan ambas la misma región S de una subvariedad orientada p-dimensional N y θ es una p-forma

definida sobre S , entonces $\int_C \theta = \int_D \theta$.

Demostración. De acuerdo con el corolario anterior podemos reemplazar a C y a D por p -cadenas equivalentes, así que podemos suponer que son irreducibles, digamos, $C = \sum_{i=1}^h (\alpha_i, \omega_i)$ y $D = \sum_{j=1}^k (\beta_j, \omega_j)$, donde α_i está definida sobre U_i y β_j está definida sobre V_j . Sean U_i^r y V_j^r los correspondientes conjuntos de puntos regulares. Es importante para la demostración recordar el hecho de que la integración sobre S puede hacerse integrando sobre el subconjunto que es la parte común de las imágenes de los α_i y β_j sobre sus conjuntos regulares, es decir, sobre

$$\left(\bigcup_i \alpha_i U_i^r \right) \cap \left(\bigcup_j \beta_j V_j^r \right) = \bigcup_{i,j} (\alpha_i U_i^r \cap \beta_j V_j^r).$$

La razón por la que podemos ignorar la parte restante de S es que los puntos no regulares corresponden a conjuntos "de medida cero" en S , es decir, los puntos no regulares no contribuyen a la integral sobre S ; los puntos frontera de U_i y V_j forman un conjunto de dimensión menor y pueden ser incluidos en cubos de medida arbitrariamente pequeña. En una vecindad de un punto singular de α_i , α_i es aproximada por su diferencial $d\alpha_i$ la cual aplica el espacio tangente en un subespacio de dimensión menor. La misma situación se tiene para β_j . Haciendo uso de estos hechos, la demostración se reduce al siguiente cálculo:

$$\begin{aligned} \int_C \theta &= \sum_i \int (\alpha_i, \omega_i) \theta \\ &= \sum_i \int_{U_i^r} \langle \alpha_i^* \theta, \omega_i \rangle_p d\mu_p \\ &= \sum_{i,j} \int_{(\alpha_i^r \cap \beta_j^r) \cap U_i^r} \langle \alpha_i^* \theta, \omega_i \rangle_p d\mu_p \\ &= \sum_{i,j} \int_{U_i^r \cap \beta_j^{-1} \alpha_i U_i^r} \langle \beta_j^* \alpha_i^{-1} \alpha_i^* \theta, \omega_j \rangle_p d\mu_p \\ &= \sum_j \int_{V_j^r} \langle \beta_j^* \theta, \omega_j \rangle_p d\mu_p \\ &= \int_D \theta. \end{aligned}$$

La cuarta igualdad en esta serie usa el teorema de cambio de variables, con el determinante jacobiano dado por la acción de $\beta_j^* \alpha_i^{-1}$ en lugar de φ^* como en la prueba del teorema 30.1. Los conjuntos $\alpha_i U_i^r \cap \beta_j V_j^r$ son todos disjuntos, $i = 1, \dots, h$, $j = 1, \dots, k$ y las inversas α_i^{-1} y β_j^{-1} están definidas sobre ellos debido a las hipótesis que hicimos en la definición de parametrización; es decir, que α_i es biyectiva sobre U_i^r y que $\alpha_i U_i^r$ son disjuntos, y similarmente para β_j , V_j^r . Así, no hay peligro de duplicación en las sumas del cálculo anterior. ##

El teorema 30.3 nos permite dar una definición de la integral de una p -forma sobre un "subconjunto parametrizable orientado" S , la clase de subconjunto que puede ser parametrizado por una p -cadena. Es evidente que tal subconjunto debe ser una subvariedad orientada p -dimensional "casi en todas partes", es decir, excepto sobre un subconjunto de medida cero, pero a diferencia de una subvariedad p -dimensional con frontera, puede tener picos interiores, aristas, etc. La definición es obvia, $\int_S \theta = \int_C \theta$, donde C es cualquier cadena que parametriza a S .

Proposición 30.4 (a) Supóngase que θ es una p -forma sobre M tal que para todo p -cubo C sobre M , $\int_C \theta = 0$. Entonces θ es idénticamente cero.
 (b) Si θ y τ son p -formas sobre M tales que para todo p -cubo C en M , $\int_C \theta = \int_C \tau$, entonces $\theta = \tau$.

Demostración. Es claro que (a) implica (b). Para demostrar (a), tómese un sistema coordenado y contruyamos pequeños p -cubos sobre los hiperplanos coordenados. ##

31 TEOREMA DE STOKES

Existe una generalización del teorema fundamental del cálculo para las integrales de p -formas exactas sobre p -cadenas en variedades. Esta generalización unifica una serie de teoremas que incluye, además de el teorema fundamental del cálculo, a los teoremas de la divergencia en \mathbb{R}^2 y en \mathbb{R}^3 , y al teorema de Stokes para integrales de superficie. Puesto que este último se parece más a la generalización que daremos, la generalización se llama teorema general de Stokes, o simplemente teorema de Stokes.

Antes de considerar el teorema de Stokes, necesitamos un resultado importante que relaciona la retracción con respecto a una aplicación suave con el operador de derivada exterior.

Lema 31.1 Si $\varphi: M \rightarrow N$ es una aplicación suave y θ es una p -forma sobre N , entonces $d\varphi^*\theta = \varphi^*d\theta$.

Demostración. Puesto que ambos operadores, d y φ^* , son locales, es decir, el valor en $m \in M$ de $\varphi^*\theta$ depende sólo de los valores de θ en una vecindad de φm , y similarmente para d , podemos considerar una vecindad coordenada. De hecho, ambos operadores son lineales, así que es suficiente demostrar el resultado en el caso en que θ es un monomio, digamos, $\theta = f dx^1 \dots dx^p$, donde x^1, \dots, x^p son coordenadas sobre N . Entonces $d\theta = df \wedge dx^1 \dots dx^p$ y $\varphi^*d\theta = \varphi^*df \wedge \varphi^*dx^1 \wedge \dots \wedge \varphi^*dx^p$, (ver lema 20.1, pag. 50). Pero para cualquier función g sobre N , tenemos $\varphi^*dg = d\varphi^*g$ (ver el comentario que hicimos antes del lema 20.1), así que $\varphi^*d\theta = d\varphi^*f \wedge d\varphi^*x^1 \wedge \dots \wedge d\varphi^*x^p$. Por otra parte, tenemos

$$\begin{aligned} \varphi^*\theta &= \varphi^*f \varphi^*dx^1 \wedge \dots \wedge \varphi^*dx^p \\ &= \varphi^*f d\varphi^*x^1 \wedge \dots \wedge d\varphi^*x^p. \end{aligned}$$

Por lo tanto, usando el hecho de que d es una derivación y que $d^2 = 0$, obtenemos

$$d\varphi^*\theta = d\varphi^*f \wedge d\varphi^*x^1 \wedge \dots \wedge d\varphi^*x^p = \varphi^*d\theta. \quad \#\#$$

Teorema 31.2 (Stokes) Sea θ una $(p-1)$ -forma definida sobre las imágenes de todos los cubos de una p -cadena C , donde $p > 0$. Entonces,

$$\int_C d\theta = \int_{\partial C} \theta.$$

Demostración. Por linealidad, es suficiente demostrar el teorema para un sólo p-cubo $C = \langle \alpha, \omega \rangle$. Sea

$$\alpha^* \theta = \sum_i (-1)^{i-1} f_i du^1 \dots du^{i-1} du^{i+1} \dots du^p.$$

Entonces, $d\alpha^* \theta = \alpha^* d\theta = (\sum_i d f_i) du^1 \dots du^p$. Podemos suponer que $\omega = du^1 \dots du^p$, ya que en el caso de tener la negativa de esta orientación, se tiene un cambio de signo en ambos miembros. Sea $U = \{ (u^1, \dots, u^p) \mid b^i \leq u^i \leq b^i + c^i \}$ el dominio de α y sea U_i el dominio de $\omega_i \epsilon$. Si damos a la i -ésima coordenada de U_i es valor $b^i + \epsilon c^i$, obtenemos una cara de U y $\omega_i \epsilon$ es esencialmente la restricción de α a esta cara. La orientación inducida es

$$\omega_i \epsilon = (2\epsilon - 1)(-1)^{i-1} du^1 \dots du^{i-1} du^{i+1} \dots du^p.$$

Extenderemos el producto interno $\langle \cdot, \cdot \rangle_{p-1}$ a $(p-1)$ -formas sobre \mathbb{R}^p mediante la condición de que ω_i sea ortonormal. Entonces se sigue que $\langle \alpha_i^* \theta, \omega_i \epsilon \rangle_{p-1} = \langle \alpha^* \theta, \omega_i \epsilon \rangle_{p-1} = (2\epsilon - 1) f_i$, y por lo tanto,

$$\begin{aligned} \int_{\partial(\alpha, \omega)} \theta &= \sum_{i \in E} \int_{\langle \alpha_i^* \theta, \omega_i \epsilon \rangle} \theta \\ &= \sum_{i \in E} \int_{U_i} (2-1) f_i |_{u^i = b^i + \epsilon c^i} d\mu_{p-1} \\ &= \sum_i \int_{U_i} f_i |_{u^i = b^i + \epsilon c^i} d\mu_{p-1}. \end{aligned}$$

Por otra parte, $\int_U \langle \alpha^* d\theta, \omega \rangle_p d\mu_p = \sum_i \int_U d f_i d\mu_p$. Aplicaremos el primer paso del teorema de Fubini a $\int_U d f_i d\mu_p$ con la i -ésima variable como variable de integración, y obtenemos

$$\begin{aligned} \int_U d f_i d\mu_p &= \int_{U_i} \int_{b^i}^{b^i + c^i} d f_i(u^1, \dots, u^i, \dots, u^p) du^i d\mu_{p-1} \\ &= \int_{U_i} (f_i(u^1, \dots, b^i + c^i, \dots, u^p) - \\ &\quad - f_i(u^1, \dots, b^i, \dots, u^p)) d\mu_{p-1}, \end{aligned}$$

en donde el segundo paso se sigue del teorema fundamental del cálculo. Ahora los términos que hemos obtenido coinciden con los de $\int_{\partial(\alpha, \omega)} \theta$. ##

El siguiente corolario es una generalización de las fórmulas clásicas de Green en \mathbb{R}^2 y \mathbb{R}^3 .

Corolario 31.3 (Integración por partes) Con las mismas hipótesis, si f es una función con valores reales definida sobre el dominio de θ ,

$$\int_C df \wedge \theta = \int_{\partial C} f \theta - \int_C f d\theta.$$

Más en general, si θ es una p -forma, τ es una q -forma, y C es una $(p+q+1)$ -cadena, entonces

$$\int_C d\theta \wedge \tau = \int_{\partial C} \theta \wedge \tau - (-1)^p \int_C \theta \wedge d\tau.$$

Demostración. Se aplica el teorema de Stokes a la fórmula $d(\theta \wedge \tau) = d\theta \wedge \tau + (-1)^p \theta \wedge d\tau$. ##

Corolario 31.4 Si θ es una $(p-1)$ -forma definida sobre variedad M compacta y orientada de dimensión p , entonces $\int_M d\theta = 0$ y $d\theta = 0$ en algún punto.

Demostración. Por el teorema 28.5 (pag. 79), sabemos que existe una p -cadena C en M que parametriza a M y para la cual ∂C es equivalente a la cadena nula O . Así, tenemos

$$\int_M d\theta = \int_C d\theta = \int_{\partial C} \theta = \int_O \theta = 0.$$

Para demostrar la última parte del corolario, podemos suponer que M es conexa, ya que en caso contrario, podemos considerar la restricción de θ a una componente conexa de M . Sea Ω un elemento de volumen para M . Entonces $d\theta = f\Omega$, donde $f \in \mathcal{F}(M)$. Para cualquier p -cubo (α_i, ω_i) de C , $\langle \alpha_i^* d\theta, \omega_i \rangle_p = f \circ \alpha_i \langle \alpha_i^* \Omega, \omega_i \rangle_p$ tiene el mismo signo en un punto regular x de α_i que $f(\alpha_i(x))$. Como

$$\sum_i \int_{(\alpha_i, \omega_i)} \langle \alpha_i^* d\theta, \omega_i \rangle_p d\mu_p = 0,$$

vemos que o f es idénticamente 0 sobre M , o bien f no tiene el mismo signo sobre M . En el último caso, la continuidad de f nos dice que f debe ser cero en algún punto de M . ##

32 TEOREMA DE FROBENIUS

En esta sección, M será una variedad suave de dimensión n . Una distribución k -dimensional sobre un subconjunto A de M es una función P que asigna a cada punto $p \in A$ un subespacio k -dimensional P_p del espacio tangente $T_p(M)$. Decimos que P es suave en A si A es abierto y para cada $p \in A$ existen k campos vectoriales suaves y linealmente independientes X_1, \dots, X_k que generan a P_m para cada m en una vecindad de p . Un campo vectorial X con dominio B está sobre P si $B \subset A$ y $X|_p \in P_p$ para cada $p \in B$. Una distribución suave P se llama involutiva (integrable o cerrada) cuando es cerrada con respecto a la operación del paréntesis de Lie, es decir, si X e Y son campos suaves cualesquiera con dominio común contenido en el dominio de P , entonces $[X, Y]$ está sobre P . Una subvariedad V de M es una subvariedad integral de P si V está contenida en el dominio de P y $T_p(V) = P_p$, para cada p de V ; así pues, el subespacio del espacio tangente $T_p(M)$ que es $T_p(V)$ coincide exactamente con el subespacio P_p .

El teorema principal de esta sección afirma que una distribución suave tiene subvariedades integrales si y sólo si es involutiva. Veamos ante todo un poco más de terminología: si x^1, \dots, x^n es un sistema de coordenadas para M con dominio en U , definimos una sección de U como un subconjunto de U en el cual r de las funciones coordenadas x^1, \dots, x^n son constantes, con $1 \leq r \leq n$. Evidentemente cada sección de U es una subvariedad de U (o de M).

Teorema 32.1 Sea P una distribución involutiva suave y de dimensión k definida sobre M . Para cada punto $m \in M$, existe un sistema coordenado x^1, \dots, x^n con dominio U , que contiene a m , tal que los campos coordenados $\partial/\partial x^i$ generan a P en cada punto de U , para $j = 1, \dots, k$. En otras palabras, las secciones de U para las cuales x^{k+1}, \dots, x^n son constantes, son subvariedades integrales de P .

Este teorema se demuestra por inducción sobre k . El caso $k = 1$ queda cubierto por lema siguiente, y obsérvese que en este caso, cualquier distribución es automáticamente involutiva.

Lema 32.2 Si X es un campo vectorial suave sobre M , p es un punto de M y $X|_p \neq 0$, entonces existe un sistema de coordenadas y^1, \dots, y^n en un entorno U de p , tal que $X = \partial/\partial y^1$ sobre U .

Demostración del lema 32.2. Nótese que este resultado no es mas que una reformulación del lema 13.6 que demostramos en la página 38. ##

Demostración del teorema 32.1. Tomemos el punto m y sean X_1, \dots, X_k campos vectoriales suaves que generan a P en una vecindad U_1 de m . Aplicando el lema anterior, obtenemos un sistema de coordenadas y^1, \dots, y^n en m con dominio $U_2 \subset U_1$, tal que $\partial/\partial y^i = X_i$ sobre U_2 y supongamos que $y^i(m) = 0$.

Si $k = 1$, entonces el sistema de coordenadas y^1, \dots, y^n satisface la afirmación del teorema. Si $k > 1$, supongamos que el teorema es cierto para distribuciones de dimensión menor que k , y definamos la distribución $(k-1)$ -dimensional \bar{P} sobre U_2 mediante $\bar{P}_p = \{X \in P_p | X y^1 = 0, p \in U_2\}$. Esta es una distribución $(k-1)$ -dimensional y está generada por los $k-1$ campos vectoriales suaves e independientes $Y_i, i = 2, \dots, k$, definidos por $Y_i = X_i - (X_i y^1) X_1$. Es involutiva puesto que si Y y Z están sobre \bar{P} , entonces $[Y, Z]$ está sobre P y $\{Y, Z\} y^1 = Y(Z y^1) - Z(Y y^1) = 0$ sobre U_2 , por lo que $[Y, Z]$ está sobre \bar{P} .

Sea V_0 la sección de U_2 definida por $y^1 = 0$. Entonces, para todo $p \in V_0$, $\bar{P}_p \subset T_p(V_0)$ por lo que, aplicando la hipótesis de inducción a la distribución \bar{P} sobre la variedad V_0 , obtenemos un sistema de coordenadas z^1, \dots, z^n en una vecindad $U_3 \subset V_0$ del punto m , tal que $\partial/\partial z^1, \dots, \partial/\partial z^k$ generan a \bar{P} en U_3 . Definamos ahora un difeomorfismo $\Pi: U_2 \rightarrow V_0$ por $\Pi(p) = \phi^{-1}(0, y^2(p), \dots, y^n(p))$, donde ϕ es la aplicación coordenada cuyas componentes son las y^i . Sea $U_4 = \Pi^{-1}(U_3)$ y definamos las funciones x^1, \dots, x^n sobre U_4 mediante $x^1 = y^1, x^2 = z^1 \circ \Pi, \dots, x^n = z^n \circ \Pi$. Entonces, las funciones x^1, \dots, x^n forman un sistema coordenado en una vecindad $U \subset U_4$ del punto m , verificándose que $\partial/\partial x^i|_m = \partial/\partial y^i|_m$, mientras que $\partial/\partial x^2, \dots, \partial/\partial x^k$ en el punto m generan a $T_m(V_0)$.

Probemos ahora que $\partial/\partial x^1, \dots, \partial/\partial x^k$ generan a P en U , sabiendo que generan el mismo subespacio que los campos vectoriales X_1, Y_1, \dots, Y_k . Sea $Y_j = X_j$ y demostremos que $Y_i x^j = 0$ para $i = 1, \dots, k, j = k+1, \dots, n$. Puesto que $Y_i = X_i = \partial/\partial x^i$, vemos inmediatamente que $Y_i x^j = 0$ si $j \neq i$.

Como P es involutiva, deben existir ciertas funciones $g_{irs} \in \mathcal{F}(U)$, tales que para $i \leq k$ y $r \leq k$, se cumple que $[Y_i, Y_r] = \sum_{s=1}^k g_{irs} Y_s$. Así pues, para $i = 2, \dots, k$ y $j > k$, $Y_i(Y_j x^j) = [Y_i, Y_j] x^j = \sum_{s=1}^k g_{ijs} Y_s x^j$. Esto implica que las funciones $Y_i x^j$ satisfacen un sistema lineal y homogéneo de ecuaciones diferenciales ordinarias. Pero en V_0 , $x^j = z^j$ para $j > 1$ y $Y_i x^j = Y_i z^j = 0$ sobre V_0 para $j > k$, debido a la construcción de las coordenadas z^1, \dots, z^n . Por lo tanto, por la unicidad de las soluciones de los sistemas de ecuaciones del tipo anterior, tenemos que $Y_i x^j = 0$ para $i \leq k$ y $j > k$. ##

Ahora aplicaremos este teorema para demostrar el clásico teorema de Frobenius acerca de las ecuaciones diferenciales en derivadas parciales (totales). Este teorema podría enunciarse como sigue: existen unas únicas funciones solución $f_i(x^1, \dots, x^k)$ con valores prefijados en un punto, para el sistema de ecuaciones en derivadas parciales

$$\frac{\partial f_i}{\partial x^j} = A_{ij}(x^1, \dots, x^k, f_1, \dots, f_d)$$

si y sólo si para todo $j \leq k$, $r > k$ e $i > d$, se tiene

$$\frac{\partial A_{ij}}{\partial x^r} + \sum_{s=1}^d \frac{\partial A_{is}}{\partial f_s} A_{sr} = \frac{\partial A_{ir}}{\partial x^j} + \sum_{s=1}^d \frac{\partial A_{is}}{\partial f_s} A_{sj}.$$

(Esta condición se llama condición de integrabilidad y se sigue inmediatamente de la regla de la cadena y de la identidad

$$\frac{\partial^2 f_i}{\partial x^r \partial x^j} = \frac{\partial^2 f_i}{\partial x^j \partial x^r}.)$$

Teorema 32.3 (Frobenius) Para $1 \leq i \leq d$ y $1 \leq j \leq k$, sean $A_{ij}(x^1, \dots, x^k, u^1, \dots, u^d)$ funciones suaves de valores reales definidas en un conjunto abierto Q de \mathbb{R}^n , con $n = k + d$. Sea $(a; b) = (a^1, \dots, a^k, b^1, \dots, b^d)$ un punto de Q . Entonces existe un único sistema de funciones suaves con valores reales y definidas en una vecindad V de a , que satisfacen las siguientes condiciones:

- (1) $f_i(a) = b^i$ o bien $f(a) = b$, donde $f: V \subset \mathbb{R}^n \rightarrow \mathbb{R}^d$ está dada por $f(p) = (f_1(p), \dots, f_d(p))$;
- (2) Si $p \in V$, entonces $(p; f(p)) \in Q$;
- (3) Si $p \in V$, entonces $\frac{\partial f_i}{\partial x^j}(p) = A_{ij}(p; f(p))$;

si y sólo si en cada punto de Q se satisface la condición de integrabilidad

$$(4) \quad \frac{\partial A_{ij}}{\partial x^r} + \sum_{s=1}^d \frac{\partial A_{is}}{\partial u^s} A_{sr} = \frac{\partial A_{ir}}{\partial x^j} + \sum_{s=1}^d \frac{\partial A_{is}}{\partial u^s} A_{sj}.$$

Demostración. Sean e_1, \dots, e_n los campos vectoriales suaves que son ortonormales en cada punto de \mathbb{R}^n . Utilizaremos las funciones A_{ij} para definir unos campos vectoriales suaves Y_1, \dots, Y_k sobre Q , dados por $Y_r = e_r + \sum_{s=1}^d A_{sr} e_{k+s}$. Estos campos vectoriales son linealmente inde-

pendientes en cada punto de Q y por ello generarán una distribución suave P de dimensión k sobre Q . Calculemos los paréntesis de Lie

$$\begin{aligned} [Y_r, Y_q] &= [e_r + \sum_{s=1}^d A_{sr} e_{k+s}, e_q + \sum_{t=1}^d A_{tq} e_{k+t}] = \\ &= \sum_{t=1}^d \left(\frac{\partial A_{tq}}{\partial x^r} + \sum_{s=1}^d A_{sr} \frac{\partial A_{tq}}{\partial u^s} \right) e_{k+t} - \sum_{s=1}^d \left(\frac{\partial A_{sr}}{\partial x^q} + \sum_{t=1}^d A_{tq} \frac{\partial A_{sr}}{\partial u^t} \right) e_{k+s}, \end{aligned}$$

y por la condición (4), $[Y_r, Y_q] = 0$.

Entonces, la distribución P es involutiva y por el teorema 32.1, existe una subvariedad integral U de P que pasa por $(a; b)$ y tal que $U \subset Q$. Sea $\phi: U \rightarrow \mathbb{R}^k$ definida por $\phi(a'; b') = a'$, entonces $d\phi(Y_r) = e_r$ y $d\phi$ es no singular en el espacio tangente a U en $(a; b)$. Así pues, existen una vecindad V de a y un difeomorfismo $F: V \rightarrow F(V) \subset U$, tal que $F \circ \phi$ y $\phi \circ F$ son las aplicaciones identidad sobre $F(V)$ y V , respectivamente. Definimos f_1, \dots, f_d en V por medio de la ecuación $F(p) = (p; f_1(p), \dots, f_d(p))$. En estas condiciones, las funciones f_1, \dots, f_d son suaves y satisfacen (1) y (2); la condición (3) se sigue del hecho de que $dF(e_r) = Y_r$ para $r \leq k$.

La implicación del teorema en el otro sentido es trivial. ##

Como corolario del teorema de Frobenius obtenemos un resultado muy útil relativo a la existencia de sistemas de coordenadas y que es una generalización del lema 32.2.

Corolario 32.4 Sea M una variedad de dimensión n y sea X_1, \dots, X_n un sistema de campos vectoriales suaves sobre una vecindad U de $m \in M$. Entonces, existe un sistema de coordenadas x^1, \dots, x^n para una vecindad V de m con $V \subset U$, tales que $X_i = \partial/\partial x^i$ sobre V para toda i , si y sólo si $[X_i, X_j] = 0$ para toda $i, j = 1, \dots, n$.

33 DUAL DEL TEOREMA DE FROBENIUS

Muchos sistemas de ecuaciones diferenciales parciales tienen una interpretación geométrica en términos de formas diferenciales. La razón principal de esto es muy simple. Por ejemplo, sean x, y, z, p, q coordenadas para \mathbb{R}^5 y sea S una subvariedad bidimensional para la cual x y y pueden ser usadas como coordenadas. Entonces, la condición de que $p = \partial z / \partial x$ y $q = \partial z / \partial y$ sobre S , es equivalente a que la forma diferencial $dz - p dx - q dy$ sea cero sobre S .

Una ecuación diferencial parcial de primer orden está dada por una ecuación $F(x, y, z, p, q) = 0$. Esto especifica una hipersuperficie N de \mathbb{R}^5 . Una solución $z = f(x, y)$ de la ecuación diferencial parcial, determina una subvariedad bidimensional S de N , sobre la cual, $z = f(x, y)$, $p = (\partial_x f)(x, y)$ y $q = (\partial_y f)(x, y)$. Esta superficie S es una subvariedad integral de la distribución tridimensional D sobre N , dada por la ecuación $\theta = dz - p dx - q dy = 0$. Más formalmente, D está especificada por $D_n = \{t \in T_n(N) \mid \theta(t) = 0\}$. Desde luego, hay varias dificultades que nos

obligan a ser más cuidadosos: Los puntos donde $F = 0$ y $dF = 0$ simultáneamente deben ser eliminados, ya que N no siempre es una variedad en la vecindad de tales puntos; si hay un punto $n \in N$ donde dF sea proporcional a θ , entonces el subespacio D_n coincide con todo el espacio $T_n(N)$ y D no es tridimensional en todo punto; finalmente, la superficie solución S , debe ser tal que x e y puedan ser tomadas como coordenadas sobre S . No siempre ocurre que la distribución D sea involutiva.

En esta sección, tomando como base el ejemplo anterior, construiremos los objetos duales a las distribuciones (aquí, "dualidad" tiene el mismo sentido que la dualidad entre campos vectoriales y uno-formas). En particular, daremos una formulación dual del teorema 32.1 por medio del cual obtuvimos en la sección anterior el teorema clásico de Frobenius.

Una codistribución k-dimensional Δ sobre una variedad M es una función que asigna a cada $m \in UCM$ un subespacio k-dimensional Δ_m del espacio cotangente $T_m^*(M)$. Δ es suave si su dominio U es abierto y para cada $m \in U$ hay una vecindad V de m y hay uno-formas $\omega^1, \dots, \omega^k$ definidas sobre V , tales que en cada $n \in V$, el subespacio Δ_n está generado por $\omega^1|_n, \dots, \omega^k|_n$. A cada codistribución k-dimensional Δ le asociaremos una distribución $(n-k)$ -dimensional D dada por $D_m = \{t \in T_m(M) \mid \omega(t) = 0, \text{ para todo } \omega \in \Delta_m\}$, e inversamente, a cada distribución $(n-k)$ -dimensional D le asociaremos una codistribución k-dimensional Δ dada por $\Delta_m = \{\omega \in T_m^*(M) \mid \omega(t) = 0, \text{ para toda } t \in D_m\}$. Es claro que si D está asociada a Δ , entonces Δ está asociada a D . La D y la Δ asociadas de esta manera, se dice que se aniquilan mutuamente.

Una subvariedad N de M es una subvariedad integral de una codistribución Δ si N es una subvariedad integral de la distribución D asociada con Δ . Una codistribución Δ se llama involutiva (integrable o cerrada) si la distribución D asociada con Δ es involutiva.

El teorema 32.1 dice que para una distribución involutiva D existe un sistema coordenado x^i tal que D_m está generado por $\partial_1|_m, \dots, \partial_{h-1}|_m$ y las subvariedades integrales de dimensión $h-n-k$ son las secciones coordenadas $x^j = c^j$, donde c^j es constante para $j=h+1, \dots, n$. Se sigue que la codistribución asociada Δ está generada por dx^{h+1}, \dots, dx^n sobre la vecindad coordenada. Cualquiera otra uno-forma perteneciente a Δ puede expresarse como $\omega = \sum_{j=h+1}^n f_j dx^j$. La derivada exterior $d\omega = \sum_{j=h+1}^n df_j \wedge dx^j$ es una "combinación lineal" de las dx^j cuyos coeficientes son las uno-formas df_j . Si $\omega^{h+1}, \dots, \omega^n$ es otra base local para Δ , entonces $dx^j = \sum_{r=h+1}^n g_r^j \omega^r$, donde (g_r^j) es una matriz no singular de funciones suaves. Entonces, tenemos que $d\omega = \sum_{j=h+1}^n df_j \wedge dx^j = \sum_{j,r=h+1}^n (g_r^j df_j) \wedge \omega^r$, la cual es una combinación lineal de las ω^r . Así pues, una condición necesaria para que Δ sea involutiva es que $d\omega$ sea una combinación lineal de una base local ω^r para toda ω perteneciente a Δ . Que esta condición es también suficiente, es la formulación dual del teorema 32.1, que damos a continuación.

Teorema 33.1 Una codistribución suave Δ es involutiva si y sólo si para toda uno-forma ω perteneciente a Δ , la dos-forma $d\omega$ es localmente una combinación lineal $\sum_j \tau^j \wedge \omega^j$ de una base local de uno-formas ω^j para Δ , donde las τ^j son uno-formas.

Demostración. Ya hemos visto que si Δ es involutiva, entonces $d\omega$ es una combinación lineal de la forma indicada.

Supongamos que, inversamente, $d\omega$ es una tal combinación lineal siempre que $\omega \in \Delta$. En particular, para una base local ω^j de Δ , tenemos que $d\omega^j = \sum_r \tau_r^j \wedge \omega^r$ para algunas uno-formas τ_r^j . Entonces, para cualesquiera campos vectoriales $X, Y \in D$, la distribución asociada a Δ , tenemos que $2d\omega^j(X, Y) = \tau_r^j(X)\omega^r(Y) - \tau_r^j(Y)\omega^r(X) = 0$. Pero, por la fórmula invariante para d , que dimos en la sección 25(C), tenemos que

$$2d\omega^j(X, Y) = X\omega^j(Y) - Y\omega^j(X) - \omega^j([X, Y]) = -\omega^j([X, Y]),$$

ya que las derivadas de $\omega^j(Y) = 0$ y $\omega^j(X) = 0$ son cero. Así pues, $[X, Y]$ es aniquilado por una base de Δ y, por lo tanto, $[X, Y] \in D$. Hemos demostrado que D es involutiva y, por lo tanto, Δ también lo es. ##

+++++

IV. GEOMETRIA SEMI-RIEMANNIANA

34 VARIETADES SEMI-RIEMANNIANAS

La geometría del espacio euclidiano \mathbb{R}^3 deriva sus propiedades, en última instancia, del producto interno natural de \mathbb{R}^3 , del producto punto. Por medio del isomorfismo natural entre \mathbb{R}^3 y $T_p(\mathbb{R}^3)$, se puede proveer a cada espacio tangente con el producto punto. Entonces es posible realizar muy importantes operaciones geométricas como medir la longitud de un vector tangente o medir el ángulo entre dos vectores tangentes.

La teoría de superficies en \mathbb{R}^3 llegó a su forma clásica en el trabajo de Gauss, quien demostró en 1827 que la geometría intrínseca de una superficie S en \mathbb{R}^3 (es decir, la geometría percibida por los habitantes bidimensionales de S), depende sólo del producto punto aplicado a los vectores tangentes a S .

Desde el año de 1854 Riemann vió lo que se necesitaba para generalizar estos dos casos especiales y poder hablar de geometría sobre una variedad arbitraria de dimensión n : se debe definir un producto interno sobre cada espacio tangente de la variedad. Se puede pensar que ésto es como tener un medio para medir distancias infinitesimales sobre la variedad. Intuitivamente, si p y $p + dp$ son puntos muy próximos, la distancia entre ellos es la norma del "vector tangente" dp .

Con la teoría general de la relatividad de Einstein (1915), cobró importancia una generalización adicional de estas ideas, la cual, aunque es de carácter técnico, tiene grandes consecuencias: se debilita la exigencia de un producto interior definido positivo en favor de un producto escalar no degenerado.

Un tensor métrico g sobre una variedad suave M es un campo tensorial de tipo $(0,2)$ y simétrico definido sobre M que tiene índice constante.

En otras palabras, $g \in \mathcal{T}_2^0(M)$ asigna suavemente a cada punto $p \in M$ un producto escalar g_p sobre el espacio tangente $T_p(M)$, y el índice de g es el mismo para cada p .

