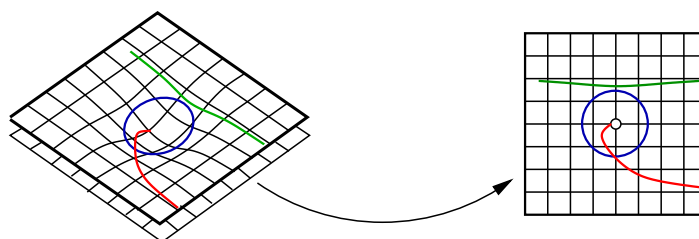


Geometría Diferencial

(teatro de variedades para estudiantes de máster)



Fernando Chamizo Lorente

Curso 2011-2012

o r e n t e F e r n a n d o
L 2011/2012 o
o z i m a h C

Índice general

1. Bases de geometría diferencial	1
1.1. El concepto de variedad	2
1.2. Vectores, covectores y tensores	9
1.3. Formas diferenciales	15
2. El teorema de Frobenius	23
2.1. Flujos de campos de vectores	24
2.2. El corchete de Lie y la derivada de Lie	27
2.3. Condiciones de integrabilidad	30
3. Geometría Riemanniana	35
3.1. Mecánica y métricas	36
3.2. Introducción a la relatividad	41
3.3. Geodésicas y conexiones	51
4. Curvatura en variedades	61
4.1. Definición y significado	62
4.2. Curvatura y topología	69
4.3. Las ecuaciones de campo	79
A. Hojas del curso 2011-2012	85
Bibliografía	95
Índice alfabético	99

Prefacio

Estas notas corresponden a la asignatura “Geometría Diferencial” impartida en el primer cuatrimestre del curso 2011-2012 dentro del Máster en Matemáticas y Aplicaciones del Departamento de Matemáticas de la Universidad Autónoma de Madrid. En cierta medida es una continuación de los cursos de grado de geometría y se superpone en parte con ellos y sobre todo con alguno de la antigua licenciatura. Por ello, especialmente en el primer capítulo, hay fragmentos tomados literalmente de [Cha08] (no se creyó necesario el permiso escrito del autor). Este primer capítulo introduce los elementos de la geometría de variedades y el álgebra tensorial. El segundo capítulo culmina, tras las definiciones necesarias de flujo de campos de vectores y corchete de Lie, con el teorema de Frobenius. El tercer capítulo trata la geometría riemanniana con cierta orientación mecánica que no se refleja habitualmente en los textos de geometría diferencial. Finalmente, el cuarto capítulo se dedica a la curvatura en variedades, explorando sus implicaciones topológicas y en relatividad general.

Varios temas fueron seleccionados en algún momento para formar parte de estas notas pero como no pertenecían al temario del curso, finalmente no llegaron a redactarse. Se enumeran aquí para referencia propia con algunas indicaciones acerca de los posibles contenidos, con vistas a la remota posibilidad de una futura edición aumentada.

1. INTEGRACIÓN EN VARIEDADES. Estudiar con algún detalle el teorema de Stokes que sólo se menciona brevemente al final de la sección 1.3. Mencionar alguna aplicación de topología diferencial [Cha08].
2. FOLIACIONES Y TEOREMA DE FROBENIUS CON FORMAS DIFERENCIALES. El enunciado del teorema de Frobenius es más elegante y de hecho más cercano al enunciado original de Frobenius usando formas diferenciales [Aub01], [Lee09]. También se podría entrar con un poco más de detalle en los grupos de Lie [NT06], [GHL87].
3. TEOREMA DE NOETHER Y LEYES DE CONSERVACIÓN. El teorema de Noether en mecánica es una bella ilustración de los métodos matemáticos de la Física. Se podría incluir el enunciado tal como está en [Arn78] y algunas reflexiones acerca de leyes de conservación [DNF85, §32] y de grupos de difeomorfismos y geodésicas.

4. COSMOLOGÍA Y TEOREMAS DE SINGULARIDAD. Ahondar un poco más en temas de cosmología y, si es posible, dar la prueba completa de algún teorema de singularidad [Nab88], [O'N83], [BE81], [Wal84], [HE73].

A la vista de los resultados, se podría decir que el curso ha sido exitoso: las calificaciones fueron altas y la asistencia continuada y numerosa en comparación con otros cursos del mismo nivel. Una parte minoritaria de la audiencia estaba compuesta por estudiantes no oficialmente matriculados que necesitaban créditos para su doctorado. La evaluación consistió primordialmente en la entrega de cuatro hojas de ejercicios, incluidas aquí como apéndice. Se podía presentar un trabajo sobre un tema libre relacionado con el curso y someterse a un examen final, también dentro del apéndice. Estas dos últimas opciones las eligieron pocos alumnos.

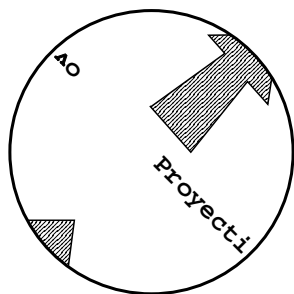
Madrid 21 de junio de 2012

Fernando Chamizo Lorente

Capítulo 1

Bases de geometría diferencial

H. Whitney comenzó un famoso artículo [Whi36] de 1936 escribiendo: *una variedad diferenciable se define en general de dos formas: como un conjunto de puntos con entornos homeomorfos a un espacio euclídeo \mathbb{R}^n , estando las coordenadas en entornos no disjuntos relacionadas por transformaciones diferenciables; o como un subconjunto de \mathbb{R}^n definido en cada punto expresando algunas coordenadas en términos de otras por medio de funciones diferenciables.* Desde entonces ha pasado mucho tiempo pero la herencia de esta dualidad tiene su reflejo en los currículos académicos actuales. Casi siempre hemos preferido ecuaciones pero la geometría de variedades, originada por Riemann y desarrollada a lo largo del siglo XX, pretende estudiar la geometría desde dentro, sin referencia a ningún espacio exterior y por ello ha hecho suya la primera de las definiciones.



Whitney obtuvo un resultado (véase el Teorema 1.1.5 más adelante) que asegura que las variedades siempre se pueden meter dentro de algún \mathbb{R}^n . En sus palabras: *la primera definición no es más general que la segunda.* Quizá al lector le parezca rebuscado entonces el empeño de estudiar siempre los objetos geométricos desde dentro. Si hay un espacio ambiente donde escribir ecuaciones ¿por qué no usarlo?

Pensemos por ejemplo en un universo dado por el disco unidad en el que los habitantes que salen por la frontera vuelven a aparecer en el punto diametralmente opuesto. Localmente ninguno de ellos notaría nada distinto a \mathbb{R}^2 , tendrían que recorrer todo un diámetro para darse cuenta por ejemplo de que su universo es cerrado. La descripción geométrica de este objeto, que no es más que el *espacio proyectivo* $\mathbb{P}^2(\mathbb{R})$, es sencilla y permite desarrollar cierta intuición que perderíamos en gran medida trabajando con las ecuaciones de una subvariedad de \mathbb{R}^4 que es el primer \mathbb{R}^n en el que se puede meter bien [dC92, p.32 Ex.5].

En resumidas cuentas, un espacio ambiente puede ser demasiado artificial para ciertos objetos. Por ejemplo, todos representamos un punto sobre la Tierra por su latitud y su longitud en vez de escribir las tres coordenadas en algún sistema de referencia del \mathbb{R}^3

circundante. En términos cosmológicos también esto es natural. Si el universo es por definición todo, ¿para qué imaginar algo que lo contiene?

Una dificultad en este enfoque es que estamos acostumbrados a ver las velocidades como vectores, flechas, que salen fuera de la variedad. Por ello las definiciones de espacio tangente y de las aplicaciones lineales y multilineales que se viven en él, requerirán una gran dosis de abstracción.

1.1. El concepto de variedad

La idea de variedad diferenciable n -dimensional es la de un objeto geométrico compuesto por parches que son similares a abiertos de \mathbb{R}^n .

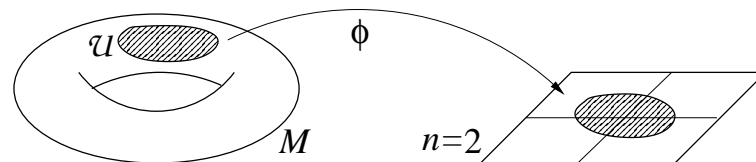
Para que este objeto capture parte de las ideas que tenemos en mente sobre continuidad y comportamiento local, postularemos que sea un espacio topológico con ciertas propiedades básicas.

Hipótesis topológica: Diremos que un espacio topológico M satisface la hipótesis topológica si tiene la propiedad de Hausdorff y admite una base numerable.

Sólo recordaremos de estas propiedades fundamentalmente técnicas que la *propiedad de Hausdorff* dice que dados dos puntos siempre hay entornos disjuntos de ellos. La existencia de una base numerable o *segundo axioma de numerabilidad* [Mun75] asegura que existe una familia numerable de abiertos tal que cualquier entorno de un punto contiene a un entorno que está en dicha familia. En el caso de \mathbb{R} se podrían tomar por ejemplo todos los intervalos con extremos racionales.

En algunos textos esta hipótesis topológica no se incluye en la definición de variedad. Su propósito es poder llevar a cabo el análisis local típico de la geometría diferencial y se manifiesta en la existencia de las particiones de la unidad (Proposición 1.1.4), las cuales permiten modificar ligeramente una función en un entorno de un punto sin que los puntos fuera de él se enteren. Esto se opone a lo que ocurre en la geometría algebraica, donde las funciones son polinómicas o racionales y una modificación en las cercanías de un punto conlleva un cambio global de la función.

Una carta nos dice la manera de allanar un parche de M en \mathbb{R}^n . Los parches allanados son como los mapas de carreteras o los de un callejero, y queremos que cubran todo el territorio aunque cada uno de ellos sólo muestre una pequeña porción.



Definición: Una *carta* n -dimensional de M es un par (\mathcal{U}, ϕ) donde \mathcal{U} es un abierto de M y ϕ es una función $\phi : \mathcal{U} \rightarrow \mathbb{R}^n$ que es homeomorfismo sobre su imagen.

Notación: A veces, con el abuso obvio, se llama carta a la función ϕ . Con frecuencia se denota con x^i la coordenada i -ésima de ϕ , es decir $\phi = (x^1, x^2, \dots, x^n)$, éstas son las llamadas *funciones coordenadas* y se dice que $(x^1(p), \dots, x^n(p))$ son las coordenadas de p en la carta (\mathcal{U}, ϕ) . El lector encontrará extraño el uso de superíndices pero cierto convenio relacionado con los tensores que veremos más adelante sugiere que es una notación apropiada. Para indicar que un abierto \mathcal{U} , típicamente de una carta, contiene al punto p escribiremos $\mathcal{U}(p)$.

Un punto puede estar tapado por varios parches, diferentes abiertos de cartas, debemos asegurarnos de que el análisis no se estropea bajando por una ϕ o por otra.

Definición: Se dice que dos cartas n -dimensionales de M , (\mathcal{U}, ϕ) y (\mathcal{V}, ψ) , son *compatibles* si $\psi \circ \phi^{-1} : \phi(\mathcal{U} \cap \mathcal{V}) \rightarrow \psi(\mathcal{U} \cap \mathcal{V})$ es C^∞ con inversa C^∞ . Se incluye como caso especial en el que \mathcal{U} y \mathcal{V} son disjuntos.

Por razones técnicas es conveniente pensar en todas las cartas n -dimensionales posibles compatibles entre sí (se habla de *atlas maximales*) y se dice que la colección correspondiente $\{(\mathcal{U}_\alpha, \phi_\alpha)\}_{\alpha \in I}$, $M = \bigcup \mathcal{U}_\alpha$, es una *estructura diferenciable n -dimensional*. Con esto llegamos al objetivo: la definición de variedad.

Definición: Se dice que un espacio topológico M con las propiedades anteriores, es una *variedad diferenciable n -dimensional* si está dotado de una estructura diferenciable de dimensión n .

Nota avanzada: Una vez que uno ha puesto la definición en términos suficientemente raros hay una pregunta extraña pero natural. ¿Es posible tener en una esfera o en otro espacio topológico de toda la vida con la topología usual diferentes estructuras diferenciables? En la esfera usual S^2 se sabe que sólo hay una estructura diferenciable pero J. Milnor probó que en la de 7 dimensiones, S^7 , la situación es muy distinta, de hecho hay 28 posibilidades. Sólo una nos resulta familiar y por ello se dice que el resto son *esferas exóticas*. Una interpretación es que las variedades C^0 (con cambios de carta continuos) son bien diferentes de las variedades C^∞ aquí definidas. Por otro lado hay un teorema que afirma que las variedades C^1 son en realidad como las C^∞ eliminando “cartas malas” (véase [Hir76] §2.10 para el enunciado preciso).

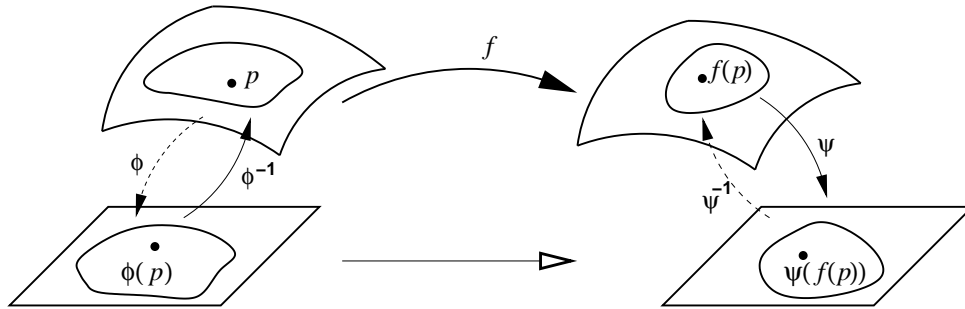
Las cartas permiten definir propiedades locales en la variedad bajando a \mathbb{R}^n . Por ejemplo, se dice que una función $f : M \rightarrow \mathbb{R}$ es C^∞ si para cada carta (\mathcal{U}, ϕ) la función $f \circ \phi^{-1} : \phi(\mathcal{U}) \rightarrow \mathbb{R}$ lo es y se define para cada $p \in \mathcal{U}$ la *derivada parcial i -ésima* en la variedad como

$$\left. \frac{\partial f}{\partial x^i} \right|_p = D_i(f \circ \phi^{-1})(\phi(p))$$

donde el símbolo D_i significa la derivada parcial usual con respecto a la i -ésima variable. En general, si M y N son variedades, se puede hablar de funciones C^∞ , $f : M \rightarrow N$ si para cada par de cartas (\mathcal{U}, ϕ) , (\mathcal{V}, ψ) , respectivamente de M y de N se cumple que $\psi \circ f \circ \phi^{-1} : \phi(\mathcal{U}) \rightarrow \psi(\mathcal{V})$.

Las funciones C^∞ entre variedades que tienen inversa C^∞ se llaman *difeomorfismos*.

La propia notación usual para las funciones coordenadas, dándoles nombre de punto, (x^1, x^2, \dots, x^n) , nos recuerda que la versión operativa de los puntos de una variedad es su reflejo en \mathbb{R}^n después de aplicar la función de una carta.



Observación: Éste es un curso de geometría diferencial, no de análisis, por ello daremos por supuesto que la regularidad no constituye ninguna obstrucción en las definiciones. A partir de ahora supondremos, sin indicarlo cada vez, que todas las funciones entre variedades que consideramos son C^∞ o al menos tantas veces diferenciables como sea necesario para que no haya problemas de regularidad.

Un concepto técnicamente más complejo de generalizar es el espacio tangente, que en el caso de subvariedades de \mathbb{R}^n tenía un significado sencillo (recuérdese el curso de Cálculo III). No es posible una mera adaptación directa porque allí los vectores tangentes eran “pelos” orientados que se salían de la subvariedad, mientras que concebimos las variedades como una entidad única, sin referencia a un posible “exterior”. Hay varias maneras de superar este obstáculo (véase [Jän01]). Aquí mencionaremos las definiciones matemáticas que corresponden a ver los vectores tangentes como velocidades de curvas y como derivadas direccionales. La segunda es más abstracta, introduciendo implícitamente el concepto de *derivación* [O’N83], pero en Geometría III se mostraba más útil en las demostraciones.

Definición: Se llama *espacio tangente* de M en un punto p al conjunto cociente $T_p(M) = \mathcal{K}_p(M) / \sim$ donde $\mathcal{K}_p(M) = \{\text{Funciones } c : (-\epsilon, \epsilon) \rightarrow M \text{ con } c(0) = p\}$ y la relación \sim identifica las funciones (curvas) tales que $(\phi \circ c_1)'(0) = (\phi \circ c_2)'(0)$ con $(\mathcal{U}(p), \phi)$ una carta. Se llama *vector tangente* de M en p a cualquiera de sus elementos.

Definición: Se llama *vector tangente* de M en p a cualquier operador \mathbb{R} -lineal $v : \mathcal{E}_p(M) \rightarrow \mathbb{R}$ que satisface $v(fg) = v(f)g(p) + f(p)v(g)$ para todo $f, g \in \mathcal{E}_p(M)$ donde $\mathcal{E}_p(M)$ es conjunto de funciones $f : \mathcal{U}(p) \rightarrow \mathbb{R}$ con $\mathcal{U}(p)$ un entorno de p . Se llama *espacio tangente* de M en un punto p al conjunto formado por los vectores tangentes.

El nexos entre ambas definiciones es que a cada $c \in \mathcal{K}_p(M)$ se le puede asignar el operador $v : f \mapsto (f \circ c)'(0)$ (véase [Gon05] §3.2).

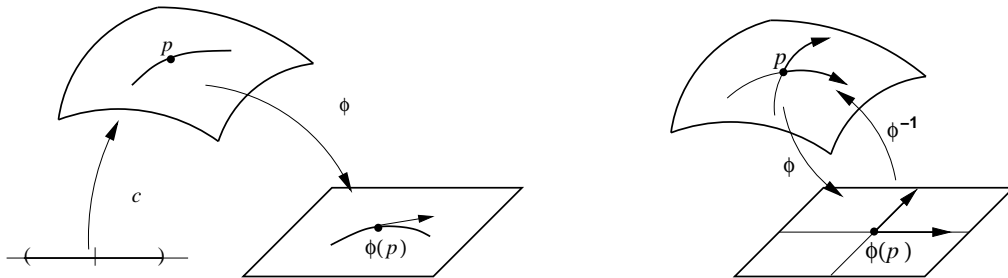
A partir de las curvas que corresponden a los ejes coordenados (una vez que bajamos a \mathbb{R}^n) se obtienen unos vectores tangentes que denotaremos con el extraño nombre $\frac{\partial}{\partial x^i} \Big|_p$. Para ser rigurosos, si $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es la base canónica de \mathbb{R}^n , fijada una carta $(\mathcal{U}(p), \phi = (x^1, \dots, x^n))$, con la primera definición se tiene

$$\frac{\partial}{\partial x^i} \Big|_p = [c_i] \quad \text{donde} \quad c_i(t) = \phi^{-1}(\phi(p) + t\vec{e}_i), \quad i = 1, 2, \dots, n.$$

Denominar a estos vectores con el mismo símbolo que el de las derivadas parciales no es casual pues con la segunda definición no son más que la derivadas parcial i -ésimas en la variedad, es decir

$$(1.1) \quad \frac{\partial}{\partial x^i} \Big|_p : f \longrightarrow \frac{\partial f}{\partial x^k} \Big|_p.$$

A menudo la notación para estos vectores tangentes se abrevia como $\partial_i \Big|_p$ o incluso como ∂_i si no se quiere indicar la carta o el punto.



Como se vio en cursos pasados:

Proposición 1.1.1 *El espacio tangente $T_p(M)$ tiene una estructura natural de espacio vectorial cuya dimensión es la de la variedad diferenciable M .*

Proposición 1.1.2 *Para cada punto p de una variedad diferenciable n -dimensional, M , el conjunto $\{\partial_1 \Big|_p, \partial_2 \Big|_p, \dots, \partial_n \Big|_p\}$ es una base de $T_p(M)$.*

Una vez que tenemos estos resultados y hemos acumulado la miseria debajo de la alfombra de la notación, nos podemos despreocupar de la dificultad y abstracción de los conceptos definidos a la hora de hacer operaciones. Podemos sumar vectores y multiplicarlos por números coordenada a coordenada como nos enseñaron en primero, escribiendo sin remordimientos cosas como: $(2\partial_1 \Big|_p + 3\partial_2 \Big|_p) + 4(\partial_1 \Big|_p - 2\partial_2 \Big|_p) = 6\partial_1 \Big|_p - 5\partial_2 \Big|_p$.

Con $f : M \rightarrow N$ podemos pasar curvas en curvas lo cual induce una aplicación $T_p(M) \rightarrow T_{f(p)}(N)$. Aunque ésta es la idea intuitiva es más sintético proceder tomando en cuenta la segunda definición de espacio tangente.

Definición: Sea $f : M \rightarrow N$. Se llama *aplicación tangente* de f en p y se denota con $df|_p$, a la aplicación lineal $T_p(M) \rightarrow T_{f(p)}(N)$ que aplica el vector $v(\cdot)$ (considerado con la segunda definición) en $v(\cdot \circ f)$.

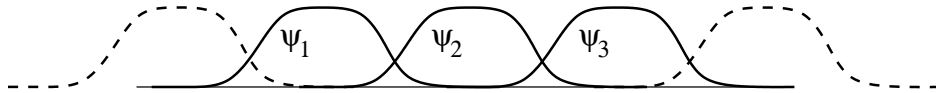
Ahora todo funciona como con la diferencial de toda la vida, siempre componiendo con las cartas.

Proposición 1.1.3 *Sea $f : M \rightarrow N$ y sean $(\mathcal{U}(p), \phi)$ y $(\mathcal{V}(f(p)), \psi)$ cartas de M y N respectivamente en los puntos indicados. La matriz de la aplicación tangente $df|_p$ en las bases $\{\partial/\partial x^1|_p, \dots, \partial/\partial x^m|_p\}$ y $\{\partial/\partial y^1|_{f(p)}, \dots, \partial/\partial y^n|_{f(p)}\}$ correspondientes a estas cartas es la matriz jacobiana de $\psi \circ f \circ \phi^{-1}$ en $\phi(p)$.*

En otras palabras, quieran lo que quieran decir los símbolos $\partial/\partial x^i$, ya sean derivadas de clases de curvas o derivaciones que actúan sobre funciones, el caso es que formalmente se transforman por medio de una matriz jacobiana, es decir, como en los otros cursos cuando el mismo símbolo tenía un significado más restringido.

A pesar de nuestra insistencia en el análisis local, en geometría diferencial también hay operaciones globales que requieren pegar adecuadamente las cartas. Supongamos por ejemplo que deseamos definir la integral sobre una variedad. Sea cual sea su definición, si bajamos a \mathbb{R}^n con cartas cuyos abiertos recubren la variedad, está claro que no basta con integrar allí y sumar los resultados. Un punto puede pertenecer a varios abiertos y en ese caso hay que ponderar su contribución para contarlos sólo una vez. Analíticamente, hay que partir la función idénticamente uno en funciones que vivan en los diferentes abiertos.

Definición: Una *partición de la unidad* subordinada a un recubrimiento abierto $\bigcup \mathcal{U}_\alpha = M$ es una colección de funciones $\{\psi_\alpha\}$, C^∞ y $0 \leq \psi_\alpha \leq 1$, tales que: 1) En cada punto $p \in M$ sólo hay un número finito de ellas con $\psi_\alpha(p) \neq 0$ y su suma es uno. 2) Cada ψ_α tiene soporte incluido en alguno de los \mathcal{U}_β .



La buena noticia acerca de las particiones de la unidad es su existencia. Este hecho está intrínsecamente ligado a la hipótesis topológica. Incluso hay una especie de equivalencia entre ambas [BC70, Prop.3.4.3], [Mun75].

Proposición 1.1.4 *Cualquier recubrimiento abierto de una variedad diferenciable M admite una partición de la unidad subordinada a él.*

Demostración: Digamos que el recubrimiento es $\{\mathcal{U}_\alpha\}$. La hipótesis topológica permite suponer, quizá considerando abiertos más pequeños incluidos en los \mathcal{U}_α , que $\overline{\mathcal{U}_\alpha}$ (el cierre) es compacto, que para todo $p \in M$ el conjunto $\{\alpha : p \in \mathcal{U}_\alpha\}$ es finito y que existe un recubrimiento $\{\mathcal{V}_\alpha\}$ con $\overline{\mathcal{V}_\alpha} \subset \mathcal{U}_\alpha$ gozando de las mismas propiedades (en nuestro caso también se puede suponer que ambas colecciones son numerables). Todo esto es un hecho topológico general que responde a las palabras mágicas de que Hausdorff, localmente compacto y base numerable implican paracompacto y existencia de refinamientos¹.

Si $(\mathcal{U}_\alpha, \phi_\alpha)$ son cartas (siempre podemos trocear los \mathcal{U}_α para que estén incluidos en los dominios de las funciones de las cartas), la compacidad asegura que $\phi_\alpha(\mathcal{V}_\alpha) \subset \mathbb{R}^n$ se puede recubrir por un número finito de bolas, $B_{\alpha,1}, B_{\alpha,2}, \dots, B_{\alpha,k(\alpha)} \subset \mathcal{U}_\alpha$. Sean las funciones $\eta_{\alpha,i} : \mathbb{R}^n \rightarrow [0, 1]$ que son idénticamente 1 en $B_{\alpha,k}$ y cero fuera de alguna bola concéntrica contenida en \mathcal{U}_α . La prueba termina considerando

$$\psi_\alpha(p) = \frac{\mu_\alpha(p)}{\sum_{\beta : p \in \mathcal{U}_\beta} \mu_\beta(p)} \quad \text{con} \quad \mu_\alpha(p) = \begin{cases} \sum_{i=1}^{k(\alpha)} \eta_{\alpha,i}(\phi_\alpha(p)) & \text{si } p \in \mathcal{U}_\alpha \\ 0 & \text{si } p \in M - \mathcal{U}_\alpha \end{cases}$$

que está bien definida porque todo p pertenece a algún \mathcal{V}_α lo que asegura que algún $\eta_{\alpha,i}(\phi_\alpha(p))$ es no nulo. \square

Las particiones de la unidad constituyen el ingrediente principal para probar que las variedades compactas se pueden considerar subvariedades de algún \mathbb{R}^n . Enunciamos aquí el resultado general de Whitney aunque sólo daremos un esquema de la prueba en el caso compacto, siguiendo [Gua08], que es bastante más sencillo que el caso general [Hir76].

Teorema 1.1.5 (de Whitney) *Una variedad diferenciable de dimensión n admite una inmersión en \mathbb{R}^{2n+1} .*

Observación: La palabra *inmersión* traduce tanto a *imbedding* como a *embedding* que son distintos en la terminología anglosajona americana. El segundo concepto es más

¹Esbozamos la prueba en el caso de variedades: Elegimos una colección numerable de cartas $\{(\mathcal{O}_i, \phi_i)\}_{i=1}^\infty$ con $\bigcup \mathcal{O}_i = M$ y $\overline{\mathcal{O}_i}$ es compacto (tómense por ejemplo preimágenes de bolas en \mathbb{R}^n y empléese que hay una base numerable). De la misma forma escogemos un recubrimiento $\{\mathcal{U}'_i\}_{i=1}^\infty$ tal que $\overline{\mathcal{U}'_i}$ es compacto y está contenido en algún \mathcal{U}_{α_i} . Para $k \geq 1$ sean el abierto $\mathcal{W}_k = \bigcup_{i=1}^{k+2} \mathcal{O}_i - \bigcup_{i=1}^{k-1} \overline{\mathcal{O}_i}$, el compacto $C_k = \bigcup_{i=1}^{k+1} \overline{\mathcal{O}_i} - \bigcup_{i=1}^k \mathcal{O}_i$ y otro abierto $\mathcal{W}'_k \supset C_k$ con $\overline{\mathcal{W}'_k} \subset \mathcal{W}_k$. Consideremos un subrecubrimiento finito de C_k por abiertos $\mathcal{U}'_j \cap \mathcal{W}'_k$. La colección $\{\mathcal{V}_j\}$ formada por todos estos abiertos (cuando varía k) cumple que $\{j : p \in \mathcal{V}_j\}$ es finito porque cada $p \in M$ está en algún $C_k \subset \mathcal{W}_k$ y $\mathcal{W}_i \cap \mathcal{W}_j = \emptyset$ si $i \geq j+3$. Para obtener el recubrimiento grande basta considerar $\mathcal{U}_{\alpha_j} \cap \mathcal{W}_k$ al recubrir C_k .

fuerte pues requiere que la aplicación sea homeomorfismo sobre su imagen. Es en ese sentido fuerte en el que se emplea aquí. Es decir, nuestra inmersión es una función cuya aplicación tangente es inyectiva (no pega direcciones distintas) y que tiene una inversa bien definida sobre su imagen (no pega puntos, ni los acerca infinitamente).

Demostración (Esbozo para el caso compacto): Por la compacidad, existe una colección finita de cartas $\{(\mathcal{O}_i, \phi_i)\}$ con $M = \bigcup \mathcal{O}_i$. Revisando el argumento topológico de la Proposición 1.1.4, vemos que da lugar a recubrimientos finitos $\{\mathcal{V}_i\}_{i=1}^N$ y $\{\mathcal{U}_i\}_{i=1}^N$. Entonces se puede suponer que la partición de la unidad es $\{\psi_i\}_{i=1}^N$ con $\text{supp } \psi_i \subset \mathcal{U}_i$ y (\mathcal{U}_i, ϕ_i) cartas. Sea $\tilde{\phi}_i(p) = \psi_i(p)\phi_i(p)$ si $p \in \mathcal{U}_i$ y cero en otro caso. Claramente $\tilde{\phi}_i : M \rightarrow \mathbb{R}^n$ es C^∞ . Consideremos $f : M \rightarrow \mathbb{R}^D$ con $D = (n+1)N$ definida por

$$f(p) = (\tilde{\phi}_1(p), \tilde{\phi}_2(p), \dots, \tilde{\phi}_N(p), \psi_1(p), \psi_2(p), \dots, \psi_N(p)).$$

Dado $p \in M$ existe i tal que $\psi_i(p) \neq 0$ y si $f(p) = f(q)$ se tendría $\psi_i(p) = \psi_i(q) \neq 0$ y $\tilde{\phi}_i(p) = \tilde{\phi}_i(q)$ y la inyectividad de ϕ_i implica $p = q$. Por topología básica, una función inyectiva definida en un conjunto compacto es homeomorfismo sobre su imagen. Por otro lado, la matriz jacobiana $D(f \circ \phi_i^{-1})$ tiene rango (máximo) n para todo $x \in \phi_i(\mathcal{U}_i)$.

En conclusión, tenemos una inmersión $f : M \rightarrow \mathbb{R}^D$ pero D está relacionado con el número de cartas y parece muy difícil de controlar. Las líneas que restan esbozan cómo demostrar que cualquier $D > 2n+1$ se puede reducir componiendo f con la proyección P sobre el hiperplano $\{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{v} = 0\}$ para algún vector unitario $\vec{v} \in S^{D-1}$.

Si $P \circ f$ dejara de ser inyectiva entonces $f(p)$ y $f(q)$ deberían determinar un vector proporcional a \vec{v} para ciertos $p, q \in M$. Dicho de otra forma, \vec{v} estaría en la imagen de la función

$$a : M \times M - \{(p, p) : p \in M\} \rightarrow S^{D-1} \quad \text{con} \quad a(p, q) = \frac{f(q) - f(p)}{\|f(q) - f(p)\|}.$$

Por otro lado, si la jacobiana $D(P \circ f \circ \phi_i^{-1})$ dejara de tener rango máximo entonces \vec{v} estaría en la imagen de $D(f \circ \phi_i^{-1})$ ya que \vec{v} genera el núcleo de $D(P) = P$. En otras palabras, se tendría que \vec{v} está en la imagen de

$$b_i : \phi_i(\mathcal{U}_i) \times S^{n-1} \rightarrow S^{D-1} \quad \text{con} \quad b_i(x, \vec{w}) = \frac{(D(f \circ \phi_i^{-1})(x))(\vec{w})}{\|(D(f \circ \phi_i^{-1})(x))(\vec{w})\|}.$$

Ahora bien, la variedad origen de a tiene dimensión $2n$ y las de las b_i tienen dimensión $2n-1$. La unión de sus imágenes en S^{D-1} será, por tanto, a lo más $2n$ dimensional y si $D-1 > 2n$ siempre hay posibilidad de escoger \vec{v} fuera de este conjunto, garantizando que la proyección reduce D .

Este argumento final con dimensiones necesita alguna justificación. Basta una versión sencilla del llamado *Lema de Sard* [Gua08, Cor.2.10], [Spi82]. \square

1.2. Vectores, covectores y tensores

En primer lugar vamos a ver algunas definiciones que no tiene que ver directamente con la geometría de variedad sino que constituyen una extensión natural del álgebra lineal estudiada en cursos anteriores.

Definición: Se dice que $f : V_1 \times V_2 \times \cdots \times V_n \longrightarrow W$, donde V_1, V_2, \dots, V_n, W son espacios vectoriales, es una *aplicación multilineal* si para todo $1 \leq i \leq n$

- a) $f(\vec{v}_1, \dots, \lambda \vec{v}_i, \dots, \vec{v}_n) = \lambda f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_n)$ con $\lambda \in \mathbb{R}$
 b) $f(\vec{v}_1, \dots, \vec{v}_i + \vec{v}_i', \dots, \vec{v}_n) = f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_n) + f(\vec{v}_1, \dots, \vec{v}_i', \dots, \vec{v}_n)$.

Es habitual que las variables de una aplicación multilineal tengan todas la misma naturaleza y por tanto $V_1 = V_2 = \cdots = V_n$. Daremos un nombre a esta situación en el caso simple en que $W = \mathbb{R}$.

Definición: Se llama *tensor n veces covariante* a cualquier aplicación multilineal de la forma $T : V \times \cdots \times V \longrightarrow \mathbb{R}$.

Ejemplo: El producto escalar usual en \mathbb{R}^m define un tensor dos veces covariante.

Ejemplo: El determinante aplicado a m vectores de \mathbb{R}^m define un tensor m veces covariante.

Ejemplo: La función que asigna a n vectores de \mathbb{R}^m el producto de sus primeras coordenadas (en la base canónica) es un tensor n veces covariante.

Al igual que en cálculo de varias variables se consideran funciones vectoriales, también podríamos definir algo así como tensores vectoriales, de la forma $f : V \times \cdots \times V \longrightarrow V$ o incluso complicar más las cosas permitiendo $f : V \times \cdots \times V \longrightarrow V \times V$, etc. Cada vector “vertical” puede pasarse a un número real (pre-)multiplicando por un vector “horizontal”, así que a cada $f : V \times \cdots \times V \longrightarrow V$ se le puede asociar $T : V^* \times V \times \cdots \times V \longrightarrow \mathbb{R}$ dada por $T(\tilde{\varphi}, \vec{v}_1, \vec{v}_2, \dots, \vec{v}_n) = \tilde{\varphi}(f(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n))$ para cada $\tilde{\varphi} \in V^*$. Además por el isomorfismo entre V y V^* , esta correspondencia es biyectiva².

En definitiva, da igual considerar los hipotéticos tensores vectoriales, sin aceptación entre los matemáticos, que considerar los tensores antes definidos pero permitiendo sustituir algunos de los factores V por V^* . Lo más breve es generalizar de esta forma la definición anterior.

Definición: Se llama *tensor r veces contravariante y s veces covariante* o *tensor de tipo (r, s)* a una aplicación multilineal $T : V^* \times \cdots \times V^* \times V \times \cdots \times V \longrightarrow \mathbb{R}$.

²Todo este párrafo se resume en lo siguiente: si tienes un vector y quieres un número, haz el producto escalar con otro vector arbitrario y si además quieres quedar bien, di que ésa es la acción de V^* sobre V .

Comparando con la definición previa, un tensor n veces covariante es un tensor de tipo $(0, n)$. Por otro lado los tensores de tipo $(n, 0)$ se dice que son n veces contravariantes. Por convenio además diremos que *una constante es un tensor de tipo* $(0, 0)$. Obsérvese que hay cierta lógica en esta notación porque una constante no depende de ningún vector.

Como ejemplo, nótese que un endomorfismo $f : V \rightarrow V$ asigna a cada vector otro vector, y según la identificación anterior da lugar a un tensor de tipo $(1, 1)$. En coordenadas, si representamos el endomorfismo como $f(\vec{v}) = A\vec{v}$ para cierta matriz cuadrada A y un elemento $\tilde{\varphi} \in V^*$ como un vector horizontal, el tensor correspondiente es $T(\tilde{\varphi}, \vec{v}) = \tilde{\varphi}(A\vec{v})$.

Al igual que hablamos de las componentes (o entradas o coeficientes) de una matriz en cierta base, nos podemos referir a las componentes de un tensor.

Definición: Supongamos que $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m\}$ es una base de V y la base dual es $\mathcal{B}^* = \{\tilde{\varphi}^1, \tilde{\varphi}^2, \dots, \tilde{\varphi}^m\} \subset V^*$. Se llaman *componentes de un tensor*, T , de tipo (r, s) , en estas bases a los números reales

$$T_{j_1 j_2 \dots j_s}^{i_1 i_2 \dots i_r} = T(\tilde{\varphi}^{i_1}, \tilde{\varphi}^{i_2}, \dots, \tilde{\varphi}^{i_r}, \vec{e}_{j_1}, \vec{e}_{j_2}, \dots, \vec{e}_{j_s}).$$

A partir de ahora pondremos especial atención en enumerar los elementos de V (los vectores) con subíndices y los de V^* (a veces llamados *contravectores*) con superíndices para que sea más claro de dónde viene cada componente de un tensor

Ejemplo: Calcular las componentes del tensor D definido por el determinante en \mathbb{R}^2 con la base usual.

Claramente $D(\vec{e}_1, \vec{e}_1) = D(\vec{e}_2, \vec{e}_2) = 0$ y $D(\vec{e}_1, \vec{e}_2) = -D(\vec{e}_2, \vec{e}_1) = 1$, por lo que sus componentes son $D_{11} = D_{22} = 0$, $D_{12} = -D_{21} = 1$. Esto está estrechamente relacionado con la igualdad (inútil)

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = (a \ b) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix}.$$

Nótese que una igualdad similar para el determinante en \mathbb{R}^m requeriría algo así como “matrices m -dimensionales” cuyos elementos serían las componentes del tensor.

Está claro que una vez fijada una base un tensor está determinado por sus componentes. Por ejemplo, el tensor T de tipo $(1, 1)$ correspondiente a un endomorfismo tiene como componente T_j^i el elemento ij de la matriz que lo define en cierta base. Para el endomorfismo identidad las componentes se suelen denotar con el símbolo δ_j^i que significa

$$\delta_j^i = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

Un vector \vec{v} también puede considerarse como un tensor de tipo $(1, 0)$ que aplica cada $\tilde{\varphi} \in V^*$ en $\tilde{\varphi}(\vec{v})$ y sus componentes en una base son simplemente sus coordenadas. De

la misma forma un elemento de V^* se puede considerar un tensor de tipo $(0, 1)$ cuyas componentes son sus coordenadas en la base dual. Consecuentemente el concepto de tensor engloba a los principales personajes del álgebra lineal del primer curso.

La notación tensorial es en principio un poco aparatosa. Por ejemplo, un tensor $(1, 3)$ muy importante es el llamado tensor de Riemann $R : V^* \times V \times V \times V \rightarrow \mathbb{R}$, que introduciremos en otro capítulo. En relatividad $\dim V = 4$ y R tiene $4 \cdot 4 \cdot 4 \cdot 4 = 256$ componentes y para aplicarlo a un elemento del dual, digamos con componentes (a_1, a_2, a_3, a_4) , y a tres vectores, cuyas coordenadas numeramos con superíndices, (b^1, b^2, b^3, b^4) , (c^1, c^2, c^3, c^4) , (d^1, d^2, d^3, d^4) , debemos escribir

$$\sum_{i=1}^4 \sum_{j=1}^4 \sum_{k=1}^4 \sum_{l=1}^4 R_{ijkl}^i a_i b^j c^k d^l$$

que, ciertamente, contiene muchos sumatorios.

Si utilizamos la notación de subíndices y superíndices (correspondientes a vectores y covectores) introducida aquí, se produce una simplificación sustancial usando el llamado *convenio de sumación de Einstein*³ que consiste en *sobreentender un sumatorio cada vez que un subíndice aparece también como superíndice*. Por ejemplo, la expresión anterior se escribe simplemente como

$$R_{ijkl}^i a_i b^j c^k d^l.$$

Las relaciones matriciales desde el punto de vista de las coordenadas, se reducen enormemente con este convenio y se vuelven más intuitivas. Así el efecto sobre las coordenadas de una aplicación lineal, digamos $\vec{y} = A\vec{x}$, se escribe

$$y^i = a_j^i x^j.$$

Y la igualdad matricial $D = ABC$ componente a componente, se reduce a

$$d_j^i = a_k^i b_l^k c_j^l.$$

Nótese lo sencillo que es de recordar apelando a una “simplificación” de índices. Lo mismo se aplica para abreviar combinaciones lineales. Por ejemplo, para decir que las coordenadas de \vec{v} en la base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m\}$ son a^1, a^2, \dots, a^m

$$\vec{v} = \sum_{j=1}^m a^j \vec{e}_j \quad \text{se abrevia como} \quad \vec{v} = a^j \vec{e}_j.$$

³Este convenio fue realmente introducido por Einstein quien bromeó al respecto diciendo: “He hecho un gran descubrimiento en Matemáticas; he suprimido el signo de sumación toda vez que la suma se haga en un índice que aparece dos veces”.

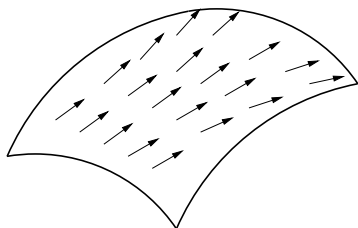
En definitiva:

Un índice duplicado arriba y abajo indica un sumatorio.

Es importante insistir en que todo funciona como si los índices repetidos se simplificasen. Por ejemplo, R_{jkl}^i es un tensor $(1, 3)$ pero como R_{jil}^i sólo depende de dos índices, j y l , es $(0, 2)$. También $a_k^i b_j^l$ representa un tensor $(2, 2)$ y $a_k^i b_j^k$ representa un tensor $(1, 1)$. Este fenómeno de igualar un índice y un subíndice y sumar en ellos, se llama *contracción*. Ahora podemos apreciar la conveniencia de pensar en las constantes como tensores de tipo $(0, 0)$. Un tensor de este tipo corresponde por ejemplo a la contracción del producto de un tensor $(0, 1)$ por otro $(1, 0)$; lo cual puede entenderse como $\tilde{\varphi}(\vec{v})$ con $\tilde{\varphi} \in V^*$, $\vec{v} \in V$, y el resultado de esta operación es constante. La contracción de un tensor está bien definida: no depende de la base en la que se lleva a cabo porque, como veremos con detalle en la tercera sección, las reglas de transformación asociadas a subíndices y superíndices (vectores y contravectores) son inversas.

Nuestra intención es llenar una variedad de tensores, uno en cada plano tangente, conservando cierta suavidad entre ellos, lo que requiere cierta noción de proximidad.