Una variedad semi-Riemanniana es una variedad diferenciable M provista con un tensor métrico g .

Así, hablando estrictamente, una variedad semi-Riemanniana es una pareja ordenada (M, g) : dos tensores métricos diferentes sobre la misma variedad determinan variedades semi-Riemannianas diferentes. No obstante, por lo general denotaremos una variedad semi-Riemanniana con el nombre de su variedad suave M , N , etc.

El valor constante ν del índice de g_p sobre una variedad semi-Riemanniana M se llama el índice de M : $0 \leq \nu \leq n = \dim M$. Si $\nu = 0$, M es una variedad Riemanniana; en este caso, cada g_p es un producto interno (definido positivo) para cada $T_p(M)$. Si $\nu = 1$ y $n \geq 2$, M se llama variedad de Lorentz.

Usaremos los paréntesis \langle, \rangle como una notación alternativa para g , y escribiremos $g(v, w) = \langle v, w \rangle \in \mathbb{R}$ para vectores tangentes, y $g(V, W) =$

$= \langle v, w \rangle \in \mathcal{F}(M)$ para campos vectoriales.

Si x^1, \dots, x^n es un sistema de coordenadas sobre $U \subset M$, las componentes del tensor métrico g sobre U son

$$g_{ij} = \langle \partial_i, \partial_j \rangle \quad (1 \leq i, j \leq n).$$

Así, para campos vectoriales $V = \sum v^i \partial_i$ y $W = \sum w^j \partial_j$,

$$g(V, W) = \langle v, w \rangle = \sum g_{ij} v^i w^j.$$

Puesto que g es no degenerado, en cada punto p de U la matriz $(g_{ij}(p))$ es invertible, y su matriz inversa se denota por $(g^{ij}(p))$. La fórmula usual para encontrar la inversa de una matriz demuestra que las funciones g^{ij} son suaves sobre U .

Como g es simétrico, tenemos que $g_{ij} = g_{ji}$ y entonces $g^{ij} = g^{ji}$ para $1 \leq i, j \leq n$. Por último, sobre U el tensor métrico se puede escribir como

$$g = \sum g_{ij} dx^i \otimes dx^j.$$

Recordemos que en el capítulo I vimos que para cada $p \in \mathbb{R}^n$ hay un isomorfismo lineal canónico entre \mathbb{R}^n y $T_p(\mathbb{R}^n)$ el cual, en términos de las coordenadas cartesianas, manda v a $v_p = \sum v^i \partial_i$. Así, el producto punto sobre \mathbb{R}^n da lugar a un tensor métrico sobre \mathbb{R}^n dado por

$$\langle v_p, w_p \rangle = v \cdot w = \sum v^i w^i.$$

De ahora en adelante, en cualquier contexto geométrico, \mathbb{R}^n denotará a la variedad Riemanniana que surge de esta manera y que se llama espacio Euclidianos de dimensión n .

Si para un entero ν con $0 \leq \nu \leq n$, cambiamos los signos de los primeros ν términos de la ecuación anterior, obtenemos un tensor métrico

$$\langle v_p, w_p \rangle = -\sum_{i=1}^{\nu} v^i w^i + \sum_{j=\nu+1}^n v^j w^j$$

de índice ν . El espacio semi-Euclidianos \mathbb{R}_ν^n que resulta de esta manera se reduce a \mathbb{R}^n si $\nu = 0$. Para $n \geq 2$, \mathbb{R}_ν^n se llama espacio n -dimensional de Minkowski; cuando $n = 4$, éste es el ejemplo más simple de un espacio-tiempo relativista.

Fijemos la notación

$$\varepsilon_i = \begin{cases} -1 & \text{si } 1 \leq i \leq \nu, \\ +1 & \text{si } \nu+1 \leq i \leq n. \end{cases}$$

Entonces, el tensor métrico de \mathbb{R}_ν^n se puede escribir como

$$g = \sum \varepsilon_i du^i \otimes du^i.$$

El significado geométrico del índice de una variedad semi-Riemanniana se deriva de la siguiente tricotomía.

<u>espacial</u>	si $\langle v, v \rangle > 0$ ó $v = 0$,
<u>nulo</u>	si $\langle v, v \rangle = 0$ y $v \neq 0$,
<u>temporal</u>	si $\langle v, v \rangle < 0$.

El conjunto de todos los vectores nulos en $T_p(\mathbb{R}_\nu^n)$ se llama cono nulo

en $p \in M$. La categoría a la que pertenece un vector tangente dado se llama su carácter causal. Esta terminología proviene de la teoría de la relatividad y, sobre todo en el caso Lorentziano, los vectores nulos se llaman también lumínicos.

Sea $q(v) = \langle v, v \rangle$ para cada vector tangente a M . En cada punto p de M , q es la forma cuadrática asociada con el producto escalar en p ; y, como sabemos, q determina completamente al tensor métrico.

Si $V \in \mathcal{L}(M)$ y $f \in \mathcal{F}(M)$, entonces $q(fV) = f^2 q(V) \in \mathcal{F}(M)$, de modo que q no es un campo tensorial. En la terminología clásica, q se llama el elemento de línea de M , y se denota por ds^2 . En términos de un sistema coordinado,

$$q = ds^2 = \sum g_{ij} dx^i dx^j.$$

Aquí la notación $dx^i dx^j$ representa el producto ordinario de funciones definidas en cada espacio tangente, así que

$$q(V) = \sum g_{ij} dx^i(V) dx^j(V) = \sum g_{ij} v^i v^j.$$

Como en la sección 23, la norma $|v|$ de un vector tangente es $|q(v)|^{1/2} = |\langle v, v \rangle|^{1/2}$, y los vectores unitarios, la ortogonalidad, y la ortonormalidad se definen como antes.

El origen de la notación ds^2 puede ser captado intuitivamente como sigue. Supongamos por simplicidad que M es Riemanniana. Si p y p' son puntos próximos con coordenadas (x^1, \dots, x^n) y $(x^1 + \Delta x^1, \dots, x^n + \Delta x^n)$ con respecto a sistema coordinado, entonces el vector tangente $\Delta p = \sum \Delta x^i \partial_i$ en p apunta aproximadamente hacia p' . Así, podemos esperar que el cuadrado de la distancia Δs de p a p' sea aproximadamente

$$|\Delta p|^2 = \langle \Delta p, \Delta p \rangle = \sum g_{ij} \Delta x^i \Delta x^j,$$

como en la fórmula $ds^2 = \sum g_{ij} dx^i dx^j$.

Dada una manera de obtener variedades nuevas a partir de variedades dadas, generalmente hay una manera correspondiente de obtener nuevas variedades semi-Riemannianas a partir de variedades semi-Riemannianas dadas.

Por ejemplo, supongamos primero que P es una subvariedad de una variedad Riemanniana M . Puesto que cada espacio tangente $T_p(P)$ de P es un subespacio de $T_p(M)$, obtenemos un tensor métrico Riemanniano g_p sobre P simplemente restringiendo la acción del tensor métrico g de M a los subespacios $T_p(P)$. Formalmente, g_p es la retracción $j^*(g)$, donde $j: P \hookrightarrow M$ es la aplicación inclusión. Por ejemplo, la n -esfera de radio $r > 0$ es la subvariedad Riemanniana de \mathbb{R}^{n+1} dada por

$$S^n(r) = \{p \mid |p| = r\}.$$

Sin embargo, cuando el tensor métrico g de M es indefinido, entonces $j^*(g)$ no es necesariamente una métrica sobre P . Es un campo tensorial simétrico suave del tipo $(0,2)$ y por lo tanto será una métrica si y sólo si cada $T_p(P)$ es un subespacio no degenerado de $T_p(M)$ con respec

to a g y el índice de $T_p(P)$ es el mismo para todo p . Así, podemos hacer la siguiente definición.

Sea P una subvariedad de una variedad semi-Riemanniana M . Si la retracción $j^*(g)$ (como antes) es un tensor métrico sobre P , entonces decimos que P es una subvariedad semi-Riemanniana de M . (Si se sabe que P es Riemanniana o de Lorentz, entonces estos términos deben aparecer en lugar del calificativo de "semi-Riemanniana" en la definición anterior.)

Consideremos ahora el caso de las variedades producto.

Lema 34.1 Sean M y N variedades semi-Riemannianas con tensores métricos g_M y g_N . Si π y σ son las proyecciones de $M \times N$ sobre M y N , respectivamente, sea

$$g = \pi^*(g_M) + \sigma^*(g_N).$$

Entonces g es un tensor métrico sobre $M \times N$. Llamaremos a $(M \times N, g)$ la variedad semi-Riemanniana producto de M y N .

Demostración. Por la definición de retracción, tenemos que si $v, w \in T_{(p,q)}(M \times N)$, entonces

$$g(v, w) = g_M(d\pi(v), d\pi(w)) + g_N(d\sigma(v), d\sigma(w)).$$

Así, vemos que g es simétrico. Para demostrar la no degeneración, supongamos que $g(v, w) = 0$ para todo $w \in T_{(p,q)}(M \times N)$. Entonces, en particular, para todo $w \in T_{(p,q)}M$ tenemos que $g_M(d\pi(v), d\pi(w)) = 0$, ya que $d\sigma(w) = 0$. Pero como π y $d\pi$ son suprayectivas, se sigue que los vectores $d\pi(w)$ recorren todo el espacio $T_p(M)$, por lo que $d\pi(v) = 0$ debido a la no degeneración de g_M . En forma similar se demuestra que $d\sigma(v) = 0$, por lo tanto, $v = 0$.

Las bases ortonormales de los espacios $T_p(M)$ y $T_q(N)$ se combinan para dar una base ortonormal del espacio $T_{(p,q)}(M \times N)$, entonces el índice de g tiene el valor constante $\text{ind } M + \text{ind } N$. ##

El mismo procedimiento se puede extender de manera obvia a productos finitos de cualquier número de variedades semi-Riemannianas. Por ejemplo, el espacio semi-Euclídeo \mathbb{R}^n es

$$\underbrace{\mathbb{R}^1 \times \dots \times \mathbb{R}^1}_v \times \underbrace{\mathbb{R}^1 \times \dots \times \mathbb{R}^1}_{n-v} = \mathbb{R}^n \times \mathbb{R}^{n-v},$$

donde, por definición, \mathbb{R}^1 es la recta real con el tensor métrico que es el negativo del producto punto usual sobre \mathbb{R}^1 .

35 ISOMETRIAS

Siempre que en la matemática aparece un nuevo objeto, se debe dar, por lo menos, un correspondiente concepto de "igualdad" entre dichos ob-

jetos. Así, tenemos las biyecciones para los conjuntos, los homeomorfismos para los espacios topológicos, los isomorfismos lineales para los espacios vectoriales y los difeomorfismos para las variedades difeomorfismos. En el caso de las variedades semi-Riemannianas, las isometrías son las que sirven para expresar la noción de equivalencia o igualdad.

Sean M y N variedades semi-Riemannianas con tensores métricos g_M y g_N . Una isometría de M a N es un difeomorfismo $\phi: M \rightarrow N$ que preserva los tensores métricos: $\phi^*(g_N) = g_M$.

Explícitamente, $\langle d\phi(v), d\phi(w) \rangle = \langle v, w \rangle$ para todo $v, w \in T_p(M)$, $p \in M$. Como ϕ es un difeomorfismo, cada aplicación diferencial $d\phi_p$ es un isomorfismo lineal; de modo que la condición de preservación de la métrica implica que cada $d\phi_p$ es una isometría lineal entre los espacios con producto escalar $T_p(M)$ y $T_{\phi(p)}(N)$. La retracción actúa de la manera usual sobre los elementos de línea y, puesto que éstos determinan a los tensores métricos, la preservación de la métrica es equivalente a: $\phi^*(g_N) = g_M$.

Es fácil verificar que

- (1) La aplicación identidad sobre una variedad semi-Riemanniana es una isometría.
- (2) La composición de isometrías es una isometría.
- (3) La aplicación inversa de una isometría es una isometría.

La interpretación de $q = ds^2$ como el cuadrado de la distancia infinitesimal, sugiere que consideremos a una isometría como si fuera un movimiento rígido, en contraste con un difeomorfismo arbitrario que puede deformar a la variedad M en el proceso de aplicarla sobre N .

Una propiedad de una variedad semi-Riemanniana que es preservada por una isometría se llama un invariante isométrico de la variedad. La geometría semi-Riemanniana se puede definir como el estudio de tales invariantes. Si existe una isometría entre M y N , decimos que M y N son isométricas. Las propiedades (1), (2) y (3) que enumeramos arriba, significan que la relación de isometría entre variedades semi-Riemannianas es una relación de equivalencia. Intuitivamente, podemos decir que las variedades isométricas tienen la misma geometría semi-Riemanniana.

Sea V un espacio con producto escalar, es decir, un espacio vectorial real con un producto escalar. Entonces V es una variedad y, tal como en el caso $V = \mathbb{R}^n$, la fórmula $\langle v, w \rangle = \langle v, w \rangle$ define un tensor métrico sobre V , que lo convierte en una variedad semi-Riemanniana.

Lema 35.1 Si $\psi: V \rightarrow W$ es una isometría lineal entre espacios con producto escalar, entonces, considerando a V y a W como variedades semi-Riemannianas, $\psi: V \rightarrow W$ es una isometría.

Demostración. Puesto que las aplicaciones lineales son suaves, el isomorfismo lineal ψ es un difeomorfismo. Si $v_p \in T_p(V)$, entonces tenemos que $d\psi(v_p) = (\psi(v))_{\psi(p)}$. Así, ψ preserva los tensores métricos, ya que

$$\begin{aligned} \langle d\psi(v_p), d\psi(w_p) \rangle &= \langle (\psi(v))_{\psi(p)}, (\psi(w))_{\psi(p)} \rangle \\ &= \langle \psi(v), \psi(w) \rangle = \langle v, w \rangle = \langle v_p, w_p \rangle. \quad \#\# \end{aligned}$$

Se sigue que si V es un espacio con producto escalar de dimensión n e índice γ , entonces, considerado como variedad semi-Riemanniana, V es isométrico a \mathbb{R}_γ^n . De hecho, el isomorfismo de coordenadas con respecto a cualquier base ortonormal de V es una isometría (lineal).

Si M es una variedad semi-Riemanniana arbitraria, su tensor métrico convierte a cada uno de los espacios tangentes en un espacio semi-Euclidiano de la misma dimensión y del mismo índice que M . Con esto vemos como la geometría semi-Riemanniana generaliza a la geometría semi-Euclidiana.

36 CONEXION DE LEVI-CIVITA

Sean V y W campos vectoriales sobre una variedad semi-Riemanniana M . Nuestro objetivo en esta sección es demostrar que es posible definir un nuevo campo vectorial sobre M , denotado por $D_V W$, y cuyo valor en cada punto p es la razón vectorial de cambio de W en la dirección de V_p . Hay una manera natural de lograr esto en \mathbb{R}_γ^n .

Sean u^1, \dots, u^n las coordenadas naturales de \mathbb{R}_γ^n . Si V y $W = \sum W^i \partial_i$ son campos vectoriales sobre \mathbb{R}_γ^n , entonces el campo vectorial

$$D_V W = \sum V(W^i) \partial_i$$

se llama la derivada covariante natural de W con respecto a V .

Puesto que esta definición hace uso de un sistema de coordenadas especial sobre \mathbb{R}_γ^n no resulta muy claro como se le puede extender para el caso de una variedad semi-Riemanniana arbitraria. Comenzaremos, por tanto, axiomatizando sus propiedades más importantes.

Una conexión D sobre una variedad suave M es una función

$$D: \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$$

tal que

- (D1) $D_V W$ es $\mathcal{X}(M)$ -lineal en V ,
- (D2) $D_V W$ es \mathbb{R} -lineal en W
- (D3) $D_V(fW) = (Vf)W + fD_V W$ para $f \in \mathcal{X}(M)$.

$D_V W$ se llama la derivada covariante de W con respecto a V para la conexión D .

El axioma (D1) asegura que $D_V W$ es un tensor con respecto a "la va-

riable" V y, de acuerdo a la proposición 17.1, para un vector tangente individual $v \in T_p(M)$ tenemos un bien definido vector tangente $D_v W \in T_p(M)$, a saber, $(D_v W)_p$ donde V es cualquier campo vectorial tal que $V_p = v$. Por otra parte, (D3) muestra que $D_v W$ no es un tensor en W .

Podemos establecer ahora nuestro objetivo en forma más precisa: demostraremos que sobre toda variedad semi-Riemanniana existe una única conexión que comparte otras dos propiedades ((D4) y (D5), ver más adelante) con la conexión natural de \mathbb{R}^n . Nuestro siguiente paso es algebraico.

Proposición 36.1 Sea M una variedad semi-Riemanniana. Si $V \in \mathcal{H}(M)$ sea V^* la uno-forma sobre M tal que

$$V^*(X) = \langle V, X \rangle \text{ para todo } X \in \mathcal{H}(M).$$

Entonces, la función $V \rightarrow V^*$ es un isomorfismo $\mathcal{F}(M)$ -lineal entre $\mathcal{H}(M)$ y $\mathcal{H}^*(M)$.

Demostración. Como V^* es $\mathcal{F}(M)$ -lineal, es realmente una uno-forma, y la función $V \rightarrow V^*$ es también $\mathcal{F}(M)$ -lineal. Que se trata de un isomorfismo, se sigue de los siguientes dos hechos:

(a) Si $\langle V, X \rangle = \langle W, X \rangle$ para todo $X \in \mathcal{H}(M)$, entonces $V = W$.

(b) Dada cualquier uno-forma $\theta \in \mathcal{H}^*(M)$ hay un único campo vectorial $V \in \mathcal{H}(M)$ tal que $\theta(X) = \langle V, X \rangle$ para todo X .

Sea $U = V - W$. Entonces la afirmación (a) equivale a mostrar que si $\langle U_p, X_p \rangle = 0$ para todo $X \in \mathcal{H}(M)$ y todo $p \in M$, entonces $U = 0$. Puesto que todo elemento de $T_p(M)$ tiene la forma X_p , el resultado se deduce de la no degeneración del tensor métrico.

La afirmación (a) es precisamente la afirmación de unicidad que se hace en (b), así que para probar (b) es suficiente encontrar V sobre una vecindad coordenada arbitraria A . (Las diferentes V localmente definidas serán consistentes en la intersección de dos vecindades coordenadas, debido a la unicidad.) Si $\theta = \sum \theta_i dx^i$ sobre A , sea $V = \sum_{j=1}^n g^{ij} \theta_j \partial_j$. Entonces, como (g_{ij}) y (g^{ij}) son matrices inversas, tenemos $\langle V, \partial_k \rangle = \sum_{j=1}^n g^{ij} \theta_j \langle \partial_j, \partial_k \rangle = \sum_{j=1}^n \theta_j g^{jk} = \sum_{j=1}^n \theta_j \delta_{jk} = \theta_k = \theta(\partial_k)$.

Se sigue de la $\mathcal{F}(M)$ -linealidad del tensor métrico que $\langle V, X \rangle = \theta(X)$ para todo X sobre A .
###

Así pues, en la geometría semi-Riemanniana podemos libremente transformar un campo vectorial en una uno-forma y vice-versa. Los pares correspondientes $V \leftrightarrow \theta$ contienen exactamente la misma información, y se dice que son métricamente equivalentes.

El siguiente resultado es tan importante que se le ha llamado a veces el milagro de la geometría semi-Riemanniana.

Teorema 36.2 Sobre una variedad semi-riemanniana M existe una única conexión D tal que

- (D4) $[V, W] = D_V W - D_W V$, y
(D5) $X \langle V, W \rangle = \langle D_X V, W \rangle + \langle V, D_X W \rangle$,

para todo $X, V, W \in \mathfrak{X}(M)$. D se llama la conexión de Levi-Civita de M , y está caracterizada por la fórmula de Koszul

$$2\langle D_V W, X \rangle = V\langle W, X \rangle + W\langle X, V \rangle - X\langle V, W \rangle \\ - \langle V, [W, X] \rangle + \langle W, [X, V] \rangle + \langle X, [V, W] \rangle.$$

Demostración. Supongamos que D es una conexión sobre M que satisface los axiomas (D4) y (D5). En el lado derecho de la fórmula de Koszul usemos (D5) en los primeros tres términos y (D4) en los otros tres. La mayoría de los términos que resultan de este modo se cancelan por parejas y queda finalmente $2\langle D_V W, X \rangle$. Así, D satisface la fórmula de Koszul y, debido a la afirmación (a) de la demostración anterior, D debe ser única.

Para demostrar la existencia, denotemos por $F(V, W, X)$ al lado derecho de la fórmula de Koszul. Para $V, W \in \mathfrak{X}(M)$ fijos, un cálculo simple demuestra que la función $X \rightarrow F(V, W, X)$ es $\mathfrak{X}(M)$ -lineal y, por lo tanto, se trata de una uno-forma. Por la proposición 36.1, sabemos que existe un único campo vectorial, que denotaremos por $D_V W$, tal que $2\langle D_V W, X \rangle = F(V, W, X)$ para todo X . Así, la fórmula de Koszul se cumple y podemos demostrar a partir de ella que D satisface los axiomas (D1) a (D5).

Por ejemplo, demostremos (D3). Para un X arbitrario, tenemos

$$2\langle D_V(fW), X \rangle = V\langle fW, X \rangle + fW\langle X, V \rangle - X\langle V, fW \rangle \\ - \langle V, [fW, X] \rangle + \langle fW, [X, V] \rangle + \langle X, [V, fW] \rangle.$$

Las funciones se pueden sacar del tensor $\langle \cdot, \cdot \rangle$ y, en cuanto a los paréntesis de Lie tenemos, por ejemplo, $[fW, X] = -XfW + f[W, X]$. Entonces la expresión en el lado derecho de la ecuación anterior se convierte en

$$Vf\langle W, X \rangle + Vf\langle X, W \rangle + Xf\langle V, W \rangle - Xf\langle V, W \rangle + fF(V, W, X) = 2\langle VfW + fD_V W, X \rangle.$$

Entonces por la afirmación (a) de la demostración anterior, $D_V(fW) = (Vf)W + fD_V W$.

Para verificar (D4) comenzamos con

$$2\langle D_V W - D_W V, X \rangle = F(V, W, X) - F(W, V, X).$$

El lado derecho se reduce a

$$\langle X, [V, W] \rangle - \langle X, [W, V] \rangle = 2\langle [V, W], X \rangle,$$

de donde se obtiene el resultado. Las otras verificaciones se hacen de manera similar. ###

Sea x^1, \dots, x^n un sistema coordenado sobre una vecindad coordenada U de una variedad semi-Riemanniana M . Los símbolos de Christoffel para este sistema coordenado son las funciones con valores reales Γ_{ij}^k sobre U tales que

$$D_{\partial_i}(\partial_j) = \sum_k \Gamma_{ij}^k \partial_k \quad (1 \leq i, j \leq n).$$

Como $[\partial_i, \partial_j] = 0$, se sigue de (D4) que $D_{\partial_i}(\partial_j) = D_{\partial_j}(\partial_i)$, de donde $\Gamma_{ij}^k = \Gamma_{ji}^k$.

La conexión D no es un tensor, de modo que los símbolos de Christoffel no obedecen la regla usual de transformación para las componentes de un tensor al cambiar de sistema de coordenadas.

Proposición 36.3 Para un sistema coordenado x^1, \dots, x^n sobre U ,

$$(1) \quad D_{\partial_i}(\sum W^j \partial_j) = \sum_k \left\{ \frac{\partial W^k}{\partial x^i} + \sum_j \Gamma_{ij}^k W^j \right\} \partial_k,$$

donde los símbolos de Christoffel están dados por

$$(2) \quad \Gamma_{ij}^k = \frac{1}{2} \sum_m g^{km} \left\{ \frac{\partial g_{im}}{\partial x^j} + \frac{\partial g_{jm}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^m} \right\}.$$

Demostración. (1) es una consecuencia inmediata de (D3). Para obtener (2), hagamos $V = \partial_i$, $W = \partial_j$, $X = \partial_k$ en la fórmula de Koszul. Como los paréntesis de Lie son cero, obtenemos

$$2\langle D_{\partial_i}(\partial_j), \partial_k \rangle = \frac{\partial}{\partial x^i}(g_{jm}) + \frac{\partial}{\partial x^j}(g_{im}) - \frac{\partial}{\partial x^m}(g_{ij}).$$

Pero por la definición de los símbolos de Christoffel,

$$2\langle D_{\partial_i}(\partial_j), \partial_k \rangle = 2 \sum_a \Gamma_{ij}^a g_{am}.$$

Multiplicando estas dos ecuaciones por $\sum_m g^{mk}$, se obtiene el resultado buscado. ##

Usando (D1) podemos calcular cualquier $D_V W$ sobre vecindades coordinadas, por medio de la primera de las fórmulas en la proposición, mientras que la segunda fórmula es la descripción coordinada de como el tensor métrico determina a la conexión de Levi-Civita.

Lema 36.4 La conexión natural D sobre \mathbb{R}_v^n es la conexión de Levi-Civita para el espacio semi Euclidiano \mathbb{R}_v^n para toda $v = 0, 1, \dots, n$. Con respecto a las coordenadas naturales sobre \mathbb{R}_v^n

- (1) $g_{ij} = \delta_{ij} \varepsilon_j$, donde $\varepsilon_j = -1$ para $1 \leq j \leq v$, y $\varepsilon_j = +1$ para $v+1 \leq j \leq n$,
 (2) $\Gamma_{ij}^k = 0$ para toda $1 \leq i, j, k \leq n$.

Demostración. (1) es esencialmente la definición del tensor métrico de \mathbb{R}_v^n . Para probar que D es la conexión de Levi-Civita de \mathbb{R}_v^n debemos verificar que satisface los axiomas (D1) a (D5). Consideremos (D5), por ejemplo. Como $\langle V, W \rangle = \sum \varepsilon_i V^i W^i$,

$$X\langle V, W \rangle = \sum \varepsilon_i X(V^i) W^i + \sum \varepsilon_i V^i X(W^i) = \langle D_X V, W \rangle + \langle V, D_X W \rangle.$$

Entonces (2) se sigue de la proposición 36.3(2), ya que las funciones g_{ij} son constantes. ##

Un campo vectorial V se llama paralelo si su derivada covariante $D_X V$ es cero para todo $X \in \mathfrak{X}(M)$. Así, la anulación de los símbolos de Christoffel en el lema anterior significa que los campos vectoriales de las coordenadas naturales sobre \mathbb{R}_v^n son paralelos.

En general, los símbolos de Christoffel de un sistema coordenado miden cuán lejos están los campos vectoriales coordenados de ser paralelos.

Ejemplo. Sean r, φ, z las coordenadas cilíndricas usuales de \mathbb{R}^3 . Realmente, (r, φ, z) es un sistema coordenado sólo sobre $\mathbb{R}^3 - H$, donde H es, por ejemplo, el semiplano $x > 0, y = 0$. En $\mathbb{R}^3 - H$ las funciones coordenadas están bien definidas y existe una aplicación coordenada inversa dada por $x = r \cos \varphi, y = r \sin \varphi, z = z$.

Por el teorema de la base tenemos que

$$\partial_r = \cos \varphi \partial_x + \sin \varphi \partial_y;$$

$$\partial_\varphi = rU, \text{ donde } U = -\sin \varphi \partial_x + \cos \varphi \partial_y;$$

$$\partial_z = \partial_z.$$

Para poder usar la notación con índices, hagamos $y^1 = r, y^2 = \varphi, y^3 = z$. Entonces $\epsilon_{11} = \epsilon_{22} = 1, \epsilon_{33} = r^2$, y $\epsilon_{ij} = 0$ para $i \neq j$, por lo tanto

$$ds^2 = dr^2 + r^2 d\varphi^2 + dz^2.$$

En particular, este es un sistema ortogonal de coordenadas; es decir, los campos vectoriales coordenados son mutuamente ortogonales.

Un cálculo directo demuestra que los símbolos de Christoffel de las coordenadas cilíndricas son todos cero con la excepción de $\Gamma_{22}^1 = -r$ y $\Gamma_{12}^2 = \Gamma_{21}^2 = 1/r$. Por lo tanto, todas las derivadas covariantes de los campos coordenados son cero, con la excepción de

$$D_{\partial_\varphi}(\partial_\varphi) = -r \partial_r,$$

$$D_{\partial_\varphi}(\partial_r) = D_{\partial_r}(\partial_\varphi) = U.$$

Estas fórmulas son consistentes con lo que podemos observar en la figura 10. Puesto que ∂_z es también un vector coordenado natural, debe ser paralelo. Similarmente, podemos esperar que $D_{\partial_z}(\partial_r) = D_{\partial_z}(\partial_\varphi) = 0$, ya que ∂_r y ∂_φ permanecen paralelos mientras el punto p se mueve en la dirección z .

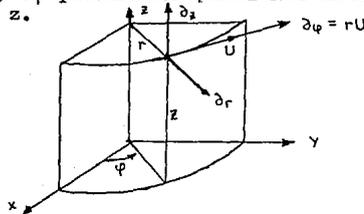


FIG. 10

La derivada covariante D_y se puede extender para operar sobre campos tensoriales de tipo arbitrario. De hecho, los axiomas (D2) y (D3) son exactamente lo que necesitamos para poder aplicar el teorema 21.4 (pag. 53) acerca de las derivaciones tensoriales.

Sea V un campo vectorial sobre una variedad semi-Riemanniana M . La

derivada covariante (de Levi-Civita) D_V es la única derivación tensorial sobre M tal que

$$D_V f = Vf \quad \text{para } f \in \mathcal{F}(M),$$

y $D_V W$ es la derivada covariante de Levi Civita para todo $W \in \mathcal{H}(M)$.

Si $A \in \mathcal{X}_s^r(M)$ entonces el campo tensorial $D_V A$ de tipo (r,s) es $\mathcal{F}(M)$ -lineal en $V \in \mathcal{H}(M)$. De hecho, por el corolario 2L.3, para verificar esta afirmación, es suficiente verificar que las derivaciones tensoriales D_{FV+gW} y $fD_V + gD_W$ coinciden sobre $\mathcal{F}(M)$ y sobre $\mathcal{H}(M)$. Pero lo primero se sigue de las definiciones y lo segundo es precisamente el axioma (D1). Esta observación sirve para justificar la siguiente definición.

La diferencial covariante de un tensor A del tipo (r,s) sobre M es el campo tensorial DA del tipo $(r,s+1)$ tal que

$$(DA)(\theta^1, \dots, \theta^r, X_1, \dots, X_s, V) = (D_V A)(\theta^1, \dots, \theta^r, X_1, \dots, X_s)$$

para todo $V, X_i \in \mathcal{H}(M)$ y $\theta^j \in \mathcal{H}^*(M)$.

En el caso excepcional $r = s = 0$ la diferencial covariante de una función f coincide con su diferencial usual $df \in \mathcal{H}^*(M)$, ya que

$$(Df)(V) = D_V f = Vf = df(V) \quad \text{para todo } V \in \mathcal{H}(M).$$

La diferencial covariante DA es simplemente una manera conveniente de recolectar todas las derivadas covariantes del tensor A .

Como en el caso de un campo vectorial, un campo tensorial A se llama paralelo si su diferencial covariante es cero, es decir, si $D_V A = 0$ para todo $V \in \mathcal{H}(M)$. Por ejemplo, usando la regla del producto (sección 2L), se sigue que el axioma (D5) es equivalente al paralelismo del tensor métrico g .

Si $A \in \mathcal{X}_s^r(M)$, las componentes de DA con respecto a un sistema coordinado se denotan por $A_{j_1 \dots j_s, k}^{i_1 \dots i_r}$. Se puede encontrar una fórmula para expresar estas componentes en términos de las componentes de A y de los símbolos de Christoffel del sistema coordinado. La fórmula general es algo complicada (ver, por ejemplo, [S y S]), pero veremos que su uso puede ser evitado. En el caso especial del sistema natural de coordenadas sobre \mathbb{R}^n , puesto que los campos vectoriales coordinados y también las diferenciales du^1, \dots, du^n son paralelas, se sigue que

$$A_{j_1 \dots j_s, k}^{i_1 \dots i_r} = (\partial/\partial u^k) A_{j_1 \dots j_s}^{i_1 \dots i_r}.$$

37 TRANSPORTE PARALELO

El caso más simple de un campo vectorial sobre una aplicación (ver la definición en la pag. 35) es un campo vectorial Z sobre una curva $\alpha: I \rightarrow M$. Z asigna suavemente a cada $t \in I$ un vector tangente a M en el punto $\alpha(t)$. Por ejemplo, la velocidad α' es un campo vectorial sobre α ,

lo mismo que la restricción V_α a la curva α de cualquier $V \in \mathfrak{X}(M)$. El conjunto $\mathfrak{X}(\alpha)$ de todos los campos vectoriales suaves sobre la curva α , es un módulo sobre $\mathcal{F}(I)$.