La manera más sintética de concretar este punto pasa por dar una estructura de variedad al conjunto



$$TM = \bigcup_p T_p(M).$$

El objeto resultante es el llamado *fibrado tangente*. Una vez que tenemos esta estructura podemos hablar de planos tangentes cercanos y de tensores cercanos. En vez de seguir este camino, sin duda

más directo e invariante y que nos introduce a la teoría de fibrados, elegiremos una definición que involucra cartas y componentes. Para ir poco a poco, llenaremos primero de “pelos” tangentes a la variedad.

Definición: Sea M una variedad n -dimensional. Un *campo de vectores* C^∞ en M es una aplicación que asigna a cada punto $p \in M$ un vector de $T_p(M)$, de manera que en cada carta se escribe como $\sum a^i(p) \partial_i|_p$ con a^i funciones C^∞ .

Para definir los tensores en variedades hay que introducir los duales de los vectores. No hace falta proceder en abstracto, nos podemos apoyar en conceptos anteriores:

Dada una carta $(\mathcal{U}(p), \phi = (x^1, \dots, x^n))$ de M tiene sentido considerar $dx^i|_p$, las aplicaciones tangentes de las funciones coordenadas como funciones de M en \mathbb{R} con la estructura de variedad obvia. Usando las definiciones de vector tangente y aplicación tangente se puede probar que

$$dx^i|_p \left(\frac{\partial}{\partial x^j} \Big|_p \right) = \delta_j^i.$$

Dicho de otra forma, $\{dx^1|_p, dx^2|_p, \dots, dx^n|_p\}$ es la base dual de $\{\frac{\partial}{\partial x^1}|_p, \frac{\partial}{\partial x^2}|_p, \dots, \frac{\partial}{\partial x^n}|_p\}$.

Definición: Dada una carta $(\mathcal{U}(p), \phi = (x^1, \dots, x^n))$ de M , al espacio vectorial sobre \mathbb{R} generado por $\{dx^1|_p, dx^2|_p, \dots, dx^n|_p\}$ se le denomina *espacio cotangente* de M en p y se denota con $T_p^*(M)$, por ser el dual de $T_p(M)$. Los elementos de $T_p^*(M)$ se llaman *uno formas* o *covectores*.

Si por ejemplo en $dx^1|_p$ no fijamos p , en cuyo caso escribiremos habitualmente dx^1 , ya tenemos un campo de uno formas definido en un abierto de la variedad. Como las uno formas son operadores lineales que actúan sobre los vectores, lo que tenemos es un campo de tensores $(0, 1)$. No hay razón para posponer más la definición general.

Definición: Sea M una variedad n -dimensional. Un *campo tensorial* C^∞ de tipo (r, s) en M , o simplemente un *tensor* de tipo (r, s) en M , es una aplicación que asigna a cada punto $p \in M$ un tensor de tipo (r, s) con $V = T_p(M)$, $V^* = T_p^*(M)$ y que en cada carta tiene componentes C^∞ .

Siguiendo el convenio que veníamos manejando en el caso $r = s = 0$, un tensor de tipo $(0, 0)$ en M le asigna a cada punto una constante, es decir, es simplemente una función C^∞ .

Las componentes de un tensor T de tipo (r, s) en una variedad definen en cada carta $(\mathcal{U}, \phi = (x^1, \dots, x^n))$ funciones C^∞ de \mathcal{U} en \mathbb{R} dadas por

$$p \in \mathcal{U} \mapsto T(p)(dx^{i_1}|_p, dx^{i_2}|_p, \dots, dx^{i_r}|_p, \frac{\partial}{\partial x^{j_1}}|_p, \frac{\partial}{\partial x^{j_2}}|_p, \dots, \frac{\partial}{\partial x^{j_s}}|_p).$$

Habitualmente expresaremos estas componentes en términos de las funciones coordenadas, que a su vez dependen del punto p .

Ejemplo: En S^1 tenemos la carta $(S^1 - \{(-1, 0)\}, \theta)$ donde $\theta = \theta(x, y)$ da el argumento (ángulo) de cada punto $(x, y) \in S^1$ en el rango $(-\pi, \pi)$. La fórmula

$$T = (x + y) \frac{\partial}{\partial \theta}$$

define un campo de vectores C^∞ en (la subvariedad) $S^1 - \{(-1, 0)\}$ porque $f : S^1 - \{(-1, 0)\} \rightarrow \mathbb{R}$ dada por $f(x, y) = x + y$ es C^∞ , ya que $f \circ \theta^{-1}(t) = \cos t + \sin t$ es C^∞ como función de $(-\pi, \pi)$ en \mathbb{R} . Como $x = \cos \theta$, $y = \sin \theta$, podemos escribir

$$T = (\cos \theta + \sin \theta) \frac{\partial}{\partial \theta}.$$

No hay un gran abuso en dar la componente en términos de la función coordenada θ pues a fin de cuentas $\theta = \theta(x, y)$. Si uno se pusiera muy pesado y quisiera ver la dependencia completa en el punto (x, y) debería escribir

$$T = (\cos \theta(x, y) + \sin \theta(x, y)) \frac{\partial}{\partial \theta} \Big|_{(x, y)}.$$

Dadas dos cartas $(\mathcal{U}, \phi = (x^1, \dots, x^m))$, $(\mathcal{U}', \phi' = (x'^1, \dots, x'^m))$ que se solapan, $\mathcal{U} \cap \mathcal{U}' \neq \emptyset$, la función $\phi \circ \phi'^{-1}$ pasa de (x'^1, \dots, x'^m) a (x^1, \dots, x^m) y por razones obvias la matriz de su diferencial se suele escribir $\partial x^i / \partial x'^j$ y su inversa $\partial x'^i / \partial x^j$. En cada carta se tendrán campos $\partial / \partial x^1, \dots, \partial / \partial x^m$, dx^1, \dots, dx^m (usando ϕ) y $\partial / \partial x'^1, \dots, \partial / \partial x'^m$, dx'^1, \dots, dx'^m (usando ϕ') que dan las bases del espacio tangente y cotangente.

Lema 1.2.1 *Con la notación anterior*

$$1) \frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i}, \quad 2) dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j.$$

Demostración: Si consideramos la aplicación tangente de la función $\text{Id} : M \rightarrow M$, 1) es consecuencia inmediata de la Proposición 1.1.3. Para dar una prueba independiente, nótese que la regla de la cadena asegura que para cada función $f : M \rightarrow \mathbb{R}$ se cumple, $D(f \circ \phi'^{-1}) = D(f \circ \phi^{-1}) \cdot D(\phi \circ \phi'^{-1})$ y empleando (1.1) se tiene el resultado, ya que las componentes de las matrices fila $D(f \circ \phi'^{-1})$ y $D(f \circ \phi^{-1})$ representan la acción de $\partial / \partial x'^j$ y $\partial / \partial x^i$ sobre f .

Para comprobar 2) basta ver que ambos miembros aplicados a cualquier $\partial / \partial x'^l$ dan el mismo resultado. Para el primer miembro éste es, por definición, δ_l^i y para el segundo

$$\frac{\partial x'^i}{\partial x^j} dx^j \left(\frac{\partial}{\partial x'^l} \right) = \frac{\partial x'^i}{\partial x^j} dx^j \left(\frac{\partial x^k}{\partial x'^l} \frac{\partial}{\partial x^k} \right) = \frac{\partial x'^i}{\partial x^j} \frac{\partial x^k}{\partial x'^l} \delta_k^j = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^k}{\partial x'^l} = \delta_l^i$$

donde en el primer paso se ha usado 1) y en el último que la primera matriz es inversa de la segunda. \square

Estas relaciones prueban que para cualquier tensor

$$T(dx'^{i_1}, \dots, dx'^{i_r}, \frac{\partial}{\partial x'^{j_1}}, \dots, \frac{\partial}{\partial x'^{j_s}})$$

coincide con

$$T\left(\frac{\partial x'^{i_1}}{\partial x^k} dx^k, \dots, \frac{\partial x'^{i_r}}{\partial x^k} dx^k, \frac{\partial x^l}{\partial x'^{j_1}} \frac{\partial}{\partial x^l}, \dots, \frac{\partial x^l}{\partial x'^{j_s}} \frac{\partial}{\partial x^l}\right)$$

Por tanto, cuando cambiamos de carta (o parametrización) las componentes de un tensor de tipo (r, s) en una variedad cambian por la fórmula

$$(1.2) \quad \boxed{T_{j_1 j_2 \dots j_s}^{i_1 i_2 \dots i_r} = \left(\frac{\partial x'^{i_1}}{\partial x^{k_1}} \cdot \frac{\partial x'^{i_2}}{\partial x^{k_2}} \cdot \dots \cdot \frac{\partial x'^{i_r}}{\partial x^{k_r}} \right) \cdot \left(\frac{\partial x^{l_1}}{\partial x'^{j_1}} \cdot \frac{\partial x^{l_2}}{\partial x'^{j_2}} \cdot \dots \cdot \frac{\partial x^{l_s}}{\partial x'^{j_s}} \right) T_{l_1 l_2 \dots l_s}^{k_1 k_2 \dots k_r}}$$

Esta fórmula es tan característica de los tensores que en muchos libros, sobre todo en los más orientados a la Física, se definen los tensores y campos de tensores como conjuntos de números o funciones sujetos a esta regla de transformación, que a veces se llama *tensorialidad* por antonomasia. No hay que asustarse con una expresión tan

compleja. En primer lugar, es fácil de recordar notando que los índices repetidos se deben “simplificar”. Y por otra parte, no tiene un significado profundo, simplemente representa lo que ocurre cuando cambiamos de base las variables de un tensor; lo que hay de singular es que los cambios de carta corresponden a cambios de base en el espacio tangente y cotangente cuya matriz es un poco fea: la jacobiana (o su inversa). Ahora podemos apreciar por qué la contracción está bien definida, basta aplicar la regla de la cadena para darse cuenta de que la contracción de un tensor se transforma como un tensor (ejercicio).

En general los tensores no se comportan bien al derivarlos componente a componente porque en (1.2) aparecerían derivadas segundas que estropean la tensorialidad. Más adelante introduciremos una derivada especial que tiene carácter tensorial. Veamos un ejemplo trivial en el que sí se puede derivar y nos debería hacer dudar del nombre “vector” gradiente.

Ejemplo: Una función $f : M \rightarrow \mathbb{R}$ es por definición un tensor de tipo $(0, 0)$, su única componente es la propia función. Sus derivadas parciales definen un tensor porque

$$\frac{\partial f}{\partial x'^j} = \frac{\partial x^l}{\partial x'^j} \frac{\partial f}{\partial x^l}.$$

Comparando con (1.2) vemos que las componentes del gradiente en variedades (que obviamente generaliza al habitual) corresponden a un tensor de tipo $(0, 1)$, no un tensor $(1, 0)$ que representaría un vector. Esto es natural porque $df = \frac{\partial f}{\partial x^1} dx^1 + \frac{\partial f}{\partial x^2} dx^2 + \dots + \frac{\partial f}{\partial x^n} dx^n$ es una uno forma.

Si todavía queda algún escéptico, tómese $f(x, y, z) = x + y + z$ definida en \mathbb{R}^3 con la carta trivial. La transformación (cambio de carta) $x' = 2x, y' = 2y, z' = 2z$ pasa el vector de $T_0(\mathbb{R}^3)$ de coordenadas $(1, 1, 1)$ al de coordenadas $(2, 2, 2)$. El gradiente de $x + y + z$ es $(1, 1, 1)$ pero el de $x'/2 + y'/2 + z'/2 (= x + y + z$ en las nuevas coordenadas) es $(1/2, 1/2, 1/2)$. Por mucho que nos empeñemos el vector gradiente no es un vector⁴ en sentido estricto.

1.3. Formas diferenciales

Los contenidos de esta sección están ligados al nombre de É. Cartan (no confundir con su hijo H. Cartan, también matemático renombrado), quien introdujo el concepto

⁴R. P. Feynman, premio Nobel de Física, dedica toda la sección 2-5 de su magnífico libro [FLS64] a demostrar al lector que el vector gradiente es un vector. ¿Dónde ha quedado el argumento de autoridad? El truco está en que Feynman sólo considera transformaciones dadas por matrices ortogonales (realmente sólo giros) y recuérdese que estas matrices cumplen $A = (A^{-1})^t$, por tanto intercambiar índices y numeradores por denominadores no tiene efecto sobre (1.2). Geométricamente el gradiente es un vector normal, y sigue siéndolo cuando sólo hacemos movimientos en \mathbb{R}^n pero como hemos visto, el gradiente no se comporta como un vector por cambios de carta generales.

de forma diferencial tal como ahora lo conocemos y además definió una nueva operación, la derivada exterior, que resulta fundamental para escribir sintéticamente algunos resultados de geometría diferencial, en especial los relacionados con la integración, como el teorema de Stokes. La estructura algebraica subyacente en la que se basó se llama álgebra exterior y fue introducida por H. Grassmann con anterioridad.

Las definiciones de partida son muy simples.

Definición: Se llama *k-forma alternada* a un tensor k veces covariante, T , que es antisimétrico en cualquier par de argumentos, es decir

$$T(\dots, \vec{v}_i, \dots, \vec{v}_j, \dots) = -T(\dots, \vec{v}_j, \dots, \vec{v}_i, \dots).$$

El conjunto de k -formas alternadas sobre un espacio vectorial V se denota con $\text{Alt}^k(V)$.

Las formas diferenciales corresponden al caso en que V es el espacio tangente de una variedad.

Definición: Una *k-forma diferencial* es un campo de k -formas alternadas en una variedad. El conjunto de k -formas diferenciales sobre una variedad M se denota con $\Omega^k(M)$.

En analogía con lo que se hacía en la teoría general de tensores, se conviene que $\text{Alt}^0(V)$ son las constantes, esto es, \mathbb{R} y por tanto $\Omega^0(M) = C^\infty(M)$. El caso $k = 1$ es también un poco singular porque las definiciones no imponen ninguna restricción y decir 1-forma alternada es lo mismo que decir tensor una vez covariante. Nótese que esto es coherente con nuestra denominación de “uno formas” en la sección anterior para referirnos a los covectores. Habitualmente se suelen representar las formas diferenciales (y también a veces las formas alternadas) con letras griegas minúsculas, especialmente ω y η . Evidentemente si $m > \dim V$ toda m -forma alternada es nula, en particular en una variedad n -dimensional $\Omega^m(M) = \{0\}$ para $m > n$.

El conjunto $\text{Alt}^k(V)$ tiene una estructura de espacio vectorial sobre \mathbb{R} con las operaciones habituales de suma y multiplicación por números reales. En el caso de $\Omega^k(M)$ esos números reales dependerán del punto sobre el que estemos considerando el espacio tangente y por tanto son funciones $f : M \rightarrow \mathbb{R}$ que, como siempre, supondremos C^∞ . La inconveniencia de que estas funciones no formen un cuerpo estropea la estructura de espacio vectorial.

Ejemplo: Hallar todos los tensores $T \in \text{Alt}^2(\mathbb{R}^3)$.

Consideremos dos vectores genéricos $\vec{v}_j = a_j^i \vec{e}_i$, $j = 1, 2$. El desarrollo de $T(\vec{v}_1, \vec{v}_2)$ da lugar a nueve términos. Olvidándonos de los tres con coeficientes nulos $T_{ii} = T(\vec{e}_i, \vec{e}_i)$ y agrupando $T_{ij} = T(\vec{e}_i, \vec{e}_j)$ con su negativo $T_{ji} = T(\vec{e}_j, \vec{e}_i)$, se llega a

$$T(\vec{v}_1, \vec{v}_2) = \begin{vmatrix} a_1^2 & a_1^3 \\ a_2^2 & a_2^3 \end{vmatrix} T_{23} + \begin{vmatrix} a_1^3 & a_1^1 \\ a_2^3 & a_2^1 \end{vmatrix} T_{31} + \begin{vmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{vmatrix} T_{12}.$$

Cada elección de las constantes T_{23} , T_{31} y T_{12} da lugar a un elemento de $\text{Alt}^2(\mathbb{R}^3)$ distinto.

Acabamos de ver que las 2-formas alternadas en \mathbb{R}^3 no son más que combinaciones lineales de determinantes formados a partir de las coordenadas de los vectores sobre los que se aplican. Esto no es un caso aislado. En cierto modo, las k -formas alternadas y diferenciales no son más que determinantes escritos con una notación conveniente.

Veamos primero esa notación y comprobaremos después que los determinantes cubren todas nuestras necesidades.

Definición: Dados $\tilde{\varphi}^1, \tilde{\varphi}^2, \dots, \tilde{\varphi}^k \in V^*$ su *producto exterior*, $\tilde{\varphi}^1 \wedge \tilde{\varphi}^2 \wedge \dots \wedge \tilde{\varphi}^k$, se define como la forma de $\text{Alt}^k(V)$ dada por

$$(\tilde{\varphi}^1 \wedge \tilde{\varphi}^2 \wedge \dots \wedge \tilde{\varphi}^k)(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) = \det(\tilde{\varphi}^i(\vec{v}_j))_{1 \leq i, j \leq k}.$$

La definición se extiende de la manera obvia a uno formas.

Proposición 1.3.1 *Sea $\{\tilde{\varphi}^1, \tilde{\varphi}^2, \dots, \tilde{\varphi}^n\}$ una base de V^* , entonces*

$$\{\tilde{\varphi}^{i_1} \wedge \tilde{\varphi}^{i_2} \wedge \dots \wedge \tilde{\varphi}^{i_k} : 1 \leq i_1 < i_2 < \dots < i_k \leq n\}$$

es una base de $\text{Alt}^k(V)$.

Demostración: Sea $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es una base V cuya base dual es $\{\tilde{\varphi}^1, \tilde{\varphi}^2, \dots, \tilde{\varphi}^n\}$. Consideremos el tensor

$$T = \sum_{i_1 < i_2 < \dots < i_k} a_{i_1 i_2 \dots i_k} \tilde{\varphi}^{i_1} \wedge \tilde{\varphi}^{i_2} \wedge \dots \wedge \tilde{\varphi}^{i_k}.$$

Si T fuera el tensor nulo para ciertos coeficientes, calculando $T(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_n})$ se tendría $a_{i_1 i_2 \dots i_k} = 0$, por tanto los elementos del conjunto son linealmente independientes.

Por otro lado, si $\omega \in \text{Alt}^k(V)$ y sus componentes en la base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ son $\omega_{i_1 i_2 \dots i_k}$ entonces eligiendo $a_{i_1 i_2 \dots i_k} = \omega_{i_1 i_2 \dots i_k}$ se tiene que T y ω tienen las mismas componentes $i_1 i_2 \dots i_k$ siempre que $i_1 < i_2 < \dots < i_k$. En el resto de los casos también deben coincidir por la antisimetría de las formas alternadas al intercambiar dos argumentos. \square

Evidentemente algo similar ocurre en $\Omega^k(M)$.

Corolario 1.3.2 *Cualquier elemento de $\Omega^k(M)$ se puede escribir como*

$$\omega = \sum_{i_1 < i_2 < \dots < i_k} f_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$$

Nota: Como explicamos en la sección anterior, las componentes $f_{i_1 i_2 \dots i_k}$ son en principio funciones de $p \in M$ pero habitualmente esa dependencia se expresa a través de las funciones coordenadas y escribiremos $f_{i_1 i_2 \dots i_k} = f_{i_1 i_2 \dots i_k}(x^1, x^2, \dots, x^n)$.

Es natural definir

$$(\tilde{\varphi}^{i_1} \wedge \tilde{\varphi}^{i_2} \wedge \dots \wedge \tilde{\varphi}^{i_k}) \wedge (\tilde{\varphi}^{j_1} \wedge \tilde{\varphi}^{j_2} \wedge \dots \wedge \tilde{\varphi}^{j_l}) = \tilde{\varphi}^{i_1} \wedge \tilde{\varphi}^{i_2} \wedge \dots \wedge \tilde{\varphi}^{i_k} \wedge \tilde{\varphi}^{j_1} \wedge \tilde{\varphi}^{j_2} \wedge \dots \wedge \tilde{\varphi}^{j_l}.$$

Con ello y la Proposición 1.3.1 o el Corolario 1.3.2 habremos extendido por la distributiva la definición del *producto exterior* a una operación

$$\wedge : \text{Alt}^k(V) \times \text{Alt}^l(V) \longrightarrow \text{Alt}^{k+l}(V) \quad \text{y} \quad \wedge : \Omega^k(M) \times \Omega^l(M) \longrightarrow \Omega^{k+l}(M).$$

Para respetar los convenios se debe interpretar que el producto exterior por números, que son 0-formas, es el producto usual (por ejemplo $2 \wedge \omega = \omega \wedge 2 = 2\omega$).

En los textos se suele dar una definición más invariante del producto exterior [Spi82]. Nuestra definición no está dentro de la estética libre de coordenadas que prefiere la geometría actual pero permite deducir sin dificultad dos propiedades básicas: la asociativa

$$\omega \wedge (\eta \wedge \lambda) = (\omega \wedge \eta) \wedge \lambda$$

y la anticonmutativa (o *superconmutativa*, si uno es físico)

$$(1.3) \quad \omega \wedge \eta = (-1)^{kl} \eta \wedge \omega,$$

donde $\omega \in \text{Alt}^k(V)$, $\eta \in \text{Alt}^l(V)$ o $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$.

Si las formas alternadas y las formas diferenciales no son más que combinaciones de determinantes, ¿por qué no escribimos simplemente esos determinantes y nos olvidamos de estas definiciones tan raras? La respuesta es que los determinantes aparecen en algunos teoremas, por ejemplo en el de Stokes, de una manera complicada y más vale inventar una notación para poder proceder simbólicamente. No hay nada nuevo en esta forma de actuar y los propios determinantes son un buen ejemplo: la relación $\det(A) \cdot \det(B) = \det(AB)$ es a la vez bonita, simple y no trivial (¿alguien recuerda la prueba?) pero si no existiera una notación especial para el determinante y escribiéramos todo el desarrollo, digamos por ejemplo en el caso 3×3 , ¿nos diría algo esa relación? ¿merecería los adjetivos anteriores?⁵

⁵La utilidad o conveniencia de una notación o un modo de cálculo no son en absoluto evidentes a priori ni siquiera para los expertos pues a veces dependen de desarrollos posteriores de las Matemáticas. Por ejemplo, cuando Grassmann creó el álgebra exterior que después fue retomada por Cartan para desarrollar la teoría de formas diferenciales, los matemáticos de su tiempo no le prestaron mucha atención, tanto es así que en los últimos años de su vida prácticamente abandonó las Matemáticas y se dedicó a la Lingüística.

En la siguiente definición no usaremos a propósito el convenio de sumación para mayor claridad.

Definición: Dado un elemento de $\Omega^k(M)$ que en una carta es de la forma

$$\omega = \sum_{i_1, i_2, \dots, i_k} f_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$$

se llama *derivada exterior* de ω al elemento $d\omega \in \Omega^{k+1}(M)$ dado por

$$d\omega = \sum_j \sum_{i_1, i_2, \dots, i_k} \frac{\partial f_{i_1 i_2 \dots i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$$

(en el caso especial $k = 0$, $df = \sum \frac{\partial f}{\partial x^j} dx^j$).

En principio no está claro con esta definición que d se aplique a todo elemento de $\Omega^k(M)$ y, todavía peor, que sea coherente con los cambios de carta. Es posible evitar este último problema comenzando la casa por el tejado con un tratamiento axiomático: se imponen las propiedades de la Proposición 1.3.3 en las que no aparecen cartas y se prueba que sólo hay un operador decente d con esas propiedades. Al calcular su expresión en coordenadas se obtiene la fórmula anterior [BG80]. Se deja al lector interesado que indague esta línea o que trate por sí mismo de probar que la definición de $d\omega$ no depende de la carta empleada. El punto crucial es que las derivadas parciales segundas cruzadas desaparecen por la anticonmutatividad del producto exterior.

Veamos dos propiedades importantes que, como hemos apuntado, determinan la derivada exterior. La primera muestra la relación entre los definiciones de producto exterior y derivada exterior y la segunda es crucial en topología diferencial [Cha08, Cap.2].

Proposición 1.3.3 Sean $\omega \in \Omega^k(M)$ y $\eta \in \Omega^l(M)$, entonces

$$1) \quad d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta, \quad 2) \quad d(d\omega) = 0.$$

Demostración: Por el Corolario 1.3.2, empleando la linealidad de d y la distributiva, podemos limitarnos al caso $\omega = f dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$, $\eta = g dx^{j_1} \wedge dx^{j_2} \wedge \dots \wedge dx^{j_l}$. Se tiene

$$\begin{aligned} d(\omega \wedge \eta) &= \frac{\partial(fg)}{\partial x^m} dx^m \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \dots \wedge dx^{j_l} \\ &= \frac{\partial f}{\partial x^m} g dx^m \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \dots \wedge dx^{j_l} \\ &\quad + (-1)^k \frac{\partial g}{\partial x^m} f dx^{i_1} \wedge \dots \wedge dx^{i_k} \wedge dx^m \wedge dx^{j_1} \wedge \dots \wedge dx^{j_l} \end{aligned}$$

donde se ha usado (1.3). El primer sumando es $d\omega \wedge \eta$ y el segundo $(-1)^k \omega \wedge d\eta$.

Es fácil ver que $d(d\omega) = 0$ se cumple para $k = 0$:

$$d(df) = d\left(\frac{\partial f}{\partial x^i} dx^i\right) = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \wedge dx^j = 0$$

porque $dx^i \wedge dx^j = -dx^j \wedge dx^i$ y las derivadas parciales cruzadas coinciden. Por otro lado, de 1) se deduce con un pequeño cálculo que $d(d(\omega \wedge \eta)) = d(d\omega) \wedge \eta - \omega \wedge d(d\eta)$. La prueba se sigue por inducción ya que el corolario anterior permite escribir toda forma como productos exteriores de funciones y diferenciales de funciones⁶. \square

Para terminar introduciremos, sin profundizar, el concepto de integral en variedades. Incluso si nos restringimos a \mathbb{R}^n , una definición invariante requiere objetos que se multipliquen por el determinante jacobiano al cambiar de coordenadas. Las formas diferenciales satisfacen este requerimiento y su razón de ser original es que constituyen los objetos matemáticos que se pueden integrar. Nótese que si

$$\omega = g dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n \in \Omega(\mathbb{R}^n) \quad \text{con } g = g(x^1, x^2, \dots, x^n)$$

y hacemos el cambio $x^j = f^j(y^1, y^2, \dots, y^n)$, $1 \leq j \leq n$, entonces $dx^j = \frac{\partial f^j}{\partial y^i} dy^i$ y sustituyendo y utilizando las propiedades de los determinantes

$$\omega = \det(Jf) g \circ f dy^1 \wedge dy^2 \wedge \cdots \wedge dy^n.$$

Entonces tiene sentido definir la integral de $\omega \in \Omega(\mathbb{R}^n)$, digamos de soporte compacto⁷ si queremos evitar problemas con la existencia, como

$$(1.4) \quad \int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} g.$$

Realmente hay un pequeño problema con el signo, sobre el que volveremos más adelante, para que esta definición sea invariante del todo por cambios de coordenadas.

Antes de seguir, generalicemos los cambios de variable incluso cuando f no es un difeomorfismo.

⁶Por ejemplo, para $k = 2$ una forma diferencial es suma de cosas del tipo $f dx^{i_1} \wedge dx^{i_2}$ y si suponemos el resultado probado para $k = 0$ y $k = 1$, se tiene $d(d(f dx^{i_1} \wedge dx^{i_2})) = d(d(f dx^{i_1})) \wedge dx^{i_2} - f dx^{i_1} \wedge d(dx^{i_2}) = 0$. Con nuestros convenios el caso especial $k = 1$, el primer paso de la inducción, no es conflictivo porque $f \wedge dx^{i_1} = f dx^{i_1}$.

⁷Cuando hablamos del soporte de $\omega = g dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n$ o de su no anulación en un punto nos referimos a los conceptos análogos para g . Esto es coherente porque los cambios de carta multiplican g por una función no nula.

Definición: Sean M y N variedades y $f : M \rightarrow N$. Se llama *imagen recíproca* (o *pullback*) a la aplicación lineal $f^* : \Omega^k(N) \rightarrow \Omega^k(M)$ dada por

$$f^*\omega(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) = \omega(df(\vec{v}_1), df(\vec{v}_2), \dots, df(\vec{v}_k)) \quad \text{para } \vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in T_p(M).$$

Nota: Esta definición se extiende sin cambios a cualquier tensor de tipo $(0, s)$.

Si $\vec{v}_j \in T_p(M)$ entonces $df(\vec{v}_j) \in T_{f(p)}(N)$, por tanto si ω está soportada en un entorno de un punto, $f^*\omega$ lo está en entornos de sus preimágenes for f . En coordenadas todo está más claro,

$$\omega = g_{i_1 i_2 \dots i_n} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_n} \Rightarrow f^*\omega = g_{i_1 i_2 \dots i_n} \circ f d(x^{i_1} \circ f) \wedge d(x^{i_2} \circ f) \wedge \dots \wedge d(x^{i_n} \circ f).$$

Si f es un difeomorfismo de \mathbb{R}^n con $\det(Jf) > 0$, la fórmula de cambio de variable es

$$(1.5) \quad \int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} f^*\omega.$$

Ahora podemos utilizar las cartas para definir la integral de una forma en un parche de una variedad. Concretamente, si (\mathcal{U}, ϕ) es una carta de una variedad m -dimensional, M , y $\omega \in \Omega^n(M)$, digamos de soporte compacto incluido en \mathcal{U} para que no haya problemas al integrar, entonces se define

$$(1.6) \quad \int_{\mathcal{U}} \omega = \int_{\phi(\mathcal{U})} (\phi^{-1})^*\omega.$$

El segundo miembro, es una integral como la de (1.4). Por (1.5) con $f = \phi \circ \tilde{\phi}^{-1}$, al cambiar la carta (\mathcal{U}, ϕ) por otra $(\mathcal{V}, \tilde{\phi})$ el resultado es el mismo si $\det(Jf) > 0$ (suponiendo que el soporte de ω está en $\mathcal{U} \cap \mathcal{V}$). Excluir la posibilidad $\det(Jf) < 0$, lleva a considerar sólo la “mitad” de las cartas. Una elección de un subconjunto (digamos maximal) de cartas tales que los cambios $\phi \circ \tilde{\phi}^{-1}$ tengan siempre jacobiano positivo, se dice que define una *orientación*. En \mathbb{R}^2 , por ejemplo, las cartas $(\mathbb{R}^2, (x, y))$ y $(\mathbb{R}^2, (y, x))$ no pertenecen a la misma orientación y darían signos distintos al definir la integral. En variedades conexas hay dos posibles orientaciones.

La extensión de la definición de integral de un abierto (1.6) a toda una variedad se lleva a cabo de la manera natural empleando las particiones de la unidad [Spi82].

Un resultado central sobre integración en variedades es el teorema de Stokes que establece una bella dualidad entre la derivada exterior y la frontera topológica. Requiere introducir el concepto de variedad con borde. Por ejemplo, el disco cerrado unidad D en \mathbb{R}^2 no es una variedad porque un punto en la frontera $\partial D = S^1$ no tiene entornos homeomorfos a un abierto de \mathbb{R}^2 , sino al semiespacio $\mathbb{R} \times [0, \infty)$. En general, se dice que M es una *variedad con borde* de dimensión n , si tiene puntos con entornos homeomorfos a un semiespacio $\mathbb{R}^{n-1} \times [0, \infty)$ que conforman una variedad ∂M de dimensión $n - 1$,

mientras que el resto de los puntos, $M - \partial M$, satisfacen la definición habitual de variedad de dimensión n . Geométricamente ∂M es el borde de M porque el semiespacio establece un plano $x^n = 0$ que no se puede rebasar.

Al estar ∂M en el “borde” de $M - \partial M$, una orientación en $M - \partial M$ condiciona una orientación en ∂M llamada *orientación inducida*. Las definiciones rigurosas de estos conceptos se pueden consultar por ejemplo en [Cha08].

Teorema 1.3.4 (Teorema de Stokes) *Sea M una variedad n -dimensional orientable con borde y $\omega \in \Omega^{n-1}(M)$ con soporte compacto, entonces*

$$\int_{\partial M} i^* \omega = \int_M d\omega$$

donde $i : \partial M \rightarrow M$ es la inclusión y ∂M tiene la orientación inducida por la de M .

Con el consiguiente abuso de notación, a veces se escribe $\int_{\partial M} \omega$ en el primer miembro en lugar de $\int_{\partial M} i^* \omega$, porque la única manera de integrar algo n -dimensional en un conjunto $n - 1$ dimensional es usar la relación entre las variables y eso es justamente lo que indica $i^* \omega$.

Los caso $n = 1, 2$ y 3 del teorema de Stokes dan todas las versiones del teorema fundamental del cálculo estudiadas en cursos anteriores. Por ejemplo, si $n = 2$ y $\omega = P dx + Q dy$ se tiene (suprimiendo la inclusión, como en esos cursos)

$$\int_{\partial M} P dx + Q dy = \int_M \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy$$

que es el *teorema de Green*.

Capítulo 2

El teorema de Frobenius

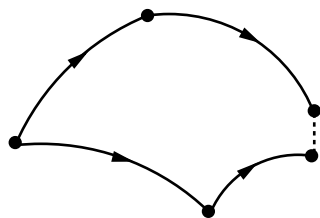
Si en el plano indicamos en cada punto una dirección (o velocidad) la trayectoria de una partícula está totalmente condicionada una vez fijado el punto de partida. En términos analíticos la ecuación diferencial $c'(t) = F(c(t))$ determina la curva solución $c = c(t)$ al especificar $c(0)$.

Un campo de direcciones tangentes da lugar entonces a curvas, sin embargo las cosas se complican en dimensiones superiores. Por ejemplo, si en \mathbb{R}^3 consideramos los campos de vectores $X = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$, $Y = \frac{\partial}{\partial y} + 2x\frac{\partial}{\partial z}$; una superficie que los tenga como vectores tangentes ∂_1, ∂_2 , se puede entender localmente como unas funciones $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, parametrizando la superficie, tales que

$$\begin{pmatrix} \partial x / \partial u \\ \partial y / \partial u \\ \partial z / \partial u \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ y \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \partial x / \partial v \\ \partial y / \partial v \\ \partial z / \partial v \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2x \end{pmatrix}.$$

Las dos primeras coordenadas de estas ecuaciones nos dicen que, salvo una traslación, $x = u$, $y = v$, por tanto podemos considerar que nuestra superficie está parametrizada por $(x, y, z(x, y))$, es decir, que es el grafo de una función. Las ecuaciones de las terceras coordenadas piden

$$\frac{\partial z}{\partial x} = y \quad \text{y} \quad \frac{\partial z}{\partial y} = 2x \quad \text{que contradicen} \quad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}.$$



Desde el punto de vista geométrico siempre podemos ordenar a una persona o a una partícula que en cada paso infinitesimal tome la dirección del vector indicado por un campo pero no podemos hacer lo mismo con superficies indicando dos direcciones (dos campos) porque hay muchas formas de avanzar y no tienen por qué llevarnos coherentemente a los mismos puntos. Si uno recuerda cómo se probaba la igualdad de las parciales cruzadas, que antes se ha violado, se empleaba que para conectar

dos vértices opuestas de un rectángulo da igual ir por los lados que están encima de la diagonal que por los que están por debajo.

La situación no es nueva desde el punto de vista analítico pues ya en cálculo de varias variables se vio que dada $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ no siempre existe una función f tal que $\nabla f = F$. Cuando tal problema tiene solución se dice que F es un *campo conservativo* y ello requiere $\text{rot } F = \vec{0}$.

El teorema de Frobenius da las condiciones necesarias y suficientes para que al especificar unas direcciones, exista localmente una subvariedad cuyo espacio tangente esté generado por ellas.

2.1. Flujos de campos de vectores

Comenzamos dando un nombre, posiblemente ya conocido, a las trayectorias obtenidas a partir de un campo de vectores.

Definición: Sea $c : I \rightarrow M$ una curva parametrizada con I un intervalo abierto y X un campo de vectores en M que se expresa como $X = X^i \frac{\partial}{\partial x^i}$ en cierta carta. Se dice que c es una *curva integral* de X si $(x^i \circ c)'(t) = X^i(c(t))$.

Muchas veces, con cierto abuso de notación, se define $c'(t)$ como un vector tangente (presuponiendo la composición con las funciones coordenadas) para así librarse de los sistemas de coordenadas locales [O'N83, Ch.1] y poder escribir $c'(t) = X|_{c(t)}$.

El teorema de existencia y unicidad de soluciones de ecuaciones diferenciales ordinarias asegura, como ya hemos notado, que fijado un punto hay (localmente) una única curva integral con $c(0) = p$ y que si dos curvas integrales $c_1 : I_1 \rightarrow M$, $c_2 : I_2 \rightarrow M$ cumplen $c_1(t_0) = c_2(t_0)$ para algún $t_0 \in I_1 \cap I_2$ entonces coinciden en todo $I_1 \cap I_2$.

Recuérdese que ecuaciones diferenciales en \mathbb{R} tan sencillas como $x' = x^2$, $x(0) = 1$ no tienen solución regular en todo \mathbb{R} , pues $x(t) = (1 - t)^{-1}$ “explota” en $t = 1$. Entonces a veces I no se puede sustituir por \mathbb{R} en la definición anterior. Dado $p \in M$, se dice que c es una *curva integral maximal* con $c(0) = p$ de un campo X si su dominio de definición I es el mayor posible en el sentido de la inclusión. La teoría dice que en una ecuación diferencial ordinaria con coeficientes regulares lo único que puede ir mal es que la solución “explote hacia el infinito”. Más concretamente en cuanto se tiene una subsucesión convergente $t_n \rightarrow l$ de valores del parámetro t que se aplica por c en una sucesión convergente, entonces c está definida en un entorno de l . En una variedad compacta no hay infinitos hacia los que escapar, ya que toda sucesión tiene una subsucesión convergente. Por tanto en variedades compactas las curvas integrales maximales tienen $I = \mathbb{R}$.

Definición: Se dice que un campo de vectores en una variedad es *completo* si las curvas integrales maximales $c = c(t)$ están definidas para todo $t \in \mathbb{R}$.

El siguiente concepto a introducir es la aplicación correspondiente a caminar durante un tiempo t prefijado por cada curva integral. La posibilidad de que $I \neq \mathbb{R}$ impide hacer una definición global.

Definición: Sea \mathcal{U} un abierto de una variedad M y sea $c_p(t)$ la curva integral maximal de X con $c_p(0) = p$. Si existe un intervalo abierto $I \ni 0$ contenido en el dominio de todas las c_p con $p \in \mathcal{U}$, entonces para cada $t \in I$ se define el *flujo local* $\Phi_t : \mathcal{U} \rightarrow M$ como $\Phi_t(p) = c_p(t)$.

Pedimos la existencia de un I que contenga algo alrededor de cero porque deseamos hacer variar t . El teorema de dependencia continua de los parámetros para ecuaciones diferenciales ordinarias asegura que tomando \mathcal{U} suficientemente pequeño (alrededor de un punto dado) siempre existe tal I y por tanto un flujo local (véanse Th.7.3 y Th.7.5 en [Wal04, §1.7]). Además $\Phi_t(p)$ es C^∞ como función de p y t .

Lema 2.1.1 *Sea I como en la definición de flujo local. Si $t_1, t_2, t_1 + t_2 \in I$ entonces*

$$\Phi_{t_1} \circ \Phi_{t_2} = \Phi_{t_1+t_2}.$$

Demostración: Sea t entre 0 y t_1 , entonces también $t, t + t_2 \in I$. Por definición

$$\Phi_t \circ \Phi_{t_2}(p) = c_{\Phi_{t_2}(p)}(t) = c_{c_p(t_2)}(t) \quad \text{y} \quad \Phi_{t+t_2}(p) = c_p(t + t_2).$$

Ambos resultados son curvas integrales que parten (para $t = 0$) del punto $c_p(t_2)$, por tanto deben coincidir y tomando $t = t_1$ se deduce el resultado. \square

Proposición 2.1.2 *Si X es un campo de vectores completo en M entonces para cada $t \in \mathbb{R}$ su flujo $\Phi_t : M \rightarrow M$ es un difeomorfismo y $\{\Phi_t\}_{t \in \mathbb{R}}$ es un grupo abeliano con la composición (se dice que es un grupo uniparamétrico de difeomorfismos).*

Demostración: Por el Lema 2.1.1, $\Phi_{-t} \circ \Phi_t = \Phi_t \circ \Phi_{-t} = \Phi_0 = \text{Id}$ y por tanto Φ_t es un difeomorfismo con función inversa Φ_{-t} . Por otro lado, la relación $\Phi_t \circ \Phi_s = \Phi_{t+s}$ permite deducir las propiedades de grupo abeliano a partir de las de $(\mathbb{R}, +)$. \square

Un grupo uniparamétrico de difeomorfismos hereda de \mathbb{R} una estructura de variedad y así es un tipo especial de *grupo de Lie*, un grupo que tiene estructura de variedad de tal modo que la operación de grupo es compatible con la estructura diferenciable. Este concepto fue introducido por S. Lie y su idea clave fue estudiar estos “grupos continuos”, que en principio pueden ser muy complicados, a través del espacio tangente en el elemento unidad dotado con una operación producto de vectores con propiedades especiales, lo que se llama un *álgebra de Lie*.

Según hemos visto, en las variedades compactas todos los campos de vectores son completos, sin embargo hay muchos ejemplos no triviales de campos completos en variedades no compactas (por ejemplo, en cualquier grupo de Lie partiendo de un vector

en un punto y aplicando todas las transformaciones del grupo se consigue siempre un campo de vectores completo). Veamos uno muy sencillo en que el grupo uniparamétrico de difeomorfismos está relacionado con la adición de velocidades en relatividad especial.

Ejemplo: Consideremos $M = (-1, 1)$ con la carta identidad y el campo de vectores $X = (1 - x^2) \frac{\partial}{\partial x}$.

Las curvas integrales $x = x(t)$ vienen determinadas por $x' = 1 - x^2$. Integrando $x'/(1 - x^2) = 1$ bajo la condición $x(0) = x_0$ se obtiene tras algunos cálculos

$$x(t) = \frac{x_0(1 + e^{-2t}) + 1 - e^{-2t}}{x_0(1 - e^{-2t}) + 1 + e^{-2t}}.$$

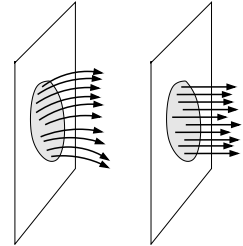
Si $x_0 = 0$ se tiene $x(t) \in M$ para todo $t \in \mathbb{R}$, de hecho $\text{Im } x = M$. Entonces hay una sola curva integral salvo traslaciones del parámetro y como está definida en \mathbb{R} , el campo de vectores es completo. Se sigue

$$\Phi_t(x) = \frac{x(1 + e^{-2t}) + 1 - e^{-2t}}{x(1 - e^{-2t}) + 1 + e^{-2t}} = \frac{x + \tanh t}{1 + x \tanh t} \quad \text{donde} \quad \tanh t = \frac{e^t - e^{-t}}{e^t + e^{-t}}.$$

La relación $\Phi_t \circ \Phi_s = \Phi_{t+s}$ en la que descansa la ley de grupo depende de la ley de adición $\tanh(t + s) = (\tanh t + \tanh s)/(1 + \tanh t \tanh s)$.