Cuando M es una variedad semi-Riemanniana, existe una manera natural de definir la razón vectorial de cambio Z' de un campo vectorial $Z \in \mathfrak{X}(\alpha)$.

Proposición 37.1 Sea $\alpha: I \rightarrow M$ una curva en una variedad semi-Riemanniana M . Entonces existe una única función $Z \rightarrow Z' = DZ/dt$ de $\mathfrak{X}(\alpha)$ a $\mathfrak{X}(\alpha)$ llamada la derivada covariante inducida y que es tal que

- (1) $(az_1 + bz_2)' = az_1' + bz_2'$ ($a, b \in \mathbb{R}$),
- (2) $(hZ)' = (dh/dt)Z + hZ'$ ($h \in \mathcal{F}(I)$),
- (3) $(V_\alpha)'(t) = D_{\alpha'(t)}(V)$ ($t \in I, V \in \mathfrak{X}(M)$).

Además,

$$(4) \quad (d/dt)\langle z_1, z_2 \rangle = \langle z_1', z_2 \rangle + \langle z_1, z_2' \rangle.$$

Demostración. Unicidad. Supongamos que existe una conexión inducida que satisface sólo las primeras tres propiedades. Podemos suponer que α está en el dominio de un sólo sistema coordenado x^1, \dots, x^n . Por el teorema de la base, si $Z \in \mathfrak{X}(\alpha)$, entonces en $\alpha(t)$,

$$Z(t) = \sum Z^i(t)x^i \partial_i = \sum (Zx^i)(t)\partial_i.$$

Denotemos a la función componente $Zx^i: I \rightarrow \mathbb{R}$ por Z^i . Por las propiedades (1) y (2), tenemos que

$$Z' = \sum \frac{dZ^i}{dt} \partial_i|_\alpha + \sum Z^i (\partial_i|_\alpha)'$$

Pero, por (3), $(\partial_i|_\alpha)' = D_{\alpha'}(\partial_i)$; así que

$$Z' = \sum \frac{dZ^i}{dt} \partial_i + \sum Z^i D_{\alpha'}(\partial_i).$$

Entonces, Z' está completamente determinada por la conexión de Levi-Civita D .

Existencia. Sobre cualquier subintervalo J de I tal que $\alpha(J)$ está contenido en una vecindad coordenada, definamos Z' por la fórmula anterior. Entonces una serie de cálculos muy simples, demuestran que se cumplen las cuatro propiedades. Debido a la unicidad, estas definiciones locales de Z' constituyen un sólo campo vectorial en $\mathfrak{X}(\alpha)$. ##

En el caso especial $Z = \alpha'$ la derivada $Z' = \alpha''$ se llama la aceleración de la curva α . A veces se usan notaciones más elaboradas para enfatizar que α'' depende de la geometría, mientras que α' sólo depende de la estructura diferenciable.

Para un campo vectorial Z sobre α la fórmula que encontramos en la demostración de la proposición anterior, nos da

$$Z' = \sum_k \left\{ \frac{dZ^k}{dt} + \sum_{ij} \Gamma_{ij}^k \frac{d(x^i \circ \alpha)}{dt} Z^j \right\} \partial_k.$$

Si $Z' = 0$, entonces Z se dice que es paralelo. Esta fórmula demuestra que la ecuación $Z' = 0$ es equivalente a un sistema lineal de ecuaciones

diferenciales ordinarias. Por el teorema fundamental de existencia y unicidad para tales sistemas, obtenemos la siguiente proposición.

Proposición 37.2 Para una curva $\alpha: I \rightarrow M$, sea $a \in I$ y $z \in T_{\alpha(a)}(M)$. Entonces existe un único campo vectorial paralelo sobre α tal que $Z(a) = z$.

Aquí hemos aprovechado el hecho de que las soluciones de un sistema lineal de ecuaciones diferenciales están definidas sobre todo el intervalo de definición de las funciones coeficientes.

Con la misma notación que en la proposición, si $b \in I$ entonces la función

$$P = P_a^b(\alpha): T_p(M) \rightarrow T_q(M)$$

que manda cada z a $Z(b)$ se llama transporte paralelo a lo largo de α de $p = \alpha(a)$ a $q = \alpha(b)$.

Lema 37.3 La aplicación transporte paralelo es una isometría lineal.

Demostración. Con la misma notación que antes, sean $v, w \in T_p(M)$ los vectores que, como en la proposición, corresponden a los campos vectoriales paralelos V y W . Como $V + W$ es también paralelo, $P(v + w) = (V + W)(b) = V(b) + W(b) = P(v) + P(w)$. Similarmente, $P(cv) = cP(v)$. Así, P es una aplicación lineal.

Si $P(v) = 0$, entonces por la unicidad en la proposición 37.2, V sólo puede ser el campo vectorial idénticamente cero sobre α . De donde $v = V(a) = 0$. Así, P es inyectiva y, como todos los espacios tangentes a M tienen la misma dimensión, P es un isomorfismo lineal.

Finalmente, para V y W como antes,

$$\frac{d}{dt} \langle V, W \rangle = \langle V', W \rangle + \langle V, W' \rangle = 0.$$

De donde se sigue que $\langle V, W \rangle$ es constante. Por lo tanto,

$$\langle P(v), P(w) \rangle = \langle V(b), W(b) \rangle = \langle V(a), W(a) \rangle = \langle v, w \rangle. \quad \#\#$$

Por lo general el transporte paralelo de p a q depende de la curva particular que se tome entre los puntos p y q . Sobre \mathbb{R}^n los campos vectoriales de las coordenadas naturales son paralelos y también lo serán sus restricciones a cualquier curva. Por lo tanto, en este caso, el transporte paralelo a lo largo de cualquier curva es simplemente el isomorfismo canónico $v_p \rightarrow v_q$. Este fenómeno se llama paralelismo distante.

38 GEODESICAS

Ahora generalizaremos la noción euclidiana de línea recta. Una geodésica en una variedad semi-Riemanniana M es una curva $\gamma: I \rightarrow M$ cuyo campo de velocidades γ' es paralelo. Equivalentemente, las geodésicas

son las curvas de aceleración cero: $\gamma'' = 0$.

Corolario 38.1 Sea x^1, \dots, x^n un sistema coordenado sobre UC M. Una curva γ en U es una geodésica de M si y sólo si sus funciones coordenadas $x^i \circ \gamma$ satisfacen

$$\frac{d^2(x^k \circ \gamma)}{dt^2} + \sum_{i,j} \Gamma_{ij}^k(\gamma) \frac{d(x^i \circ \gamma)}{dt} \frac{d(x^j \circ \gamma)}{dt} = 0$$

para $1 \leq k \leq n$.

De hecho, estas expresiones son las componentes de γ'' con respecto a los campos vectoriales $\partial_1, \dots, \partial_n$.

Al tratar con curvas, con frecuencia es conveniente usar una notación abreviada, escribiendo las funciones coordenadas de γ como x^i en lugar de $x^i \circ \gamma$. En cualquier contexto razonable no debe haber confusión entre estas funciones definidas sobre el dominio I de γ y las funciones coordenadas sobre UC M. Las ecuaciones de la geodésica son entonces

$$\frac{d^2 x^k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} = 0 \quad (1 \leq k \leq n).$$

El teorema de existencia y unicidad para las ecuaciones diferenciales ordinarias nos lleva al siguiente resultado local.

Lema 38.2 Si $v \in T_p(M)$, existe un intervalo I alrededor del 0 y existe una única geodésica $\gamma: I \rightarrow M$ tal que $\gamma'(0) = v$.

La última ecuación implica desde luego que $\gamma(0) = p$; decimos que γ es una geodésica que comienza en p con velocidad inicial v.

Lema 38.3 Sean $\alpha, \beta: I \rightarrow M$ un par de geodésicas. Si existe un número $a \in I$ tal que $\alpha'(a) = \beta'(a)$, entonces $\alpha = \beta$.

Demostración. Supongamos que la conclusión es falsa; entonces existe un $t_0 \in I$ tal que $\alpha(t_0) \neq \beta(t_0)$, con digamos $t_0 > a$. Así, el conjunto $\{t \in I \mid t > a \text{ y } \alpha(t) \neq \beta(t)\}$ tiene una máxima cota inferior b, para la cual $b > a$. Afirmamos que $\alpha'(b) = \beta'(b)$. Esto sucede si $a = b$. Si $b > a$, entonces α y β concuerdan sobre el intervalo (a, b). Las expresiones coordenadas demuestran que las funciones $t \rightarrow \alpha'(t)$ y $t \rightarrow \beta'(t)$ de (a, b) en el haz tangente TM son continuas (de hecho, suaves). Así pues, cuando t tiende a b por la izquierda

$$\alpha'(b) = \lim_{t \rightarrow b^-} \alpha'(t) = \lim_{t \rightarrow b^-} \beta'(t) = \beta'(b).$$

Puesto que $t \rightarrow \alpha(t+b)$ y $t \rightarrow \beta(t+b)$ son también geodésicas, el lema 38.2 muestra que $\alpha = \beta$ sobre algún intervalo alrededor de b. Pero esto contradice la definición de b. ##

Proposición 38.4 Dado cualquier vector tangente $v \in T_p(M)$, existe una única geodésica γ_v en M tal que

- (1) La velocidad inicial de γ_v es v ; es decir, $\gamma_v'(0) = v$.
 (2) El dominio I_v de γ_v es el más grande posible. Así que si $\alpha: J \rightarrow M$ es una geodésica con velocidad inicial v , entonces $J \subset I$ y $\alpha = \gamma_v|_J$.

Demostración. Sea G la colección de todas las geodésicas $\gamma: I_\gamma \rightarrow M$ con velocidad inicial v . (Por el lema 38.2, este conjunto G es no vacío.) El lema 38.3 demuestra que α y β de la colección G coinciden sobre $I_\alpha \cap I_\beta$. Por lo tanto, la colección G define consistentemente una sola curva γ_v sobre el intervalo $I = \cup I_\gamma$. Es evidente que γ_v tiene las propiedades requeridas. ##

Debido a la propiedad (2) de la proposición anterior, la geodésica γ_v se dice que es maximal o que es geodésicamente inextensible. La notación γ_v será usada con frecuencia.

Imaginemos por un momento a M como si fuera una superficie en \mathbb{R}^3 y supongamos que $p \in M$ es una partícula restringida a permanecer sobre M . Una vez que se ha dado a p una velocidad inicial, su movimiento está completamente determinado y su trayectoria será una geodésica sobre M .

Una variedad semi-Riemanniana M para la cual toda geodésica maximal está definida sobre toda la recta real se llama geodésicamente completa o completa. Nótese que si un sólo punto p se remueve de la variedad completa M , entonces la variedad $M - p$ ya no es completa, puesto que las geodésicas que pasaban por el punto p ahora se ven obligadas a detenerse.

Ejemplo. Geodésicas de un espacio semi-Euclidiano. Para las coordenadas naturales, los símbolos de Christoffel son cero, así que las ecuaciones de la geodésica se reducen a

$$\frac{d^2(u^i \circ \gamma)}{dt^2} = 0 \quad (1 \leq i \leq n).$$

Por lo cual, $u^i(\gamma(t)) = p^i + tv^i$ para toda t , donde p^i y v^i son constantes arbitrarias. En notación vectorial tenemos, $\gamma(t) = p + tv$. Por lo tanto, las geodésicas de \mathbb{R}_ν^n son líneas rectas. En particular, \mathbb{R}_ν^n es geodésicamente completo.

Puesto que su campo de velocidades es paralelo, una geodésica γ tiene un comportamiento muy uniforme. Toda curva constante en M es trivialmente una geodésica, pero si $\gamma'(t) \neq 0$ para un sólo t , entonces γ' nunca se hace cero. Así pues, una geodésica no puede disminuir su velocidad y detenerse.

Una curva α en M se llama espacial si todos sus vectores velocidad $\alpha'(s)$ son espaciales; en forma análoga se definen las curvas temporales y nulas. Una curva arbitraria no tiene porqué tener uno de estos caracteres causales; no obstante, una geodésica γ sí tiene un carácter causal definido, ya que su campo de velocidades es paralelo y el transporte paralelo preserva el carácter causal de los vectores.

Lema 38.5 Sea $\gamma: I \rightarrow M$ una geodésica que no es constante. Una reparametrización $\delta \circ h: J \rightarrow M$ es una geodésica si y sólo si $h: J \rightarrow I$ es de la forma $h(t) = at + b$, con $a, b \in \mathbb{R}$.

Demostración. Para cualquier curva δ , $(\delta \circ h)'(t) = (dh/dt)\delta'(h(t))$. De donde,

$$(\delta \circ h)''(t) = \frac{d^2h}{dt^2} \delta'(h(t)) + \frac{dh}{dt} \delta''(h(t)).$$

Como δ es una geodésica, $\delta'' = 0$, y δ no constante implica que δ' nunca es cero. Así, $\delta \circ h$ es una geodésica $\Leftrightarrow (\delta \circ h)'' = 0 \Leftrightarrow d^2h/dt^2 = 0 \Leftrightarrow h(t) = at + b$. ##

Este resultado muestra que las parametrizaciones de las geodésicas tienen significado geométrico. Si al reparametrizar una curva ésta se convierte en una geodésica, decimos que se trata de una pregeodésica.

Si un sistema de ecuaciones diferenciales ordinarias de segundo orden tiene funciones suaves como coeficientes, entonces sus soluciones son funciones suaves no sólo con respecto al parámetro, sino también con respecto a los valores iniciales y con respecto a los valores iniciales de las primeras derivadas. Aplicando este hecho a las ecuaciones diferenciales de las geodésicas, obtenemos el siguiente

Lema 38.6 Sea v un vector tangente a M , es decir, un elemento del haz tangente TM . Entonces existe una vecindad N de v en TM y un intervalo I alrededor de 0 , tales que $(w, s) \rightarrow \gamma_w(s)$ es una función suave bién definida de $N \times I \rightarrow M$, en donde $\gamma_w(s)$ es la geodésica con velocidad inicial w .

Una ecuación diferencial de segundo orden para la función incógnita y se puede convertir en un par de ecuaciones diferenciales de primer orden si tomamos a y y' como una nueva variable. Por esencialmente el mismo procedimiento, las geodésicas en M pueden representarse como curvas integrales de algún campo vectorial sobre el haz tangente TM .

Proposición 38.7 Existe un campo vectorial G sobre TM tal que la proyección $\pi: TM \rightarrow M$ establece una correspondencia uno a uno entre las curvas integrales de G y las geodésicas de M .

Demostración. Si $v \in TM$, sea G_v la velocidad inicial de la curva $s \rightarrow \gamma_v(s)$ en TM . Se sigue, usando el lema anterior, que G es un campo vectorial suave sobre TM .

(a) Si δ es una geodésica en M , entonces δ' es una curva integral de G . Para demostrar esta afirmación, sea $\alpha(s) = \delta'(s)$, para toda s . Para un t fijo y arbitrario, sea $w = \delta'(t)$ y sea $\beta(s) = \gamma_w'(s)$. Por el lema 38.3, $\delta'(t+s) = \gamma_w'(s)$. Tomando las velocidades en M , obtenemos $\alpha'(t+s) = \gamma_w''(s) = \beta(s)$. Entonces tomando velocidades en TM , obtenemos $\alpha'(t+s) = \beta'(s)$. En particular, $\alpha'(t) = \beta'(0) = G_w = G_{\alpha(t)}$, lo que prueba (a).

(b) Si α es una curva integral de G , entonces $\pi \circ \alpha$ es una geodésica en M .

Si $v = \alpha'(0)$ entonces por la afirmación (a), $s \mapsto \gamma'_v(s)$ es también una curva integral de G y, lo mismo que α , comienza en v . Entonces, la unicidad de las curvas integrales implica que, inicialmente al menos, $\pi \circ \alpha = \pi \circ \gamma'_v = \gamma'_v$. Para t arbitrario, sea δ la curva integral de G que comienza en $\alpha(t)$. Por el lema 13.3 (pag. 36), $\alpha(t+s) = \delta(s)$, por lo tanto $\pi \alpha(t+s) = \pi \delta(s) = \gamma'_{\delta(t)}(s)$, lo cual prueba (b).

Las identidades $\pi \circ \gamma' = \gamma'$ y $(\pi \circ \alpha)' = \alpha$ demuestran que las aplicaciones $\alpha \mapsto \pi \circ \alpha$ y $\gamma' \mapsto \gamma'$ son inversas. ##

39 LA APLICACION EXPONENCIAL

En cada punto o de una variedad semi-Riemanniana M coleccionaremos a todas las geodésicas que comienzan en o mediante una sola aplicación.

Si $o \in M$, sea E_o el conjunto de todos los vectores $v \in T_o(M)$ tales que la geodésica inextensible γ_v esté definida al menos sobre $[0, 1]$. La aplicación exponencial de M en o es la función

$$\exp_o: E_o \rightarrow M$$

tal que $\exp_o(v) = \gamma_v(1)$ para todo $v \in E_o$.

Obviamente, E_o es el más grande subconjunto de $T_o(M)$ sobre el cual \exp_o puede ser definida. Si M es completa, entonces $E_o = T_o(M)$ para todo punto o de M .

Fijemos $v \in T_o(M)$ y $t \in \mathbb{R}$; entonces la geodésica $s \mapsto \gamma_v(ts)$ tiene v como la velocidad inicial $t\gamma'_v(0) = tv$. Por lo tanto, $\gamma_{tv}(s) = \gamma_v(ts)$ para todo s y t tales que alguno de los lados de esta ecuación (y entonces ambos) esté bien definido. En particular, si $v \in E_o$ entonces

$$\exp_o(tv) = \gamma_{tv}(1) = \gamma_v(t).$$

Así pues, la aplicación exponencial transforma las rectas que pasan por el origen de $T_o(M)$ en las geodésicas de M que pasan por o .

Proposición 39.1 Para cada punto $o \in M$ existe una única vecindad U de o en $T_o(M)$ sobre la cual la aplicación exponencial \exp_o es un difeomorfismo sobre una vecindad V de o en M . (Ver fig. 11).

Demostración. Se sigue del lema 38.6 que \exp_o es una aplicación suave bien definida sobre una vecindad de o en $T_o(M)$. Afirmamos que su diferencial

$$d \exp_o: T_o(T_o(M)) \rightarrow T_o(M)$$

es el isomorfismo canónico $v_o \mapsto v$. Por definición, $v_o = \rho'(0)$, donde $\rho(t) = tv$; como hicimos notar antes, $\exp_o(tv) = \gamma_v(t)$. Así

$$d \exp_o(v_o) = d \exp(\rho'(0)) = (\exp_o \circ \rho)'(0) = \gamma'_v(0) = v.$$

La proposición se sigue del teorema de la función inversa. ##

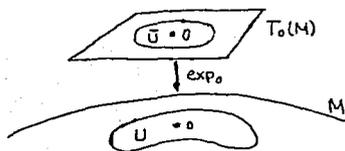


FIG. 11

Un subconjunto S de un espacio vectorial se llama con forma de estrella alrededor de 0 si $v \in S$ implica que $tv \in S$ para todo $t \in [0, 1]$. Entonces, S es una unión de segmentos radiales de recta que parten de 0. Si U y \bar{U} son como en la proposición anterior y \bar{U} tiene forma de estrella alrededor de 0, entonces U se llama una vecindad normal de o. Ahora veremos que en cierto sentido, U también tiene forma de estrella alrededor de o .

Proposición 39.2 Si U es una vecindad normal de $o \in M$, entonces para cada punto $p \in U$ existe una única geodésica $\sigma: [0, 1] \rightarrow U$ que va de o a p en U . Además, $\sigma'(0) = \exp^{-1}(p) \in \bar{U}$.

Demostración. Por definición \bar{U} es una vecindad en forma de estrella de 0 en $T_0(M)$ tal que $\exp_0: \bar{U} \rightarrow U$ es un difeomorfismo sobre U . Para $p \in U$, sea $v = \exp_0^{-1}(p) \in \bar{U}$. Puesto que \bar{U} tiene forma de estrella, el rayo $\rho(t) = tv$ ($0 \leq t \leq 1$) está en \bar{U} . Así, el segmento de geodésica $\sigma = \exp_0 \circ \rho$ está en U y va de o a p . (Se dice que σ es radial.) En el origen de $T_0(M)$, $d \exp_0$ es el isomorfismo canónico $T_0(T_0(M)) \approx T_0(M)$. Pero $\rho'(0) = v$, por lo tanto,

$$\sigma'(0) = d \exp_0(\rho'(0)) = d \exp(v) = v.$$

Supongamos que $\tau: [0, 1] \rightarrow U$ es una geodésica arbitraria en U que va de o a p . Si $w = \tau'(0)$, entonces las geodésicas $t \rightarrow \exp_0(tw)$ y tienen la misma velocidad inicial y, por lo tanto, son iguales.

El segmento radial $t \rightarrow tw$ ($0 \leq t \leq 1$) no sale en ningún momento de \bar{U} , pues si así fuera, habría un $0 \leq t_0 \leq 1$ tal que $t_0 w \in \bar{U}$ pero $\exp_0(t_0 w) \in U - \tau([0, 1])$. Así pues, $w \in \bar{U}$. Pero $\exp_0(w) = \tau(1) = p = \exp_0(v)$ y como \exp_0 es inyectiva, $w = v$. Por lo tanto, por la unicidad de las geodésicas, $\tau = \sigma$.
##

Esta demostración muestra también que una vecindad normal U de o determina de manera única la vecindad \bar{U} en $T_0(M)$.

Una geodésica quebrada es una curva suave por pedazos cuyos subsegmentos suaves son geodésicas. Por ejemplo, una geodésica quebrada en \mathbb{R}^2 es simplemente una curva poligonal o línea quebrada.

Lema 39.3 Una variedad semi-Riemanniana M es conexa si y sólo si cualesquiera dos puntos de M pueden ser unidos por una geodésica quebrada.

Demostración. Supongamos que M es conexa y fijemos un $p \in M$. Sea C el conjunto de todos los puntos que pueden ser conectados con p por medio

por medio de una geodésica quebrada. Para $q \in M$, sea U una vecindad normal. Si $q \in C$, entonces es claro que $U \subset C$. Pero también si $q \in M-C$, entonces $U \cap M = C$. Así, la hipótesis de conexidad implica que $M = C$. La afirmación inversa es obvia. ##

Sobre cualquier vecindad normal U de $o \in M$ existe un tipo especial de sistema coordenado que es particularmente simple. Sea e_1, \dots, e_n una base ortonormal para $T_o(M)$, de modo que $\langle e_i, e_j \rangle = \delta_{ij}$. El sistema normal de coordenadas $\xi = (x^1, \dots, x^n)$ determinado por e_1, \dots, e_n asigna a cada punto $p \in U$ las componentes vectoriales con respecto a la base de los e_i , del vector correspondiente a p :

$$\exp_o^{-1}(p) \in \bar{U} \subset T_o(M).$$

En resumen,

$$\exp_o^{-1}(p) = \sum x^i(p) e_i \quad (p \in U).$$

Por tanto, si f^1, \dots, f^n es la base dual a e_1, \dots, e_n , entonces $x^i \circ \exp_o = f^i$ sobre \bar{U} .

Proposición 39.4 Si x^1, \dots, x^n es un sistema normal de coordenadas en $o \in M$, entonces para todo i, j, k

$$(1) \quad g_{ij}(o) = \delta_{ij} \quad ; \quad (2) \quad \Gamma_{ij}^k(o) = 0.$$

Demostración. Con la misma notación que arriba hemos usado, si $v \in T_o(M)$ escribimos $v = \sum a^i e_i$. Como $\exp_o(tv) = \gamma_v(t)$,

$$x^i(\gamma_v(t)) = f^i(tv) = t f^i(v) = t a^i.$$

Por lo tanto, $v = \gamma_v'(0) = \sum a^i \partial_i|_o$. Tomando $a^i = \delta_{ij}$, vemos que $e_j = \partial_j|_o$, de donde se obtiene (1).

La expresión para $x^i \circ \gamma_v$ demuestra que las ecuaciones diferenciales para la geodésica γ_v se reducen a

$$\sum \Gamma_{ij}^k(\gamma_v(t)) a^i a^j = 0 \quad \text{para toda } k.$$

En particular, $\sum \Gamma_{ij}^k(o) a^i a^j = 0$ se cumple para todo $a = (a^1, \dots, a^n) \in \mathbb{R}^n$. Para k fijo, esto expresa el hecho de que una cierta forma cuadrática sobre \mathbb{R}^n es idénticamente cero. Por lo tanto, de la igualdad de polarización, se sigue que la correspondiente forma bilineal simétrica es idénticamente cero, es decir, $\Gamma_{ij}^k(o) = 0$. ##

Si recordamos el lema 36.4 (pag. 103), podemos decir que en el punto o (y, por lo general, sólo en él) el tensor métrico y los símbolos de Christoffel de un sistema normal de coordenadas son semi-Euclidianos. Puesto que los tensores son objetos de naturaleza puntual, este hecho es de sorprendente utilidad en los cálculos con tensores. Por ejemplo, supongamos un problema que involucra a la diferencial covariante DA de $A \in \mathcal{F}_2(M)$. En coordenadas arbitrarias las componentes de DA son muy complicadas, pero si para cada punto $o \in M$ usamos un sistema normal de coordenadas en o , entonces $A_{j,k,m}^i(o) = ((\partial/\partial x^m) A_{j,k}^i)(o)$ tal como si

se tratara de las coordenadas naturales en un espacio semi-Euclidiano.

En las cercanías de o , las fórmulas de la proposición 39.4 sólo son aproximadamente verdaderas, pero mientras más cerca estemos del punto o más se parece la variedad M a $T_o(M) \approx \mathbb{R}^n$. Pero esta aproximación no puede ser llevada demasiado lejos. Por ejemplo, en o las primeras derivadas de Γ_{ij}^k no son cero en general (sin embargo, sí son cero las primeras derivadas de g_{ij}).

Ejemplo. Las aplicaciones exponenciales para \mathbb{R}^n . De acuerdo al ejemplo que dimos en la sección anterior, la geodésica con velocidad inicial $v_p \in T_p(\mathbb{R}^n)$ es la línea recta $t \rightarrow p + tv$. Así, la aplicación exponencial en p manda v_p a $p + v$ (que es la punta de la flecha v_p). Se sigue que $\exp_p: T_p(\mathbb{R}^n) \rightarrow \mathbb{R}^n$ es un difeomorfismo, ya que es la composición del isomorfismo canónico $T_p(\mathbb{R}^n) \approx \mathbb{R}^n$, con la translación $x \rightarrow p + x$. De hecho, cuando el espacio con producto escalar $T_p(\mathbb{R}^n)$ está provisto con su tensor métrico usual, estas dos aplicaciones son isometrías, por lo que \exp_p es una isometría.

40 CURVATURA

En la teoría de superficies en \mathbb{R}^3 que comenzó a desarrollarse a fines del siglo XVIII, se definió una noción de curvatura que da una descripción muy razonable de la forma de una superficie considerada como un subconjunto de \mathbb{R}^3 . Fue Gauss quien demostró con su "theoremata egregium", que esta curvatura Gaussiana es un invariante isométrico de la propia superficie, y que es independiente del hecho de que la superficie sea un subconjunto de \mathbb{R}^3 . Este teorema condujo a Riemann a la invención de la geometría Riemanniana, cuyo tema dominante es la generalización de la curvatura Gaussiana al caso de las variedades Riemannianas arbitrarias. No es necesario hacer cambios significativos para poder extender estas ideas a la variedades semi-Riemannianas.

Las derivadas de Lie satisfacen la identidad $L_{[X,Y]} = [L_X, L_Y]$, donde como es usual el lado derecho equivale a $L_X L_Y - L_Y L_X$. (Es suficiente verificar esta identidad para funciones y campos vectoriales.) Entonces si $[X, Y] = 0$ (como, por ejemplo, para los campos vectoriales coordenados) de la identidad se sigue que L_X y L_Y conmutan. En contraste, estos resultados fallan en general para la derivada covariante D_X . La magnitud de esta falla es medida por un tensor que juega un papel central en toda la geometría diferencial.

Lema 40.1 Sea M una variedad semi-Riemanniana con conexión de Levi-Civita denotada por D . La función $R: \mathfrak{K}(M)^3 \rightarrow \mathfrak{K}(M)$ dada por

$$R_{XY}Z = D_{[X,Y]}Z - [D_X, D_Y]Z$$

es un campo tensorial de tipo $(1, 3)$ sobre M , llamado tensor de curvatura Riemanniana de M .

Demostración. Como hicimos notar en la página 44, la función R puede ser interpretada como un elemento de $\mathcal{F}_3(M)$, siempre y cuando R sea $\mathcal{F}(M)$ -multilineal. Puesto que la \mathbb{R} -multilinealidad es obvia, sólo debemos demostrar que es posible "factorizar hacia afuera a las funciones". Por ejemplo, como $[X, fY] = (Xf)Y + f[X, Y]$, tenemos que

$$\begin{aligned} R_{X, fY} Z &= D_{[X, fY]} Z - D_X D_{fY} Z + D_{fY} D_X Z \\ &= (Xf)D_Y Z + fD_{[X, Y]} Z - D_X (fD_Y Z) + fD_Y D_X Z \\ &= (Xf)D_Y Z - (Xf)D_Y Z + fR_{XY} Z = fR_{XY} Z. \quad \#\# \end{aligned}$$

La operación del paréntesis de Lie no es un tensor, tampoco lo es la derivada covariante, pero cuando se les combina como en el lema anterior, ellas producen el tensor R . La notación alternativa $R(X, Y)Z$ para $R_{XY}Z$ es conveniente cuando X y Y se reemplazan por expresiones más complicadas.

Como demostramos en la parte II el tensor R puede ser considerado como una función \mathbb{R} -multilineal aplicada a vectores individuales. Si $x, y \in T_p(M)$, el operador lineal

$$R_{xy} : T_p(M) \rightarrow T_p(M)$$

que envía cada z a $R_{xy}z$ se llama operador de curvatura. Las siguientes identidades son las simetrías de la curvatura.

Proposición 40.2 Si $x, y, z, v, w \in T_p(M)$, entonces

- (1) $R_{xy} = -R_{yx}$,
- (2) $\langle R_{xy}v, w \rangle = -\langle R_{xy}w, v \rangle$,
- (3) $R_{xy}z + R_{yz}x + R_{zx}y = 0$,
- (4) $\langle R_{xy}v, w \rangle = \langle R_{vw}x, y \rangle$.

Las primeras dos identidades demuestran que el tensor de curvatura contiene varias antisimetrías. En particular (2) dice que los operadores de curvatura son antiadjetos. La ecuación (3) se llama la primera identidad de Bianchi, nótese que en ella los vectores aparecen cíclicamente permutados. La simetría por parejas dada por la ecuación (4), será una consecuencia de las otras tres identidades.

Demostración. Puesto que la derivada covariante D_X y el paréntesis de Lie son operaciones de naturaleza local, será suficiente trabajar en cualquier vecindad de un punto p . Debido a que las identidades que queremos demostrar son identidades tensoriales, los vectores tangentes x, y, \dots pueden ser extendidos a campos vectoriales X, Y, \dots sobre una vecindad de manera conveniente. En este caso escogeremos los campos de manera que todos sus paréntesis de Lie sean cero. (Esto se puede lograr escogiendo un sistema de coordenadas en el cual las componentes de los campos sean constantes.) Entonces, $R_{XY}Z$ se reducirá a $D_Y(D_X Z) - D_X(D_Y Z)$.

- (1) Siempre que el paréntesis de Lie $[A, B] = AB - BA$ tiene sentido, es antisimétrico en A y B. Así que (1) es inmediato de la definición de curvatura.
- (2) Por la identidad de polarización (ver secciones 22 y 23), sólo tenemos que demostrar que $\langle R_{XY}V, V \rangle = 0$. Pero, usando (D5) (sección 36)

$$\begin{aligned}\langle R_{XY}V, V \rangle &= \langle D_Y D_X V, V \rangle - \langle D_X D_Y V, V \rangle \\ &= Y \langle D_X V, V \rangle - \langle D_X V, D_Y V \rangle - X \langle D_Y V, V \rangle + \langle D_Y V, D_X V \rangle \\ &= \frac{1}{2} YX \langle V, V \rangle - \frac{1}{2} XY \langle V, V \rangle = 0,\end{aligned}$$

ya que $[X, Y] = 0$.