Intuitivamente parece claro que siempre con un cambio de coordenadas se pueden deformar las curvas integrales en el entorno de un punto para que apunten en una dirección prefijada. Visto de otro modo, siempre cambiando las coordenadas podemos conseguir que localmente un campo de vectores sea constante.

La demostración consiste geoméricamente en cortar un haz de curvas integrales con el “plano” (la hipersuperficie) que resulta al congelar todas las variables menos una y añadir como primera variable el propio parámetro de las curvas integrales.



Lema 2.1.3 Dado un punto p y un campo de vectores X con $X(p) \neq 0$, existe una carta $(\mathcal{U}(p), \phi)$ con la cual $X = \partial_1$.

Demostración: Haremos la prueba en \mathbb{R}^n . El caso de variedades generales es similar componiendo con las funciones coordenadas. Sea $X = X^i \frac{\partial}{\partial x^i}$. Con una traslación y una aplicación lineal se puede suponer que p es el origen y que $X(p) = \frac{\partial}{\partial x^1}$.

Consideremos la función $F(\vec{y}) = (f^1(\vec{y}), f^2(\vec{y}), \dots, f^n(\vec{y}))$ con $\vec{y} = (y^1, \dots, y^n)$ definida en un entorno de p como la solución de la ecuación diferencial

$$(2.1) \quad \begin{cases} \frac{\partial f^i}{\partial y^1}(\vec{y}) = X^i(F(\vec{y})) \\ F(0, y^2, y^3, \dots, y^n) = (0, y^2, y^3, \dots, y^n). \end{cases}$$

A pesar del uso de derivadas parciales, es una ecuación diferencial ordinaria pues sólo se deriva con respecto a y^1 . Por la teoría general, $F \in C^\infty$ en un entorno del origen. Nótese que (2.1) no es más que la ecuación de las curvas integrales que parten del punto $(0, y^2, y^3, \dots, y^n)$. En términos del flujo, $F(\vec{y}) = \Phi_{y^1}(0, y^2, y^3, \dots, y^n)$.

Por la primera ecuación de (2.1) y nuestra hipótesis, $\frac{\partial f^i}{\partial y^1}(p) = \delta_1^i$, mientras que la segunda ecuación asegura $\frac{\partial f^i}{\partial y^j} = \delta_j^i$ para $j \geq 2$. Entonces $DF(p)$ es la matriz identidad y el teorema de la función inversa asegura que F define un cambio de coordenadas. Utilicemos el sistema de coordenadas (y^1, \dots, y^n) con $x^i = f^i(\vec{y})$, es decir, $\vec{y} = F^{-1}(\vec{x})$. Se tiene

$$\frac{\partial}{\partial y^1} = \frac{\partial x^i}{\partial y^1} \frac{\partial}{\partial x^i} = X^i(F(\vec{y})) \frac{\partial}{\partial x^i} = X$$

y por tanto en este nuevo sistema de coordenadas $X = \partial_1$. \square

2.2. El corchete de Lie y la derivada de Lie

Recordemos que según la definición abstracta moderna, los vectores en variedades son operadores que actúan sobre funciones como una suerte de derivadas. Entonces tiene sentido considerar la composición XY de dos campos de vectores X e Y , aunque el resultado no sea un campo de vectores. El análogo de la diferencia entre las derivadas parciales cruzadas es el *conmutador* de los operadores y éste resulta ser un objeto importante, que en nuestro contexto (y en el de las álgebras de Lie) recibe un nombre especial.

Definición: Dado un par de campos de vectores X e Y se define su *corchete de Lie* $[X, Y]$ como el operador que actúa sobre $f \in C^\infty$ como

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

Ejemplo: Para los campos $X = \frac{\partial}{\partial x} + y \frac{\partial}{\partial z}$, $Y = \frac{\partial}{\partial y} + 2x \frac{\partial}{\partial z}$ en \mathbb{R}^3 se tiene

$$\begin{aligned} [X, Y](f) &= \left(\frac{\partial}{\partial x} + y \frac{\partial}{\partial z} \right) \left(\frac{\partial f}{\partial y} + 2x \frac{\partial f}{\partial z} \right) - \left(\frac{\partial}{\partial y} + 2x \frac{\partial}{\partial z} \right) \left(\frac{\partial f}{\partial x} + y \frac{\partial f}{\partial z} \right) \\ &= \frac{\partial^2 f}{\partial x \partial y} + y \frac{\partial^2 f}{\partial z \partial y} + 2 \frac{\partial f}{\partial z} + 2x \frac{\partial^2 f}{\partial x \partial z} + 2xy \frac{\partial^2 f}{\partial z^2} \\ &\quad - \left(\frac{\partial^2 f}{\partial y \partial x} + \frac{\partial f}{\partial z} + y \frac{\partial^2 f}{\partial y \partial z} + 2x \frac{\partial^2 f}{\partial z \partial z} + 2xy \frac{\partial^2 f}{\partial z^2} \right) = \frac{\partial f}{\partial z}. \end{aligned}$$

Entonces $[X, Y] = \frac{\partial}{\partial z}$.

No es difícil ver que la cancelación de las derivadas segundas no es casual. Ocurre en todos los ejemplos y entonces el resultado es de nuevo un campo de vectores. Aunque esto

no sea muy profundo, lo enunciamos para resaltar su importancia. Además la expresión obtenida es más conveniente para cálculos explícitos que la definición directa.

Proposición 2.2.1 *Dados dos campos de vectores $X = X^i \partial_i$ e $Y = Y^i \partial_i$, su corchete de Lie es también un campo de vectores que actúa como*

$$[X, Y](f) = (X(Y^j) \partial_j - Y(X^j) \partial_j)(f)$$

Demostración: Lo más rápido es escribir el ejemplo anterior en general. Llamando x^1, \dots, x^n a las funciones coordenadas:

$$[X, Y](f) = X^i \frac{\partial}{\partial x^i} \left(Y^j \frac{\partial f}{\partial x^j} \right) - Y^i \frac{\partial}{\partial x^i} \left(X^j \frac{\partial f}{\partial x^j} \right),$$

y derivando los productos

$$[X, Y](f) = X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j} - Y^i \frac{\partial X^j}{\partial x^i} \frac{\partial f}{\partial x^j} - Y^i X^j \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

Por la igualdad de las derivadas parciales cruzadas, se obtiene la fórmula del enunciado. Esto prueba que $[X, Y]$ es un campo de vectores, pues es combinación lineal de ∂_i . \square

El corchete de Lie se comporta bien con respecto a las aplicaciones lineales inducidas en el espacio tangente por difeomorfismos entre variedades.

Proposición 2.2.2 *Sean M y N variedades y $f : M \rightarrow N$ un difeomorfismo¹. Para cada par de campos de vectores X e Y en M , se tiene*

$$[dfX, dfY]|_{f(p)} = df([X, Y]|_p) \quad \text{para todo } p \in M.$$

Demostración: Por la definición de la aplicación tangente, si $g : N \rightarrow \mathbb{R}$

$$(dfX)(g) \circ f(p) = (dfX)|_{f(p)}(g) = df(X|_p)(g) = X|_p(g \circ f).$$

Aplicando esta relación repetidas veces

$$\begin{aligned} [dfX, dfY]|_{f(p)}(g) &= (dfX)|_{f(p)}((dfY)(g)) - (dfY)|_{f(p)}((dfX)(g)) \\ &= X|_p((dfY)(g) \circ f) - Y|_p((dfX)(g) \circ f) \\ &= X|_p(Y(g \circ f)) - Y|_p(X(g \circ f)) \end{aligned}$$

y esto es $[X, Y]|_p(g \circ f)$ o equivalentemente $df([X, Y]|_p)$ aplicado a g . \square

¹La condición de difeomorfismo se puede rebajar un poco [Aub01]. Esencialmente lo que se necesita es que dfX y dfY se puedan extender a campos vectoriales si no lo fueran. Agradezco a A. Martín Zamora la corrección de un enunciado anterior erróneo.

Definición: Dado un par de campos de vectores X e Y se define la *derivada de Lie* de Y a lo largo de X en un punto p como $F'(0)$ donde $F(t) = d\Phi_{-t}(Y|_{\Phi_t(p)})$ y Φ_t es el flujo local de X . Normalmente se suele denotar mediante $L_X Y$.

Aquí $F'(0)$ se entiende naturalmente como el límite del cociente incremental.

Lo que mide la derivada de Lie es la variación de Y al moverse por una curva integral de X porque $F(0) = Y|_p$ y $F(\epsilon)$ es el resultado de “trasladar” el vector $Y|_{\Phi_\epsilon(p)}$. Según V.I. Arnold la derivada de Lie es la “derivada del pescador” porque es la derivada (velocidad) que mediría un pescador en una barca siguiendo el flujo del río.

Proposición 2.2.3 Sean X e Y campos de vectores en una variedad y sea Φ_t el flujo local de X en un entorno de un punto p , entonces $L_X Y = [X, Y]$.

Demostración: Supongamos $X(p) \neq 0$. Según el Lema 2.1.3 podemos suponer que trabajamos en una carta $(\mathcal{U}, \phi = (x^1, \dots, x^n))$ tal que $X = \partial/\partial x_1$, por tanto el flujo, en coordenadas locales, aplica (x^1, x^2, \dots, x^n) en $(x^1 + t, x^2, \dots, x^n)$ y su matriz jacobiana es la identidad. Así pues, $Y = Y^i \partial_i$ implica, con la notación de la definición anterior, $F(t) = Y^i(x^1 + t, x^2, \dots, x^n) \partial_i$. Por consiguiente $F'(0) = \partial_1 Y^i \partial_i$ que coincide con $[X, Y]$.

Si $X(p) = 0$ pero X no es idénticamente nulo en un entorno de p , entonces un argumento de continuidad sobre $[X, Y]|_p$ y $F'(0)$ prueba el resultado. Por otra parte, si X es nulo en todo un entorno de p , el resultado es trivial (las curvas integrales son constantes). \square

El corchete de Lie cuantifica de alguna forma la conmutatividad de los flujos locales. A este respecto, recuérdese la relación mencionada al comienzo del capítulo, entre la coincidencia de las parciales cruzadas y la independencia del camino.

Proposición 2.2.4 Sean Φ_t y Ψ_s flujos locales de los campos vectoriales X e Y , respectivamente, en el entorno de un punto. Se cumple $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$ si y sólo si $[X, Y] = 0$.

Demostración: Sea F como en la última definición, por el Lema 2.1.1 y la regla de la cadena, para $q = \Phi_s(p)$ se sigue $F(t + s) = d\Phi_{-s} d\Phi_{-t}(Y|_{\Phi_t(q)})$. Por la proposición anterior, $F'(0 + s) = 0$ y se tiene que F es constante. Recíprocamente, si F es constante, entonces $L_X Y = [X, Y]$ se anula. Notando que Φ_0 es la identidad, $[X, Y] = 0$ equivale a $Y|_p = d\Phi_{-t}(Y|_{\Phi_t(p)})$ (con p variable). El flujo de Y es Ψ_s , por tanto basta comprobar que $\Phi_{-t} \circ \Psi_s \circ \Phi_t$ es el flujo de $d\Phi_{-t}(Y|_{\Phi_t(\cdot)})$ para terminar la prueba, ya que el Lema 2.1.1 asegura que Φ_{-t} es la inversa de Φ_t .

Consideremos una carta $(\mathcal{U}, \phi = (x^1, x^2, \dots, x^n))$ y digamos que en ella la matriz de $d\Phi_{-t}$ (que es la jacobiana de $\phi \circ \Phi_{-t} \circ \phi^{-1}$) tiene elementos a_j^i y que las componentes de

Y son Y_i . Por la definición de flujo, $\frac{d}{ds}(x^i \circ \Psi_s) = Y^i \circ \Psi_s$. Sea $c_p(s) = \Phi_{-t} \circ \Psi_s \circ \Phi_t(p)$, entonces por la regla de la cadena

$$\frac{d}{ds}(x^i \circ c_p(s)) = \frac{d}{ds}((x^i \circ \Phi_{-t} \circ \phi^{-1}) \circ \phi \circ \Psi_s \circ \Phi_t(p)) = a_j^i Y^j \circ \Psi_s \circ \Phi_t(p) = a_j^i Y^j \circ \Phi_t(c_p(s)).$$

En definitiva, hemos probado que $(x^i \circ c_p)'(s)$ coincide con la i -ésima componente de $d\Phi_{-t}(Y|_{\Phi_t(c_p(s))})$ o equivalentemente que $\Phi_{-t} \circ \Psi_s \circ \Phi_t$ es el flujo de $d\Phi_{-t}(Y|_{\Phi_t(\cdot)})$. \square

Ejemplo: Ya sabíamos que los campos $X = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$, $Y = \frac{\partial}{\partial y} + 2x\frac{\partial}{\partial z}$ en \mathbb{R}^3 tenían corchete de Lie no nulo. Comprobemos que sus flujos no conmutan.

Las curvas integrales de X responden a las ecuaciones diferenciales $x'(t) = 1$, $y'(t) = 0$, $z'(t) = y$ cuya solución es $x(t) = t + x_0$, $y(t) = y_0$, $z(t) = y_0t + z_0$. Así pues el flujo local es $\Phi_t(x, y, z) = (t + x, y, yt + z)$. De la misma forma las curvas integrales de Y vienen dadas por $x(s) = x_0$, $y(s) = s + y_0$, $z(s) = 2x_0s + z_0$, dando lugar al flujo $\Psi_s(x, y, z) = (x, s + y, 2xs + z)$. De aquí

$$\begin{cases} \Phi_t \circ \Psi_s(x, y, z) = (t + x, s + y, st + yt + 2xs + z) \\ \Psi_s \circ \Phi_t(x, y, z) = (t + x, s + y, 2st + yt + 2xs + z) \end{cases}$$

que tienen terceras coordenadas distintas. Geométricamente, para $t_0 \neq 0$, $\Phi_{t_0}(\vec{0})$ y $\Psi_{t_0}(\vec{0})$ recorren t_0 unidades en los ejes x e y , y $\Psi_{t_0}(\Phi_{t_0}(\vec{0}))$ y $\Phi_{t_0}(\Psi_{t_0}(\vec{0}))$ son dos rectas que se cruzan cuando t varía. Esta imposibilidad de cerrar el cuadrilátero implica la no conmutatividad.

2.3. Condiciones de integrabilidad

Retomando los comentarios al comienzo del capítulo, si en cada punto nos dicen en qué dirección dirigirnos, el camino quedará determinado. Asignar un módulo a esa dirección sólo cambiará la velocidad a la que recorreremos dicho camino. En términos matemáticos, mutliplicar por una constante no nula en cada punto (por una función) los vectores de un campo tiene el efecto de reparametrizar las curvas integrales.

Siguiendo la filosofía de considerar objetos geométricos sin referencia a coordenadas o parametrizaciones particulares, para definir subvariedades de una variedad M que sean tangentes a campos de vectores, debemos fijarnos en los subespacios vectoriales de $T_p(M)$ que generan, más que en los propios campos en sí. Esto lleva al concepto de *distribución*, un nombre, por cierto, no muy afortunado.

Definición: Sea M una variedad n -dimensional. Una *distribución* Δ de dimensión k es una forma de asignar a cada $p \in M$ un subespacio k dimensional $\Delta_p \subset T_p(M)$, de forma que en un entorno $\mathcal{U}(p)$ venga generado por campos de vectores $\{X_1, X_2, \dots, X_k\}$.

La definición parece repetitiva. Insistir sobre los campos de vectores (por supuesto C^∞) es sólo una forma de impedir que la asignación $p \mapsto \Delta_p$ sea en cierto modo discontinua o poco regular. Se dice que los campos X_1, X_2, \dots, X_k , como en la definición anterior, conforman una *base local* de Δ .

No es posible en general tomar una misma base local válida para todo $p \in M$. Por ejemplo, el *teorema de la bola de pelo* [Hir76] [Mun75] dice, con este lenguaje, que no se puede elegir $\mathcal{U}(p) = S^2$ para ninguna distribución de dimensión 1 en S^2 .

Dada una distribución Δ el análogo de las curvas integrales será ahora una subvariedad cuyo espacio tangente en cada punto p coincida con Δ_p . Éste es un buen lugar para recordar qué es una *subvariedad*. Esencialmente es una variedad N dentro de otra M . En términos matemáticos, pedimos que la inclusión $i : N \rightarrow M$ sea una inmersión, es decir, que las estructuras diferenciales sean coherentes. Se podría generalizar el concepto cambiando i por cualquier inmersión inyectiva.

Definición: Sea N una subvariedad de M con $i : N \rightarrow M$ la inclusión. Se dice que N es una *subvariedad integral* de una distribución Δ en M si $di(T_p(N)) = \Delta_p$ para todo $p \in N$.

Nota: Algunos autores [Boo75] sólo piden $di(T_p(N)) \subset \Delta_p$ pero, al menos en nuestro contexto, esta generalidad no tiene recompensa (véase [War83, §1.58]).

El ejemplo del principio del capítulo y las explicaciones que le siguen sugieren que hay distribuciones que tiene subvariedades integrales y otras que no. Las del primer tipo se recogen en la siguiente definición de manera indirecta.

Definición: Se dice que Δ una distribución de dimensión k es *completamente integrable* si para cada punto p existe una carta $(\mathcal{U}(p), \phi)$ con la cual $\{\partial_1, \partial_2, \dots, \partial_k\}$ es una base local de Δ .

Lema 2.3.1 *Una distribución Δ en M es completamente integrable si y sólo si para cada $p \in M$ hay alguna subvariedad integral de Δ con $p \in N$.*

Demostración: Con la notación de la definición anterior, si Δ es completamente integrable y $\phi = (x^1, x^2, \dots, x^n)$, consideremos

$$(2.2) \quad N = \{q \in M : x^{k+1}(q) = x^{k+1}(p), \dots, x^n(q) = x^n(p)\}.$$

Entonces N es una subvariedad integral de Δ , pues $di(\partial_i) = \partial_i$.

Por otro lado, si N es una variedad integral de Δ entonces siempre existe un cambio de coordenadas que permite escribirla en la forma (2.2) [O'N83, Ch.1.28] (esto es el teorema de la inmersión) y por tanto para todo $p \in N$ se tiene que $\{\partial_1, \partial_2, \dots, \partial_k\}$ es una base local de Δ . \square

El teorema de Frobenius establece bajo qué condiciones una distribución es completamente integrable. En la demostración se atisbará que tras un cambio de coordenadas son similares a la exigida en la Proposición 2.2.4. Tras la relación entre flujos y corchetes, la siguiente definición debería traer a la memoria las parciales cruzadas que aparecieron en la discusión inicial.

Definición: Se dice que Δ una distribución de dimensión k es *involutiva* si para cada base local $\{X_1, X_2, \dots, X_k\}$ existen funciones c_{ij}^k tales que

$$[X_i, X_j] = c_{ij}^r X_r \quad \text{para } 1 \leq i, j \leq k.$$

Nótese que se ha empleado el convenio de sumación y se debe sobreentender un sumatorio en $1 \leq r \leq k$. En cada entorno, basta comprobar la condición anterior para una base local, pues gracias a la Proposición 2.2.2, por cambios de coordenadas los vectores y los corchetes de Lie se transforman linealmente de la misma forma y por tanto la relación lineal de la definición anterior se transforma en otra similar.

Teorema 2.3.2 (teorema de Frobenius) *Una distribución es involutiva si y sólo si es completamente integrable.*

Para probar este resultado seguimos aquí esencialmente los pasos de [Lun92]. Separamos primero un caso fácil.

Proposición 2.3.3 *Sea Δ una distribución que en cierta carta admite una base del tipo $\{\partial_1, \partial_2, \dots, \partial_k, X\}$. Si Δ es involutiva entonces es completamente integrable*

Demostración: Escribamos $X = a^1 \partial_1 + a^2 \partial_2 + \dots + a^n \partial_n$. Se puede suponer $a^1 = a^2 = \dots = a^k = 0$ porque en otro caso basta restar a X el vector $a^1 \partial_1 + a^2 \partial_2 + \dots + a^k \partial_k$ que es combinación lineal de los primeros elementos de la base. En un entorno de un punto alguna de las otras coordenadas de X es no nula, quizá intercambiando las variables, digamos que es a^{k+1} . Dividiendo entre esta coordenada se puede suponer $X = \partial_{k+1} + a^{k+2} \partial_{k+2} + \dots + a^n \partial_n$.

Por hipótesis, $[\partial_i, X] \in \Delta$, $1 \leq i \leq k$, y la particular forma de X implica que $[\partial_i, X] = 0$ y que las funciones a^{k+2}, \dots, a^n no dependen de las k primeras funciones coordenadas.

Por el Lema 2.1.3, existe un cambio de coordenadas tal que $X = \partial_{k+1}$, de hecho este cambio se puede elegir dejando las primeras k variables invariantes (pues no participan en X , nótese que en (2.1) si $X^i = 0$ y el resto de los X^j no dependen de i -ésima coordenada, entonces $f^i(\vec{y}) = y^i$). Con ello hemos conseguido la base $\{\partial_1, \partial_2, \dots, \partial_k, \partial_{k+1}\}$ buscada. \square

El plan de la demostración del teorema de Frobenius es emplear inducción y el paso de k a $k + 1$ sólo requerirá utilizar la versión fácil dada por la Proposición 2.3.3.

Demostración del teorema de Frobenius: Claramente, si una distribución es completamente integrable entonces es involutiva, porque $[\partial_i, \partial_j] = 0$. Nos concentramos entonces en la implicación directa.

Como hemos sugerido, procedemos por inducción en la dimensión k de la distribución Δ . Para $k = 1$ basta aplicar el Lema 2.1.3.