- (3) Supongamos que $F: \mathcal{M}(M)^3 \rightarrow \mathcal{M}(M)$ es una función que sólo es \mathbb{R} -multilineal, y designemos por $\sum F(X, Y, Z)$ a la suma sobre las permutaciones cíclicas de X, Y, Z:

$$F(X, Y, Z) + F(Y, Z, X) + F(Z, X, Y).$$

Una permutación cíclica de X, Y, Z deja invariante a $\sum F(X, Y, Z)$. Por lo tanto,

$$\begin{aligned}\sum R_{XYZ} &= \sum D_Y D_X Z - \sum D_X D_Y Z \\ &= \sum D_X D_Z Y - \sum D_X D_Y Z = \sum D_X [Z, Y] = 0.\end{aligned}$$

- (4) Esta última verificación es un ejercicio en combinatoria. Por (3) tenemos que $\langle \sum R_{XYZ}, W \rangle = 0$, donde ahora \sum actúa sobre cualesquiera tres campos vectoriales que estén asociados con R. Sumemos sobre las cuatro permutaciones cíclicas de Y, V, X, W y después escribamos explícitamente cada una de las sumas \sum , con lo cual se obtienen doce términos. Usando (1) y (2), ocho de estos términos se cancelan por parejas, quedando

$$2\langle R_{XY}V, W \rangle + 2\langle R_{VW}X, Y \rangle = 0.$$

Por lo tanto, $\langle R_{XY}V, W \rangle = \langle R_{VW}X, Y \rangle$. ##

Las simetrías del tensor de curvatura R conducen a una menos obvia simetría para su diferencial covariante DR, que se conoce como la segunda identidad de Bianchi. Por definición, DR es un campo tensorial de tipo (1,4) que podemos reinterpretar como una función que asigna a cada cuatro campos vectoriales, un campo vectorial dado por $(D_Z R)_{XY}V = (D_Z R)(X, Y)V$. Como siempre, esto tiene sentido también para vectores tangentes individuales, de modo que cada uno de los términos de la identidad que daremos en seguida, es un operador lineal sobre $T_p(M)$.

Proposición 40.3 (Segunda identidad de Bianchi). Si $x, y, z \in T_p(M)$, entonces

$$(D_Z R)(x, y) + (D_X R)(y, z) + (D_Y R)(z, x) = 0.$$

Demostración. Como en la demostración anterior extendamos los vectores

x, y, z a campos vectoriales X, Y, Z sobre una vecindad de p . Esta vez tomaremos un sistema de coordenadas normal en el punto p , para el cual los campos vectoriales tengan componentes constantes. Así, no sólo se anulan todos los paréntesis de Lie de los campos, sino que en el punto p (en donde todos los símbolos de Christoffel se anulan) las nueve derivadas covariantes que involucran sólo a X, Y, Z se anulan, debido a nuestras suposiciones y a la fórmula (1) de la proposición 36.3. Usando la regla del producto, vemos que al aplicar $(D_{\bar{z}}R)(X, Y)$ a un campo V se obtiene

$$D_{\bar{z}}(R(X, Y)V) - R(D_{\bar{z}}X, Y)V - R(X, D_{\bar{z}}Y)V - R(X, Y)D_{\bar{z}}V.$$

En el punto p los dos términos de enmedio son cero, por lo que si ahora dejamos de escribir el campo vectorial superfluo V , obtenemos

$$(D_{\bar{z}}R)(X, Y) = [D_{\bar{z}}, R(X, Y)] = [D_{\bar{z}}, [D_Y, D_X]] \quad \text{en } p.$$

Pero la identidad de Jacobi es tan válida aquí como para el caso de los paréntesis de Lie de campos vectoriales. Así que sumando la fórmula anterior sobre las permutaciones cíclicas de X, Y, Z , obtenemos el resultado $\sum (D_{\bar{z}}R)(X, Y) = 0$ en p . ##

Lema 40.4 Sea x^1, \dots, x^n un sistema coordinado sobre la vecindad $U \subset M$, entonces

$$R_{\partial_k \partial_l}(\partial_j) = \sum_i R_{ijkl}^i \partial_i, \quad \text{sobre } U,$$

en donde las componentes de R están dadas por

$$R_{ijkl}^i = \frac{\partial}{\partial x^l} \Gamma_{jk}^i - \frac{\partial}{\partial x^k} \Gamma_{jl}^i + \sum_m \Gamma_{lm}^i \Gamma_{kj}^m - \sum_m \Gamma_{km}^i \Gamma_{lj}^m.$$

Demostración. Para los campos vectoriales coordinados, tenemos que $[\partial_k, \partial_l] = 0$, por lo cual,

$$R_{\partial_k \partial_l}(\partial_j) = D_{\partial_l}(D_{\partial_k} \partial_j) - D_{\partial_k}(D_{\partial_l} \partial_j).$$

El primer término en el lado derecho es

$$D_{\partial_l}(\sum_m \Gamma_{kj}^m \partial_m) = \sum_m \frac{\partial}{\partial x^l} \Gamma_{kj}^m \partial_m + \sum_{m,r} \Gamma_{kj}^m \Gamma_{lm}^r \partial_r.$$

Reetiquetando un par de índices mudos, se obtiene

$$\sum_i \left\{ \frac{\partial}{\partial x^l} \Gamma_{kj}^i + \sum_m \Gamma_{lm}^i \Gamma_{kj}^m \right\} \partial_i.$$

Entonces, restando la expresión correspondiente con k y l intercambiados, se obtiene el resultado. ##

Por medio de la fórmula (2) de la proposición 36.3, que da una expresión explícita para los símbolos de Christoffel, podemos obtener la correspondiente expresión explícita para las componentes del tensor de curvatura, en términos de las componentes del tensor métrico. Aún en casos simples, estos cálculos son muy tediosos y proporcionan un mínimo de información. La manera más práctica de calcular la curvatura de

una variedad semi-Riemanniana M , es usar resultados teóricos para aprovechar las características distintivas de la variedad M . Varios ejemplos de este procedimiento aparecerán más adelante.

41 CURVATURA SECCIONAL

El tensor de curvatura Riemanniano es bastante complicado; en esta sección consideraremos una función de valores reales mucho más simple y que sirve para determinar completamente a R .

Un subespacio bidimensional del espacio tangente $T_p(M)$ se llama un plano tangente a M en p . Para vectores tangentes v, w definamos

$$Q(v, w) = \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle^2.$$

En la sección 23 vimos que un plano tangente Π es no degenerado si y sólo si $Q(v, w) \neq 0$ para una y , entonces para todas las bases v, w de Π . El valor absoluto $|Q(v, w)|$ es el cuadrado del área del paralelogramo cuyos lados son v y w . $Q(v, w)$ es positiva si $g|_{\Pi}$ es definida y es negativa si $g|_{\Pi}$ es indefinida (para comprobar esta afirmación, tómese una base ortonormal para Π .)

Lema 41.1 Sea Π un plano tangente no degenerado a M en p . El número

$$K(v, w) = \langle R_{vw} v, w \rangle / Q(v, w)$$

es independiente de la elección de la base v, w para Π , y se llama la curvatura seccional $K(\Pi)$ de Π .

Demostración. Cualesquiera dos bases para Π están relacionadas por las ecuaciones

$$\begin{aligned} v &= ax + by \\ w &= cx + dy, \end{aligned}$$

donde el determinante de los coeficientes $ad - bc$ es distinto de cero. Un cálculo directo demuestra que

$$\langle R_{vw} v, w \rangle = (ad - bc)^2 \langle R_{xy} x, y \rangle,$$

y

$$Q(v, w) = (ad - bc)^2 Q(x, y). \quad \#\#$$

Así, la curvatura seccional K de M es una función real-valuada sobre el conjunto de los planos tangentes no degenerados de M .

Por definición, R determina a K ; para demostrar que K determina a R , necesitamos un resultado técnico acerca de los productos escalares indefinidos.

Lema 41.2 Dados vectores v, w en un espacio con producto escalar, existen vectores v', w' , arbitrariamente cercanos a v y a w , respectivamente, y que generan un plano no degenerado.

Demostración. Podemos suponer que v, w son linealmente independientes ya que cualquier par de vectores puede ser aproximado por un par de vectores linealmente independientes. Obviamente, podemos suponer que el plano generado por v y w es degenerado, de ahí que el producto escalar deba ser indefinido. Si v es un vector nulo, sea x un vector tal que $\langle v, x \rangle \neq 0$; si v es no nulo, tomemos un $x \neq 0$ con un caracter causal opuesto al de v . En ambos casos se tiene que $Q(v, x) < 0$.

Basta ahora con demostrar que para todo número $\delta \neq 0$ suficientemente pequeño, los vectores v y $w + \delta x$ generan un plano no degenerado. El desarrollo de $Q(v, w + \delta x)$ nos da una expresión de la forma

$$2b\delta + \delta^2 Q(v, x).$$

Si $b \neq 0$, esta expresión será distinta de cero, ya que para δ pequeño, el término que contiene a δ^2 prácticamente no cuenta. Si $b = 0$, la expresión tampoco es cero, ya que $Q(v, x) < 0$. ##

Proposición 41.3 Si $K = 0$ en $p \in M$, entonces $R = 0$ en p .

Demostración. En forma más explícita el enunciado de la proposición se puede hacer como sigue: si $K(\Pi) = 0$ para todo plano no degenerado en $T_p(M)$, entonces $R_{xyz} = 0$ para todo $x, y, z \in T_p(M)$. Consideraremos tres casos.

(1) $\langle R_{vw}v, w \rangle = 0$ para todo $v, w \in T_p(M)$.

Si v y w generan un plano no degenerado, entonces $\langle R_{vw}v, w \rangle = 0$. Pero por el lema anterior, cualquier par de vectores es un límite de tales vectores. Puesto que $\langle R_{vw}x, y \rangle$ es multilineal, debe ser continua en $T_p(M)^4$, de lo cual se sigue (1).

(2) $R_{vw}v = 0$ para todo $v, w \in T_p(M)$.

Para un x arbitrario tenemos

$$\langle R_{v, w+x}v, w+x \rangle = \langle R_{vw}v, w \rangle + \langle R_{vx}v, w \rangle + \langle R_{vw}v, x \rangle + \langle R_{vx}v, x \rangle.$$

Tres de estos términos se anulan por el resultado (1). La simetría por parejas asegura que los otros dos términos son iguales, de modo que $\langle R_{vw}v, x \rangle = 0$ para toda x . Esto último prueba el resultado (2).

(3) $R_{vw}x = R_{wx}v$ para todo $v, w, x \in T_p(M)$.

Ahora tenemos

$$R_{v+x, w}(v+x) = R_{vw}v + R_{xw}v + R_{vw}x + R_{xw}x.$$

Nuevamente, tres de estos términos se anulan por (2), y el resultado (3) se sigue por la antisimetría de R en sus subíndices.

Para concluir la demostración, observamos que según la afirmación (3), $R_{vw}x$ no cambia cuando se hace una permutación cíclica de los vectores v, w, x . Entonces, la primera identidad de Bianchi implica que $R_{vw}x = 0$ para todo v, w, x ; por lo tanto $R = 0$ en p . ##

Una vareidad semi-Riemanniana M para la cual el tensor de curvatura R es cero en todo punto se llama plana. De acuerdo con la proposición

anterior, M es plana si y sólo si la función de curvatura seccional K es idénticamente cero. Por ejemplo, todo espacio semi-Euclidiano \mathbb{R}^3 es plano: para las coordenadas naturales, todos los símbolos de Christoffel son cero, por lo cual $R = 0$, de acuerdo con el lema 40.4.

La demostración anterior realmente prueba algo más. Diremos que una función multilínea $F: T_p(M)^4 \rightarrow \mathbb{R}$ es de tipo curvatura si F tiene todas las simetrías que hemos demostrado que posee la función

$$(v, w, x, y) \rightarrow \langle R_{vw}x, y \rangle.$$

En la demostración anterior sólo usamos estas simetrías, por lo cual tenemos que para cualquier función de tipo curvatura F , si $F(v, w, v, w) = 0$ para todo $v, w \in T_p(M)$ que generan un plano no degenerado, entonces $F = 0$. Se sigue que K determina a R en el siguiente sentido.

Corolario 41.4 Sea F una función tipo curvatura sobre $T_p(M)$ tal que

$$K(v, w) = \frac{F(v, w, v, w)}{\langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle^2}$$

siempre que v y w generen un plano no degenerado. Entonces,

$$R_{vw}x, y = F(v, w, x, y)$$

para todo $v, w, x, y \in T_p(M)$.

Demostración. La función diferencia $\Delta(v, w, x, y) = F(v, w, x, y) - \langle R_{vw}x, y \rangle$ también es tipo curvatura. Por hipótesis, $\Delta(v, w, x, y) = 0$ si v y w generan un plano no degenerado. Entonces, por el comentario que antecede a este corolario, $\Delta = 0$. ##

Una variedad semi-Riemanniana M tiene curvatura constante si su función de curvatura seccional es constante.

El corolario anterior nos conduce a una fórmula simple para R cuando K es constante.

Corolario 41.5 Si M tiene curvatura constante $K = C$, entonces

$$R_{xy}z = C(\langle z, x \rangle y - \langle z, y \rangle x).$$

Demostración. Un cálculo rutinario demuestra que la fórmula

$$F(x, y, v, w) = C(\langle v, x \rangle \langle y, w \rangle - \langle v, y \rangle \langle x, w \rangle)$$

define una función de tipo curvatura en cada punto y , además, $F(x, y, x, y) = CQ(x, y)$. Así que si x y y generan un plano no degenerado,

$$K(x, y) = C = \frac{F(x, y, x, y)}{Q(x, y)},$$

y el resultado se sigue del corolario anterior. ##

En el caso Riemanniano, esta fórmula para la curvatura tiene una interpretación geométrica simple: si x, y es una base ortonormal para un

plano II, entonces R_{xy} es cero sobre Π^{\perp} , y sobre Π es la rotación que manda x a y y y a $-x$, seguida de la multiplicación por el escalar C .

42 SUPERFICIES SEMI-RIEMANNIANAS

Sea M una superficie semi-Riemanniana, es decir, una variedad semi-Riemanniana de dimensión 2. Para un sistema coordenado u, v sobre M , las componentes del tensor métrico se denotan tradicionalmente por

$$E = g_{11} = \langle \partial_u, \partial_u \rangle, \quad F = g_{12} = g_{21} = \langle \partial_u, \partial_v \rangle, \quad G = g_{22} = \langle \partial_v, \partial_v \rangle$$

donde para poder usar la notación con índices se hace $u = u^1$ y $v = u^2$. El elemento de línea será entonces,

$$ds^2 = Edu^2 + 2Fdu dv + Gdv^2,$$

$$\text{y } Q = Q(\partial_u, \partial_v) = EG - F^2.$$

Entonces, por la proposición 36.3, vemos que los símbolos de Christoffel están dados por

$$Q\Gamma_{11}^1 = \begin{vmatrix} E_{\mu}/2 & F \\ F_{\mu} - (E_{\nu}/2) & G \end{vmatrix}, \quad Q\Gamma_{11}^2 = \begin{vmatrix} E & E_{\mu}/2 \\ F & F_{\mu} - (E_{\nu}/2) \end{vmatrix},$$

$$Q\Gamma_{12}^1 = \begin{vmatrix} E_{\nu}/2 & F \\ G_{\mu}/2 & G \end{vmatrix}, \quad Q\Gamma_{12}^2 = \begin{vmatrix} E & E_{\nu}/2 \\ F & G_{\mu}/2 \end{vmatrix},$$

$$Q\Gamma_{22}^1 = \begin{vmatrix} F_{\nu} - (G_{\mu}/2) & F \\ G_{\nu}/2 & G \end{vmatrix}, \quad Q\Gamma_{22}^2 = \begin{vmatrix} E & F_{\nu} - (G_{\mu}/2) \\ F & G_{\nu}/2 \end{vmatrix}.$$

Las ecuaciones de las geodésicas se obtiene substituyendo estos símbolos de Christoffel en las ecuaciones

$$\begin{aligned} u'' + \Gamma_{11}^1 u' + 2\Gamma_{12}^1 u'v' + \Gamma_{22}^1 v'^2 &= 0 \\ v'' + \Gamma_{11}^2 u'^2 + 2\Gamma_{12}^2 u'v' + \Gamma_{22}^2 v'^2 &= 0. \end{aligned}$$

Como M es bidimensional, $T_p(M)$ es el único plano tangente en p . Entonces, la curvatura seccional K se reduce a una función real-valuada sobre M , llamada la curvatura Gaussiana de M . Encontrar fórmulas explícitas para K es complicado en el caso general, así que consideraremos un caso especial que es muy útil.

Proposición 42.1 Sean u, v un sistema ortogonal de coordenadas sobre una superficie semi-Riemanniana, de manera que $F = \langle \partial_u, \partial_v \rangle = 0$.

$$(1) \quad D_{\partial_u} \partial_u = \frac{E_{\mu}}{2E} \partial_u - \frac{E_{\nu}}{2G} \partial_v, \quad D_{\partial_v} \partial_v = -\frac{G_{\mu}}{2E} \partial_u + \frac{G_{\nu}}{2G} \partial_v,$$

$$D_{\partial_u} \partial_v = D_{\partial_v} \partial_u = \frac{E_{\nu}}{2E} \partial_u + \frac{G_{\mu}}{2G} \partial_v.$$

(2) Sea $e = |E|^{1/2}$ y sean $g = |G|^{1/2}$, ε_1 el signo de E y ε_2 el signo de G . Entonces,

$$K = -\frac{1}{e g} \left[\varepsilon_1 \left(\frac{g_{\mu}}{e} \right)_{\mu} + \varepsilon_2 \left(\frac{e_{\nu}}{g} \right)_{\nu} \right].$$

Demostración. (1) Haciendo $F = 0$ en las fórmulas anteriores para los símbolos de Christoffel, se obtiene el resultado.

(2) Por definición, $K = \langle R_{\partial_\mu \partial_\nu}(\partial_\mu), \partial_\nu \rangle / EG$. Para calcular el denominador de esta expresión, podemos usar

$$\langle D_{\partial_\mu} D_{\partial_\nu} \partial_\mu, \partial_\nu \rangle = \frac{\partial}{\partial u} \langle D_{\partial_\nu} \partial_\mu, \partial_\nu \rangle - \langle D_{\partial_\nu} \partial_\mu, D_{\partial_\mu} \partial_\nu \rangle,$$

la cual, por (1) se convierte en $G_{\mu\mu}/2 - (E_\nu)^2/4E - (G_\mu)^2/4G$. Hay otro término análogo. Entonces la curvatura tiene una fórmula que puede verificarse calculando su lado derecho. ##

43 CAMBIO DE TIPO Y CONTRACCION METRICA

En lenguaje tensorial, la proposición 36.1 (pag. 101) afirma que para una variedad semi-Riemanniana M , existe un isomorfismo $\mathcal{F}(M)$ -lineal entre $\mathcal{X}_0^1(M)$ y $\mathcal{X}_1^0(M)$. Este isomorfismo se puede extender fácilmente a espacios tensoriales de tipo arbitrario en la manera siguiente. Fijemos enteros $1 \leq a \leq r$ y $1 \leq b \leq s$. Si $A \in \mathcal{X}_s^r(M)$, entonces el valor de $\downarrow_b^a A \in \mathcal{X}_{s+1}^{r-1}(M)$ sobre uno-formas arbitrarias y campos vectoriales arbitrarios es

$$\begin{aligned} (\downarrow_b^a A)(\theta^1, \dots, \theta^{r-1}, X_1, \dots, X_{s+1}) &= \\ &= A(\theta^1, \dots, \underbrace{X_b^a, \dots, \theta^{r-1}, X_1, \dots, X_{b-1}, X_{b+1}, \dots, X_{s+1}}_{a\text{-ésima posición}}, \dots) \end{aligned}$$

donde X_b^a es la uno-forma métricamente equivalente a X_b . Así pues, en el lado derecho extraemos el b -ésimo campo vectorial e insertamos la uno-forma que le es métricamente equivalente en la a -ésima posición entre las uno-formas.

Por ejemplo, sea A un campo tensorial de tipo $(2,2)$. Entonces $B = \downarrow_2^1 A$ es el campo tensorial de tipo $(1,3)$ tal que $B(\theta, X, Y, Z) = A(Y^*, \theta, X, Z)$ para todas las uno-formas θ y campos vectoriales X, Y, Z . Para un sistema coordenado, la uno-forma dual a ∂_i es $\sum g_{ij} dx^j$. De donde se sigue que

$$B_{jkl}^i = B(dx^i, \partial_j, \partial_k, \partial_l) = A(\sum_m g_{km} dx^m, dx^i, \partial_j, \partial_l) = \sum_m g_{km} A_{jl}^{mi}.$$

Así pues, el operador hace uso del tensor métrico para convertir el primer superíndice en el segundo subíndice.

El resultado del operador $\downarrow_b^a : \mathcal{X}_s^r(M) \rightarrow \mathcal{X}_{s+1}^{r-1}(M)$ se conoce en la literatura clásica como bajar un índice. Se trata claramente de un operador $\mathcal{F}(M)$ -lineal y es, de hecho, un isomorfismo $\mathcal{F}(M)$ -lineal ya que posee un operador inverso \uparrow_b^a el cual, con una notación semejante a la anterior, extrae la a -ésima uno-forma e inserta el campo vectorial que le es métricamente equivalente en la b -ésima posición entre los campos vectoriales. Para un sistema coordenado, el campo vectorial métricamente equivalente a dx^i es $\sum g^{ij} \partial_j$. Si B es un tensor de tipo $(1,3)$, entonces

$$(\uparrow_b^a B)_{kl}^{ij} = \sum_q g^{iq} B_{kql}^j,$$

en donde el tensor métrico transforma el segundo subíndice en el primer superíndice. Así, la operación \uparrow_b^a se conoce clásicamente como subir un

índice. Como ejemplo, en un sistema coordinado, tenemos que

$$(\downarrow_2 \downarrow_1 A)_{kl}^{ij} = \sum_P g^{iP} (\downarrow_1 A)_{kPl}^j = \sum_{pm} g^{iP} \varepsilon_{pm} A_{kl}^{mj} = \sum_m \delta_m^i A_{kl}^{mj} = A_{kl}^{ij}.$$

Este cambio en el tipo de los tensores es tan natural que a veces ocurre en la práctica sin ser notado. Un caso importante es el de un tensor A de tipo (1, s) dado como una función $\mathcal{A}(M)$ -multilineal $A; \mathcal{A}(M)^s \rightarrow \mathcal{A}(M)$. Es entonces particularmente simple bajar el único índice contravariante a, digamos, la primera posición covariante:

$$(\downarrow_1 A)(V, X_1, \dots, X_s) = \langle V, A(X_1, \dots, X_s) \rangle.$$

De hecho, el lado izquierdo es por definición $A(V^*, X_1, \dots, X_s)$, el cual es, por la interpretación que dimos en la sección 16, $V^*(A(X_1, \dots, X_s)) = \langle V, A(X_1, \dots, X_s) \rangle$.

Todos los tensores que se obtienen a partir de uno dado mediante las operaciones de subir y bajar índices se llaman metricamente equivalentes. Todos ellos contienen la misma información y, por lo tanto, pueden ser considerados como diferentes manifestaciones del mismo objeto.

La versión clásica en términos de sistemas coordinados de la geometría diferencial multidimensional fue desarrollada mucho antes que su versión invariante y, para poder armonizar ambas versiones, hay que poner atención a algunos detalles. Cuando el tensor de curvatura $R: \mathcal{A}(M)^3 \rightarrow \mathcal{A}(M)$ se escribe en la forma ordinaria como una función de tres campos vectoriales, el patrón clásico de colocación de los índices que aparece en el lema 40.4 demanda que $R(Z, X, Y) = R_{XY}Z$. Las componentes del tensor $\downarrow_1 R$ están dadas entonces por

$$R_{ijkl} = (\downarrow_1 R)(\partial_i, \partial_j, \partial_k, \partial_l) = \langle \partial_l, R_{\partial_k \partial_l}(\partial_j) \rangle = \sum \varepsilon_{im} R_{jkl}^m.$$

Esta es la manera usual para bajar el índice contravariante de R, es te convenio se expresa en la notación clásica escribiendo R_{ijkl} en lugar de R_{jkl}^i .

Sobre una variedad suave, la contracción opera sobre un índice contravariante y sobre un índice covariante y convierte a un tensor de tipo (r, s) en uno de tipo (r-1, s-1). Pero en una variedad semi-Riemanniana podemos tomar la contracción métrica de dos índices covariantes, primero elevando cualquiera de los dos y después contrayendo de la manera usual. Así, para $1 \leq a < b \leq s$ y r arbitrario, la contracción métrica $C_{ab}: \mathcal{T}_s^r(M) \rightarrow \mathcal{T}_{s-2}^r(M)$ está dada en coordenadas por

$$(C_{ab} A)_{j_1 \dots j_{s-2}}^{i_1 \dots i_r} = \sum_{pq} g^{pq} A_{j_1 \dots j_{s-2}}^{i_1 \dots i_r p \dots q \dots j_{s-2}}.$$

Por ejemplo, si A es un tensor de tipo (1, 3), entonces

$$(C_{12} A)_j^i = \sum_{pq} g^{pq} A_{pqj}^i.$$

Similarmente, en el caso contravariante, para $1 \leq a < b \leq r$ y s arbitraria obtenemos

$$C^{ab}: \mathcal{T}_s^r(M) \rightarrow \mathcal{T}_s^{r-2}(M)$$

con una fórmula coordinada análoga a la anterior, pero cambiando el papel de los índices covariantes por los índices contravariantes (entonces

ces g_{ij} aparece en lugar de g^{ij}). Cuando no sea importante especificar los índices, todas las contracciones se denotarán por C.

Lema 43.1 Las derivadas covariantes D_Y y la diferencial covariante D conmutan con los operadores de elevación y de descenso y con las contracciones.

Demostración. Puesto que la operación de elevar un índice es la inversa de la operación de descenso del mismo índice, es suficiente considerar esta última. Por un argumento simple de permutaciones vemos que sólo es necesario considerar \downarrow_i^a . La expresión coordinada para $\downarrow_i^a A$ muestra que este tensor es la contracción ordinaria C_i^a aplicada a $g \otimes A$. Siendo una derivación tensorial, D_Y conmuta con la contracción ordinaria. Puesto que el tensor métrico es paralelo,

$$D_Y(\downarrow_i^a A) = D_Y(C_i^a(g \otimes A)) = C_i^a(g \otimes D_Y A) = \downarrow_i^a(D_Y A).$$

Por lo tanto, D_Y también conmuta con la contracción métrica.

Cálculos formales semejantes, demuestran los resultados correspondientes para la diferencial covariante D. ##

44 CAMPOS DE REFERENCIALES

Una base ortonormal para un espacio tangente $T_p(M)$ se llama una referencial de M en p. Si $n = \dim M$ entonces un conjunto de n campos vectoriales unitarios y mutuamente ortogonales se llama un campo referencial, ya que asigna una referencial a cada punto. Por ejemplo, sobre \mathbb{R}^n los campos vectoriales de las coordenadas naturales forman un campo referencial.

En general, no siempre existe un campo referencial definido sobre toda la variedad M , pero veremos en un momento que ellos siempre existen localmente. Por medio de un desarrollo ortonormal, cualquier campo vectorial V puede expresarse en términos de un campo referencial como sigue:

$$V = \sum \epsilon_i \langle V, E_i \rangle E_i, \text{ donde } \epsilon_i = \langle E_i, E_i \rangle.$$

De donde,

$$\langle V, W \rangle = \sum \epsilon_i \langle V, E_i \rangle \langle W, E_i \rangle.$$

En el origen o de un sistema normal de coordenadas los vectores coordinados son ortonormales. Se sigue que mientras sólo estén involucradas operaciones puntuales, las fórmulas para los campos referenciales serán consecuencia de las correspondientes fórmulas con coordenadas. Por ejemplo, consideremos la contracción métrica C_{ab} de $A \in \mathcal{X}_5^0(M)$. Con respecto a un campo referencial,

$$(C_{ab} A)(X_1, \dots, X_{s-2}) = \sum \epsilon_m A(X_1, \dots, \overset{\text{LUGAR b}}{E_m}, \dots, X_{s-2}).$$

LUGAR a

Para demostrar esta ecuación tensorial basta trabajar en un sólo punto

o, el origen de un sistema normal de coordenadas tal que $\partial_i|_o = E_i|_o$. Debido a la multilinealidad, basta con que los campos X_i sean los ∂_i . Pero entonces la fórmula se sigue de la expresión coordenada para C_{ab} , ya que en el punto o, tanto g_{ij} como g^{ij} se convierten en δ_{ij} y δ^{ij} .

Similarmente, para un campo tensorial del tipo (1,s) dado por $A: \mathcal{H}(M)^s \rightarrow \mathcal{H}(M)$, tenemos

$$(C^b_a)(X_1, \dots, X_{s-1}) = \sum_m E_m \langle E_m, A(X_1, \dots, \overset{\text{LUGAR } b}{E_m}, \dots, X_{s-1}) \rangle.$$

Estos comentarios son válidos sólo para el álgebra tensorial, ya que en el análisis tensorial la ventaja de $\langle E_i, E_j \rangle = \delta_{ij}$ sobre $\langle \partial_i, \partial_j \rangle = g_{ij}$ queda anulada por la desventaja de que, a diferencia de $[\partial_i, \partial_j]$, los paréntesis de Lie $[E_i, E_j]$ no tienen por qué ser cero.

Un campo referencial sobre una curva $\alpha: I \rightarrow M$ es un conjunto de campos vectoriales unitarios E_1, \dots, E_n sobre α que son mutuamente ortogonales. También podemos escoger a los campos $E_i \in \mathcal{H}(\alpha)$ de tal manera que sean paralelos.

Corolario 44.1 Si $\alpha: I \rightarrow M$ es una curva y e_1, \dots, e_n es una referencial en $\alpha(0)$, entonces existe un único campo referencial paralelo a lo largo de α tal que $E_i(0) = e_i$ para $1 \leq i \leq n$.

Demostración. Sabemos que existe para cada i un campo vectorial paralelo a lo largo de α tal que $E_i(0) = e_i$. Puesto que el transporte paralelo a cualquier $t \in I$ es una isometría lineal, E_1, \dots, E_n es efectivamente un campo referencial paralelo. ##

Se sigue que sobre M existen localmente los campos referenciales. En efecto, dada cualquier referencial e_1, \dots, e_n en un espacio tangente $T_p(M)$, tomemos una vecindad coordenada normal U de o y extendamos la referencial dada a un campo referencial E_1, \dots, E_n sobre U por transporte paralelo a lo largo de las geodésicas radiales. La teoría de ecuaciones diferenciales garantiza que los campos vectoriales E_i son suaves.

La triple ventaja de ortonormalidad, paralelismo y definición global hacen que el uso de campos referenciales paralelos sobre una curva sea superior al uso de coordenadas.

45 ALGUNOS OPERADORES DIFERENCIALES

Sobre una variedad semi-Riemanniana M se pueden dar las generalizaciones naturales de los conocidos operadores del cálculo vectorial de \mathbb{R}^n , tales como la divergencia, el gradiente y el laplaciano.

El gradiente $\text{grad } f$ de una función $f \in \mathcal{F}(M)$ es el campo vectorial métricamente equivalente a la diferencial $df \in \mathcal{H}^*(M)$. Así,

$$\langle \text{grad } f, X \rangle = df(X) = Xf \quad \text{para todo } X \in \mathcal{H}(M).$$

En términos de un sistema coordinado, $df = \sum (\partial f / \partial x^i) dx^i$, por lo que

$$\text{grad } f = \sum_{i,j} g^{ij} \frac{\partial f}{\partial x^i} \partial_j.$$

En particular, para las coordenadas naturales de un espacio semi-Riemanniano, tenemos $\text{grad } f = \sum \varepsilon_i (\partial f / \partial u^i) \partial_i$, la cual se reduce a la fórmula usual en \mathbb{R}^3 .

Para un tensor A , la contracción del nuevo índice covariante que aparece en su diferencial covariante DA con uno de sus índices originales se llama divergencia $\text{div } A$ de A . Consideraremos dos casos especiales, para los cuales la divergencia es única.

(1) Si V es un campo vectorial, entonces $\text{div } V = C(DV) \varepsilon \in \mathcal{F}(M)$. Así, para un campo referencial, tenemos

$$\text{div } V = \sum \varepsilon_i \langle D_{E_i} V, E_i \rangle,$$

y para un sistema coordinado,

$$\text{div } V = \sum V^i{}_{;i} = \sum_i \left\{ \frac{\partial V^i}{\partial x^i} + \sum_j \Gamma_{ij}^i V^j \right\}.$$

Por lo tanto, para las coordenadas naturales de \mathbb{R}_v^n , $\text{div } V = \sum \partial V^i / \partial u^i$, la cual sobre \mathbb{R}^3 es la fórmula usual de la divergencia.

(2) Si A es un tensor simétrico de tipo $(0,2)$, entonces $\text{div } A = C_{13}(DA) = C_{23}(DA) \varepsilon \in \mathcal{K}^*(M)$.

Para un campo referencial, $(\text{div } A)(X) = \sum \varepsilon_i (D_{E_i} A) E_i(X)$, mientras que para coordenadas, tenemos

$$(\text{div } A)_i = \sum_{r,s} g^{rs} A_{r;s} = \sum_s A^s{}_{;s}.$$

El hessiano de una función $f \in \mathcal{F}(M)$ es su segunda diferencial covariante $H^f = D(Df)$.

Lema 45.1 El hessiano H^f de f es un campo tensorial simétrico de tipo $(0,2)$ tal que

$$H^f(X, Y) = XYf - (D_X Y)f = \langle D_X(\text{grad } f), Y \rangle.$$

Demostración. Puesto que $Df = df$,

$$\begin{aligned} H^f(X, Y) &= D(df)(X, Y) = D_Y(df)(X) = Y(df(X)) - df(D_Y X) \\ &= YXf - (D_Y X)f. \end{aligned}$$

Como $XY - YX = [X, Y] = D_X Y - D_Y X$, podemos cambiar X por Y en la fórmula anterior, mostrando con esto que H^f es simétrico. Finalmente,

$$\langle D_X(\text{grad } f), Y \rangle = X \langle \text{grad } f, Y \rangle - \langle \text{grad } f, D_X Y \rangle = H^f(X, Y). \quad \#\#$$

El laplaciano Δf de una función $f \in \mathcal{F}(M)$ es la divergencia de su gradiente: $\Delta f = \text{div}(\text{grad } f) \in \mathcal{F}(M)$.

Lema 45.2 El laplaciano de f es la contracción métrica de su hessiano.

Demostración. Como la diferencial covariante conmuta con el operador que eleva índices, tenemos que

$$\Delta f = \text{div}(\text{grad } f) = \text{CD}(\text{grad } f) = \text{CD}(\uparrow; \text{df}) = \text{C}\uparrow; \text{Ddf} = (\text{C}\uparrow;)\text{H}^f = \text{C}_{1,2}(\text{H}^f). \quad \#\#$$

Para un sistema coordenado, las componentes del hessiano se pueden obtener del lema 45.1. Así, el laplaciano de f tiene la siguiente expresión coordenada

$$\Delta f = \sum_{i,j} g^{ij} H_{ij} = \sum_{i,j} g^{ij} \left\{ \frac{\partial^2 f}{\partial x^i \partial x^j} - \sum_k \Gamma_{ij}^k \frac{\partial f}{\partial x^k} \right\}.$$

Para las coordenadas naturales de (\mathbb{R}^3) , las componentes de H^f son simplemente las segundas derivadas parciales $\partial^2 f / \partial u^i \partial u^j$, y $\Delta f = \sum \varepsilon_i \partial^2 f / \partial (u^i)^2$ la cual se reduce a la fórmula usual para \mathbb{R}^3 .

46 CURVATURA DE RICCI Y CURVATURA ESCALAR

Veremos en esta sección que la contracción del tensor de curvatura de Riemann nos lleva a obtener nuevos y más simples invariantes de curvatura.

Sea R el tensor de curvatura Riemanniana de M . El tensor de curvatura de Ricci Ric de la variedad M es la contracción $\text{C}_3^1(R) \in \mathfrak{F}_2^0(M)$, cuyas componentes con respecto a un sistema coordenado son $\text{Ric}_{ij} = \sum R_{ijm}^m$.

Debido a las simetrías de R , las únicas contracciones de R que son diferentes de cero son $\pm \text{Ric}$.

Lema 46.1 El tensor de curvatura de Ricci Ric es simétrico, y está dado en un campo referencial por

$$\text{Ric}(X, Y) = \sum_m \varepsilon_m \langle R_{X E_m} Y, E_m \rangle,$$

en donde, como siempre, $\varepsilon_m = \langle E_m, E_m \rangle$.

Demostración. Como hicimos notar antes, la notación clásica de los índices, nos obliga a escribir $R(X, Y, E_m) = R_{Y E_m} X$. Así,

$$\text{Ric}(X, Y) = (\text{C}_3^1 R)(X, Y) = \sum \varepsilon_m \langle E_m, R(X, Y, E_m) \rangle = \sum \varepsilon_m \langle R_{Y E_m} X, E_m \rangle.$$

Entonces, la simetría por parejas de R nos da la fórmula requerida y muestra que Ric es simétrico. #\#

Si el tensor de Ricci es idénticamente cero, se dice que M es plana según Ricci. Una variedad plana es plana según Ricci, pero veremos después que la afirmación inversa no es cierta en general.

Puesto que la curvatura seccional determina al tensor de curvatura R , también debe determinar a Ric . Debido a la simetría de Ric , podemos usar la identidad de polarización para reconstruir los valores de Ric en cada punto p , a partir de sus valores $\text{Ric}(u,u)$ sobre los vectores unitarios en p . Pero si e_1, \dots, e_n es una referencial en p tal que $u = e_i$, entonces, por el lema anterior,

$$\text{Ric}(u,u) = \sum \varepsilon_m \langle R_{u e_m}(u), e_m \rangle = \langle u, u \rangle \sum K(u, e_m).$$

Así pues, $R(u,u)$ es, excepto por el signo $\langle u, u \rangle = \pm 1$, la suma de las curvaturas seccionales de cualesquiera $n-1$ planos no degenerados ortogonales que pasan por el vector u .

La curvatura escalar S de M es la contracción $G(\text{Ric}) \in \mathcal{F}(M)$ de su tensor de Ricci.

En coordenadas, tenemos

$$S = \sum g^{ij} R_{ij} = \sum g^{ij} R_{ijk}^k.$$

La contracción con respecto a un campo referencial da

$$S = \sum_{i \neq j} K(E_i, E_j) = 2 \sum_{i < j} K(E_i, E_j).$$

La siguiente consecuencia de la segunda identidad de Bianchi es de gran importancia en la formulación matemática de la teoría general de la relatividad.

Corolario 46.2 $dS = 2 \text{div Ric}$.

Demostración. Si σ denota la suma sobre las permutaciones cíclicas de X, Y, Z , la segunda identidad de Bianchi es $\sigma(D_Z R)_{XY} = 0$. Para obtener su correspondiente expresión coordenada, apliquémosla a

$$(D_{\partial_r} R)_{\partial_k \partial_l}(\partial_j) = \sum R_{jkl;r}^i \partial_i$$

para obtener

$$R_{jkl;r}^i + R_{jlr;k}^i + R_{jrk;l}^i = 0.$$

Cambiando las posiciones de r y k en el tercer término, hay un cambio de signo, y tomando la contracción con respecto a i y a r , obtenemos

$$\sum_r R_{jkl;r}^r + \sum_r R_{jlr;k}^r - \sum_r R_{jkr;l}^r = 0,$$

la cual se escribe como

$$\sum_r R_{jkl;r}^r + R_{jl;k} - R_{jk;l} = 0.$$

Ahora tomando la contracción métrica con respecto a j y a k :

$$\sum_{r,i,j,k} g^{jk} R_{jkl;r}^r + \sum g^{jk} R_{jl;k} - S_{;l} = 0.$$

En el primer término, podemos escribir R_{jkl}^r como $\sum_l g^{lm} R_{mjkl}$ y usar las simetrías de R para expresar este término como $\sum R_{l;m}^m$. Por tanto,

$$2 \sum R_{l;m}^m = S; l.$$

El lado izquierdo de esta ecuación es la expresión coordinada para $2\text{div}(\text{Ric})$; el lado derecho es la expresión coordinada de $DS = dS$. ##

Una variedad semi-Riemanniana M se llama variedad de Einstein si es tal que $\text{Ric} = cg$, para una cierta constante C .

Proposición 46.3 Si M es conexa, $n = \dim M \geq 3$ y $\text{Ric} = fg$, con $f \in \mathcal{F}(M)$, entonces M es una variedad de Einstein.

Demostración. Tomando la contracción de la ecuación $\text{Ric} = fg$, obtenemos

$$S = C(\text{Ric}) = C(fg) = fC(g) = nf,$$

de donde, $dS = ndf$. Por otra parte,

$$\text{div}(\text{Ric}) = \text{div}(fg) = C_{13}(D(fg)) = C_{13}((df) \otimes g),$$

ya que $D(fg) = (df) \otimes g + f \otimes Dg$, pero $Dg = 0$, pues g es paralelo.

Así pues, $(dS)_i = n \partial f / \partial x^i$ y $\text{div}(\text{Ric})_i = \partial f / \partial x^i$. Por lo tanto, del corolario 46.2 obtenemos

$$(n-2) \frac{\partial f}{\partial x^i} = 0 \quad \text{y como } n \geq 3, \quad \frac{\partial f}{\partial x^i} = 0 \text{ sobre cualquier sistema coordinado de } M.$$

Como M es conexa, las relaciones anteriores implican que f es constante sobre M . ##

Proposición 46.4 Si M es una variedad de Einstein de dimensión 3, entonces M tiene curvatura seccional constante.

Demostración. Sea e_1, e_2, e_3 una base ortonormal para $T_p(M)$. Sabemos que (ver pag. anterior)

$$\text{Ric}(e_1, e_1) = \varepsilon_1 K(e_1, e_2) + \varepsilon_1 K(e_1, e_3),$$

$$\text{Ric}(e_2, e_2) = \varepsilon_2 K(e_2, e_1) + \varepsilon_2 K(e_2, e_3),$$

$$\text{Ric}(e_3, e_3) = \varepsilon_3 K(e_3, e_1) + \varepsilon_3 K(e_3, e_2),$$

en donde $\varepsilon_i = \langle e_i, e_i \rangle$ y $K(e_i, e_j) = K(e_j, e_i)$ es la curvatura seccional para el plano no degenerado generado por e_i y e_j . Así pues,

$$\text{Ric}(e_1, e_1) + \text{Ric}(e_2, e_2) - \text{Ric}(e_3, e_3) = 2K(e_1, e_2) = 2K(p).$$

Pero como $\text{Ric} = cg$, tenemos que

$$2K(p) = c \langle e_1, e_1 \rangle + c \langle e_2, e_2 \rangle - c \langle e_3, e_3 \rangle = c(\varepsilon_1 + \varepsilon_2 - \varepsilon_3),$$

lo cual demuestra la constancia de K . ##

Lema 46.5 (Schur) Si M es una variedad semi-Riemanniana conexa con $\dim M \geq 3$, y para cada $p \in M$, K es constante sobre los planos no degenera-

dos de $T_p(M)$, entonces K es constante.

Demostración. Sea e_1, \dots, e_n una referencial en p tal que $e_i = u$, entonces (ver pag. 128),

$$\text{Ric}(u, u) = \langle u, u \rangle \sum_i K(u, e_i).$$

Pero como K no depende del plano no degenerado que se tome en $T_p(M)$,

$$\text{Ric}(u, u) = nK(p)\langle u, u \rangle.$$

Entonces, M es tal que $\text{Ric} = fg$, con $f \in \mathcal{F}(M)$ dada por $f(p) = nK(p)$. Se obtiene el resultado aplicando la proposición 46.3. ##

47 ISOMETRIAS LOCALES

Los conceptos de la geometría semi-Riemanniana que hemos definido hasta ahora son invariantes isométricos: cada uno es, en un sentido apropiado, preservado por isometrías. Esto es fácil de comprender, ya que hemos usado en su construcción las herramientas de la teoría de variedades diferenciables y el tensor métrico, y tanto la estructura diferenciable como el tensor métrico son preservados por las isometrías. Por ejemplo, la conexión de Levi-Civita es preservada en el siguiente sentido.

Proposición 47.1 Si $\phi: M \rightarrow N$ es una isometría, entonces $d\phi(D_X Y) = D_{d\phi X}(d\phi Y)$ para todo $X, Y \in \mathcal{K}(M)$.

Demostración. Recordemos que el valor del campo vectorial transferido $d\phi(X)$ en el punto $\phi(p)$ es $d\phi(X_p)$.

Como ϕ es un difeomorfismo, para cualquier $p \in M$ existen sistemas coordenados ξ en p y η en $\phi(p)$ que son preservados por ϕ , es decir, $y^i(\phi q) = x^i(q)$ para q en una vecindad de p . Se sigue inmediatamente que los campos vectoriales coordenados y las derivadas parciales son preservados por ϕ . Entonces, las componentes de X y de $Y = d\phi(X)$ son preservadas ya que

$$Y^i(\phi q) = Y_{\phi q}^i(y^i) = (d\phi X_q)^i y^i = X_q(y^i \circ \phi) = X_q(x^i) = X^i(q).$$

Como ϕ es también una isometría, las componentes del tensor métrico son preservadas: $g_{ij}^N(\phi q) = g_{ij}^M(q)$. Por lo tanto, las fórmulas que dan los símbolos de Christoffel en términos de las derivadas parciales de las componentes del tensor métrico, muestran que la conexión de Levi-Civita es preservada. ##

El carácter local de esta demostración sugiere que pueden existir resultados análogos de invariancia para aplicaciones de tipo más general que las isometrías.

Una aplicación suave $\phi: M \rightarrow N$ entre variedades semi-Riemannianas se

llama una isometría local si cada aplicación diferencial $d\phi: T_p(M) \rightarrow T_{\phi(p)}(N)$ es una isometría lineal.

En vista del teorema de la función inversa, podemos reformular la definición anterior en la siguiente forma: Cada punto $p \in M$ tiene una vecindad U tal que $\phi|U$ es una isometría de U sobre una vecindad de $\phi(p)$ en N . Esta formulación, equivalente a la anterior, es la que justifica el nombre de isometría local.

Ejemplo. Sea S^1 el círculo unitario en \mathbb{R}^2 . La aplicación exponencial $\exp: \mathbb{R} \rightarrow S^1$ enrolla a la línea recta alrededor del círculo mediante $t \rightarrow (\cos t, \sin t)$. Considerando a S^1 como una subvariedad Riemanniana de \mathbb{R}^2 , vemos que \exp es una isometría local. (Basta verificar que como curva en \mathbb{R}^2 , \exp tiene velocidad unitaria.)

De acuerdo con el criterio anterior para las isometrías, vemos que todo objeto de carácter local (o puntual) que es preservado por isometrías, será preservado automáticamente por las isometrías locales. Si $\phi: M \rightarrow N$ es una isometría local, estos son algunos ejemplos:

(1) La derivada covariante inducida. Si Y es un campo vectorial sobre una curva α en M , entonces $(d\phi Y)' = d\phi(Y')$, donde $(d\phi Y)(s) = d\phi(Y(s))$ para toda s . (La demostración es una ligera variante de la de la proposición 47.1.)

(2) Transporte paralelo. Sea P la aplicación transporte paralelo de $\alpha(a)$ a $\alpha(b)$ a lo largo de la curva α en M . Sea \bar{P} la aplicación transporte paralelo de $\phi\alpha(a)$ a $\phi\alpha(b)$ a lo largo de $\phi\alpha$ en N . Entonces, $d\phi_{\alpha(b)} \circ P = \bar{P} \circ d\phi_{\alpha(a)}$. Por lo tanto, se sigue que:

(3) Geodésicas. Si γ es una geodésica en M , entonces $\phi \circ \gamma$ es una geodésica en N . Así, $\phi \circ \gamma'_v = \gamma'_v | I_v$, ya que en ambos miembros tenemos geodésicas con la misma velocidad inicial. De aquí se sigue que:

(4) Aplicaciones exponenciales. $\phi \circ \exp_p = \exp_{\phi(p)} \circ d\phi_p$, siempre que alguno de los dos lados de esta ecuación esté definido.

(5) Tensores de curvatura Riemanniana. $d\phi(R(x,y)z) = R(d\phi x, d\phi y)(d\phi z)$. Esto se sigue inmediatamente de la definición, ya que una isometría local preserva tanto los paréntesis de Lie como las conexiones de Levi-Civita. De aquí tenemos que:

(6) Curvatura seccional. $K_N(d\phi II) = K_M(II)$ para todos los planos II no degenerados sobre M . Curvatura de Ricci: $\phi^*(Ric_N) = Ric_M$. Curvatura escalar: $S_N \circ \phi = S_M$. El último resultado usa el hecho de que la operación de contracción métrica es de naturaleza puntual y es, por lo tanto, preservada por isometrías locales.

Una isometría local está determinada de manera única por su aplicación diferencial en un sólo punto.

Proposición 47.2 Sean $\phi, \psi: M \rightarrow N$ isometrías locales en donde M es una variedad semi-Riemanniana conexa. Si existe un punto $p \in M$ tal que $d\phi_p =$

$= d\psi_p$, entonces $\phi = \psi$.

Demostración. Sea $A = \{q \in M \mid d\phi_q = d\psi_q\}$. Por continuidad, A es un conjunto cerrado en M . Como por hipótesis A es no vacío, basta con demostrar que A es abierto. Afirmamos que si $q \in A$, entonces cualquier vecindad normal U de q está contenida en A . Si $r \in U$, existe un vector $v \in T_q(M)$ tal que $\gamma_v(1) = \exp_q(v) = r$. Por lo tanto,

$$\phi(r) = \phi(\gamma_v(1)) = \gamma_{d\phi_v}(1) = \gamma_{d\psi_v}(1) = \psi(\gamma_v(1)) = \psi(r).$$

Así, $\phi = \psi$ sobre U ; por lo tanto, $d\phi_r = d\psi_r$ para todo $r \in U$. ##

Una aplicación suave $\psi: M \rightarrow N$ entre variedades semi-Riemannianas se llama conforme si $\psi^*(g_N) = h g_M$ para alguna función $h \in \mathcal{F}(M)$ tal que $h > 0$ o $h < 0$. El siguiente caso especial es muy útil.

Un difeomorfismo $\psi: M \rightarrow N$ entre variedades semi-Riemannianas tal que $\psi^*(g_N) = c g_M$ para alguna constante $c \neq 0$, se llama una homotecia de coeficiente c .

Así, $\langle d\psi(v), d\psi(w) \rangle = c \langle v, w \rangle$ para todo $v, w \in T_p(M)$, $p \in M$: todos los productos escalares son "alargados" por la misma constante c . Si $c > 0$ [$c < 0$], entonces ψ se llama una homotecia positiva [negativa] de factor escalar $|c|^{1/2}$. Por ejemplo, en $(\mathbb{R}^{n+1})^{1/2}$ la multiplicación de los vectores por el escalar b/a es una homotecia positiva entre $S^n(a)$ y $S^n(b)$ con factor escalar b/a .

Vemos que una isometría es simplemente una homotecia con $c = 1$. Si $c = -1$, diremos que ψ es una anti-isometría.

Lema 47.3 Las homotecias preservan la conexión de Levi-Civita.

Demostración. Si $\psi: M \rightarrow N$ es una homotecia con coeficiente c , designemos por N^* a la variedad suave N provista con el nuevo tensor métrico $c g_N$. Entonces, $\psi: M \rightarrow N^*$ es una isometría y, por lo tanto, preserva a la conexión de Levi-Civita. Así, sólo falta demostrar que $c g_N$ y g_N determinan la misma conexión de Levi-Civita. Pero esto último se sigue inmediatamente de la fórmula de Koszul, ya que el tensor métrico aparece exactamente una vez en cada uno de sus términos y, por lo tanto, se cancela el coeficiente c . ##

(1) Puesto que una homotecia preserva la conexión de Levi-Civita D , debe también preservar todas las nociones que se derivan sólo de D , como es el caso de la conexión inducida sobre una curva, el transporte paralelo, las geodésicas, la curvatura Riemanniana R , y la curvatura de Ricci. (Lo último se debe a que Ric se obtiene de R mediante una contracción no métrica.)

(2) En contraste, las curvaturas seccional y escalar no son invariantes bajo homotecias. De hecho, si $\psi: M \rightarrow N$ tiene coeficiente c , es fácil verificar que

$$K_N(d\psi II) = \frac{1}{c} K_M(II) \quad y \quad S_N \circ \psi = \frac{1}{c} S_M.$$

(3) Es claro que un coeficiente de homotecia $c > 0$ preserva el carácter causal de los vectores tangentes (y de las curvas). Sin embargo, si $c < 0$, entonces el carácter causal es revertido: v temporal $\Rightarrow d\psi v$ espacial; v espacial $\Rightarrow d\psi v$ temporal, y v nulo $\Rightarrow d\psi v$ nulo.

La operación de transformar la variedad semi-Riemanniana M con tensor métrico g , en la misma variedad suave con el nuevo tensor métrico $-g$, se llama revertir la métrica de M . El efecto que esta operación tiene sobre la geometría de la variedad M debe quedar claro de los comentarios anteriores. Debido a ésto, podemos preferir considerar variedades Riemannianas en lugar de variedades con métrica definida negativa. Análogamente, el caso de índice $\nu = n-1 \geq 1$ puede ser considerado como si fuera el caso de la geometría Lorentziana $\nu = 1$.

48 VARIEDADES CON CONEXION

Básica en toda la matemática es la noción, que en estas notas hemos usado informalmente, de conjunto con estructura. Para cada tipo de estructura matemática, hay una noción correspondiente de equivalencia (o isomorfismo), que es una biyección que, en un sentido adecuado, preserva la estructura. Cada tipo de estructura define una rama de la matemática como el estudio de aquellos conceptos que son preservados por la correspondiente noción de equivalencia. Por ejemplo, un grupo es un conjunto con la estructura operación del grupo. La noción de equivalencia es la noción usual de isomorfismo de grupos y la rama de la matemática que estudia aquellos conceptos que son invariantes bajo los isomorfismos es la teoría de grupos.

En el caso de la geometría semi-Riemanniana, tenemos toda una jerarquía de estructuras:

Rama de la matemática	Conjunto con estructura	Estructura	Equivalencia
Teoría de conjuntos	Conjunto	(Ninguna)	Biyección
Topología	Espacio topológico	Topología	Homeomorfismo
Teoría de variedades	Variedad Diferenciable	Estructura Diferenciable, Topología	Difeomorfismo
Geometría semi-Riemanniana	Variedad semi-Riemanniana	Topología, Atlas, Tensor métrico	Isometría

En esta sección consideraremos brevemente variedades con una estructura que puede considerarse intermedia entre la estructura diferenciable determinada por un atlas y la estructura semi-Riemanniana determinada por el tensor métrico. Estas serán las variedades con conexión.

Si bien en la teoría de la relatividad los objetos de interés son las variedades semi-Riemannianas γ , más concretamente, las variedades de Lorentz con índice $\nu = 1$ y dimensión 4, hay varias teorías físicas que pretenden generalizar a la relatividad, y en las cuales es importante el estudio de las variedades con conexión. Un ejemplo famoso es la "teoría del campo no simétrico", presentada por Einstein (en 1954) como un intento de unificar la gravitación y el electromagnetismo en una sola teoría. (Ver [E], en la bibliografía.)

Actualmente, son de interés para los físicos, variedades con otras estructuras diferentes de la semi-Riemanniana, tales como las variedades con estructura conforme (determinada por las geodésicas nulas), las variedades con estructura proyectiva (determinada por las geodésicas temporales), o los espacios de Weyl y de Finsler, etc. (Ver, por ejemplo, [E, P y S].)

Ya vimos en la sección 36 la manera en que la métrica g de una variedad semi-Riemanniana determina una conexión especial, la conexión de Levi-Civita, sobre la variedad. En esta sección estudiaremos variedades diferenciables provistas con una conexión general que no depende de la existencia de un tensor métrico.

Sea M una variedad n -dimensional suave. Una conexión (conexión afín o derivación covariante) en M , es una función $D: \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$ que a cada par de campos vectoriales $X, Y \in \mathfrak{X}(M)$ les asigna un campo vectorial $D_X Y \in \mathfrak{X}(M)$, y que satisface las siguientes propiedades:

- (1) $D_X(Y + Z) = D_X Y + D_X Z$
- (2) $D(X + Y)(Z) = D_X Z + D_Y Z$
- (3) $D_{fX} Y = f D_X Y$
- (4) $D_X(fY) = (Xf)Y + f D_X Y$,

en donde $X, Y, Z \in \mathfrak{X}(M)$ y $f \in \mathcal{F}(M)$.

Una variedad con conexión es una pareja ordenada (M, D) , en donde M es una variedad diferenciable y D es una conexión sobre M . Esta distinción es necesaria, pues sobre la misma variedad se pueden dar dos conexiones distintas, obteniéndose así dos objetos diferentes.

Las propiedades anteriores implican que $D_X Y$ es un tensor con respecto a X , pero $D_X Y$ no es un tensor con respecto a Y . Debido a esto, el vector $(D_X Y)_m$, en cada punto $m \in M$, depende únicamente de X_m y de los valores de Y sobre una curva en M cuya velocidad en m sea X_m .

Sea E_1, \dots, E_n un campo referencial definido en una vecindad de $m \in M$. Entonces, en dicha vecindad podemos escribir $X_m = \sum a^i(m) E_i|_m$ e $Y = \sum b^i E_i$.

Usando las propiedades de D , tenemos

$$(D_X Y)_m = (D_X(\sum b^j E_j))_m = \sum \{ (X_m b^j) E_j|_m + b^j(m) \sum a^i(m) (D_{E_i} E_j)_m \}.$$

Así pues, $a^i(m)$, $b^j(m)$ y $X_m b^j$ determinan a $D_X Y$ completamente si los campos $D_{E_i} E_j$ son conocidos. Sobre el dominio del campo referencial E_i , \dots, E_n podemos escribir

$$D_{E_i} E_j = \sum_k G_{ij}^k E_k.$$

Las funciones $G_{ij}^k \in \mathcal{F}(M)$ se llaman componentes de la conexión D con respecto al campo referencial E_i . En el caso de la conexión de Levi-Civita, las componentes son los símbolos de Cristoffel.

Sea σ una curva en M cuyo campo de velocidades es T . Un campo vectorial $Y \in \mathcal{K}(\sigma)$ sobre σ es paralelo a lo largo de σ si $D_T Y = 0$ sobre σ . La curva σ se llama geodésica si $D_T T = 0$ sobre σ . Así pues, se puede decir que una curva es geodésica si su aceleración es cero, es decir, si su campo de velocidades permanece paralelo a lo largo de la curva. Si tomamos un sistema coordinado en una vecindad de la variedad con conexión M , las ecuaciones $D_T Y = 0$ y $D_T T = 0$ tienen expresiones coordinadas enteramente análogas a las que se obtuvieron en el caso de la conexión de Levi-Civita, con la única diferencia de que ahora aparecen las componentes G_{ij}^k de la conexión D en lugar de los símbolos de Cristoffel. Nuevamente, mediante la teoría de existencia y unicidad de las ecuaciones diferenciales ordinarias, se pueden obtener los siguientes resultados.

Proposición 48.1 Sea $\sigma: [a, b] \rightarrow M$ una curva con campo de velocidades T . Para cada vector $v \in T_{\sigma(a)}(M)$, existe un único campo $Y(t) \in \mathcal{K}(\sigma)$ tal que $Y(a) = v$, y $Y(t)$ es paralelo a lo largo de σ . La aplicación $P_a^t: T_{\sigma(a)}(M) \rightarrow T_{\sigma(t)}(M)$ definida por $P_a^t(v) = Y(t)$ es un isomorfismo lineal que se llama transporte paralelo a lo largo de σ desde $\sigma(a)$ hasta $\sigma(b)$.

Proposición 48.2 Sean $m \in M$ y $v \in T_m(M)$. Entonces, para cada $b \in \mathbb{R}$ existe un número real $r > 0$ y existe una única geodésica γ definida en $[b-r, b+r]$, tal que $\gamma(b) = m$ y $\gamma'(b) = v$.

La teoría de existencia y unicidad de las ecuaciones diferenciales ordinarias nos da realmente mucho más que la conclusión de la proposición anterior. En particular, si $\gamma(t; m, b, v)$ es la geodésica cuya existencia y unicidad afirma la proposición anterior, entonces la aplicación γ es suave con respecto a todos los parámetros t , m , b y v .

El tensor de torsión de una conexión D sobre M , está dado por una función $\text{Tor}: \mathcal{K}(M) \times \mathcal{K}(M) \rightarrow \mathcal{K}(M)$ definida por

$$\text{Tor}(X, Y) = D_X Y - D_Y X - [X, Y].$$

Nótese que aunque $D_X Y$ y $[X, Y]$ no son tensores, la combinación de ellos

en Tor sí es un tensor, ya que puede comprobarse fácilmente que Tor es $\mathcal{F}(M)$ -bilineal y que $\text{Tor}(X, Y) = -\text{Tor}(Y, X)$. Si se estudia más de una conexión, se escribirá Tor_D para el tensor de torsión correspondiente a la conexión D . Por cierto, Tor es un campo tensorial sobre M del tipo $(1, 2)$. Si $\text{Tor}_D = 0$ idénticamente sobre M , diremos que D es simétrica o libre de torsión. Por la propiedad $(D4)$ de la conexión de Levi-Civita, vemos que ésta es una conexión libre de torsión.

Tal como en el caso de la conexión de Levi-Civita, podemos hacer la siguiente definición. El tensor de curvatura de la variedad con conexión (M, D) está dado por la función $R: \mathcal{K}(M)^3 \rightarrow \mathcal{K}(M)$ definida por

$$R_{XY}Z = R(X, Y)Z = D_{[X, Y]}Z - [D_X, D_Y]Z.$$

Con la misma técnica que usamos en el lema 40.1, se demuestra que R es una función $\mathcal{F}(M)$ -trilineal y que define, por lo tanto, un campo tensorial de tipo $(1, 3)$ sobre M .

La noción de aplicación que conserva las conexiones se sigue de modo natural. Sean (M, D) y (M', D') variedades con conexión. Una aplicación suave $\phi: M \rightarrow M'$ conserva las conexiones si $d\phi(D_X Y) = D'_{d\phi X}(d\phi Y)$ para todo vector X y para todo campo vectorial Y . Obsérvese que el miembro derecho de esta ecuación está bien definido puesto que $d\phi Y$ es un campo vectorial bien definido sobre una curva con velocidad inicial $d\phi X$. Una aplicación suave $\psi: M \rightarrow M'$ se dice que conserva las geodésicas si $\psi \circ \gamma$ es una geodésica en M' para cada geodésica γ en M . Trivialmente, toda aplicación que conserva las conexiones, conserva las geodésicas.

En una variedad con conexión tenemos las siguientes estructuras: topología, estructura diferenciable y conexión. Para una aplicación $\phi: M \rightarrow M'$ que conserva las conexiones, bien puede suceder que la topología y la estructura diferenciable no se conserven. De modo que una aplicación $\phi: M \rightarrow M'$ que conserva las conexiones no puede por sí sola darnos la adecuada noción de equivalencia en la teoría de las variedades con conexión, también llamada geometría afín de variedades. Para esto necesitamos que ϕ sea también un difeomorfismo. Así pues, un isomorfismo entre variedades con conexión (M, D) y (M', D') , será un difeomorfismo $\phi: M \rightarrow M'$ que conserva las conexiones.

Una variedad diferenciable M en la cual se ha especificado un conjunto de curvas C , llamadas geodésicas de M , se llama variedad con estructura proyectiva y se denota por (M, C) , si

- (1) Para todo $p \in M$ y todo $v \in T_p(M)$, existe $\gamma \in C$ tal que $\gamma(0) = p$ y $\gamma'(0) = v$.
- (2) Para cualquier sistema coordenado x^1, \dots, x^n en una vecindad U de M y cualquier curva $\gamma \in C$, existen funciones $\Pi_{ij}^k \in \mathcal{F}(M)$ tales que

$$\frac{d^2(x^k \circ \gamma)}{dt^2} + \Pi_{ij}^k \frac{d(x^i \circ \gamma)}{dt} \frac{d(x^j \circ \gamma)}{dt} = 0 \quad \text{sobre } U.$$

Así pues, es claro que toda variedad semi-Riemanniana es una variedad con conexión y que toda variedad con conexión es una variedad con estructura proyectiva. Como antes, no toda aplicación $\psi: M \rightarrow M'$ que conserva las geodésicas, podrá darnos la adecuada noción de equivalencia en esta geometría proyectiva de variedades. Para ésto, $\psi: M \rightarrow M'$ debe ser también un difeomorfismo. Así pues, un isomorfismo entre variedades proyectivas (M, C) y (M', C') , será un difeomorfismo $\psi: M \rightarrow M'$ que conserva las geodésicas.

Sabemos que la estructura diferenciable de una variedad M determina a la topología de M y que la métrica de una variedad semi-Riemanniana determina a la conexión de Levi-Civita. También sabemos que una variedad con conexión posee una estructura proyectiva determinada por la conexión. Pero, ¿En que medida la conexión determina a la estructura diferenciable? ¿Como se relacionan la estructura proyectiva y la topología? ¿Como se relacionan la métrica y la topología? Todas estas son cuestiones interesantes desde el punto de vista matemático y que, cada vez más, encuentran aplicación en física (ver [E, P y S]), pero que aquí sólo podemos plantear.

Proposición 48.3 Sea $\phi: M \rightarrow M'$ un difeomorfismo y sea D' una conexión sobre M' . Entonces, existe una única conexión D sobre M para la cual ϕ es una aplicación que conserva las conexiones.