Sea Δ de dimensión $k + 1$, con una base dada por los vectores $X_j = a_i^j \partial_j$ con $1 \leq i \leq k + 1$. Al ser linealmente independientes siempre podemos suponer, quizá intercambiando los nombres de las variables, que la matriz $A = (a_i^j)_{1 \leq i, j \leq k+1}$ es no singular. Sean b_i^j los elementos de A^{-1} , entonces $Y_i = b_i^j X_j$, $1 \leq i \leq k + 1$, constituyen una base de Δ . Además $Y_i = \partial_i + \sum_{j=k+2}^n c_i^j \partial_j$ para ciertos c_i^j , donde n es la dimensión de la variedad. La condición $[Y_i, Y_j] \in \Delta$ implica $[Y_i, Y_j] = 0$, porque en $[Y_i, Y_j]$ no aparece Δ_i con $1 \leq i \leq k + 1$.

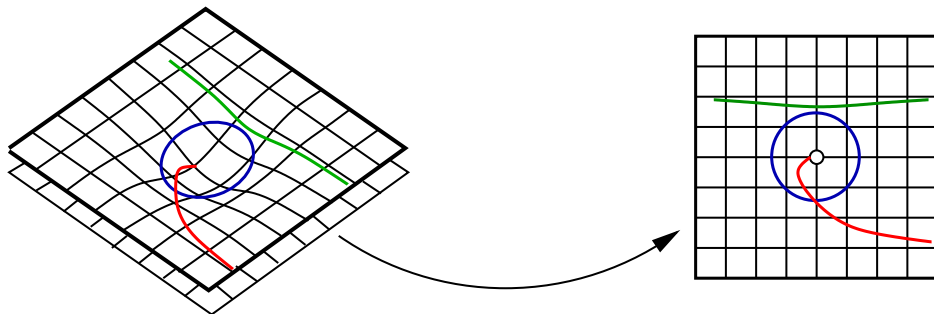
La distribución generada por $\{Y_1, Y_2, \dots, Y_k\}$ es por tanto involutiva y la hipótesis de inducción permite encontrar un cambio de coordenadas de forma que esté generada por $\{\partial_1, \partial_2, \dots, \partial_k\}$. La prueba se termina aplicando la Proposition 2.3.3 a la base de Δ dada por $\{\partial_1, \partial_2, \dots, \partial_k, Y_{k+1}\}$. \square

Ejemplo: Volviendo una vez más al ejemplo inicial, sabemos que para $X = \frac{\partial}{\partial x} + y \frac{\partial}{\partial z}$, $Y = \frac{\partial}{\partial y} + 2x \frac{\partial}{\partial z}$ se cumple $[X, Y] = \frac{\partial}{\partial z}$. Claramente $[X, Y]$ no está en el subespacio generado por X e Y , por consiguiente el teorema de Frobenius asegura que no existe ninguna subvariedad de \mathbb{R}^3 que tenga a X e Y como vectores tangentes en cada punto.

Capítulo 3

Geometría Riemanniana

Imaginemos una hondonada suave en el terreno. Los objetos más lentos que pasasen cerca quedarían atrapados, mientras que los más rápidos se desviarían menos de su trayectoria natural rectilínea. El problema mecánico es intrínsecamente bidimensional, pues el movimiento de las partículas tiene lugar en una superficie. Sin embargo la mecánica de Newton más elemental (la de $F = ma$), lleva a considerar las tres coordenadas x, y, z de nuestro \mathbb{R}^3 ambiente y a introducir las fuerzas de reacción del suelo en la dirección de la normal, que al igual que las tensiones de cuerdas y otras fuerzas poco intuitivas¹ de los cursos básicos de mecánica, confunden a los principiantes.



Supongamos unos seres que vivieran en un plano subterráneo paralelo al de la hondonada y que pudieran percibir las proyecciones (sombras) de las trayectorias. Posiblemente concluirían que existe un “sol” en el centro de la hondonada que ejerce una fuerza gravitatoria atractora, pues todo lo que pasa cerca del presunto sol desvía su trayectoria hacia él. Por supuesto, esta fuerza gravitatoria es ficticia y se explica a partir de una deformación geométrica que es susceptible de ser medida intrínsecamente. Éste es el punto de vista de la relatividad general, que se deshace de la gravedad como una fuerza

¹El principio de acción y reacción dice por ejemplo que si empujamos una pared, la pared nos responde empujándonos a nosotros. Por más que los profesores nos dieran la razón convincente de que si no fuera así, la pared se tendría que mover, lo cierto es que resulta extraño pensar que a la pared le entren ganas de empujarnos.

real y explica sus efectos afirmando que el espacio-tiempo se curva en las cercanías de una masa.

La búsqueda de elegir libremente las coordenadas y la naturalidad de un tratamiento intrínseco, acerca la mecánica a la geometría diferencial. En esta dirección, hay una solución elegante y físicamente muy interesante del problema mecánico que involucra longitudes de vectores tangentes. Matemáticamente, una vez fijada una manera de medir vectores en el espacio tangente de una variedad, lo que se llama una *métrica*, tendremos trayectorias naturales de las partículas libres, que llamaremos *geodésicas*. Inventando una derivada especial adaptada a la métrica, la *derivada covariante*, se recupera el *principio de inercia* que dice que en ausencia de fuerzas la derivada de la velocidad (la aceleración) es cero, de esta forma las geodésicas son el análogo de las rectas de la geometría euclídea.

Una métrica expresa cierta deformación con respecto a la geometría euclídea a la que más adelante daremos carácter tensorial. Las bases de todo ello, y en realidad de la propia geometría diferencial, las asentó B. Riemann en una famosa lección inaugural sin apenas fórmulas (véase [Spi79, Ch.4]). En su honor se llama *geometría riemanniana* a la geometría de las variedades dotadas de una métrica. Más de sesenta años después, A. Einstein formuló la relatividad general; una teoría geométrica de la gravitación que postula que la masa y la energía deforman el espacio-tiempo dotándole de una estructura de variedad cuatridimensional con una métrica bien distinta de la euclídea.

3.1. Mecánica y métricas

Supongamos una partícula de masa m en un campo con energía potencial $U(x, y, z)$. Según los cursos básicos de Física, su ecuación de movimiento viene dada por la solución de

$$(3.1) \quad m\ddot{x} + \frac{\partial U}{\partial x} = 0, \quad m\ddot{y} + \frac{\partial U}{\partial y} = 0, \quad m\ddot{z} + \frac{\partial U}{\partial z} = 0$$

donde emplearemos la notación de los físicos, heredada de los trabajos de Newton, consistente en señalar una o dos derivadas con respecto al tiempo con uno o dos puntos sobre la función dada. Consideremos ahora el *lagrangiano*, la diferencia entre la energía cinética y la potencial:

$$L = L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = T - U = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - U(x, y, z).$$

Entonces las ecuaciones de Newton (3.1) son equivalentes a

$$(3.2) \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) = \frac{\partial L}{\partial q^i} \quad \text{para } i = 1, 2, 3 \text{ con } q^1 = x, q^2 = y, q^3 = z.$$

Los matemáticos formalistas se mostrarán reacios a escribir una derivada con respecto a la derivada de una función pero el sentido está claro, simplemente por ejemplo $\partial L/\partial \dot{x}$ significa la parcial de L con respecto a la cuarta variable y sustituir en su lugar \dot{x} .

En principio parece que no hay ninguna ventaja en escribir (3.2) en lugar de (3.1), sin embargo el siguiente resultado básico de cálculo de variaciones nos dará una nueva visión.

Proposición 3.1.1 *Dados $a, b \in \mathbb{R}$ y $\vec{c}, \vec{d} \in \mathbb{R}^n$ sea $\mathcal{C} = \{F = (q^1, \dots, q^n) : q^j \in C^2([a, b]), F(a) = \vec{c}, F(b) = \vec{d}\}$. Supongamos que $\int_a^b L$ con $L = L(t, F(t), \dot{F}(t))$ alcanza un máximo o un mínimo en \mathcal{C} para cierta F , entonces F es solución de las ecuaciones diferenciales de Euler-Lagrange:*

$$(3.3) \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^k} \right) = \frac{\partial L}{\partial q^k} \quad k = 1, 2, \dots, n.$$

Demostración: Si la integral alcanza un extremo en \mathcal{C} para $F = F_0(t)$ entonces para cualquier función regular $\alpha = \alpha(t)$ con $\alpha(a) = \alpha(b) = \vec{0}$ se cumple que la función real

$$f(\epsilon) = \int_a^b L(t, F_0(t) + \epsilon \alpha(t), \dot{F}_0(t) + \epsilon \dot{\alpha}(t)) dt$$

alcanza un extremo en $\epsilon = 0$. Nuestros conocimientos del primer curso de cálculo, llevan a $f'(0) = 0$. Derivando bajo el signo integral y con una integración por partes

$$0 = \int_a^b \left(\frac{\partial L}{\partial q^k} \alpha^k + \frac{\partial L}{\partial \dot{q}^k} \dot{\alpha}^k \right) = \int_a^b \left(\frac{\partial L}{\partial q^k} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^k} \right) \right) \alpha^k$$

donde α^k son las componentes de α . como éstas son arbitrarias, la única posibilidad para que la integral sea siempre nula es que se cumpla (3.3). \square

La gran ventaja es que las ecuaciones de Euler-Lagrange manifiestan que cierto funcional es estacionario y tal condición es independiente de las coordenadas. Por ejemplo los máximos y mínimos (y los valores estacionarios en general) de $f(x)$ son los mismos que los de $f(3x - \sin x)$ o $f(-x - x^3)$. Es decir, aunque no lo parezca, (3.2) sigue describiendo el movimiento de la partícula si empleamos por ejemplo coordenadas cilíndricas $q^1 = r, q^2 = \theta, q^3 = z$. La idea de que las leyes físicas se puedan interpretar sin referencia a coordenadas específicas, en términos de los extremos de cierto funcional es muy atractiva y de una forma u otra ha sobrevivido a las revoluciones relativista y cuántica del siglo XX.

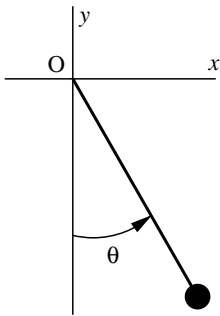
Tanto (3.1) como (3.2) se generalizan sin dificultad al caso de N partículas, simplemente la dimensión pasa de 3 a $3N$ y U quizá refleje también interacciones entre las partículas. Establecer ligaduras significa forzar relaciones entre las variables haciendo

que el *espacio de configuración*, el conjunto donde potencialmente se mueven las variables, pase de \mathbb{R}^{3N} a una subvariedad suya. Por ejemplo, una partícula en \mathbb{R}^3 restringiría su movimiento a S^1 si tiene la ligadura de ser una cuenta de collar. Desde el punto de vista físico-teórico, las ligaduras equivalen a poner una barrera de potencial, es decir, a declarar que se necesitaría demasiada energía para salirse de esa subvariedad tomando U muy grande fuera de ella [Arn78, 17]. Por tanto es una hipótesis física natural suponer que las ecuaciones de Euler-Lagrange (3.3) dan las ecuaciones de movimiento con $L = T - U$ (energía cinética y potencial) también en el caso en que hay ligaduras, expresando T y U en términos de q^i y \dot{q}^i , $1 \leq i \leq n$ con q^i variables independientes que describen la configuración del sistema. A estas variables se les llama *coordenadas generalizadas* mientras que n , la dimensión de la subvariedad, se conoce como número de *grados de libertad*.

En el ámbito de la mecánica clásica se llama a la hipótesis anterior *principio de mínima acción* o *principio de Hamilton*. El enunciado que se encuentra para este principio en algunos libros de Física [LL76], es del tipo:

Principio de mínima acción. Si en los instantes $t = t_1$ y $t = t_2$ el sistema ocupa posiciones que se caracterizan por dos conjuntos de valores de las coordenadas, entonces entre estas posiciones el sistema se moverá de manera que la acción $S = \int_{t_1}^{t_2} L dt$ tendrá el menor valor posible.

Literalmente el principio no es correcto y más bien debería decir que la acción es estacionaria o mínima “localmente” (en intervalos de tiempo pequeños). Se puede encontrar una buena introducción elemental en [FLS64, Ch.19].



Aunque en geometría sólo nos intereseamos en el caso $U = 0$, a modo de ilustración analizaremos el caso del péndulo simple, en el que hay un potencial gravitatorio.

El bien conocido esquema es una varilla inextensible sin masa de longitud l que puede girar en el plano alrededor de uno de sus extremos y tiene una partícula de masa m en el otro. La energía cinética de esta partícula es $\frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$ y la potencial mgy . ahora bien, las coordenadas (x, y) de la partícula no son variables independientes: el sistema sólo tiene un grado de libertad. La manera natural de describirlo es con el ángulo que forma la varilla con la vertical. Se tiene $x = l \sin \theta$, $y = -l \cos \theta$ y el lagrangiano en términos de la coordenada

generalizada θ es

$$L(\theta, \dot{\theta}) = \frac{1}{2}m((l\dot{\theta} \cos \theta)^2 + (l\dot{\theta} \sin \theta)^2) + mgl \cos \theta = \frac{1}{2}ml^2\dot{\theta}^2 + mgl \cos \theta$$

y por tanto

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) = ml^2\ddot{\theta}, \quad \frac{\partial L}{\partial \theta} = -mgl \sin \theta.$$

Entonces las ecuaciones de Euler-Lagrange (3.3) conducen a la conocida ecuación del péndulo, $\ddot{\theta} + \frac{g}{l} \sin \theta = 0$, sin necesidad de pensar en la tensión de la varilla.

Si nos olvidamos de U , que refleja que el sistema está inmerso en un campo de fuerzas, el lagrangiano en cada punto es una forma cuadrática definida positiva en \dot{q}^i . En términos geométricos, en cada punto indica una manera de medir vectores tangentes, o equivalentemente (por polarización) un producto escalar entre ellos. Motivados por la relatividad, en primera instancia exigimos la no degeneración en vez de la positividad a este producto escalar.

Definición: Un campo tensorial G de tipo $(0, 2)$ en una variedad se dice que es una *métrica* si sus componentes g_{ij} forman una matriz simétrica no singular en cada punto.

Definición: Una *variedad semiriemanniana* (o *variedad pseudoriemanniana*) es una variedad dotada de una métrica y se dice que es *variedad riemanniana* si la métrica es definida positiva en cada punto.

Si en una carta $(\mathcal{U}, \phi = (x^1, \dots, x^n))$ una métrica tiene componentes g_{ij} , para cada par de campos de vectores $X = X^i \partial_i$, $Y = Y^i \partial_i$ se tiene, por definición $G(X, Y) = g_{ij} X^i Y^j$. Si queremos referirnos a a métrica G sin mencionar las componentes de los vectores a los que se aplica, en pura ortodoxia notacional se escribe $g_{ij} dx^i \otimes dx^j$ ya que $dx^i(X) = X^i$ y $dx^i(Y) = Y^i$, como covectores, y \otimes significa el producto del resultado de aplicar dos tensores. Esta operación se define en general para cualquier par de tensores².

Ejemplo: En \mathbb{R}^n con la carta trivial la métrica que corresponde al producto escalar usual es $g_{ij} dx^i \otimes dx^j$ con $g_{ij} = 1$ si $i = j$ y 0 en otro caso. ésta es la *métrica usual* y la más empleada, pero hay una infinidad de formas de convertir \mathbb{R}^n en una variedad semiriemanniana.

La forma clásica de expresar una métrica, a veces denostada por los matemáticos, reemplaza productos tensoriales por productos habituales en un sentido formal. Así la métrica usual en \mathbb{R}^2 se escribiría $dx^2 + dy^2$ y la métrica $dx \otimes dx + dx \otimes dy + dy \otimes dx + 10 dy \otimes dy$ sería $dx^2 + 2 dx dy + 10 dy^2$. También es habitual el nombre común ds^2 en vez de G , incluso si la métrica no es definida positiva.

Al igual que el lagrangiano $L = T$ (energía cinética) determina las ecuaciones de movimiento intrínsecamente, sin estar forzados a emplear un sistema de coordenadas externo; también una métrica determina unas curvas destacadas en una variedad riemanniana. En ambos casos, lo que se obtiene es una generalización de las rectas. En el primer caso por analogía con el *principio de inercia* (en ausencia de fuerzas el movimiento es rectilíneo y uniforme) y en el segundo por la propiedad minimizante de la distancia que estudiaremos más adelante.

²Si T es un tensor de tipo (r, s) y S es un tensor de tipo (u, v) , se llama *producto tensorial* de T y S al tensor $T \otimes S$ de tipo $(r+u, s+v)$ cuyo valor es $T(\tilde{\varphi}^1, \dots, \tilde{\varphi}^r, \vec{v}_1, \dots, \vec{v}_s) \cdot S(\tilde{\varphi}^{r+1}, \dots, \tilde{\varphi}^{r+u}, \vec{v}_{s+1}, \dots, \vec{v}_{s+v})$ en $(\tilde{\varphi}^1, \dots, \tilde{\varphi}^{r+u}, \vec{v}_1, \dots, \vec{v}_{s+v})$.

Definición: Dada una variedad semiriemanniana M , se dice que una curva parametrizada $c : I \rightarrow M$ es una *geodésica* si en cada carta (\mathcal{U}, ϕ) con $\text{Im } c \cap \mathcal{U} \neq \emptyset$, las funciones $(\phi \circ c)(t) = (x^1(t), \dots, x^n(t))$ satisfacen las ecuaciones de Euler-Lagrange para $L = g_{ij}\dot{x}^i\dot{x}^j$ con g_{ij} las componentes de la métrica de M .

El diccionario geométrico de la mecánica es entonces:

Mecánica		Geometría
espacio de configuración	\rightarrow	variedad riemanniana M
grados de libertad	\rightarrow	dimensión de M
coordenadas generalizadas	\rightarrow	funciones coordenadas
energía cinética	\rightarrow	métrica en M
ecuaciones de movimiento	\rightarrow	geodésicas

Ejemplo: Sea G la métrica usual en \mathbb{R}^2 , entonces $L = g_{ij}\dot{q}^i\dot{q}^j = \dot{x}^2 + \dot{y}^2$ con $q^1 = x$, $q^2 = y$. Los cálculos para las ecuaciones de Euler-Lagrange son:

$$\frac{\partial L}{\partial \dot{x}} = 2\dot{x}, \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{d}{dt}(2\dot{x}) = 2\ddot{x}, \quad \frac{\partial L}{\partial x} = 0$$

y lo mismo con y . Entonces las ecuaciones de Euler-Lagrange son:

$$\ddot{x} = 0, \quad \ddot{y} = 0$$

que se resuelven como $(x(t), y(t)) = (x_0, y_0) + t(a_0, b_0)$. Esto concuerda con el principio de inercia. En coordenadas polares, un cálculo prueba $L = \dot{r}^2 + r^2\dot{\theta}^2$ y las ecuaciones de Euler-Lagrange son:

$$(3.4) \quad \begin{cases} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = \frac{\partial L}{\partial r} & \Rightarrow \quad \ddot{r} = r\dot{\theta}^2, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta} & \Rightarrow \quad r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0. \end{cases}$$

Estas ecuaciones tan complicadas todavía representan las mismas trayectorias rectilíneas. Por ejemplo, podemos comprobar que $r(t) = 1/\text{sen } \theta(t)$ con $\text{cot } \theta(t) = t$, correspondiente a la recta horizontal antes mencionada, es solución.

Como anticipo de una breve introducción a la relatividad, consideremos una métrica que no es definida positiva y en la que intencionadamente una de las funciones coordenadas se llama t , lo que obliga a un pequeño cambio notacional.

Ejemplo: En \mathbb{R}^2 , denominando a las funciones coordenadas (t, x) consideramos la métrica $G = -e^{x^2} dt^2 + dx^2$. LLamando τ al parámetro de las geodésicas, éstas quedan determinadas por las ecuaciones de Euler-Lagrange

$$\frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{t}} \right) = \frac{\partial L}{\partial t} \quad \text{y} \quad \frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x} \quad \text{con} \quad L = -e^{x^2} \dot{t}^2 + \dot{x}^2.$$

La primera ecuación implica $e^{x^2} \dot{t}^2 = K = \text{cte}$ y sustituyendo en la segunda

$$\ddot{x} + K^2 x e^{-x^2} = 0.$$

Esta fórmula implica que si $x(0) > 0$ y $\dot{x}(0) = 0$ entonces $\ddot{x}(0) < 0$. En términos mecánicos, si x representa el espacio, una partícula que parte del reposo en \mathbb{R}^+ se acelerará hacia el origen. La situación es simétrica en \mathbb{R}^- y entonces podemos explicar la gravedad atractiva de un sol en el origen de la recta real \mathbb{R} , convenciendo a sus habitantes de que su tiempo y espacio tienen conjuntamente la métrica anterior. Dentro de este esquema, el tiempo t es sólo una función coordenada más del espacio de configuración que se relaciona a lo largo de cada geodésica $\gamma = \gamma(\tau)$ con el parámetro τ .

3.2. Introducción a la relatividad

La relatividad especial adquirió gracias a H. Minkowski una formulación geométrica elegante en términos de una métrica. Dicho sea de paso, tal formulación no fue inicialmente del agrado de Einstein pero irónicamente fue el primer paso que permitió que años después formulase su teoría geométrica de la gravedad, la relatividad general. Por ello y por la importancia de la relatividad dentro de la cultura científica, nos entretendremos aquí con algunas deducciones matemáticas en la base de la relatividad especial.

Antes de comenzar y para el resto del capítulo, cada vez que tratemos magnitudes físicas supondremos que hemos elegido unidades, llamadas relativistas, tales que la velocidad de la luz c es 1.

$$\text{unidades relativistas} \quad \longleftrightarrow \quad c = 1.$$

Como en el sistema internacional $c = 299\,792\,458 \text{ m/s}$ (en 1983 se redefinió el metro para que c tuviera este valor exacto) no es una constante adimensional, imponiendo $c = 1$ se pueden medir metros en segundos y viceversa: $299\,792\,458 \text{ m} = 1 \text{ s}$, haciendo hincapié en la idea de que espacio y tiempo son todo uno y se habla del *espacio-tiempo*; idea que, por cierto, es de Minkowski³ aunque a menudo se atribuye a Einstein.