Demostración. Tomemos un $X \in T_m(M)$ y sea Y un campo vectorial definido en una vecindad de m . Puesto que ϕ es un difeomorfismo, $d\phi Y$ es un campo vectorial definido en una vecindad de $\phi(m)$. Definamos

$$D_X Y = d\phi^{-1}(D' d\phi Y).$$

Es fácil verificar que la D así definida es una conexión. ##

Si cada geodésica $\gamma(t)$ puede ser extendida de tal modo que sea una geodésica para todo $t \in \mathbb{R}$, entonces la correspondiente variedad proyectiva o con conexión se llama completa.

Para problemas locales relativos a una conexión, pueden transformarse las propiedades de D en ciertas propiedades de formas diferenciales, como veremos a continuación.

Sea D una conexión sobre una variedad suave n -dimensional M . Sea U un conjunto abierto (por ejemplo el dominio de un sistema coordenado) en M y sea E_1, \dots, E_n un sistema de campos vectoriales linealmente independientes y definidos en U . Sean $\omega^1, \dots, \omega^n$ las uno-formas sobre U que son la base dual de E_1, \dots, E_n en cada punto de U . Definiremos las n^2 uno-formas de conexión ω_j^i en U asociadas a D y con respecto al campo base, mediante

$$D_X E_j = \sum_{i=1}^n \omega_j^i(X) E_i.$$

Las ω_j^i son efectivamente uno-formas, ya que de las propiedades (2) y

(3) de la conexión se sigue que ω_j^i es $\mathcal{F}(M)$ -lineal. También vemos que si X es un campo vectorial suave sobre U , entonces $D_X E_j$ es un campo vectorial suave sobre U y, por consiguiente, $\omega_j^i(X) = \omega^i(D_X E_j) \in \mathcal{F}(M)$.

Los tensores de torsión y curvatura pueden también expresarse a través de las formas diferenciales asociadas con el campo base. Definimos las dos-formas T^i y R_j^i en U mediante

$$\begin{aligned} T(X, Y) &= \sum T^i(X, Y) E_i, \\ R(X, Y) E_j &= \sum R_j^i(X, Y) E_i, \end{aligned}$$

donde las propiedades de un tensor antisimétrico se deducen para T^i y R_j^i a través de las propiedades de T y R .

Las formas $\omega^i, \omega_j^i, T^i, R_j^i$, están relacionadas por medio de las ecuaciones estructurales de Cartan que equivalen a las definiciones de los tensores de torsión y curvatura. Nosotros lo expresaremos todo simplemente en función del campo base. Sean X e Y campos vectoriales suaves en U . Entonces,

$$\begin{aligned} \sum T^i(X, Y) E_i &= D_X Y - D_Y X - [X, Y] = \\ &= D_X (\sum \omega^j(Y) E_j) - D_Y (\sum \omega^j(X) E_j) - \sum \omega^j([X, Y]) E_j = \\ &= \sum_j (X \omega^j(Y) - Y \omega^j(X) - \omega^j([X, Y])) E_j + \sum_i (\omega^i(Y) \omega_j^i(X) - \omega^i(X) \omega_j^i(Y)) E_i. \end{aligned}$$

Iguando componentes, tenemos

$$T^i(X, Y) - (\sum_j \omega_j^i \wedge \omega^j)(X, Y) = X \omega^i(Y) - Y \omega^i(X) - \omega^i([X, Y]).$$

Puesto que el miembro de la izquierda en esta ecuación es una 2-forma, también debe ser una dos-forma el miembro de la derecha tomado en su conjunto. De hecho, recordando lo visto en la sección 25, vemos que es la derivada exterior $d\omega^i$ de ω^i calculada sobre X e Y .

Podemos escribir ahora la primera ecuación estructural de Cartan:

$$d\omega^i = - \sum_j \omega_j^i \wedge \omega^j + T^i.$$

Partiendo de la definición de $R(X, Y)$, y por un cálculo completamente análogo al anterior, se obtiene la segunda ecuación estructural de Cartan:

$$d\omega_j^i = - \sum_k \omega_k^i \wedge \omega_j^k + R_j^i.$$

Sea ahora U el dominio de un sistema coordenado x^1, \dots, x^n , entonces $\partial_1, \dots, \partial_n$ es el campo base asociado al sistema coordenado x^1, \dots, x^n . Entonces, $\omega^i = dx^i$ y las formas asociadas ω_j^i, T^i , y R_j^i , definen las funciones G_{jk}^i, T_{jk}^i y R_{jkh}^i , respectivamente mediante

$$\begin{aligned} G_{jk}^i &= \omega_j^i(\partial_k) & \text{de donde} & \quad \omega_j^i = \sum_k G_{jk}^i dx^k, \\ T_{jk}^i &= T^i(\partial_j, \partial_k) & \text{de donde} & \quad T^i = \sum_{jk} T_{jk}^i dx^j \otimes dx^k, \\ R_{jkh}^i &= R_j^i(\partial_k, \partial_h) & \text{de donde} & \quad R_j^i = \sum_{kn} R_{jkn}^i dx^k \otimes dx^h. \end{aligned}$$

Por las ecuaciones de estructura, tenemos

$$T_{jk}^i = (d^2x^i + \sum_r G_{rs}^i dx^r \wedge dx^s)(\partial_j, \partial_k) = G_{jk}^i - G_{kj}^i,$$

ya que $d^2x^i = 0$, $dx^r(\partial_j) = \delta_j^r$; y

$$\begin{aligned} R_{jkh}^i &= (d\omega_j^i + \sum_r G_{rs}^i dx^r \wedge dx^s) \wedge (G_{jt}^r dx^t)(\partial_k, \partial_h) = \\ &= \partial_k \omega_j^i(\partial_h) - \partial_h \omega_j^i(\partial_k) + \sum_r (G_{rk}^i G_{jh}^r - G_{rh}^i G_{jk}^r) = \\ &= \frac{\partial G_{jh}^i}{\partial x^k} - \frac{\partial G_{jk}^i}{\partial x^h} + \sum_r (G_{rk}^i G_{jh}^r - G_{rh}^i G_{jk}^r), \end{aligned}$$

que son las clásicas expresiones coordenadas de estos tensores.

Finalmente, consideraremos algunos resultados acerca de la relación que guardan entre sí dos conexiones definidas sobre la misma variedad.

Sea M una variedad diferenciable y sean D y D' conexiones sobre M . Definimos una función $B: \mathcal{M}(M) \times \mathcal{M}(M) \rightarrow \mathcal{M}(M)$ mediante $B(X, Y) = D'X Y - D_X Y$. La función B es $\mathcal{X}(M)$ -bilineal. Esto se sigue inmediatamente de las propiedades (2) y (3) de las conexiones, y si $f \in \mathcal{F}(M)$, entonces $B(X, fY) = (Xf)Y + fD'X Y - (Xf)Y - fD_X Y = fB(X, Y)$. Por lo tanto, B puede ser interpretada como un campo tensorial sobre M de tipo (1,2) que se llama el tensor diferencia de las conexiones D y D' .

Sea $B(X, Y) = S(X, Y) + A(X, Y)$ la típica descomposición de una función bilineal en sus partes simétrica y antisimétrica, es decir,

$$S(X, Y) = \frac{1}{2}\{B(X, Y) + B(Y, X)\} \quad y$$

$$A(X, Y) = \frac{1}{2}\{B(X, Y) - B(Y, X)\}.$$

Podemos expresar la parte antisimétrica A en términos de los tensores de torsión T y T' de las conexiones D y D' respectivamente, pues

$$\begin{aligned} 2A(X, Y) &= D'X Y - D_X Y - D'Y X + D_Y X = T'(X, Y) + [X, Y] - T(X, Y) - [X, Y] \\ &= T'(X, Y) - T(X, Y). \end{aligned}$$

Proposición 48.4 Las siguientes afirmaciones son equivalentes:

- (a) Las conexiones D y D' definen las mismas geodésicas en M .
- (b) $B(X, X) = 0$ para todo $X \in \mathcal{X}(M)$.
- (c) $S = 0$.
- (d) $B = A$.

Demostración. (a) implica (b): Tomemos un vector $X_m \in T_m(M)$ y sea γ la geodésica con velocidad inicial X_m . Sea $X \in \mathcal{X}(\gamma)$ tal que $X|_{\gamma(0)} = X_m$ y tal que X sea el campo de velocidades de γ . Entonces,

$$B(X, X) = D'X X - D_X X = 0 - 0 = 0,$$

puesto que γ es una geodésica para ambas conexiones.

(b) implica (a): Sea γ una geodésica para D con campo de velocidades X .

Entonces, $D'_X X = B(X, X) + D_X X = 0$, luego γ es una geodésica para D' .

(b) equivale a (c): Puesto que S es simétrico, por la identidad de polarización, sabemos que S queda determinado por la correspondiente forma cuadrática $S(X, X)$. Pero $B(X, X) = 0$ si y sólo si $S(X, X) = 0$.

(c) equivale a (d): Ya que $B = S + A$. ##

Proposición 48.5 Las conexiones D y D' sobre M son iguales si y sólo si determinan las mismas geodésicas en M y tienen los mismos tensores de torsión.

Demostración. Que la primera parte implica la segunda es trivial.

Recíprocamente, si las geodésicas son las mismas, entonces $S = 0$ y si los tensores de torsión son iguales, entonces $A = 0$, por tanto $B = 0$ y $D = D'$. ##

Proposición 48.6 Dada una conexión D' en M , existe y es única una conexión D en M que determina las mismas geodésicas que D' pero que tiene torsión nula.

Demostración. Existencia. Sea $D_X Y = D'_X Y - (1/2)T'(X, Y)$. Es fácil ver que D satisface las propiedades que definen a una conexión. Aquí $B = (1/2)T' = A$, puesto que el tensor de torsión es antisimétrico. Así pues, $S = 0$, con lo que D y D' tienen las mismas geodésicas. Además, $T = T' - 2A$, con lo que D tiene torsión nula.

Unicidad. La unicidad es consecuencia de la proposición anterior. ##

Así pues, si efectuamos una partición entre las conexiones sobre una variedad en clases de equivalencia colocando a dos conexiones que definan las mismas geodésicas en la misma clase, entonces en cada clase existe una única conexión que es libre de torsión (torsión nula). Además, dada una conexión arbitraria D y un tensor de tipo $(1, 2)$ T' antisimétrico en sus índices covariantes, existe una única conexión D' con torsión T' que determina las mismas geodésicas que D . De la demostración anterior, tenemos que $T'(X, Y) = 2(D'_X Y - D_X Y)$. Esto proporciona una interpretación geométrica del tensor de torsión de una conexión, como una medida de la diferencia que hay entre la derivada covariante de una conexión dada y la derivada covariante de la correspondiente conexión libre de torsión que determina las mismas geodésicas.

+++++

V. GEOMETRIA LORENTZIANA Y RELATIVIDAD.

49 CARACTER CAUSAL

El estudio de las variedades de Lorentz, que son variedades semi-Riemannianas de índice 1 y dimensión ≥ 2 , se puede llamar geometría Lorentziana. Para estudiar los espacios tangentes de una variedad de Lorentz en términos abstractos, definimos un espacio vectorial de Lorentz como un espacio con producto escalar de índice 1 y dimensión ≥ 2 .

La noción de carácter causal de vectores tiene en este contexto una generalización natural a subespacios vectoriales.

Sea W un subespacio de un espacio vectorial de Lorentz V , y sea g el producto escalar de V . Para W hay tres posibilidades mutuamente excluyentes:

(1) $g|_W$ es definido positivo; es decir, W es un espacio con producto interno. En este caso se dice que W es espacial.

(2) $g|_W$ es no degenerado de índice 1. En este caso se dice que W es temporal.

(3) $g|_W$ es degenerado. En este caso se dice que W es luminoso o nulo.

La categoría a la que pertenece W se llama su carácter causal.

Esta definición es consistente con la definición del carácter causal de los vectores que hemos dado antes, en el sentido de que el carácter causal de un vector individual v es el mismo que el carácter causal del subespacio $\{tv \mid t \in \mathbb{R}\}$ generado por v . (El subespacio cero, al igual que el vector cero, es espacial.)

El siguiente es un resultado simple pero muy útil.

Lema 49.1 Si z es un vector temporal en un espacio vectorial de Lorentz V , entonces el subespacio z^\perp es espacial y V es la suma directa $\mathbb{R}z + z^\perp$, en donde z^\perp es el subespacio de los vectores ortogonales a z y $\mathbb{R}z$ es el subespacio generado por z .

Demostración. El subespacio $\mathbb{R}z$ es no degenerado con índice 1. Por lo tanto, por el lema 23.2, z^\perp es no degenerado y $V = \mathbb{R}z + z^\perp$ es una suma directa. Así, $\text{ind } V = \text{ind } \mathbb{R}z + \text{ind } z^\perp$, lo cual implica que $\text{ind } z^\perp$ es cero. Por lo tanto, z^\perp es espacial. ##

Este argumento muestra, más en general, que un subespacio W es temporal si y sólo si W^\perp es espacial. Puesto que $(W^\perp)^\perp = W$, las palabras temporal y espacial pueden ser intercambiadas en esta afirmación. Se sigue entonces, que W es luminoso si y sólo si W^\perp es luminoso.

Los subespacios espaciales W son los más fáciles de manejar, ya que, por ejemplo, todo subespacio de W es también espacial y la desigualdad de Schwarz está a nuestra disposición: $|\langle v, w \rangle| \leq \|v\| \|w\|$, y la igualdad se cumple cuando y sólo cuando v y w son colineales.

Consideremos ahora un criterio para decidir cuando un subespacio es temporal, omitiendo el caso trivial $\dim W = 1$.

Lema 49.2 Sea W un subespacio de dimensión ≥ 2 de un espacio vectorial de Lorentz. Entonces las siguientes afirmaciones son equivalentes:

- (1) W es temporal, por lo cual es en sí mismo un espacio de Lorentz.
- (2) W contiene al menos un par de vectores nulos linealmente independientes.
- (3) W contiene al menos un vector temporal.

Demostración. (1) implica (2): Sea e_1, \dots, e_n una base ortonormal para W con e_1 temporal. Entonces $e_1 + e_2$ y $e_1 - e_2$ son vectores nulos independientes.

(2) implica (3): Es fácil ver que si v y w son vectores nulos independientes, entonces $g(v, w) \neq 0$. Por lo tanto, $v + w$ ó $v - w$ es un vector temporal.

(3) implica (1). Si z es un vector temporal en W , entonces $W^\perp \subset z^\perp$ y el último subespacio es espacial. Pero entonces $W = (W^\perp)^\perp$ es temporal. ##

Lema 49.3 Para un subespacio W de un espacio vectorial de Lorentz, las siguientes afirmaciones son equivalentes:

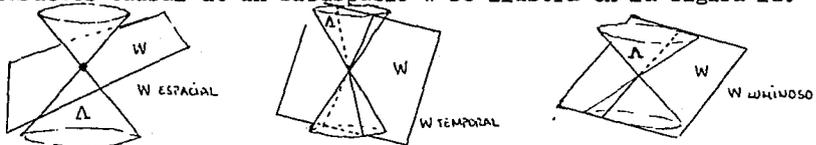
- (1) W es luminoso, es decir, degenerado.
- (2) W contiene un vector nulo, pero ningún vector temporal.
- (3) $W \cap \Lambda = L - O$, donde L es un subespacio unidimensional y Λ es el cono nulo de V .

Demostración. (1) implica (2): Como W es degenerado, debe contener a un vector nulo. Por el lema anterior, W no puede contener a un vector temporal.

(2) implica (3): Como W contiene a un vector nulo, $W \cap \Lambda$ es no vacío. Por el lema anterior, si W contuviera a dos vectores nulos independientes, se seguiría que W contiene a un vector temporal.

(3) implica (1): W no puede ser espacial y, nuevamente por el lema anterior, W no puede ser temporal. Por lo tanto, W es luminoso. ##

El carácter causal de un subespacio W se ilustra en la figura 12.



Sea P una subvariedad de una variedad de Lorentz. Si para todo $p \in P$ el subespacio $T_p(P)$ tiene el mismo carácter causal en $T_p(M)$, entonces ese carácter causal se le atribuye a la propia subvariedad P . Así, las subvariedades semi-Riemannianas de M son o espaciales o temporales. El cono nulo en R_1^n es un ejemplo de subvariedad luminosa, por eso también se llama cono de luz. Desde luego, una subvariedad arbitraria no tiene en general un carácter causal definido.

El carácter causal de los vectores tangentes, las curvas y las sub-variedades es preservado no sólo por isometrías locales, sino también por aplicaciones conformes con $h > 0$. Si $h < 0$, entonces hay una reversión de la causalidad.

La teoría de la relatividad unifica tres dimensiones espaciales y una dimensión temporal por medio de una variedad de Lorentz de cuatro dimensiones. En lugar de considerar al tiempo como "la cuarta dimensión", al tratar con variedades de Lorentz de varias dimensiones, es más consistente considerarlo como la 0-ésima dimensión. Así, para las coordenadas naturales de \mathbb{R}^n usaremos con frecuencia la indexación relativista $t = u^0, u^1, \dots, u^{n-1}$. Similarmente, una base ortonormal para un espacio vectorial de Lorentz será denotada por e_0, e_1, \dots, e_{n-1} en donde e_0 denota al vector temporal. (Sin embargo, es costumbre que en los diagramas de la teoría de la relatividad el eje del tiempo se dibuje como un eje vertical.)

50 CONOS DE TIEMPO Y ORIENTACION EN EL TIEMPO

Sea \mathcal{X} el conjunto de todos los vectores temporales de un espacio vectorial de Lorentz V . Para $u \in \mathcal{X}$

$$C(u) = \{v \in \mathcal{X} \mid \langle u, v \rangle < 0\}$$

es el cono de tiempo de V que contiene a u . El cono de tiempo opuesto es

$$C(-u) = -C(u) = \{v \in \mathcal{X} \mid \langle u, v \rangle > 0\}.$$

Puesto que u^\perp es espacial, \mathcal{X} es la unión de estos dos conos de tiempo disjuntos.

Lema 50.1 Los vectores temporales v y w de un espacio vectorial de Lorentz están en el mismo cono de tiempo si y sólo si $\langle v, w \rangle < 0$.

Demostración. Mostraremos que si $v \in C(u)$ y w es temporal, entonces $w \in C(u)$ si y sólo si $\langle v, w \rangle < 0$. Puesto que $C(u/|u|) = C(u)$, podemos suponer que u es un vector temporal unitario.

Escribamos $v = au + v'$, $w = bu + w'$, en donde $v', w' \in u^\perp$ y $a, b \in \mathbb{R}$. Puesto que estos son vectores temporales, $|a| > |v'|$ y $|b| > |w'|$. Ahora $\langle v, w \rangle = -ab + \langle v', w' \rangle$, donde por la desigualdad de Schwarz, $|\langle v', w' \rangle| \leq |v'| |w'| < |ab|$.

Como $v \in C(u)$, $a > 0$. Por lo tanto, $\text{sgn} \langle v, w \rangle = \text{sgn}(-ab) = -\text{sgn}(b)$, lo cual nos da el resultado. ##

Se sigue que para vectores temporales,

$$u \in C(v) \iff v \in C(u) \iff C(v) = C(u).$$

Además los conos de tiempo son convexos, ya que si $v, w \in C(u)$ y $a \geq 0$, $b \geq 0$ (no ambos cero), entonces es fácil verificar que $av + bw \in C(u)$.

Muchas de las propiedades de los espacios con producto interno tienen análogos novedosos en el caso Lorentziano. Por ejemplo, en un espacio con producto interno la desigualdad de Schwarz permite la definición del ángulo φ entre v y w como el único número $0 \leq \varphi \leq \pi$ tal que $\cos \varphi = \langle v, w \rangle / |v||w|$. Un análogo Lorentziano de este resultado es el siguiente.

Proposición 50.2 Sean v y w vectores temporales en un espacio vectorial de Lorentz. Entonces

(1) $|\langle v, w \rangle| \geq |v||w|$, la igualdad se cumple si y sólo si v y w son colineales.

(2) Si v y w están en el mismo cono de tiempo de V , entonces hay un único número $\varphi \geq 0$, llamado el ángulo hiperbólico entre v y w , tal que

$$\langle v, w \rangle = -|v||w| \cosh \varphi.$$

Demostración. (1) Escribamos $w = av + w'$, con $a \in \mathbb{R}$ y $w' \in v^\perp$. Como w es temporal,

$$\langle w, w \rangle = a^2 \langle v, v \rangle + \langle w', w' \rangle < 0.$$

Entonces,

$$\langle v, w \rangle^2 = a^2 \langle v, v \rangle = (\langle w, w \rangle - \langle w', w' \rangle) \langle v, v \rangle \geq \langle w, w \rangle \langle v, v \rangle = |v|^2 |w|^2,$$

ya que $\langle w', w' \rangle \geq 0$ y $\langle v, v \rangle < 0$. Evidentemente la igualdad se cumple si y sólo si $\langle w', w' \rangle = 0$, lo cual es equivalente a $w' = 0$, es decir, $w = av$.

(2) Si v y w están en el mismo cono de tiempo, entonces $\langle v, w \rangle < 0$, por lo cual

$$-\langle v, w \rangle / |v||w| \geq 1,$$

y el resultado se sigue de las propiedades del coseno hiperbólico. ##

Como la desigualdad de Schwarz cambia de sentido en este contexto, también cambia la desigualdad del triángulo.

Corolario 50.3 Si v y w son vectores temporales en el mismo cono de tiempo, entonces $|v| + |w| \leq |v + w|$, y se cumple la igualdad si y sólo si v y w son colineales.

Demostración. Como $\langle v, w \rangle < 0$, entonces la contraria de la desigualdad de Schwarz nos da $|v||w| \leq -\langle v, w \rangle$. Por lo tanto,

$$(|v| + |w|)^2 = |v|^2 + 2|v||w| + |w|^2 \leq -\langle v + w, v + w \rangle = |v + w|^2.$$

Esto se convierte en igualdad si y sólo si $|v||w| = -\langle v, w \rangle$. Pero el último término es $|\langle v, w \rangle|$, de modo que la proposición anterior nos da el criterio de colinealidad. ##

Para nuestra intuición Euclidiana sólo puede ser perturbador al principio que un segmento de línea recta ya no sea la ruta más corta entre dos puntos. No obstante, el resultado anterior es fundamental en la

geometría Lorentziana y en sus aplicaciones a la teoría de la relatividad.

La existencia de los conos de tiempo hace surgir una cuestión fundamental de naturaleza global acerca de una variedad de Lorentz arbitraria. En cada espacio tangente de Lorentz $T_p(M)$ hay dos conos de tiempo y no hay una manera intrínseca de distinguir uno del otro. Elegir uno de ellos es orientar en el tiempo a $T_p(M)$. La cuestión es esta: ¿Es posible orientar en el tiempo a cada uno de los espacios tangentes de M de una manera adecuadamente continua?

Sea τ una función definida en M que asigna a cada punto $p \in M$ un cono de tiempo τ_p en $T_p(M)$. τ es suave si para cada $p \in M$ existe un campo vectorial suave V sobre una vecindad U de p tal que $V_q \in \tau_q$ para cada $q \in U$. Tal función suave se llama una orientación temporal de M . Si M admite una orientación temporal, se dice que M es orientable en el tiempo. Entonces, elegir una orientación temporal específica sobre M es orientar en el tiempo a M .

Por ejemplo, el espacio de Minkowski \mathbb{R}^n es orientable en el tiempo; su orientación temporal usual es la que contiene al campo vectorial coordenado ∂_0 correspondiente a las coordenadas naturales u_0, \dots, u_{n-1} .

Lema 50.4: Una variedad de Lorentz M es orientable en el tiempo si y sólo si existe un campo vectorial temporal $X \in \mathfrak{X}(M)$.

Demostración. Si tal campo X existe, entonces (como en el caso de \mathbb{R}^n), asignando a cada $p \in M$ el cono de tiempo que contiene a X_p , nos da la orientación temporal sobre M .

Recíprocamente, sea τ una orientación temporal de M . Como τ es suave, cada punto de M tiene una vecindad U sobre la cual está definido un campo vectorial temporal X_U cuyo valor en cada $p \in U$ pertenece a τ_p . Ahora, sea $\{f_\alpha | \alpha \in A\}$ una partición suave de la unidad subordinada a la cubierta de M formada por todas las vecindades U . Así, cada $\text{supp } f_\alpha$ está contenido en algún miembro $U(\alpha)$ de la cubierta. Como las funciones f_α son no negativas y los conos de tiempo son convexos, el campo vectorial $X = \sum_{\alpha} f_\alpha X_U(\alpha)$ es temporal. ##

Para una variedad de Lorentz no hay relación entre la orientabilidad topológica que habíamos definido en la sección 11 y la orientabilidad en el tiempo. Por ejemplo, no es difícil asignar a la banda de Möbius (que es no orientable) una métrica de Lorentz con la cual se obtiene una variedad de Lorentz orientable en el tiempo.

En un espacio vectorial de Lorentz, un vector que es no espacial (y que es, por tanto, nulo o temporal), se llama también causal. Para un vector temporal v , el conjunto $\bar{C}(v)$ de todos los vectores causales w tales que $\langle v, w \rangle < 0$, se llama el cono causal que contiene a v . Los conos causales tienen propiedades muy semejantes a las de los conos de tiempo. En una variedad de Lorentz una curva causal es aquella en la que el campo de velocidades es causal.

51 RELATIVIDAD ESPECIAL

A fines del siglo pasado ya era claro que existían serias dificultades en la física clásica Newtoniana, las cuales se centraban en las propiedades de la luz. Varios progresos en la resolución de estas dificultades fueron alcanzados por Lorentz, Poincaré y otros, pero la primera solución completa fue dada por Einstein en 1905 con la publicación de su teoría especial de la relatividad. La esencia matemática de esta teoría era una manera novedosa de transformar las coordenadas de espacio y tiempo; en 1908 Minkowski demostró que esto ocurre de la manera más natural si el espacio \mathbb{R}^3 y el tiempo \mathbb{R}^1 se unen en un espaciotiempo \mathbb{R}^4 . "De ahora en adelante, el espacio por sí mismo y el tiempo por sí mismo", escribió (ver [E,L,W,H]), "están condenados a desvanecerse como meras sombras y sólo una clase de unión de los dos preservará una realidad independiente."

La relatividad especial llega a varios resultados, muy extraños desde el punto de vista Newtoniano que, no obstante, se obtienen con bastante facilidad de la geometría del espaciotiempo de Minkowski \mathbb{R}^4 . Entre otros, tenemos:

No hay manera de determinar si dos diferentes eventos ocurren al mismo tiempo o en el mismo lugar.

No existe una noción de velocidad absoluta para una partícula material; de hecho, no hay manera de determinar si una partícula con aceleración cero se está moviendo o no.

Por otra parte, la velocidad de la luz en el vacío es una constante, independiente del movimiento de su fuente.

Los relojes móviles se atrasan; los cuerpos móviles se acortan en la dirección de su movimiento.

Con objeto de poder comparar con los resultados relativistas, definiremos brevemente algunos elementos del movimiento Newtoniano.

El espacio Newtoniano es un espacio Euclidiano tridimensional E , es decir, una variedad Riemanniana isométrica a \mathbb{R}^3 con el producto punto.

Puesto que en la naturaleza no están dados los ejes coordenados, esta definición es mejor que simplemente declarar que el espacio Newtoniano es \mathbb{R}^3 .

Por simplicidad, el tiempo Newtoniano será modelado mediante la recta real \mathbb{R}^1 , aún cuando sólo las diferencias $s-t$ tienen sentido físico, no así los tiempos individuales s, t .

Informalmente, una partícula es un objeto material de una pequeñez despreciable en comparación con las distancias típicas del problema que se está considerando. Así, un electrón es una partícula en comparación con un átomo, y una galaxia es una partícula en comparación con el universo. Las partículas tienen masa (carga, etc.) y son capaces de moverse; este movimiento está descrito por una función de tiempo a espacio.

Una partícula Newtoniana es una curva $\alpha: I \rightarrow E$ en el espacio Newtoniano, donde I es un intervalo en el tiempo Newtoniano.

Esta definición es consistente con la terminología que ya hemos empleado: como velocidad para $\alpha' = d\alpha/dt$, rapidez para $v = |\alpha'|$, y aceleración para $\alpha'' = d^2\alpha/dt^2$.

Fundamentalmente, la masa es constante y positiva en la física Newtoniana, pero pueden existir partículas más complicadas, que pueden servir para modelar, digamos, un cohete que quema combustible, en las que la masa es una función del tiempo.

Si $\alpha: I \rightarrow E$ es una partícula Newtoniana de masa m , entonces definimos:

- (1) El impulso de α es el campo vectorial sobre α dado por $m\alpha' \in \mathfrak{K}(\alpha)$. el impulso escalar es la función $m|\alpha'|$ sobre I .
- (2) La fuerza sobre α es el campo vectorial $d(m\alpha')/dt \in \mathfrak{K}(\alpha)$.
- (3) La energía cinética de α es la función $mv^2/2$ sobre I , donde $v = |\alpha'|$.

Cuando m es constante, entonces (2) toma la forma familiar: "fuerza es igual a masa por aceleración" (segunda ley de Newton).

La definición del espacio Newtoniano muestra que éste admite sistemas de coordenadas distinguidos.

Un sistema Euclidiano de coordenadas para E es una isometría $\xi: E \rightarrow \mathbb{R}^3$.

Como ξ es en particular un difeomorfismo, es realmente un sistema de coordenadas para la variedad E . Para cualquier sistema coordenado, tenemos $d\xi(\partial/\partial x^i) = \partial/\partial u^i$, por tanto, ξ es Euclidiano si y sólo si $g_{ij} = \delta_{ij}$ para $1 \leq i, j \leq 3$. Las coordenadas euclidianas son muy eficientes en problemas que involucran líneas rectas. Por ejemplo, las geodésicas tienen las coordenadas $x^i(\delta(t)) = a^i t + b^i$, los vectores distantes paralelos tienen las mismas componentes euclidianas, y la distancia de p a q está dada por la fórmula pitagórica usual

$$\text{dist}(p, q) = \left(\sum_{i=1}^3 (x^i(q) - x^i(p))^2 \right)^{1/2}.$$

Así pues, una vez que se han introducido las coordenadas Euclidianas, el espacio Newtoniano E se convierte en \mathbb{R}^3 .

52 ESPACIO TIEMPO NEWTONIANO

Supongamos que α es una partícula Newtoniana moviéndose en una curva L difeomorfa a \mathbb{R}^1 . Entonces estamos acostumbrados a dibujar su gráfica $\{(t, \alpha(t)) | t \in I\}$ y a interpretar la pendiente de esta gráfica como la rapidez de la partícula, a la razón de cambio de la pendiente como la aceleración de α , etc. La partícula, siendo por definición una curva, se ha convertido en una subvariedad unidimensional de $\mathbb{R}^2 = (\text{tiempo } \mathbb{R}^1) \times X$ (espacio \mathbb{R}^1). Esto puede hacerse con mayor generalidad en la forma siguiente.

rapidez de $c - c/10$. Las observaciones astronómicas de las estrellas dobles deberían revelar estas diferencias en la rapidez de la luz, sin embargo, la rapidez de la luz resulta ser siempre igual a la constante c .

(3) En la teoría Newtoniana una partícula sólo tiene dos posibilidades: o está en reposo (la curva es constante) o no lo está. Supongamos que ϱ es un cohete en el espacio y que está lejos de toda influencia externa. Si ϱ no está acelerando, ¿Cómo se puede determinar si ϱ está o no en reposo? Un intento natural para contestar esta pregunta sería medir la posición de ϱ con respecto a un objeto que estuviera en reposo. Pero uno a uno, los posibles candidatos fallan: la tierra, el sol, las estrellas "fijas" y el hipotético eter. Tampoco puede la tripulación de ϱ realizar un experimento en el interior de ϱ para obtener la respuesta. Tales experimentos involucran mediciones con respecto a ϱ y no pueden dar información acerca del movimiento absoluto.

En resumen, la física Newtoniana trata a la luz en forma relativa (como en (2)) cuando debería tratarla en forma absoluta (1), y trata al movimiento en forma absoluta cuando debería tratarlo en forma relativa (3).

Hay un modo directo de eliminar la dificultad Newtoniana (1). La velocidad de una partícula Newtoniana α al tiempo t puede leerse de su línea de universo W en términos del ángulo entre el eje del tiempo y la línea tangente a W en $(t, \alpha(t))$. Las direcciones tangentes de los rayos de luz (que siempre tienen rapidez constante c), determinan un cono en todo espacio tangente de $R^1 \times E$. Requiriendo que las partículas materiales siempre tengan sus líneas tangentes en el interior de este cono, asegurará que su rapidez sea inferior a c .

Teniendo a nuestra disposición la geometría semi-Riemannina, se ve que una manera más natural de obtener tales conos de luz, es cambiar el signo de la componente temporal del tensor métrico de $R^1 \times E$, con lo cual obtenemos un espaciotiempo de Minkowski con sus conos nulos en cada espacio tangente. Veremos que un intento razonable de reconstruir la mecánica Newtoniana en este contexto, elimina automáticamente las dificultades (2) y (3) y nos lleva a la teoría especial de la relatividad.

Unidades Geométricas. Un sistema muy eficiente de unidades físicas se obtiene tomando la velocidad de la luz c y la constante gravitacional G y haciendo que sean iguales al número adimensional 1. Todas las otras unidades se convierten entonces en potencias de una sola unidad que se escoge libremente, por ejemplo, segundos, metros, etc. Los factores de conversión entre las unidades geométricas y cualquier otro sistema convencional de unidades se obtienen de un modo obvio a partir de los valores c_{CONV} y G_{CONV} de c y G en el sistema convencional. Por ejemplo,

$$c_{CGS} = 3 \times 10^{10} \text{ cm/seg};$$

$$G_{CGS} = 6.67 \times 10^{-8} \text{ cm}^3/\text{g seg}^2.$$

En las unidades geométricas, la distancia y el tiempo se miden con las mismas unidades, como en el bien conocido caso de medir la distancia en años(luz): el tiempo requerido para que la luz recorra esa distancia. Inversamente, el tiempo puede medirse en unidades de distancia: la distancia que viaja la luz en el tiempo dado. Se sigue que la velocidad v (rapidez) estará dada por un número adimensional, con $v = v_{\text{conv}} / c_{\text{conv}}$ para cualquier sistema convencional.

Por ejemplo, la distancia tierra-sol $x_{\text{CGS}} = 1.5 \times 10^{13}$ cm, puede también darse como $x_{\text{CGS}} / c_{\text{CGS}} = 500$ seg. Un cohete con rapidez $v = 0.01$ tiene una rapidez $v_{\text{CGS}} = (0.01)c_{\text{CGS}} = 3 \times 10^8$ cm/seg, es decir, 3000km/seg.

En unidades geométricas, la masa también se mide en las mismas unidades que la distancia y el tiempo. $G_{\text{CGS}} / (c_{\text{CGS}})^2 = 7.42 \times 10^{-29}$ cm/g, por lo tanto, por ejemplo, la masa del sol, $m_{\text{CGS}} = 2 \times 10^{33}$ g, se convierte en 14.8×10^4 cm, como 1.5 km. Usando segundos como unidad geométrica, la masa del sol es

$$\frac{14.8 \times 10^4 \text{ cm}}{3 \times 10^{10} \text{ cm/seg}} = 4.9 \times 10^{-6} \text{ seg.}$$

Las unidades geométricas, que están basadas por completo en las dos constantes fundamentales c y G , tienen un significado físico del que carecen la mayoría de los sistemas convencionales de unidades. Ciertamente, una rapidez como $v = v_{\text{conv}} / c_{\text{conv}}$ es más informativa que un número de pies por segundo o de kilómetros por año. Usando unidades geométricas, también tiene sentido, por ejemplo, decir que la masa del sol es pequeña comparada con su radio ($m = 1.5 \text{ km} \ll r = 7 \times 10^5 \text{ km}$).

53 ESPACIOTIEMPO DE MINKOWSKI

Un espaciotiempo es una variedad Lorentziana conexa, tetradimensional y orientada en el tiempo. (Frecuentemente, el requisito de que un espaciotiempo esté orientado en el tiempo, se debilita a pedir que sea orientable en el tiempo.)

El estudio de los espaciotiempos en general y de sus relaciones con la materia, corresponde a la teoría general de la relatividad. Por el momento, sólo consideraremos el caso de la teoría especial de la relatividad.

Un espaciotiempo de Minkowski M es un espaciotiempo que es isométrico a \mathbb{R}^4 .

En las siguientes secciones, M denotará siempre a un espaciotiempo de Minkowski. Como en cualquier espaciotiempo, la orientación temporal de M se llama el futuro y su negativa el pasado. Un vector tangente, contenido en uno de los conos causales del futuro, se llama dirigido hacia el futuro. Una curva causal se llama dirigida hacia el futuro si todos sus vectores velocidad están dirigidos hacia el futuro.

Como en el caso del espaciotiempo Newtoniano, es conveniente tener una terminología intuitiva; así, los puntos de M se llaman eventos y las partículas serán líneas de universo (parametrizadas). Por otra parte, en contraste con $\mathbb{R}^1 \times E$ y con \mathbb{R}^4 , no existe una función de tiempo canónica sobre M . En este caso hay muchos tiempos:

Una partícula material en M es una curva temporal dirigida hacia el futuro $\alpha: I \rightarrow M$ tal que $|\alpha'(\tau)| = 1$ para todo $\tau \in I$. El parámetro τ se llama el tiempo propio de la partícula.

Como en el caso Newtoniano una partícula modela la historia de un objeto que en un contexto dado se puede considerar despreciablemente pequeño. Podemos imaginar que cada partícula está provista de un reloj (mecánico, atómico, biológico, ...) que mide su tiempo propio. Como para el tiempo Newtoniano, sólo los intervalos de tiempo propio tienen sentido físico. Por ejemplo, si α es una partícula material de $\alpha(a) = p$ a $\alpha(b) = q$, entonces $b - a$ es el tiempo propio que ha transcurrido entre los eventos p y q . Para cualquier número d , $\tau \mapsto \alpha(\tau + d)$ puede considerarse como la misma partícula, pero que ha reajustado su reloj, ya que en este caso, el intervalo de tiempo propio entre los eventos no cambia.

Una partícula luminosa en M es una geodésica nula dirigida hacia el futuro $\gamma: I \rightarrow M$

La física moderna identifica tres tipos de partículas luminosas: los fotones (partículas luminosas por excelencia); los neutrinos (partículas elementales que, tal vez, no sean exactamente luminosas); y los todavía hipotéticos gravitones (que quedan fuera del marco de la relatividad especial).

Cualquier partícula $\beta: I \rightarrow M$ es una curva regular, y su imagen $\beta(I)$ es una subvariedad unidimensional de M , llamada la línea de universo de β .

Como en el caso Newtoniano, las partículas en M tienen masa , la cual es positiva para las partículas materiales pero, como veremos, necesariamente debe ser cero para las partículas luminosas. Que la luz se mueva en geodésicas es una hipótesis fundamental de la relatividad, y puesto que $\langle \gamma', \gamma' \rangle = 0$ para una partícula luminosa, no se puede parametrizar a γ por medio de un tiempo propio. Es como si γ por tener masa igual a cero, no pudiera llevar un reloj integrado.

Una partícula en M que es una geodésica, se dice que cae libremente . En general, "caer libremente" significa moverse bajo la influencia de la gravedad solamente. Pero como veremos en las últimas secciones de estas notas, el hecho de que el espaciotiempo de Minkowski sea plano limita a la relatividad especial a situaciones en donde la gravitación sea despreciable; por ejemplo, en la teoría de partículas elementales, en donde dominan el electromagnetismo y las fuerzas nucleares; en el espacio vacío lejos de fuentes significativas de gravitación, etc.

Un sistema coordenado ξ en M se llama de Lorentz o inercial si es una isometría $\xi: M \rightarrow \mathbb{R}^4$ que preserva la orientación en el tiempo.

Como en el caso Newtoniano, se sigue inmediatamente que un sistema coordenado $\xi: M \rightarrow \mathbb{R}^4$ es inercial si y sólo si $g_{ij} = \delta_{ij} \varepsilon_j$, $\varepsilon = (-1, 1, 1, 1)$ y ∂_0 está dirigido hacia el futuro.

Lema 53.1 Dada una referencial e_0, e_1, e_2, e_3 en $T_p(M)$ tal que e_0 está dirigido hacia el futuro, existe un único sistema inercial de coordenadas ξ tal que $\partial_i|_p = e_i$ para $0 \leq i \leq 3$.

Demostración. De hecho, el sistema normal de coordenadas determinado por la referencial dada es el sistema inercial buscado, además de que $\exp_p(\sum x^i(q)e_i) = q$ para todo $q \in M$. Si η fuera otro sistema inercial, entonces $\xi^{-1} \circ \eta$ es una isometría de M que deja fija la referencial dada. Así pues, por la proposición 47.2, $\xi^{-1} \circ \eta =$ identidad sobre M ; es decir, $\eta = \xi$. ##

En la física prerrelativista, el verdadero papel jugado por las coordenadas no estaba claro. Sin una clara noción de variedad, sólo era razonable suponer que la física del espaciotiempo Newtoniano estaba ligada a la forma que tomaban sus leyes en términos de las coordenadas Euclidianas. Una de las grandes contribuciones de Einstein fue insistir en que todas las leyes de la física deban tener una expresión independiente de la elección del sistema coordenado (principio de covariancia general). Es por ésto que el cálculo tensorial es tan adecuado para expresar las leyes de la física y, en su forma moderna, permite evitar el uso de las coordenadas. De hecho, introducir un sistema particular de coordenadas en un contexto dado, hace surgir el nuevo problema de distinguir entre las propiedades intrínsecas y las propiedades debidas al sistema coordenado. Desde luego, este problema se simplifica usando coordenadas bien adaptadas a los datos intrínsecos del problema.

54 GEOMETRIA MINKOWSKIANA

Puesto que el espaciotiempo de Minkowski es isométrico a \mathbb{R}^4 sabemos que (1) para cualesquiera puntos $p, q \in M$ existe una única geodésica tal que $\sigma(0) = p$ y $\sigma(1) = q$, (2) hay una isometría natural $T_p(M) \approx T_q(M)$, llamada paralelismo distante, y (3) cada aplicación exponencial $\exp_p: T_p(M) \rightarrow M$ es una isometría. Así, M visto desde p es geoméricamente lo mismo que $T_p(M)$ visto desde 0 .

M es una vecindad normal de cada uno de sus puntos. Así, para cada $p, q \in M$, el vector desplazamiento $\vec{p}q = \sigma'(0)$ está bien definido, donde σ es la geodésica que une los puntos p y q . Nótese que $\exp_p(\vec{p}q) = q$.

En términos de un sistema inercial de coordenadas ξ , los vectores distantes paralelos tienen las mismas componentes, entonces

$$pq = \sum_i (x^i(q) - x^i(p)) \partial_i.$$

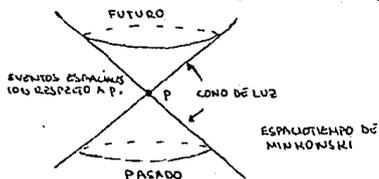
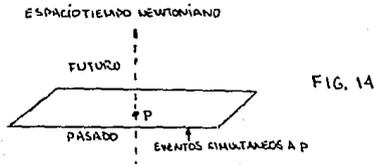
Por medio del comentario anterior estamos en posibilidad de llevar la noción de causalidad de los espacios tangentes de M a la propia M . Para un evento $p \in M$ el cono del futuro de p es $\{q \in M \mid \vec{pq}$ es temporal y está dirigido hacia el futuro}. Este es un cono sólido cuya frontera, excepto por p , es el cono de luz del futuro de p , es decir, $\{q \in M \mid \vec{pq}$ es nulo y está dirigido hacia el futuro}. La unión de estos dos conjuntos es el cono causal del futuro de p . Los conjuntos análogos para el pasado de p se definen similarmente.

El cono de luz $\Lambda(p) \subset M$ de p es la unión de los conos de luz del futuro y del pasado de p . Un punto q que no pertenece a ninguno de los conos causales de p se llama espacial con respecto a p ; es decir, \vec{pq} es espacial.

Ahora resulta claro el porque se usa el termino "causalidad". Es natural decir que un evento p puede influir en un evento q si y sólo si existe una partícula entre p y q . Por la definición de las partículas materiales y luminosas, se sigue que:

- (1) Los únicos eventos que pueden ser influidos por un evento p son aquellos que pertenecen al cono causal del futuro de p .
- (2) Los únicos eventos que pueden influir sobre p son aquellos que pertenecen al cono causal del pasado de p .

Así, todos los eventos que son espaciales con respecto a p , no pueden ni inflir ni ser influidos por p . En la figura 14, se ilustran las relaciones de causalidad para los espaciotiempos Newtoniano y de Minkowski, en donde se usa la convención de dibujar el futuro hacia arriba y los rayos luminosos hacen ángulos de 45° con el eje del tiempo (ver el lema 54.1, más adelante).



La causalidad relativista contrasta mucho con la causalidad Newtoniana, donde para un evento $p = (x_0, t_0)$ el pasado y el futuro llenan todo el espaciotiempo, excepto por el plano tridimensional $t = t_0$ de los eventos simultáneos con p . Como los cohetes Newtonianos no están restringidos por la rapidez de la luz, pueden viajar de x_0 a cualquier lugar distante x en un tiempo arbitrariamente corto $t - t_0$.

Para $p, q \in M$ el número $pq = |\vec{pq}| > 0$ se llama la separación entre p y q .

En términos de un sistema inercial de coordenadas, tenemos

$$pq = |-(x^0q - x^0p)^2 + \sum_1^3 (x^jq - x^jp)^2|^{1/2}.$$

Como el espacio y el tiempo quedan unidos en el espaciotiempo M , la separación proporciona mayor información que el correspondiente concepto de distancia en el espacio Euclidiano.

Comentario. Significado físico de la separación. Sean $p, q \in M$.

- (1) Si \vec{pq} es temporal y dirigido hacia el futuro, entonces pq es el tiempo propio de la única partícula material que cae libremente entre p y q . (Una nave espacial que cae libremente mide pq como el tiempo transcurrido entre el evento p y el evento q .)
- (2) \vec{pq} es nulo $\Leftrightarrow pq = 0 \Leftrightarrow$ hay una partícula luminosa que pasa por p y q .
- (3) Si pq es espacial, entonces $pq \geq 0$ es la distancia medida de p a q por cualquier observador que cae libremente $\perp \vec{pq}$. (Hemos adelantado aquí, algo de la terminología de la siguiente sección.)

Los planos k -dimensionales de M son las imágenes bajo cualquier isometría $\mathbb{R}_1^4 \rightarrow M$ de los planos k -dimensionales de \mathbb{R}_1^4 . Estos planos son subvariedades de M isométricas a \mathbb{R}^k o a \mathbb{R}_1^k . Si V es un subespacio de $T_p(M)$, entonces

$$P = \exp_p(V) = \{q \in M \mid \vec{pq} \in V\}$$

es el único plano k -dimensional en M tal que $T_p(P) = V$.

Finalmente, consideraremos algo de la trigonometría hiperbólica del espaciotiempo de Minkowski.

Lema 54.1 Si \vec{op} es espacial y \vec{oq} es temporal, entonces cualesquiera dos de las siguientes afirmaciones implica a la tercera:

- (1) \vec{pq} es luminoso,
- (2) $\vec{op} \perp \vec{oq}$,
- (3) $op = oq$.

Demostración. Moviendo el vector \vec{pq} por paralelismo al punto o se obtiene $\vec{pq} = \vec{oq} - \vec{op}$. Tomando la norma, se obtiene

$$\pm pq^2 = -oq^2 - 2\langle \vec{oq}, \vec{op} \rangle + op^2,$$

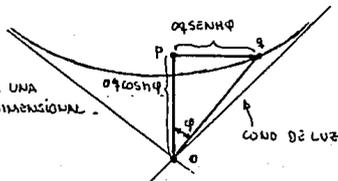
de donde se sigue el resultado. ##

La noción de ángulo hiperbólico puede ser definida también para M en la siguiente forma: si p y q están en el mismo cono de tiempo de o , entonces el ángulo hiperbólico $\varphi = \angle poq$ es el ángulo hiperbólico entre \vec{op} y \vec{oq} .

Proposición 54.2 Sea p y q eventos en el mismo cono de tiempo de o , y tales que $\vec{op} \perp \vec{pq}$ (ver fig. 15). Entonces,

- (1) $oq^2 = op^2 - pq^2$.
- (2) $op = oq \cosh \varphi$, $pq = oq \sinh \varphi$, donde $\varphi = \angle poq$.

FIG. 15
LA LÍNEA CURVA REPRESENTA A UNA
SUBVARIEDAD HIPERBÓLICA TRIDIMENSIONAL.



Demostración. (1) Como en el lema anterior, moviendo el vector espacial \vec{pq} a o nos da $\vec{oq} = \vec{op} + \vec{pq}$. Entonces, tomando normas se obtiene, $-oq^2 = -op^2 + pq^2$.

(2) Sean u y v vectores temporales unitarios en la dirección de \vec{op} y \vec{oq} respectivamente. Entonces $\langle \vec{op}, \vec{oq} \rangle = op \ oq \langle u, v \rangle = -op \ oq \operatorname{cosh} \phi$. Pero también, $\langle \vec{op}, \vec{oq} \rangle = \langle \vec{op}, \vec{op} + \vec{pq} \rangle = -op^2$. Así, $op = oq \operatorname{cosh} \phi$, y por (1), $pq^2 = oq^2(\operatorname{cosh}^2 \phi - 1) = oq^2 \operatorname{senh}^2 \phi$. Pero como $\phi \geq 0$, $\operatorname{senh} \phi \geq 0$ y entonces $pq = oq \operatorname{senh} \phi$. ##

Así, la fórmula pitagórica es reemplazada por (1), y las proyecciones ortogonales están dadas por las funciones hiperbólicas en lugar de las funciones circulares. Nótese que la proyección temporal op es siempre $\geq oq$ y lo mismo es cierto para la proyección espacial pq .

55 OBSERVADORES

Un observador en M es simplemente una partícula material en M , esta terminología sugiere un nuevo papel para las partículas. Sea ξ un sistema inercial para M . El eje x^0 de ξ es la línea de universo de un observador que cae libremente ω ; la parametrización natural de ω es tal que $x^0 \omega(t) = t$, de modo que t es el tiempo propio para ω . Consideraremos que los números producidos por ξ son mediciones tomadas por el observador ω . Por el lema 53.1, todo observador que cae libremente tiene muchos de tales sistemas inerciales asociados.

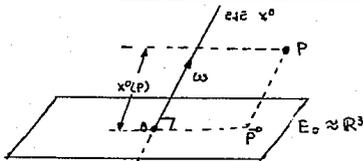


FIG. 16

La sección coordenada $x = 0$ de es un espacio Euclidiano E_0 que identificaremos con \mathbb{R}^3 por medio de la isometría natural

$$q \leftrightarrow (x^1(q), x^2(q), x^3(q)).$$

En M no hay una manera natural de definir ni el tiempo ni el espacio separadamente, pero un sistema inercial produce una descomposición artificial como sigue.

Sea ξ un sistema inercial para M . Para cada evento $p \in M$, el número $x^0(p)$ se llama ξ -tiempo de p y el punto $\vec{p} = (x^1(p), x^2(p), x^3(p)) \in \mathbb{R}^3$ se llama la ξ -posición de p (ver fig. 16).

Ahora sea $\alpha: I \rightarrow M$ una partícula, ya sea material o luminosa. Para cada valor del parámetro $s \in I$ (que es el tiempo propio si α es material), el ξ -tiempo del evento $\alpha(s)$ es $t = x^0(\alpha(s))$ y su ξ -posición es $(x^1(\alpha(s)), x^2(\alpha(s)), x^3(\alpha(s))) \in \mathbb{R}^3 \approx E_0$.

Como α es causal (no espacial) y está dirigida hacia el futuro,

$$\frac{d(x^0 \circ \alpha)}{ds} = -\langle \alpha', \partial_0 \rangle > 0.$$

Por tanto, $x^0 \circ \alpha$ es un difeomorfismo de I sobre un intervalo $J \subset \mathbb{R}^1$. Sea $u: J \rightarrow I$ la función inversa. Entonces al ξ -tiempo $t \in J$, la ξ -posición de α es

$$\vec{\alpha}(t) = (x^1 \circ \alpha \circ u(t), x^2 \circ \alpha \circ u(t), x^3 \circ \alpha \circ u(t)).$$

Así, las mediciones de la partícula α en M producen una curva $\vec{\alpha}: J \rightarrow \mathbb{R}^3 \approx E_0$ llamada la ξ -partícula Newtoniana asociada con α . Esta es lo que el observador ω observa de α .

La relación entre α y $\vec{\alpha}$ es una guía para desarrollar la relatividad especial. Al aplicar los conceptos Newtonianos a $\vec{\alpha}$ vemos como encontrar sus análogos relativistas para α ; al reinterpretar los conceptos relativistas en términos de $\vec{\alpha}$ vemos como la nueva teoría modifica a la vieja.

Por convención, los parámetros t y s están relacionados por $t = x^0 \circ \alpha(s)$ e inversamente por $s = u(t)$, y sólo por estas funciones. Así, tiene sentido escribir $dt/ds = d(x^0 \circ \alpha)/ds > 0$ y, por la regla de la cadena,

$$\frac{d\vec{\alpha}}{dt} = \frac{d\vec{\alpha}/ds}{dt/ds}.$$

Lema 55.1 Sea γ una partícula luminosa en M . Para un sistema inercial ξ , la partícula Newtoniana asociada $\vec{\gamma}$ es una línea recta en \mathbb{R}^3 con rapidez 1.

Demostración. Como γ es una geodésica en M , $\xi \circ \gamma$ es una geodésica en \mathbb{R}^4 . Así γ es tal que

$$x^i \circ \gamma(s) = a^i s + b^i \quad (0 \leq i \leq 3).$$

Por tanto, la proyección $\vec{\gamma}(s) = (x^1 \circ \gamma(s), x^2 \circ \gamma(s), x^3 \circ \gamma(s))$ sobre $\mathbb{R}^3 \approx E_0$ es una línea recta, y su reparametrización sigue esta misma línea recta. Denotaremos por $\vec{\gamma}(t)$ a la reparametrización. Como el vector

$$\frac{d\vec{\gamma}}{ds} = \sum_i a^i \partial_i = \frac{dt}{ds} \partial_0 + \sum_{j=1}^3 \frac{d(x^j \circ \gamma)}{ds} \partial_j$$

es nulo, y dt/ds es positiva, se sigue que

$$\frac{dt}{ds} = \left| \frac{d\vec{\gamma}}{ds} \right|.$$

Así la partícula asociada $\vec{\alpha}$ tiene la rapidez

$$v = \left| \frac{d\vec{\alpha}}{dt} \right| = \frac{|d\vec{\alpha}/ds|}{dt/ds} = 1. \quad \#\#$$

En particular se sigue que la luz tiene la rapidez constante 1 con respecto a todo observador que cae libremente.

Consideremos ahora el caso de una partícula material, de modo que el parámetro s se convierte en el tiempo propio τ .

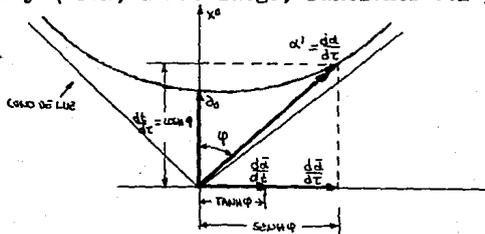
Proposición 55.2 Sea ξ un sistema inercial en M . Si $\vec{\alpha}: J \rightarrow \mathbb{R}^3$ es la ξ -partícula Newtoniana asociada a la partícula material $\alpha: I \rightarrow M$, entonces

(1) La rapidez $|d\vec{\alpha}/dt|$ de $\vec{\alpha}$ es $v = \tanh \varphi$ donde φ es el ángulo hiperbólico entre $\alpha' = d\alpha/d\tau$ y el vector coordenado ∂_0 de ξ . En particular, $0 \leq v < 1$.

(2) El tiempo propio τ de α y su ξ -tiempo t están relacionados por

$$\frac{dt}{d\tau} = \frac{d(x^0 \circ \alpha)}{d\tau} = \cosh \varphi = \frac{1}{(1-v)^{1/2}} \geq 1.$$

Aquí v y φ son, desde luego, funciones del parámetro de α .



Demostración. Como α' y ∂_0 son temporales y dirigidos hacia el futuro, existe un único ángulo hiperbólico $\varphi \geq 0$ entre ellos el cual está determinado por $-\langle \alpha', \partial_0 \rangle = \cosh \varphi \geq 1$. Ver fig. 17. Como $\alpha' = \sum (d(x^i \circ \alpha)/d\tau) \partial_i$, tenemos que

$$\frac{dt}{d\tau} = \frac{d(x^0 \circ \alpha)}{d\tau} = -\langle \alpha', \partial_0 \rangle = \cosh \varphi,$$

y la expresión coordenada para $\langle \alpha', \alpha' \rangle = -1$, se convierte en

$$-\left(\frac{dt}{d\tau}\right)^2 + \left|\frac{d\vec{\alpha}}{d\tau}\right|^2 = -1.$$

Como $\varphi \geq 0$ se sigue que

$$\left|\frac{d\vec{\alpha}}{d\tau}\right| = (\cosh^2 \varphi - 1)^{1/2} = \sinh \varphi \geq 0.$$

Así, $\vec{\alpha}$ tiene rapidez

$$v = \left| \frac{d\vec{\alpha}}{dt} \right| = \frac{d\vec{\alpha}/d\tau}{dt/d\tau} = \frac{\sinh \varphi}{\cosh \varphi} = \tanh \varphi.$$

Una identidad de la trigonometría hiperbólica nos da $\varphi = (1-v^2)^{-1/2}$. $\#\#$

Las siguientes interpretaciones ponen más énfasis en el observador que cae libremente ω que en el sistema inercial ξ asociado con él.

(1) Tiempo. Para cualquier sistema inercial ξ asociado con ω , el hiperplano coordenado E_t dado por $x^0 = t$ es el único plano tridimensional que pasa por $\omega(t)$ y es perpendicular a ω . Así, x^0 es el mismo para toda elección de ξ . En efecto, x^0 impone el tiempo propio de ω a todo M , donde E_t consiste de todos los eventos que ω considera que son simultáneos con $\omega(t)$.

(2) Espacio. Para el propio observador ω , la ξ -partícula Newtoniana asociada $\bar{\omega}$ es constante; Así E_0 se llama el espacio de reposo de ω . Para cualesquiera s, t , la proyección ortogonal $E_s \rightarrow E_t$ envía $p \in E_s$ a un único punto $q \in E_t$ tal que $\bar{p}q$ es paralelo a ω . Consecuentemente, los E_t son espacios Euclidianos canónicamente isométricos, y cada uno de ellos puede servir como espacio de reposo para ω .

(3) Rapidez. En la proposición anterior, como ω es la curva coordenada x^0 , ω' es siempre distantesamente paralelo a ∂_0 . Así, $\varphi(\tau)$ es el ángulo hiperbólico entre $\alpha'(\tau)$ y ω' . La función $v = |d\bar{x}/dt|$ da la rapidez de α con respecto a ω , y la función $\psi = \tanh^{-1} v$, aunque mide la rapidez, tradicionalmente se conoce como el parámetro de velocidad de α con respecto a ω . (A veces $\alpha' = d\alpha/d\tau$ se llama la cuadrivelocidad de α para distinguirla de la velocidad relativa $d\bar{x}/dt$ en el espacio euclidiano.)

(4) Dilatación del tiempo. Para una partícula con tiempo propio τ las ecuaciones en (2) de la proposición anterior, muestran que mientras más rápido se mueve la partícula con respecto al observador (es decir, mientras más grande es v), más lento corre el reloj (τ) de la partícula con respecto al reloj (t) del observador. De ahí que los relojes móviles se atrasan.

(5) Distancia. Que $\bar{p}q$ sea ortogonal a un observador ω que cae libremente significa que $x^0(p) = x^0(q)$. Por tanto, p y q están en el mismo hiperplano E_t , y su separación coincide con la distancia euclidiana ordinaria:

$$pq = \left(\sum_{j=1}^3 (x^j(p) - x^j(q))^2 \right)^{1/2}.$$

Así, la distancia entre eventos tiene sentido sólo para observadores que consideran a dichos eventos como simultáneos.

56 ALGUNOS EFECTOS RELATIVISTAS

En las secciones anteriores ha quedado claro que cada observador que cae libremente en M , tiene su propia noción de tiempo y espacio. Muchos de los resultados aparentemente paradójicos de la relatividad surgen al tratar de comparar las conclusiones a las que llegan dos diferentes observadores.

Por ejemplo, si ω_1 y ω_2 son no paralelos, ellos tienen distintos espacios de reposo. Así, si $\bar{p}q$ es ortogonal a ω_1 , pero no a ω_2 , los even-

tos p y q son simultáneos para ω_1 pero no para ω_2 . Análogamente, si \vec{p} es paralelo a ω_1 , entonces no puede ser paralelo para ω_2 y, proyectando sobre los espacios de reposo, los eventos p y r ocurren en la misma posición para ω_1 pero en distintas posiciones para ω_2 .

Usando el paralelismo distante, la noción de rapidez relativa puede ser generalizada de la manera siguiente. Si α y β son partículas materiales, entonces el ángulo hiperbólico $\varphi = \chi(\alpha'(\sigma), \beta'(\tau))$ es su parámetro instantáneo de velocidad y $v = \tanh \varphi$ es su rapidez instantánea relativa. Para una partícula que cae libremente β , los observadores que caen libremente y son paralelos a β la consideran como si estuviera en reposo, mientras que para otros observadores β puede tener cualquier rapidez constante $0 < v < 1$. En particular, dos observadores que caen libremente se consideran a sí mismos como moviéndose con rapidez relativa constante $= \tanh \chi(\alpha', \beta')$.

Las dificultades con el movimiento Newtoniano que mencionamos antes, no surgen aquí: sólo el movimiento relativo está definido, la luz se mueve con la rapidez constante 1 para todo observador que cae libremente, y todas las partículas materiales se mueven con rapidez relativa $v < 1$ (todo esto sucede en el espaciotiempo vacío modelado por M). La dicotomía esencial no es ahora entre reposo y movimiento, sino entre caída libre y aceleración, ya que si β'' no es idénticamente cero, ningún observador que cae libremente puede considerar que β está en reposo.

A. Adición relativista de velocidades. Una nave espacial ϱ abandona una estación espacial σ (ambas caen libremente) con una rapidez relativa $v_1 > 0$. Una cápsula μ es expulsada de ϱ en el plano de ϱ y σ con rapidez constante v_2 relativa a ϱ . Aquí $v_2 > 0$ quiere decir que μ se mueve en el mismo sentido que σ y $v_2 < 0$ quiere decir que μ se mueve en sentido contrario a σ . ¿Cuál es la rapidez v de μ con respecto a σ ?

La respuesta Newtoniana es, desde luego, $v = v_1 + v_2$, pero la respuesta Einsteiniana es diferente.

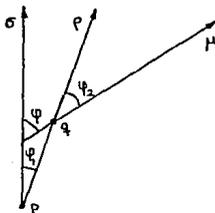


FIG. 18

La figura 18 ilustra el caso $v_2 > 0$. El evento p es la partida de ρ de σ ; el evento q es la partida de μ de ρ . Así, $v_1 = \tanh \varphi_1$ y $v_2 = \tanh \varphi_2$. Por paralelismo distante, ρ' está entre σ' y μ' , por tanto, por la aditividad de los ángulos hiperbólicos (que se demuestra fácilmente), el ángulo $\varphi = \chi(\sigma', \mu')$ es $\varphi_1 + \varphi_2$. Así, la adición de los pa-

rámetros de velocidad reemplaza a la adición Newtoniana de rapidez. De hecho, como

$$v = \tanh \varphi = \tanh(\varphi_1 + \varphi_2) = \frac{\tanh \varphi_1 + \tanh \varphi_2}{1 + \tanh \varphi_1 \tanh \varphi_2}.$$

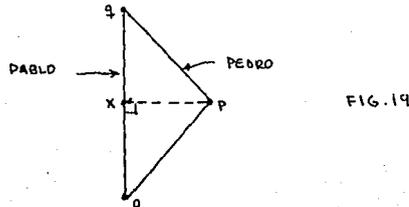
encontramos que

$$v = \frac{v_1 + v_2}{1 + v_1 v_2}.$$

La misma fórmula se cumple para $v_2 < 0$.

B. La "paradoja" de los gemelos. Al cumplir 21 años, Pedro deja a su gemelo Pablo para realizar un viaje en una nave espacial que cae libremente y viaja con una rapidez relativa de $v = 24/25$. El viaje dura 7 años del tiempo propio de Pedro. Luego, regresa en otro viaje con las mismas condiciones que dura otros 7 años de la vida de Pedro. Al regresar, Pedro tiene entonces 35 años de edad, pero Pablo tiene 71.

Para calcular la edad de Pablo, dibujemos una perpendicular desde el punto de retorno p hasta la línea de universo de la estación espacial que cae libremente en la que permanece Pablo. (ver fig. 19)



Por la proposición 54.2, tenemos que

$$ox = \text{opcosh } \varphi = \frac{7}{(1 - (24/25)^2)^{1/2}} = 25.$$

Por simetría, $xq = 25$. Así, la edad de Pablo cuando regresa Pedro es $21 + 2(25) = 71$ años.