³Minkowski afirmó: “A partir de ahora el espacio por sí mismo y el tiempo por sí mismo, están condenados a desvanecerse en meras sombras y sólo una especie de unión de ambos conservará una realidad independiente” [LEMWed].

La relatividad especial está basada en el famoso artículo de 1905 “Sobre la electrodinámica de cuerpos en movimiento” (véase [LEMWed]), posiblemente el más conocido y citado de Einstein. En él se resuelve una paradoja sobre los fenómenos electromagnéticos de una forma físicamente muy radical. Estos fenómenos vienen determinados por dos campos de vectores \vec{E} y \vec{B} dependientes del tiempo que satisfacen las *ecuaciones de Maxwell* [FLS64], que en ausencia de cargas y corrientes son:

$$(3.5) \quad \operatorname{div} \vec{E} = 0, \quad \operatorname{div} \vec{B} = 0, \quad \operatorname{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \operatorname{rot} \vec{B} = \frac{\partial \vec{E}}{\partial t}.$$

Una consecuencia de estas ecuaciones, fundamental en el mundo contemporáneo, es que existen las ondas electromagnéticas. Este hecho se comprobó años después de que J.C. Maxwell enunciase sus ecuaciones, que son la traducción de hechos experimentales simples con cálculo vectorial [FLS64] [AW01, Ch. 1]. No es casual que Maxwell diera unas de las primeras pruebas del teorema de Stokes. Maxwell también conjeturó que la luz es una onda electromagnética, como se comprobó después, porque de (3.5) se deduce que estas ondas viajan con velocidad c (uno en unidades relativistas).

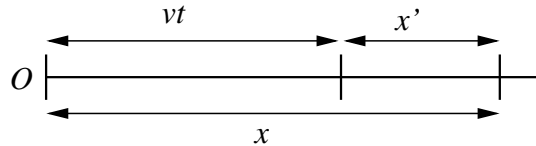
Proposición 3.2.1 *Si \vec{E} y \vec{B} satisfacen las ecuaciones de Maxwell (3.5) entonces cada una de sus componentes satisface la ecuación de ondas*

$$(3.6) \quad \frac{\partial^2 u}{\partial t^2} = \Delta u \quad \text{donde} \quad \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Demostración: Derivando con respecto al tiempo la última ecuación y sustituyendo la penúltima, $\partial^2 \vec{E} / \partial t^2 = -\operatorname{rot} \operatorname{rot} \vec{E}$. La relación del cálculo vectorial [AW01] $\Delta \vec{F} = \nabla(\operatorname{div} \vec{F}) - \operatorname{rot} \operatorname{rot} \vec{F}$ implica que \vec{E} es una solución vectorial de la ecuación de ondas.

La simetría $\vec{E} \mapsto -\vec{B}$, $\vec{B} \mapsto \vec{E}$ de (3.5) implica que lo mismo ocurre para \vec{B} . \square

La paradoja a la que nos hemos referido es que un observador que viaja a velocidad v por el eje X mediría un espacio $x' = x - vt$:



Sin embargo esta transformación no deja invariante la ecuación de ondas (3.6). Por ejemplo, transforma la solución $u = \sin(x - t) + \sin(x + t)$ en $u = \sin(x' + vt - t) + \sin(x' + vt + t)$ que ya no satisface (3.6) cambiando x por x' .

Einstein no dudó de que las ecuaciones de Maxwell (3.5) fueran correctas y tuvo la revolucionaria idea de que $x' = x - vt$ era incorrecta y que los únicos cambios de sistema de referencia admisibles eran los cambios lineales de espacio y tiempo que preservan

la ecuación de ondas⁴, en particular el tiempo también debe cambiar. Se llega así a la conclusión radical de que las longitudes e intervalos de tiempo son relativos, dependen del observador.

Si consideramos cambios en los que no participen y y z , lo cual corresponde naturalmente al caso antes mencionado de un observador que se mueve por el eje X , veremos que la invariancia de la ecuación de ondas implica la de cierta métrica.

Proposición 3.2.2 *Si la ecuación de ondas (3.6) es invariante por cierto cambio de variables lineal en t y en x , entonces tal cambio de variables también deja invariante la métrica de Minkowski*

$$(3.7) \quad ds^2 = -dt^2 + dx^2 + dy^2 + dz^2.$$

Demostración: Digamos que el cambio viene dado por $t = at' + bx'$, $x = ct' + dx'$. Al imponer que $u = \text{sen}(t - y)$ siga siendo solución de (3.6) tras el cambio, se obtiene $a^2 - b^2 = 1$ y procediendo de forma análoga con $\text{sen}(t - x)$ y $\text{sen}(t + x)$ se obtiene $ac = bd$ y $d^2 - c^2 = 1$. Al aplicar el cambio traspuesto (intercambiando b y c) sobre ds^2 , queda

$$ds^2 = -(a^2 - b^2)dt'^2 + 2(bd - ac)dt'dx' + (d^2 - c^2)dx'^2 + dy^2 + dz^2,$$

que con la información anterior es $-dt'^2 + dx'^2 + dy^2 + dz^2$. Por otro lado, la matriz de componentes $D = (g_{ij})$ de ds^2 cumple $D = D^{-1}$ y entonces para cualquier matriz no singular A , la invariancia por A , $A^t D A = D$, equivale a la invariancia por A^t , $A D A^t = D$, sin más que invertir y despejar. \square

La métrica de Minkowski (3.7), formalmente es igual que la usual cambiando t por $t\sqrt{-1}$. Ahora bien, sabemos que las transformaciones que dejan fija la métrica usual son los movimientos del plano. Si nos limitamos al caso de movimientos directos que preservan el origen (una traslación o cambiar de signo uno de los ejes, no añade mucho físicamente) entonces se tienen los giros y el cambio $x \mapsto t\sqrt{-1}$, $y \mapsto x$ que pasa la métrica usual a la de Minkowski, los convertirá en transformaciones que preservan (3.7).

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos \alpha & -\text{sen } \alpha \\ \text{sen } \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \xrightarrow[\substack{x \mapsto t\sqrt{-1} \\ y \mapsto x}]{=} \begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} \cos \alpha & \sqrt{-1} \text{sen } \alpha \\ \sqrt{-1} \text{sen } \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}.$$

⁴Einstein afirmó en su artículo: “está claro que las ecuaciones deben ser *lineales* por las propiedades de homogeneidad que atribuimos al espacio y al tiempo”. Intuitivamente, esto significa que si por ejemplo t' no fuera lineal en x entonces $\partial t'/\partial x \neq \text{cte}$ y el tiempo (en movimiento) transcurriría a diferentes velocidades en diferentes puntos, con lo que podríamos diseñar un experimento con diferente resultado dependiendo del sitio en el que se hiciese. Para no faltar a la verdad, es necesario aclarar que Einstein en su artículo no trabajó con la ecuación de ondas, sino con la consecuencia de que la velocidad de las ondas electromagnéticas, en particular de la luz, debería ser la misma, uno en nuestras unidades, para todos los observadores inerciales.

Para que el cambio sea real, queda forzado $\alpha = i\beta$ con $\beta \in \mathbb{R}$ y entonces las transformaciones buscadas preservando (3.7) son $t' = t \cosh \beta - x \sinh \beta$ y $x' = -t \sinh \beta + x \cosh \beta$. Para ver el significado físico de β , observemos que por definición de velocidad relativa v , la recta $x = vt$ se debe aplicar en $x' = 0$. Es decir, $\tanh \beta = v$ y se obtiene, usando $1 - \tanh^2 \beta = 1/\cosh^2 \beta$, que las ecuaciones anteriores equivalen a las *transformaciones de Lorentz*

$$(3.8) \quad t' = \frac{t - vx}{\sqrt{1 - v^2}}, \quad x' = \frac{x - vt}{\sqrt{1 - v^2}}.$$

Como su nombre indica, estas transformaciones fueron introducidas por H.A. Lorentz antes de que Einstein escribiera su famoso artículo, incluso G. FitzGerald había propuesto la segunda en 1889 (recordamos que Einstein nació en 1879). Sin embargo fue Einstein quien las dotó de su pleno significado físico actual. Para apreciarlo, deshagamos las unidades relativistas. Notemos que t y t' tienen unidades de tiempo, escribimos convencionalmente $[T]$, y x y x' de espacio $[L]$. Para que todo cuadre, vx debe ser $[T]$ en lugar de $[L^2T^{-1}]$ y $\sqrt{1 - v^2}$ debe ser adimensional. Entonces, sin unidades relativistas

$$t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}}, \quad x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}.$$

Incluso a la velocidad de un avión, v/c^2 y $1/\sqrt{1 - v^2/c^2}$ difieren de 0 y de 1 en menos de 10^{-12} , por ello estas fórmulas son una ínfima perturbación de las relaciones clásicas $t' = t$, $x' = x - vt$ para las velocidades habituales de objetos macroscópicos.

Einstein comprobó en su famoso artículo que las transformaciones de Lorentz dejan invariantes no sólo a la ecuación de ondas sino a las ecuaciones de Maxwell, un hecho también avanzado por Lorentz y H. Poincaré. El empleo de tensores permite hacer esta invariancia manifiesta [AW01, 4.6] [FLS64, 26], como más tarde notó el propio Einstein.

Limitarnos a observadores que se mueven por el eje X es una simplificación no esencial, pues siempre podemos pasar a este caso con una rotación de los ejes espaciales x, y, z , lo cual preserva (3.7). En suma, los cambios admisibles de sistema de referencia en la relatividad especial son los movimientos del *espacio de Minkowski* definido como \mathbb{R}^4 dotado con la métrica (3.7).

La relatividad general de alguna forma parte de la relatividad especial pero no es una generalización suya sino una teoría geométrica de la gravitación. Fue creada por Einstein en 1915 y a veces se señala como su aportación más original. En las líneas del ejemplo de la introducción, la relatividad general explica la gravitación como una curvatura, una deformación en la manera de medir, más que como una fuerza en sí misma. Cerca de un cuerpo muy masivo las longitudes son más pequeñas y por eso los cuerpos parecen acelerarse hacia ella. Sin entrar en las motivaciones de la teoría (véase [MTW73]), lo cierto es que explica algunas desviaciones infinitesimales respecto a lo que predice la

teoría de la gravitación creada por I. Newton que permaneció inamovible durante más de dos siglos.

Una curvatura del espacio-tiempo la entendemos como una modificación de la métrica de Minkowski. El esquema seguido en el caso clásico sugiere que el movimiento de las partículas está representado por geodésicas, pero en este caso, ni siquiera con la métrica de Minkowski nos interesan todas ellas.

Consideremos las geodésicas del espacio de Minkowski que parten del origen y tiene $y = z = 0$, éstas vienen dadas por $t = a\tau$, $x = b\tau$ con a y b constantes y τ el parámetro de la geodésica. Claramente $a = 0$, $b \neq 0$ no se realiza físicamente como el movimiento de una partícula: en el instante cero está en todo el eje X . En general, $v = b/a$ indica la velocidad y la imposibilidad de tomar $|v| \geq 1$ en (3.8) (nada va más rápido que la luz), restringe el rango a $|b| < |a|$ que corresponde a valores negativos cuando la métrica de Minkowski se aplica a los vectores tangentes de la geodésica. Multiplicar τ por una constante no afecta a v y si normalizamos de manera que el resultado negativo de la métrica de Minkowski al que nos hemos referido sea -1 , se obtiene $t(\tau) = \tau/\sqrt{1-v^2}$, $x(\tau) = -v\tau/\sqrt{1-v^2}$. Estas fórmulas coinciden con las transformaciones de Lorentz para un observador que mide tiempo τ y espacio 0 , es decir, que viaja con la partícula. En resumen, bajo la normalización de que los vectores tangentes “midan” -1 , el parámetro τ de la geodésica se interpreta físicamente como el llamado *tiempo propio*, el tiempo medido desde la partícula material cuya trayectoria representa la geodésica. Quizá es más sencillo entender su significado intuitivo pensando que se cumple

$$-\Delta\tau^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2,$$

donde Δ indica incremento, y que $t = \tau$ corresponde al caso $\Delta x = \Delta y = \Delta z$ en el que no hay incrementos de espacio. El caso límite $|b| = |a|$ daría $|v| = 1$ que corresponde a un rayo de luz y el vector tangente tendría “longitud” nula con la métrica de Minkowski.

La siguiente definición contempla estos conceptos cuando se trabaja con una perturbación de la métrica de Minkowski.

Definición: Se dice que una variedad semiriemanniana M de dimensión 4 es una *variedad de Lorentz* si en cada punto la signatura de su métrica G coincide con la de la métrica de Minkowski (3.7). Se dice que γ es una *geodésica nula* si $G(\vec{v}, \vec{v}) = 0$ para cada uno de sus vectores tangentes \vec{v} y se dice que es una *geodésica temporal* si $G(\vec{v}, \vec{v}) < 0$. En este último caso cuando $G(\vec{v}, \vec{v}) = -1$ se dice que está parametrizada por el *tiempo propio*.

Observación: Más adelante (Corolario 3.3.4) veremos que a lo largo de cualquier geodésica $G(\vec{v}, \vec{v})$ permanece constante, por tanto el paso de una geodésica temporal a otra parametrizada por el tiempo propio es sólo un cambio de unidades (lineal) en el parámetro. La definición de variedad de Lorentz se extiende a variedades de dimensión

n con signatura $(n - 1, 1)$. Es decir, tales que al diagonalizar la métrica en cada punto haya $n - 1$ signos positivos y uno negativo.

El postulado básico de la relatividad general es que el espacio-tiempo es una variedad de Lorentz con el siguiente diccionario geométrico:

Relatividad		Geometría
espacio-tiempo	→	variedad de Lorentz M
gravedad	→	métrica en M
rayos luminosos	→	geodésicas nulas
partículas materiales	→	geodésicas temporales
tiempo propio	→	parámetro de geodésicas temporales normalizadas

Evidentemente la pregunta es cómo se halla la métrica que corresponde a la gravitación. La respuesta más sincera es que resolviendo una ecuación tan complicada que es casi inútil en la práctica y no puede competir con el modelo de Newton. Volveremos sobre ello más adelante. Por ahora veremos sin justificación, la métrica introducida por K. Schwarzschild que según la teoría, corresponde a la deformación del espacio-tiempo en el exterior de una masa esférica estática y homogénea despreciando la influencia de otras masas.

Definición: Sea $M = \mathbb{R} \times (\mathbb{R}^3 - \{0\})$ con las funciones coordenadas (t, r, θ, φ) siendo r, θ, ϕ las coordenadas esféricas habituales. Se llama *métrica de Schwarzschild* a

$$-\left(1 - \frac{r_0}{r}\right) dt^2 + \left(1 - \frac{r_0}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2$$

donde r_0 es una constante, llamada *radio de Schwarzschild*, que físicamente representa $2Gm$ con G la constante de gravitación universal y m la masa⁵.

Observación: Para $r_0 = 0$ (masa cero), la métrica de Schwarzschild no es más que la métrica de Minkowski en esféricas.

Para estudiar movimientos radiales, los que en términos espaciales sólo dependen del radio r (la “distancia” al origen), nos podemos olvidar de θ y φ y para mayor simplicidad considerar que el espacio-tiempo está representado por $(t, r) \in \mathbb{R} \times \mathbb{R}^+$.

Definición: En $M = \{(t, r) \in \mathbb{R} \times \mathbb{R}^+\}$ se llama *métrica de Schwarzschild bidimensional* a

$$(3.9) \quad -\left(1 - \frac{r_0}{r}\right) dt^2 + \left(1 - \frac{r_0}{r}\right)^{-1} dr^2.$$

⁵En breve veremos que ésta es la única interpretación posible de r_0 si queremos recuperar la concordancia aproximada con la teoría de Newton.

Una primera sorpresa es que aparte de la singularidad natural en $r = 0$, que también existe en la fuerza de Newton, haya otra en $r = r_0$. ¿Por qué no nos percatamos de ella? En el sistema internacional $G = 6.670 \cdot 10^{-11} Nm^2/kg^2$, y si queremos calcular $2GM$ con unidades de longitud debemos deshacer las unidades relativistas, tomando $2GM/c^2$ que para las masas de la Tierra, Júpiter y el Sol da como valores de r_0 respectivamente, $8.87mm$, $2.82m$ y $2.96km$. Entonces para notar esa singularidad la Tierra debería tener el tamaño de una canica, Júpiter debería caber en nuestra habitación o el Sol tendría que poderse recluir en una ciudad.

Los avances astronómicos han mostrado que tales objetos, tan diferentes de los astros que nos rodean y bien conocidos en ciencia ficción, existen realmente. Cuando el radio de una estrella se reduce por un colapso más allá de r_0 entonces la singularidad aparece y se dice que es un *agujero negro* [HE73], [HT90]. La esfera $r = r_0$ en la que la métrica es singular se llama *horizonte de sucesos* del agujero negro. En rigor habría que omitir esta esfera de la variedad que representa el espacio-tiempo aunque veremos más adelante que la singularidad no es tan crítica como parece.

Una cuestión básica es cómo cae un objeto que parte del reposo por efecto de la gravedad. El problema matemático consiste en calcular las geodésicas de (3.9).

Proposición 3.2.3 *Si $\gamma(\tau) = (t(\tau), r(\tau))$ es una geodésica de (3.9) parametrizada por el tiempo propio y tal que $r(0) = R_0 > r_0$, $r'(0) = 0$, entonces*

$$r(\tau) = R_0 \cos^2\left(\frac{1}{2}V(K_0\tau)\right) \quad \text{donde} \quad K_0 = \frac{2}{R_0} \sqrt{\frac{r_0}{R_0}}$$

y $V(x)$ es la función inversa de $f(x) = x + \sin x$.

Observación: Para masas y distancias “normales” K_0 es muy pequeño y entonces $V(K_0\tau)$ es $K_0\tau/2$ con gran aproximación, de donde se infiere que la aceleración inicial es $r''(0) \sim -R_0 K_0^2/8$. Si se quiere que esto se parezca a la fórmula de Newton $-Gm/R_0^2$, se debe cumplir $R_0 K_0^2 = 8GM/R_0^2$ que después de operar lleva a $r_0 = 2GM$, por eso la única manera de interpretar físicamente r_0 es como $2Gm$ si se quiere que la nueva teoría concuerde en el límite con la newtoniana.

Demostración: Las componentes de la métrica no dependen de t , entonces

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{t}}\right) = 0 \quad \Rightarrow \quad \left(1 - \frac{r_0}{r}\right)\dot{t} = \text{cte}$$

y por la definición del tiempo propio $-(1 - r_0/r)\dot{t}^2 + (1 - r_0/r)^{-1}\dot{r}^2 = -1$. Combinando estas dos ecuaciones se tiene $\dot{r}^2 = r_0(r^{-1} - \text{cte})$ y $r'(0) = 0$ implica que esta constante es R_0^{-1} . En resumen, hay que resolver la ecuación diferencial ordinaria

$$\left(\frac{dr}{d\tau}\right)^2 = r_0\left(\frac{1}{r} - \frac{1}{R_0}\right)$$

bajo la condición inicial $r(0) = 0$. Con un cambio de variable $v = V(K_0\tau)$ o equivalentemente $\tau = (v + \sin v)/K_0$, la ecuación anterior es por la regla de la cadena

$$\left(\frac{dr}{dv}\right)^2 = \frac{r_0}{K_0^2}(1 + \cos v)^2\left(\frac{1}{r} - \frac{1}{R_0}\right),$$

y por sustitución directa $r(v) = R_0 \cos^2(v/2) = R_0(1 + \cos v)/2$ es solución de esta ecuación. \square

Consideremos ahora los rayos luminosos en las direcciones radiales.

Proposición 3.2.4 *Si $\gamma(\tau) = (t(\tau), r(\tau))$ es una geodésica nula de (3.9) con $r(0) = R_0 > r_0$, $t(0) = 0$, entonces*

$$\pm t = r - R_0 + r_0 \log\left(\frac{r - r_0}{R_0 - r_0}\right)$$

donde el signo coincide con el de $r'(0)/t'(0)$, es decir, es positivo si es saliente y negativo si es entrante.

Demostración: Por ser una geodésica nula se tiene

$$-\left(1 - \frac{r_0}{r}\right)\left(\frac{dt}{d\tau}\right)^2 + \left(1 - \frac{r_0}{r}\right)^{-1}\left(\frac{dr}{d\tau}\right)^2 = 0$$

y por la regla de la cadena $(dt/dr)^2 = (1 - r_0/r)^{-2}$. empleando las condiciones iniciales se tiene $\pm t = \int_{R_0}^r (1 - r_0/x)^{-1} dx$ y la integral es elemental. \square

Un subproducto de la prueba es que la “velocidad” dr/dt de un rayo de luz es $(1 - r_0/r)$, por consiguiente se ralentiza cuanto más próximo está al horizonte de sucesos y en el límite la velocidad es nula. Esto concuerda con la imagen clásica de un agujero negro como una estrella con una gravedad tan poderosa que ni la luz puede escapar. Curiosamente P.S. Laplace ya teorizó sobre tales objetos a finales del XVIII (véase [HE73]).