Es inútil objetar que este fenómeno (no es una paradoja), se debe a la dificultad de construir relojes suficientemente exactos. Un "reloj" es meramente un nombre para un medidor de tiempo, sea que éste esté basado en el ritmo de los átomos, en dispositivos mecánicos, o en procesos biológicos. Mientras los gemelos están separados, el corazón de Pablo ha latido 50/14 veces más que el corazón de Pedro.

Una mejor objeción es que si Pedro debe sobrevivir al viaje, las esquinas en o , p y q deben ser suavizadas para mantener baja la aceleración. Pero aún así el fenómeno permanece.

Corolario 56.1 Sea $\alpha: [a, b] \rightarrow M$ una partícula material de p a q en M . El tiempo propio transcurrido $\Delta\tau = b - a$ es a lo más pq , la igualdad se cumple si y sólo si α cae libremente.

C. Contracción de Lorentz-Fitzgerald. Sean α y β partículas materiales paralelas que caen libremente en M . Podemos considerar que α y β son los puntos extremos de una regla $[\alpha, \beta]$ que cae libremente en M .

Sea ω un observador que cae libremente cuyo espacio de reposo es E_ω que pasa por $\omega(0)$. Como α y β son paralelas, sus partículas Newtonianas asociadas $\vec{\alpha}$ y $\vec{\beta}$ se mueven a lo largo de líneas rectas paralelas en E_ω ambas con rapidez constante v . Por tanto ω ve que la regla es un segmento de recta $[\vec{\alpha}(t), \vec{\beta}(t)]$ moviéndose paralelamente a sí misma en E_ω . La longitud de la regla como medida por ω es la distancia constante L_ω de $\vec{\alpha}(t)$ a $\vec{\beta}(t)$ en E_ω . Para un observador sobre la regla, digamos α , la regla Newtoniana asociada $[\vec{\alpha}, \vec{\beta}]$ está en reposo en su espacio de reposo E_α y él mide la longitud propia L . Esta es la longitud de cualquier vector que va de α a β y que es ortogonal a ambos.

Proposición 56.2 Como arriba, sea $[\alpha, \beta]$ una regla que cae libremente con rapidez v relativa a un observador ω . En E_ω , (1) si $[\vec{\alpha}, \vec{\beta}]$ se mueve en la dirección perpendicular a su eje, entonces $L_\omega = L$. (2) Si $[\vec{\alpha}, \vec{\beta}]$ se mueve en la dirección de su eje, entonces $L_\omega = L\sqrt{1 - v^2}$.

Demostración. Evidentemente, un desplazamiento paralelo de ω no afecta las mediciones tomadas por ω , así, podemos suponer que $\omega(0) = \alpha(0)$. (1) Sea $A = [\vec{\alpha}(0), \vec{\beta}(0)]$ la posición inicial de la regla Newtoniana en E . Es suficiente demostrar que A también pertenece al espacio de reposo E_α de α a través de $\alpha(0)$, ya que entonces ω y α miden la misma longitud para A . Como $A \in E_\omega$, $A \perp \omega$. Por hipótesis, $A \perp \vec{\alpha}$. Pero α está en el plano determinado por ω y $\vec{\alpha}$, por tanto, $A \perp \alpha$; es decir, $A \in E_\alpha$. (2) En este caso α , β , y ω están en el mismo plano temporal y en la figura 20 aparecen las líneas en las que los espacios de reposo E_ω y E_α cortan a dicho plano.

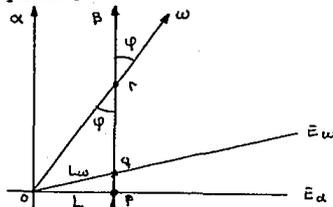


FIG. 20

En la figura, $v = \tanh \varphi$, op es la longitud propia L de la regla, y oq es L_ω . Por la proposición 54.2, en el triángulo opr , $L = ro \sinh \varphi$, y en el triángulo roq , $L_\omega = ro \cosh \varphi$. Así, $L = L_\omega \cosh \varphi$, de donde $L_\omega = L\sqrt{1 - v^2}$. ##

Este fenómeno de acortamiento en la dirección del movimiento, fue conjeturado independientemente por Lorentz y Fitzgerald unos años antes de que Einstein publicara la teoría especial de la relatividad.

57 IMPULSO Y ENERGIA

Si $\alpha: I \rightarrow M$ es una partícula material de masa m , su campo vectorial de impulso-energía es el campo vectorial $P = m d\alpha/d\tau \in \mathcal{K}(\alpha)$.

Para entender esta definición, consideremos lo que ve un observador ω que cae libremente en términos de sus nociones de tiempo y espacio. Para un sistema inercial asociado con el observador ω , las componentes de P son

$$P^i = m \frac{d(x^i \circ \alpha)}{d\tau} \quad (0 \leq i \leq 3),$$

donde, como es usual, τ es el tiempo propio de α . Introduciendo el tiempo propio t del observador, se obtiene

$$P^i = m \frac{d(x^i \circ \alpha)}{dt} \frac{dt}{d\tau}, \quad \text{donde} \quad \frac{dt}{d\tau} = \frac{1}{\sqrt{1-v^2}}.$$

Las componentes espaciales P^1, P^2, P^3 , describen un campo vectorial

$$\vec{P} = \frac{m}{\sqrt{1-v^2}} \frac{d\vec{\alpha}}{dt}$$

sobre la partícula Newtoniana asociada $\vec{\alpha}$ en $E_0 \approx \mathbb{R}^3$.

Esta es una extensión razonable del campo vectorial de impulso Newtoniano, ya que para bajas velocidades, el factor de dilatación temporal $(1-v^2)^{-1/2}$ está próximo a 1. Sin embargo, la componente temporal de P es algo nuevo.

$$P^0 = m \frac{d(x^0 \circ \alpha)}{d\tau} = m \frac{dt}{d\tau} = \frac{m}{\sqrt{1-v^2}}.$$

Por el teorema del binomio de Newton,

$$P^0 = m + \frac{1}{2}mv^2 + O(v^4).$$

El segundo término de esta ecuación es la energía cinética Newtoniana de $\vec{\alpha}$, y Einstein identificó a P^0 como la energía total de la partícula α como medida por el observador ω , concluyendo en particular que la masa es meramente otra forma de energía. Específicamente, m es la energía en reposo E_{REP} , ya que $P^0 = m$ cuando $v = 0$. Convirtiendo a unidades convencionales, se obtiene la famosa ecuación $E_{\text{REP}} = mc^2$. En resumen, podemos hacer la siguiente definición:

Sea α una partícula material de masa m en M . Si ω es un observador que cae libremente, entonces el impulso de α con respecto a ω es el campo vectorial Euclidiano

$$\vec{P} = \frac{m}{\sqrt{1-v^2}} \frac{d\vec{\alpha}}{dt}$$

sobre $\vec{\alpha}$. La energía de α con respecto a ω es la componente temporal de $P = m d\alpha/d\tau \in \mathcal{K}(\alpha)$. Finalmente, el impulso escalar de α con respecto a ω es la función $\rho = |\vec{P}|$.

Del mismo que el espacio y el tiempo están unidos en la relatividad, así también el impulso y la energía están unidos y sólo se les separa artificialmente al considerar un observador particular.

Usando el paralelismo distante, la definición anterior puede ser expresada concisamente como $P = E \partial_0 + \vec{P}$, donde

(1) ∂_0 es el vector coordenado temporal de cualquier sistema inercial asociado con ω ; es decir, ∂_0 es el campo vectorial paralelo sobre α que es distantesamente paralelo a ω' , y

(2) el vector espacial \vec{P} , ortogonal a ∂_0 , ha sido movido desde $\vec{\alpha}$ en el espacio de reposo hasta la propia α por paralelismo distante.

Corolario 57.1 Sea α una partícula material de masa m en M . Con respecto a un observador que cae libremente, la energía E , el impulso escalar ρ , y la rapidez $v = \tanh \psi$ de α están relacionados por

$$(1) E^2 = m^2 + \rho^2.$$

$$(2) E = m / \sqrt{1 - v^2} = m \cosh \psi = -\langle P, \partial_0 \rangle.$$

$$(3) \rho = m \sinh \psi = mv / \sqrt{1 - v^2}.$$

$$(4) \rho/E = \tanh \psi = v.$$

Demostración. Como $P = m\alpha' = E\partial_0 + \vec{P}$, con $\vec{P} \perp \partial_0$, estas afirmaciones se siguen inmediatamente de la proposición 54.2, pero usando la trigonometría hiperbólica en cada espacio tangente en lugar de en la propia M , como habíamos hecho en dicha proposición. ##

Sea $\{\alpha_i\}$ una sucesión de partículas materiales cuyas líneas de universo tienden a la de una partícula luminosa γ . Para cualquier observador que cae libremente, las rapidezes relativas tienden a 1: $v_i \rightarrow 1$, por tanto, las ecuaciones (1) y (4) del corolario anterior dan $E^2 = m^2 + \rho^2$ y $E = \rho$ confirmando que las partículas luminosas deben tener masa cero. La luz transporta energía e impulso, esto es indiscutible, pero como no tiene masa, necesitamos una nueva definición para su impulso-energía.

El campo vectorial de impulso-energía de una partícula luminosa γ es su cuadrivelocidad: $P = \gamma' = d\gamma/ds$.

Entonces, para cualquier observador ω que cae libremente P se puede partir en energía E con respecto a ω e impulso P con respecto a ω , escribiendo $P = E\partial_0 + \vec{P}$ con $\vec{P} \perp \partial_0$, tal como en el caso de una partícula material. Pero como $P = \gamma'$ es luminoso, la energía y el impulso escalar son iguales: $E = \rho = -\langle \gamma', \partial_0 \rangle$. Más aún, como γ y el observador ω son geodésicas, $E = \rho$ es constante y \vec{P} es paralelo.

En unidades geométricas, la energía E y el impulso escalar ρ tienen la misma unidad común de distancia, tiempo o masa. Para una partícula material, la unidad de masa para E y ρ está clara de las fórmulas del corolario 57.1. Para una partícula luminosa γ , el parámetro geodésico s es un número adimensional, por tanto, la fórmula

$$p = E = d(x^0 \cdot \lambda) / ds$$

muestra que la unidad es la misma que para el tiempo.

Para convertir a un sistema convencional de unidades, usamos los valores convencionales c_{conv} y G_{conv} . Por ejemplo, en unidades cgs, un gramo de energía, es decir, la energía en reposo de una masa de un gramo es

$$E_{\text{cgs}} = (c_{\text{cgs}})^2 g = 9 \times 10^{20} \text{ g cm}^2 / \text{seg}^2 (= \text{erg}).$$

El carácter ondulatorio de la luz se mide como sigue: Un fotón de energía E con respecto a un observador, tiene frecuencia $\nu = E/h$, donde h es la constante de Planck. Como es usual, la frecuencia por la longitud de onda λ es la rapidez c . En unidades geométricas h es $1.8 \times 10^{-27} \text{ seg}^2$ y $\lambda \nu = 1$. Puesto que la frecuencia y la longitud de onda se derivan de la energía, ellas también dependerán del observador que esté realizando las mediciones. Así, lo que es luz visible para un observador, pueden ser ondas de radio para otro, y rayos x para un tercero.

58 APLICACIONES BIPARAMÉTRICAS Y CAMPOS DE JACOBI

Sea M una variedad y sea A un subconjunto abierto del plano \mathbb{R}^2 con la propiedad de que cualquier línea horizontal o vertical que interseque a A lo haga en un intervalo. Una aplicación biparamétrica es una aplicación suave $x: A \rightarrow M$. Así, x se descompone en dos familias de curvas paramétricas:

La curva u-paramétrica $v = v_0$ de x es $u \rightarrow x(u, v_0)$.

La curva v-paramétrica $u = u_0$ de x es $v \rightarrow x(u_0, v)$.

Las velocidades parciales

$$x_u = dx(\partial_u), \quad x_v = dx(\partial_v)$$

son campos vectoriales sobre x . Evidentemente, $x_u(u_0, v_0)$ es el vector velocidad en u_0 de la curva u-paramétrica $v = v_0$, y análogamente para $x_v(u, v)$.

Si x está en el dominio de un sistema coordenado x^1, \dots, x^n , entonces sus funciones coordenadas $x^i = x^i \circ x$ ($1 \leq i \leq n$) son funciones de valores reales definidas sobre A , y tenemos

$$x_u = \sum \frac{\partial x^i}{\partial u} \delta_i, \quad x_v = \sum \frac{\partial x^i}{\partial v} \delta_i.$$

Hasta ahora, M ha podido ser una variedad suave sin mayor estructura. Supongamos ahora que M es semi-Riemanniana. Si Z es un campo vectorial suave sobre x , sus derivadas covariantes parciales se definen como

$Z_u = DZ/\partial u$, la derivada covariante de Z a lo largo de las curvas u-paramétricas, y

$Z_v = DZ/\partial v$, la derivada covariante de Z a lo largo de las curvas v-paramétricas.

Explícitamente, $Z_u(u_0, v_0)$ es la derivada covariante en u_0 del campo vectorial $u \rightarrow Z(u, v_0)$ sobre la curva $u \rightarrow x(u, v_0)$.

En términos de coordenadas, $Z = \sum Z^i \partial_i$, donde cada $Z^i = Z(x^i)$ es una función de valores reales definida sobre A . Entonces, por la fórmula que sigue a la proposición 37.1 (pag. 106), tenemos

$$Z_u = \sum_k \left\{ \frac{\partial Z^k}{\partial u} + \sum_{i,j} \Gamma_{ij}^k Z^i \frac{\partial x^j}{\partial u} \right\} \partial_k.$$

En el caso especial $Z = x_u$, la derivada $Z_u = x_{uu}$ da las aceleraciones de las curvas u -paramétricas, mientras que x_{vv} da las aceleraciones de las curvas v -paramétricas.

Proposición 58.1 (1) Si x es una aplicación biparamétrica en una variedad semi-Riemanniana M , entonces $x_{uv} = x_{vu}$. (2) Si Z es un campo vectorial sobre x , entonces

$$Z_{uv} - Z_{vu} = R(x_u, x_v)Z,$$

donde R es el tensor de curvatura Riemanniana de M .

Demostración. (1) Si usamos la misma notación coordenada que antes, tenemos

$$x_{uv} = \sum_k \left\{ \frac{\partial^2 x^k}{\partial v \partial u} + \sum_{i,j} \Gamma_{ij}^k \frac{\partial x^i}{\partial u} \frac{\partial x^j}{\partial v} \right\} \partial_k.$$

Esta fórmula es simétrica en u y v , ya que los símbolos de Christoffel son simétricos en i y j .

(2) Un cómputo con coordenadas de $Z_{uv} - Z_{vu}$ hace aparecer al tensor de curvatura en forma semejante a como apareció en la demostración del lema 40.4. ##

Aquí (1) expresa en términos de aplicaciones biparamétricas el axioma (D4) de la conexión de Levi-Civita, mientras que (2) expresa nuevamente el hecho de que la curvatura mide la falla de la conmutatividad en la diferenciación covariante.

Una curva puede ser comparada con curvas cercanas a ella usando la siguiente definición.

Una variación de un segmento de curva $\alpha: [a, b] \rightarrow M$ es una aplicación biparamétrica

$$x: [a, b] \times (-\delta, \delta) \rightarrow M,$$

tal que $x(u) = x(u, 0)$ para toda $a \leq u \leq b$.

Las curvas u -paramétricas de una variación se llaman longitudinales y las curvas v -paramétricas se llaman transversas. La curva base de x es α . Típicamente, estamos interesados en las curvas longitudinales y el número no es importante.

El campo vectorial V sobre α dado por $V(u) = x_v(u, 0)$ se llama el cam-

po vectorial de variación de x . Cada $V(u)$ es la velocidad inicial de la curva transversa $v \rightarrow x(u, v)$; así, para $\delta > 0$ suficientemente pequeño, el campo vectorial V es una medida infinitesimal de la variación x .

Si toda curva longitudinal de x es una geodésica, x se llama una variación geodésica o una familia uniparamétrica de geodésicas.

Si δ es una geodésica, un campo vectorial Y sobre δ que satisface la ecuación diferencial de Jacobi $Y'' = R_{\gamma} \cdot \gamma'(\delta')$ se llama campo vectorial de Jacobi.

Proposición 58.2 El campo vectorial de variación de una variación geodésica es un campo de Jacobi.

Demostración. Como cada curva longitudinal es una geodésica, $x_{uu} = 0$. Así, por la proposición 58.1, tenemos

$$x_{vuu} = x_{uvu} = x_{uuv} + R(x_v, x_u)x_u = R(x_v, x_u)x_u.$$

Por tanto, x_v satisface la ecuación de Jacobi sobre toda curva longitudinal, en particular sobre la curva base, en donde x_v es precisamente el campo vectorial de variación. ##

Debido a este resultado, la ecuación de Jacobi se conoce también como la ecuación de desviación geodésica. Hay una interpretación física de este resultado que es muy importante. Si consideramos que una variación geodésica x de δ representa a una familia uniparamétrica de partículas que caen libremente, entonces el campo de variación V da la posición, con respecto a δ , de las partículas arbitrariamente próximas. Así, la derivada V' da la velocidad relativa y V'' da la aceleración relativa. Asignando a estas partículas masa unitaria, podemos interpretar la ecuación de Jacobi $V'' = R(V, \delta')\delta'$ por medio de la segunda ley de Newton con el vector de curvatura $R(V, \delta')\delta'$ en el papel de fuerza. Esta fuerza se llama fuerza de marea y veremos en la próxima sección que juega un papel clave en la interpretación de la gravitación como curvatura del espaciotiempo.

59 RELATIVIDAD GENERAL

Mientras que la relatividad especial es una teoría enteramente satisfactoria en su rango de aplicabilidad, no hay manera de que pueda incluir a la gravitación. Por otra parte la ley de gravitación de Newton, que se refiere al espacio y no al tiempo, ha dejado de ser válida con la unión relativista de estas dos nociones. En los años posteriores a 1905, Einstein llegó al convencimiento de que la gravitación debía expresarse en términos de curvatura. Para 1915 él ya había encontrado la manera de hacerlo, y es así que en la relatividad general el espaciotiempo plano de Minkowski deja su sitio a espaciotiempos de curvatura arbitraria.

En esta sección daremos una muy breve reseña de los fundamentos de la teoría general de la relatividad.

A. La relatividad especial como un caso especial de la relatividad general. De hecho, la relatividad especial es la relatividad general del espaciotiempo de Minkowski. Se puede demostrar que el espaciotiempo de Minkowski, es excepto por isometrías, la única variedad Lorentziana tetradimensional plana que es completa y simplemente conexa. Así, cualquier resultado de la relatividad general puede ser verificado en el espaciotiempo de Minkowski e, inversamente, es natural extender a la relatividad general aquellos resultados que no dependen de las propiedades distintivas de \mathbb{R}^4 . En particular, los eventos, las partículas materiales y luminosas, el tiempo propio, el campo vectorial de impulso-energía y los observadores, se definen en la relatividad general del mismo modo que en la teoría especial.

En regiones suficientemente grandes del universo, la gravedad no puede ser ignorada. Así, la relatividad especial es en el mejor de los casos una teoría local; el hecho físicamente significativo es que el espaciotiempo de Minkowski sea un espacio plano, no así que sea completo o simplemente conexo. La relatividad general se basa en el estudio de variedades Lorentzianas de curvatura arbitraria y, debido a esto, abre la posibilidad de hacer estudios globales (topológicos) de los espaciotiempos.

B. La relatividad general es localmente aproximada por la relatividad especial. Si p es un evento en un espaciotiempo M , entonces la relatividad especial tiene sentido en el espacio tangente $T_p(M) \approx \mathbb{R}^4$, y la aplicación exponencial \exp_p proporciona un medio para hacer la comparación. Hemos visto que en vecindades suficientemente pequeñas, y mientras no intervenga la curvatura, $T_p(M)$ es una buena aproximación geométrica de M . Históricamente esta aproximación, casi siempre en la forma de coordenadas normales, ha sido muy importante para asignar significado físico a la geometría de M . En particular, como en la relatividad especial, a un vector temporal unitario $u \in T_p(M)$ dirigido hacia el futuro, lo llamaremos un observador instantáneo en p . La descomposición ortogonal $T_p(M) = \mathbb{R}u + u^\perp$ parte el espacio tangente en el eje temporal del observador $\mathbb{R}u$ y en el espacio de reposo u^\perp . Como antes, si α es una partícula con $\alpha(t_0) = p$, entonces $\alpha'(t_0) = au + x$, $a \in \mathbb{R}$ y $x \in u^\perp$. Corrigiendo a por la dilatación temporal a , se obtiene la velocidad instantánea x/a de α medida por u . (Así, la rapidez $|x|/a$ es 1 para la luz y es menor que 1 para las partículas materiales.) Similarmen-
te, si P es el impulso-energía de α en p , la expresión $P = Eu + \vec{P}$, $\vec{P} \in u^\perp$, da la energía E y el impulso \vec{P} de α en p como medidos por u .

Como muestra esta terminología, la relatividad general también se puede comparar con la física Newtoniana. Esquemáticamente, el alcance de las tres teorías está dado en la siguiente tabla. Cuando los datos de un problema tienen limitaciones Newtonianas, la relatividad general da aproximadamente resultados Newtonianos.

	Gravitación	rapidez
Relatividad general	Arbitraria	Arbitraria
Relatividad especial	despreciable	Arbitraria
Física Newtoniana	débil	baja

C. La gravedad domina en gran escala.

Entre las fuerzas nucleares, el electromagnetismo, y la gravedad, es ésta última la interacción más débil. Sin embargo, el alcance de las interacciones nucleares es tan corto que éstas son ignoradas a priori por una teoría que modela al mundo real por medio de variedades suaves.

En una escala un poco mayor, la repulsión eléctrica entre dos electrones supera enormemente a su atracción gravitacional. Pero en el electromagnetismo hay fuerzas atractivas y repulsivas y así, en grandes agregados de partículas, los efectos electromagnéticos se pueden cancelar. En cambio, la gravitación siempre da lugar a fuerzas atractivas y es acumulativa. Así pues, en gran escala, la gravedad es la interacción dominante. Aun cuando el electromagnetismo puede ser estudiado dentro de la relatividad general, la aplicación principal de ésta es al estudio de la gravedad.

D. La caída libre es geodésica; La materia curva al espaciotiempo.

La caída libre es un movimiento realizado sólo bajo la influencia de la gravedad. La física Newtoniana distingue dos casos: el movimiento acelerado y el movimiento no acelerado. Por ejemplo, consideremos dos naves espaciales idénticas S_1 y S_2 . Supongamos que S_1 se mueve en la vecindad de una estrella gigante roja y que S_2 está también en caída libre pero lejos de cualquier galaxia. De acuerdo a Newton, S_1 tendrá un movimiento acelerado bajo la atracción gravitacional de la estrella y seguirá una órbita curvada en el espacio y por tanto, una órbita curvada en el espaciotiempo Newtoniano. En contraste, S_2 se moverá en una línea recta a velocidad constante, por tanto, su línea de universo será una geodésica recta. Pero si las naves están tripuladas, ninguna de las tripulaciones puede determinar por medio de experimentos físicos realizados a bordo, en cual de las dos naves se encuentra. Esto es lo que se conoce como principio de equivalencia. En ambas naves los objetos no perturbados permanecen en reposo y, si se les empuja, sus movimientos relativos no difieren de nava a nava. Como en el caso de la dicotomía Newtoniana, "reposo contra velocidad constante", Einstein se rehusó a admitir una distinción teórica entre estados que no son experimentalmente diferentes y declaró que toda caída libre tiene una línea de universo que es geodésica en el espaciotiempo. El efecto gravitacional de la estrella no es curvar la línea de universo de S_1 , sino curvar el espaciotiempo en el cual la línea de universo es una geodésica.

E. Gravedad como curvatura.

Si la estrella en el párrafo anterior es una bola uniforme, entonces hay un experimento simple que permite a las tripulaciones determinar su

situación. (En el párrafo anterior suponíamos que el radio de la estrella era prácticamente infinito, en este párrafo ya no hacemos esta suposición.) Es suficiente soltar algunas canicas y observar su comportamiento. En S_1 las canicas permanecerán en reposo, pero en S_2 las canicas comenzarán a moverse. Supongamos por simplicidad que S_2 está cayendo directamente hacia la estrella y que las canicas están colocadas como en la figura 21.

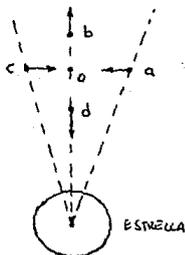


FIG 21

Entonces, a y c se mueven hacia o , mientras que b y d se mueven alejándose de o . La explicación Newtoniana es que a y c están cayendo directamente hacia el centro de la estrella, mientras que por la ley de gravitación de Newton, d acelera hacia la estrella con más intensidad que o y o con más intensidad que b , por lo cual b se retrasa.

En la explicación relativista de este fenómeno, las canicas que caen libremente son modeladas por geodésicas temporales en el espaciotiempo. Como vimos en la sección anterior, la posición relativa de los vecinos de la canica λ , esta dada por campos de Jacobi Y sobre λ_0 . Los cambios en la posición relativa resultan de la aceleración relativa Y'' , la cual, por la ecuación de Jacobi está dada por $R(Y, \lambda')\lambda_0'$. De este modo, vemos cómo la curvatura, en su papel de fuerza de marea, reemplaza a la noción Newtoniana de gravitación.

F. Fuentes de gravitación.

"Materia" es un término indefinido que usamos como sinónimo de la sustancia del universo. En la física Newtoniana la única fuente de gravitación es la masa de la materia. En la relatividad, la gravitación surge del impulso-energía de la materia, al cual la masa es sólo un contribuyente más. Para una forma particular de materia modelada en un espaciotiempo M , el contenido de impulso-energía queda descrito infinitesimalmente por un campo tensorial llamado tensor de impulso-energía T definido sobre M . Como no hay una definición general de materia, no puede haber una fórmula general para construir a T , pero hay algunas reglas empíricas. Sea u un observador instantáneo en $p \in M$. Sobre u^{λ} la parte espacial de T generaliza al clásico tensor de stress como medido por u ; como él, T es un campo tensorial simétrico de tipo $(0,2)$. La densidad de energía medida por u es $T(u,u)$ y para todas las formas conoci-

das de materia es positiva. Finalmente, la ley de conservación del impulso-energía se expresa infinitesimalmente por $\text{div } T = 0$.

Dos espaciotiempos cuya materia ha sido modelada se llaman físicamente equivalentes si existe una isometría entre dichos espaciotiempos que preserva los modelos de materia. En particular, dicha isometría preservará los respectivos tensores de impulso-energía.

60 LA ECUACION DE EINSTEIN

La materia es gravitacionalmente significativa sólo como portadora de impulso-energía, así que para conocer su efecto como productora de gravitación (o curvatura) debemos fijar nuestra atención sobre el correspondiente tensor de impulso-energía T . Pero, ¿Cómo está relacionado T al tensor de curvatura R ? Demandando simplicidad ante todo, Einstein propuso la fórmula $G = kT$, donde G es alguna variante del tensor de curvatura de Ricci y k es una constante. Einstein probó varias posibilidades para G , contrastándolas con el problema de la precesión del perihelio de mercurio. Pero gracias a las poderosas herramientas matemáticas que hemos coleccionado a lo largo de estas notas, es fácil ahora escoger un obvio candidato para G . Si $\text{div } T$ debe ser cero, entonces $G = kT$ implica que $\text{div } G = 0$. Pero para la curvatura de Ricci, el corolario 46.2 afirma que $\text{div Ric} = \frac{1}{2}dS$. Así, restando un medio de la curvatura escalar $S = C(\text{Ric})$ de Ric , debemos obtener el resultado deseado.

Por tanto, definimos el tensor gravitacional de Einstein de un espaciotiempo M como $G = \text{Ric} - \frac{1}{2}Sg$.

Proposición 60.1 (1) G es un campo tensorial simétrico de tipo $(0,2)$ con divergencia cero.

(2) $\text{Ric} = G - \frac{1}{2}C(G)g$.

Demostración. Tanto Ric como g son tensores simétricos de tipo $(0,2)$, por tanto, G es simétrico de tipo $(0,2)$. Un cálculo directo demuestra que $\text{div}(Sg) = dS$. Entonces,

$$\text{div } G = \text{div}(\text{Ric} - \frac{1}{2}Sg) = \frac{1}{2}(dS - dS) = 0.$$

(2) Como $C(g) = 4$, $C(G) = C(\text{Ric}) - \frac{1}{2}SC(g) = -S$. Así, la definición de G nos da

$$\text{Ric} = G + (\frac{1}{2}S)g = G - \frac{1}{2}C(G)g. \quad \#\#$$

Por (2) se sigue que el tensor de Einstein y el tensor de Ricci contienen la misma información. De hecho, como $S = C(\text{Ric})$, cada uno tiene la misma expresión formal en términos del otro.

El postulado fundamental de la relatividad general es la ecuación de Einstein:

Si M es un espaciotiempo que contiene materia con tensor de

impulso-energía T , entonces

$$G = 8\pi T,$$

donde G es el tensor gravitacional de Einstein.

Se han dado muchos argumentos para exponer la plausibilidad de la ecuación de Einstein, casi todos se basan en la comparación con la física Newtoniana a bajas velocidades y gravitación débil. La constante universal 8π (en unidades geométricas), se determina por medio de tales comparaciones (ver, por ejemplo, [H y E], sección 3.4).

Pero debe quedar claro que la relatividad general no puede deducirse; lo mismo que las cuatro leyes de Newton, se trata de una serie de postulados de Einstein acerca del funcionamiento del universo macroscópico.

La ecuación de Einstein implica que el tensor de impulso-energía es un campo tensorial simétrico del tipo $(0,2)$ con divergencia cero. Por otra parte, la ecuación $\text{div } T = 0$ proporciona leyes dinámicas que rigen el comportamiento de la materia que produce a T . Intuitivamente, podemos decir que

$G = 8\pi T$	nos dice cómo la materia determina la curvatura;
$\text{div } T = 0$	nos dice cómo la curvatura mueve a esta materia.

Pero recuérdese que, como en el párrafo E de la sección anterior, es todo el tensor de curvatura de Riemann, no sólo el tensor de Ricci, el que controla el movimiento relativo de las partículas de prueba, aquellas cuyo impulso-energía hace una contribución despreciable al tensor de impulso-energía del espaciotiempo.

Si $T = 0$, es decir, si M es plana según Ricci, entonces M se llama espaciotiempo vacío.

+++ :+++++++

BIBLIOGRAFIA

EL MATERIAL PRESENTADO EN ESTAS NOTAS PROVIENE DE LOS SIGUIENTES LIBROS:

- 1.- R. K. Sachs, y H. Wu. "General Relativity for Mathematicians", Springer-Verlag, 1977.
- 2.- R. L. Bishop y S. I. Goldberg. "Tensor Analysis on Manifolds", Dover, 1980.
- 3.- N. J. Hicks. "Notas sobre geometría diferencial", editorial Hispano Europea, 1974.
- 4.- Boothby, W. "An introduction to Differentiable Manifolds and Riemannian Geometry", Academic Press, 1975
- 5.- R. Penrose. "Techniques of Differential Topology in Relativity", Regional Conference Series in Applied Mathematics, Vol. 7, SIAM publications, 1972.
- [H y E] 6.- S. Hawking y G. F. Ellis. "The Large Scale Structure of Space-Time", Cambridge University Press, 1973.

Los siguientes libros son citados en el texto:

- [E] A. Einstein "The Meaning of Relativity", Princeton University Press, fifth edition, 1974.
- [E,L,W y M] A. Einstein, H. A. Lorentz, H. Weyl, y H. Minkowski. "The Principle of Relativity", Dover, 1953.
- [E,P y S] J. Ehlers, F. A. E. Pirani, y A. Schild. "The Geometry of Free fall and Light Propagation", General relativity papers in honour of J. L. Synge, Oxford University Press, 1972.
- [H] I. N. Herstein. "Topics in Algebra", Wiley, 1975.
- [H y Y] J. G. Hocking y G. S. Young, "Topology", Addison Wesley, 1961.
- [I y V] C. Imaz y Z. Vorel. "Ecuaciones Diferenciales Ordinarias", Limusa, 1975.
- [L] S. Lang. "Introduction to Differentiable Manifolds", Wiley, 1962.
- [S] E. G. Simmons. "Introduction to Topology and Modern Analysis", McGraw-Hill, 1963.
- [S y S] J.L. Synge y A. Schild. "Tensor Calculus", Dover, 1978.
- [Sp] M. Spivak. "Calculo en Variedades", Editorial Reverté, 1975.

+++++