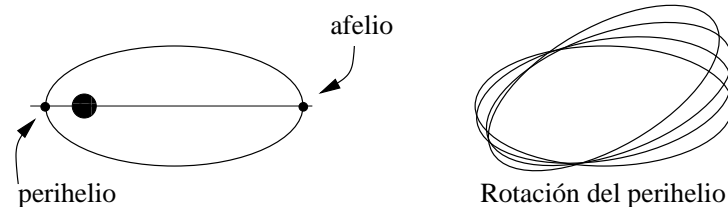
Por otro lado la Proposición 3.2.3 implica que una partícula sólo necesita un tiempo finito para llegar al horizonte de sucesos porque $\cos^2\left(\frac{1}{2}V(K_0\tau)\right)$ toma todos los valores entre 0 y 1. En otras palabras: una partícula puede alcanzar el horizonte de sucesos de un agujero negro en un segundo medido por un observador subido a ella pero los observadores estáticos exteriores si tuvieran instrumentos infinitamente precisos percibirían indefinidamente las señales previas a su entrada en el horizonte de sucesos sin detectar que ésta se efectúa. Esta situación choca fuertemente con nuestro sentido común y Einstein trató de evitarla en [Ein39] probando que bajo ciertas hipótesis que pudieran darse en la formación de cúmulos estelares, el horizonte $r = r_0$ debería estar “tapado” con masa sugiriendo así que los agujeros negros eran un modelo matemático sin realidad

física. Desde el punto de vista geométrico es posible remediar los problemas y paradojas que plantea la singularidad en $r = r_0$ utilizando unas sorprendentes cartas que permiten eliminarla (véase [HE73] §5.5).

Parte de las confusiones en la relatividad general se deben a interpretar t y r como “verdaderos” tiempos y espacios medidos por todos los observadores. La relatividad del tiempo es particularmente contraintuitiva y adquiere una dimensión más profunda que en la relatividad especial por presentarse incluso para observadores estáticos.

Terminaremos esta sección deduciendo un efecto que durante algún tiempo fue el único débil apoyo experimental a la relatividad general.

Según enunció J. Kepler en su primera ley y probó Newton a partir de su ley de gravitación, los planetas se mueven siguiendo órbitas elípticas con el Sol en uno de los focos. Los puntos de la órbita más cercano y más lejano al Sol se llaman *perihelio* y *afelio*, respectivamente. Debido a la influencia de otros planetas y objetos astronómicos las órbitas no son exactamente elípticas. En el siglo XIX se estudió con suma precisión la órbita de Mercurio, el planeta más afectado por la gravedad del Sol, y se observó que no era una elipse estática sino que el perihelio iba rotando levemente de revolución en revolución cierto ángulo que con los datos actuales es de $574''$ por cada siglo. Sorprendentemente una minúscula parte de esta rotación, concretamente $43''$ por siglo no era debida a la influencia de otros planetas. A pesar de ser un problema menor, pues habría que esperar casi 10000 años para detectar una variación de un grado, intrigó a algunos físicos.



Probaremos que la métrica de Schwarzschild tiene órbitas casi-elípticas que presentan una rotación del perihelio coincidiendo con la esperada en el caso de Mercurio. Con este propósito no podemos emplear la simplificación (3.9) porque ahora buscamos geodésicas confinadas en el plano de la eclíptica $\theta = \pi/2$ con r y φ variando. Nos ocuparemos primero de la relación entre estas dos funciones sin fijarnos en la dependencia en τ .

Proposición 3.2.5 *Las geodésicas de la métrica de Schwarzschild incluidas en el plano $\theta = \pi/2$ verifican que r y φ están relacionados mediante*

$$\left(\frac{dr}{d\varphi}\right)^2 = Ar^4 + Br^3 - r^2 + r_0r$$

donde A y B son constantes que dependen de las condiciones iniciales y de r_0 .

Demostración: Las componentes de la métrica de Schwarzschild no dependen de t ni de φ , entonces por las ecuaciones de Euler-Lagrange

$$\frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{t}} \right) = \frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{\varphi}} \right) = 0 \quad \Rightarrow \quad \left(1 - \frac{r_0}{r}\right) \frac{dt}{d\tau} = C_1, \quad r^2 \frac{d\varphi}{d\tau} = C_2$$

con C_1 y C_2 constantes. Sustituyendo los valores de \dot{t} y $\dot{\varphi}$ (y $\dot{\theta} = 0$) en L se deduce

$$-C_1^2 \left(1 - \frac{r_0}{r}\right)^{-1} + \left(1 - \frac{r_0}{r}\right)^{-1} \left(\frac{dr}{d\tau}\right)^2 + C_2^2 r^{-2} = -1$$

y operando

$$\left(\frac{dr}{d\tau}\right)^2 = (C_1^2 - 1) + \frac{r_0}{r} - \frac{C_2^2}{r^2} + \frac{r_0 C_2^2}{r^3}.$$

Dividiendo entre $(d\varphi/d\tau)^2 = C_2^2/r^4$ se obtiene el resultado. \square

Seguiremos ahora [FN79] con un argumento debido a C. Møller que es bastante directo y general no necesitando la hipótesis de que la órbita sea casi circular lo cual no sería aplicable a Mercurio.

El cambio $u = 1/r$ en la proposición anterior conduce a

$$(3.10) \quad (u')^2 = A + Bu - u^2 + r_0 u^3.$$

El perihelio y el afelio corresponden a los valores extremos de u , digamos u_p y u_a , en lo que la derivada se anula, por consiguiente el polinomio P del segundo miembro de (3.10) es divisible por $(u - u_p)(u - u_a)$. Además P tiene como suma de sus raíces r_0^{-1} por las fórmulas de Vieta y esto es suficiente para determinar la tercera raíz, permitiendo escribir (3.10) como

$$\frac{u'}{\pm \sqrt{P(u)}} = 1 \quad \text{donde} \quad P(u) = (u - u_a)(u_p - u)(1 - r_0(u + u_a + u_p))$$

y el signo será positivo si u es creciente ($u' > 0$) y negativo en caso contrario.

Partiendo del afelio, al pasar al perihelio siguiente u crece mientras que cuando se pasa del perihelio anterior al afelio presente u decrece. Con esta idea en mente, integrando la ecuación anterior se obtienen las siguientes fórmulas para la variación del ángulo:

$$\varphi_{\text{per.sig.}} - \varphi_{\text{afe.}} = \int_{u_a}^{u_p} \frac{du}{\sqrt{P(u)}}, \quad \varphi_{\text{afe.}} - \varphi_{\text{per.ant.}} = \int_{u_p}^{u_a} \frac{du}{-\sqrt{P(u)}}.$$

Sumando ambas fórmulas tenemos que la variación del ángulo entre dos perihelios consecutivos es

$$\Delta = 2 \int_{u_a}^{u_p} \frac{du}{\sqrt{P(u)}} - 2\pi.$$

Para Mercurio las medidas astronómicas indican $u_a = 1.43 \cdot 10^{-11}m$ y $u_p = 2.19 \cdot 10^{-11}m$ y el r_0 correspondiente al Sol (la masa que genera la métrica de Schwarzschild) es, como habíamos mencionado, $2.96 \cdot 10^3m$. Calculando numéricamente la integral (con ayuda de un ordenador), se obtiene $\Delta = 5.04 \cdot 10^{-7}$. Teniendo en cuenta que Mercurio tarda 0.24 años en dar una vuelta alrededor del Sol, cada siglo habrá dado 416.67 vueltas y la variación del ángulo se multiplicará por este número, siendo

$$\text{Variación secular} = 2.10 \cdot 10^{-4}rad = 43.32''$$

lo que coincide con gran precisión con la cantidad observada experimentalmente. En realidad no es necesario utilizar el ordenador para estimar Δ , con aproximaciones de Taylor sencillas se puede deducir

$$\Delta \approx 3\pi r_0(u_a + u_p)/2.$$

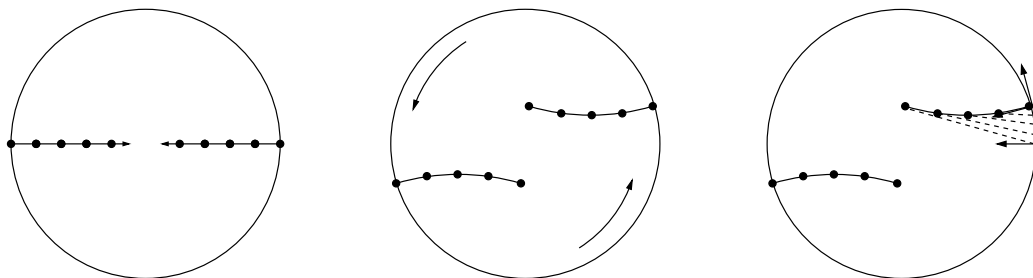
En particular el efecto está en relación inversa con la distancia al Sol.

3.3. Geodésicas y conexiones

Para estudiar la variación de un vector no sólo debemos tener en cuenta la variación de sus componentes sino la del propio sistema de coordenadas. Consideremos por ejemplo en \mathbb{R}^2 una base $B = \{\vec{e}_1, \vec{e}_2\}$ que va rotando con el tiempo, entonces la derivada del vector $\vec{v} = (a(t), b(t)) = a(t)\vec{e}_1 + b(t)\vec{e}_2$ no es $(a'(t), b'(t))$ sino $a'(t)\vec{e}_1 + b'(t)\vec{e}_2 + a(t)\vec{e}'_1 + b(t)\vec{e}'_2$. Como B es una base, existen Γ_j^k con $j, k \in \{1, 2\}$ tales que $\vec{e}'_j(t) = \Gamma_j^1\vec{e}_1(t) + \Gamma_j^2\vec{e}_2(t)$ y la derivada en la base B es $(a', b') + (a\Gamma_1^1 + b\Gamma_2^1, a\Gamma_1^2 + b\Gamma_2^2)$.

Con esta definición la derivada es *absoluta* en el sentido de que se transforma tensorialmente (como un vector, el vector aceleración). En mecánica, estos términos que se añaden a la derivada de las componentes en sistemas que rotan da lugar a la llamada *aceleración de Coriolis*. Una de las demostraciones más claras de su existencia es considerar en una plataforma circular dos chorros de agua que apuntan al centro desde puntos diametralmente opuestos. Al hacer girar la plataforma, los chorros se desvían dejando de converger en el centro.

Para un observador en el borde, la velocidad del chorro sigue la dirección del eje radial y es constante, por ello la desviación es sorprendente desde su punto de vista (lo mismo que lo es la aceleración de Coriolis desde la superficie de la Tierra).



Un observador exterior nota que hay una velocidad tangencial inicial imprimida por el giro y la trayectoria de cada gota es la recta que debería ser (no exactamente radial por efecto de la velocidad tangencial) pero cada una parte de puntos diferentes por la variación del sistema de coordenadas.

Si deseamos derivar campos de vectores en una variedad, la dependencia será en el punto más que en un parámetro t y debemos especificar en la dirección en la que movernos. Si el campo es $V = V^k \partial_k$ y elegimos la dirección ∂_i , entonces un razonamiento como el anterior sugiere que esta “derivada direccional absoluta” debiera tener como componente k -ésima

$$\frac{\partial V^k}{\partial x^i} + \Gamma_{ij}^k V^j$$

para ciertos Γ_{ij}^k que indican la variación del sistema de coordenadas $\phi = (x^1, x^2, \dots, x^n)$. La elección de unos Γ_{ij}^k que den resultados coherentes en cartas compatibles es lo que da lugar a una conexión afín: una manera absoluta (tensorial) de derivar en una variedad. Sin embargo una definición a través de estas funciones sería demasiado fea (aunque posible [Spi79, p.233]) y se prefiere habitualmente otra más sintética y operativa pero menos transparente.

Definición: Una *conexión afín* en una variedad M es una aplicación que asigna a cada par de campos de vectores X, Y un tercero, denotado mediante $\nabla_X Y$, tal que se verifican las siguientes propiedades

- a) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$
- b) $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$
- c) $\nabla_X (fY) = f\nabla_X Y + X(f)Y$

donde f y g son funciones arbitrarias $M \rightarrow \mathbb{R}$.

Tomando f muy concentrada alrededor de un punto y $g = 0$, se deduce de las propiedades anteriores que $\nabla_X Y$ en cada punto p sólo depende de $X|_p$ y del comportamiento local de Y alrededor de p . La terminología “conexión” proviene de T. Levi-Civita que en un trabajo pionero introdujo un concepto de paralelismo generalizado en variedades que permitía *conectar* vectores paralelos de espacios tangentes cercanos.

Ahora reconciliamos esta definición con los comentarios previos.

Lema 3.3.1 *Sea una conexión afín ∇ y sean X y V campos de vectores que se expresan como $X = X^i \partial_i$ y $V = V^i \partial_i$ con una carta dada $(\mathcal{U}, \phi = (x^1, \dots, x^n))$, entonces*

$$\nabla_X V = \left(\frac{\partial V^k}{\partial x^i} + \Gamma_{ij}^k V^j \right) X^i \partial_k$$

donde Γ_{ij}^k es la componente k -ésima de $\nabla_{\partial_i} \partial_j$.

Demostración: Usando las propiedades de la definición de conexión afín, $\nabla_X V = X^i \nabla_{\partial_i} (V^j \partial_j) = X^i (V^j \nabla_{\partial_i} \partial_j + \frac{\partial V^j}{\partial x^i} \partial_j)$ y basta renombrar la variable muda j como k en el último sumando. \square

Cuando uno tiene que hacer cálculos explícitos con conexiones, por ejemplo en relatividad, la notación se vuelve un poco tediosa y existe una especie de taquigrafía bastante extendida que hace más pintorescos los textos de cálculo tensorial.

Se escribe $f_{,i}$ en lugar de $\frac{\partial f}{\partial x^i}$ cuando no se quiere hacer explícito el sistema de coordenadas y $V^k_{;i}$ en lugar de $V^k_{,i} + \Gamma^k_{ij} V^j$.

Definición: Dado un campo de vectores $V = V^k \partial_k$ en una variedad M , se llama *derivada covariante* de V al tensor de tipo $(1, 1)$ que tiene componentes $V^k_{;i}$. Si $c : I \rightarrow M$ es una curva parametrizada, se llama *derivada covariante* de V a lo largo de c , y se denota $\frac{DV}{dt}$, al vector $\nabla_u V$ en $c(t)$ con u el vector tangente a c en $c(t)$.

A veces a la derivada covariante a lo largo de una curva o a la derivada covariante en general, se le llama *derivada absoluta*.

Por la definición libre de coordenadas de ∇ y el Lema 3.3.1, realmente las $V^k_{;i}$ se transforman como componentes de un tensor bajo cambios de carta. Esto obliga a que los Γ^k_{ij} tengan que cumplir una regla de transformación no tensorial (véase la Prop. 4.5 de [Aub01] y su prueba).

Si escribimos $x^i(t)$ y $V^k(t)$ para indicar $(x^i \circ c)(t)$ y $V^k(x^1(t), \dots, x^n(t))$, entonces aplicando la regla de la cadena al Lema 3.3.1

$$(3.11) \quad \frac{DV}{dt} = \left(\frac{dV^k(t)}{dt} + \Gamma^k_{ij}(c(t)) V^j(t) \frac{dx^i(t)}{dt} \right) \partial_k \Big|_{c(t)}.$$

En particular para determinar DV/dt basta con que V esté definido en la imagen de c , no hace falta tener realmente un campo vectorial en toda la variedad. Si $DV/dt = 0$ se dice que $V(t)$ es el *transporte paralelo* de $V(0)$ a $V(t)$ a lo largo de c . Éste es el concepto generalizado de paralelismo al que antes nos hemos referido.

De todas las conexiones afines posibles, hay sólo una que es natural en una variedad semiriemanniana. Implícitamente, sin repetirlo cada vez, consideraremos que es esta conexión afín la que usamos a partir de ahora.

Teorema 3.3.2 *Sea M una variedad semiriemanniana con métrica $G = g_{ij} dx^i \otimes dx^j$. Existe una única conexión ∇ , llamada conexión de Levi-Civita, tal que*

$$\nabla_{\partial_i} \partial_j = \nabla_{\partial_j} \partial_i \quad y \quad \frac{d}{dt} G(V, W) = G\left(\frac{DV}{dt}, W\right) + G\left(V, \frac{DW}{dt}\right)$$

donde $V = V(t)$ y $W = W(t)$ son campos de vectores arbitrarios definidos sobre una curva. Además los Γ_{ij}^k dando la k -ésima componente de $\nabla_{\partial_i}\partial_j$, llamados símbolos de Christoffel, responden a la fórmula

$$\Gamma_{ij}^k = \frac{1}{2}g^{mk}(g_{mi,j} + g_{jm,i} - g_{ij,m})$$

donde (g^{ij}) es la matriz inversa de (g_{ij}) .

Demostración: Supongamos primero la existencia de ∇ . De $\Gamma_{ij}^k = \Gamma_{ji}^k$ y las propiedades de una conexión, se deduce (véase la definición de $[X, Y]$ en el capítulo anterior)

$$(3.12) \quad \nabla_X Y - \nabla_Y X = [X, Y].$$

Por otro lado, considerando curvas integrales de un campo vectorial X , la hipótesis sobre la derivada de un producto escalar equivale a

$$(3.13) \quad X(G(Y, Z)) = G(\nabla_X Y, Z) + G(Y, \nabla_X Z).$$

Con unos cálculos tediosos pero simples, utilizando (3.12) y (3.13) permutando las variables, se sigue [dC92, p.55]

$$(3.14) \quad \begin{aligned} 2G(Z, \nabla_Y X) &= X(G(Y, Z)) + Y(G(Z, X)) - Z(G(X, Y)) \\ &\quad - G([X, Z], Y) - G([Y, Z], X) - G([X, Y], Z). \end{aligned}$$

Tomando $Y = \partial_i$, $X = \partial_j$, $Z = \partial_k$, esta igualdad afirma

$$g_{kl}\Gamma_{ij}^l = \frac{1}{2}(g_{ik,j} + g_{kj,i} - g_{ji,k}).$$

Usando la simetría $g_{ij} = g_{ji}$ y multiplicando por la matriz inversa de la métrica, se obtiene la fórmula esperada.

Para la existencia, notemos primero que por dualidad (las componentes de una métrica forman una matriz no singular), la función $Z \mapsto G(Z, X)$ determina el vector X , por tanto (3.14) define $\nabla_Y X$. Además, gracias a la relación $[X, fY] = f[X, Y] + X(f)Y$, no es difícil ver que realmente ∇ tiene las propiedades de una conexión afín.

Con el cálculo antes realizado, $G(\partial_k, \nabla_{\partial_i}\partial_j) = G(\partial_k, \nabla_{\partial_j}\partial_i)$ de donde $\nabla_{\partial_i}\partial_j = \nabla_{\partial_j}\partial_i$. Finalmente, a partir de (3.14) se obtiene (3.13) que, como hemos dicho, equivale a la hipótesis sobre la derivada de un producto escalar. \square

Para una subvariedad M de \mathbb{R}^n la métrica natural, *inducida por la usual* (a veces también llamada *métrica usual*) es $G(\vec{v}, \vec{w}) = E(di(\vec{v}), di(\vec{w}))$ donde $i : M \hookrightarrow \mathbb{R}^n$ es la inclusión y $E = (dx^1)^2 + \dots + (dx^n)^2$ es la métrica euclídea habitual. En esta situación la derivada covariante no es más que la proyección sobre el espacio tangente de la derivada

usual. De esta forma, la primera propiedad se sigue de la igualdad de las derivadas parciales cruzadas y la segunda de la regla para derivar un producto.

En términos generales, la segunda condición pedida a la conexión de Levi-Civita es una suerte de *compatibilidad* (de hecho a veces se denomina así) entre la derivada covariante y la usual al derivar productos escalares. La primera condición es algo más difícil de interpretar geoméricamente⁶. Sólo diremos que tiene que ver con la existencia de un sistema de coordenadas con el que $V_{;i}^k(p) = V_{;i}^k(p)$ en un punto p [Spi79, Ch.5] lo cual equivale a que los símbolos de Christoffel se anulen en p o a que se pueda encontrar un *sistema normal de coordenadas* en el que los vectores canónicos $\partial_i|_p$ son ortonormales, en el sentido de que $G(\partial_i|_p, \partial_j|_p) = 0$ si $i \neq j$ y $G(\partial_i|_p, \partial_j|_p) = \pm 1$.

Recordemos que habíamos definido las geodésicas como curvas estacionarias para cierto funcional de energía, con la conexión afín asociada a la métrica ahora también las podemos entender como las curvas con aceleración (derivada del vector tangente) nula, obteniéndose así un principio de inercia generalizado. De nuevo las geodésicas se muestran como una generalización de las rectas parametrizadas.

Proposición 3.3.3 *Las geodésicas en una variedad semiriemanniana M son las curvas parametrizadas $c : I \rightarrow M$ para las que $DV/dt = 0$ donde V es el campo de vectores tangentes de c . Es decir, si (\mathcal{U}, ϕ) es una carta y $\phi \circ c(t) = (q^1(t), \dots, q^n(t))$, las geodésicas son las soluciones de las ecuaciones diferenciales*

$$\frac{d^2 q^k}{dt^2} + \Gamma_{ij}^k \frac{dq^i}{dt} \frac{dq^j}{dt} = 0.$$

Demostración: Todo lo que hay que hacer es comprobar que para el lagrangiano $L = g_{ij} \dot{q}^i \dot{q}^j$ las ecuaciones de Euler-Lagrange son $\ddot{q}^k + \Gamma_{ij}^k \dot{q}^i \dot{q}^j = 0$.

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^k} \right) = \frac{d}{dt} (2g_{kj} \dot{q}^j) = 2g_{kj,i} \dot{q}^i \dot{q}^j + 2g_{kj} \ddot{q}^j, \quad \frac{\partial L}{\partial q^k} = g_{ij,k} \dot{q}^i \dot{q}^j.$$

Las ecuaciones de Euler-Lagrange se pueden escribir, por tanto, como

$$2g_{kj} \ddot{q}^j - g_{ij,k} \dot{q}^i \dot{q}^j = -2g_{kj,i} \dot{q}^i \dot{q}^j.$$

En el segundo miembro podemos renombrar arbitrariamente los índices de sumación i y j . Si los intercambiamos y sumamos las ecuaciones resultantes, se deduce

$$g_{kj} \ddot{q}^j + \frac{1}{2} (g_{kj,i} + g_{ki,j} - g_{ij,k}) \dot{q}^i \dot{q}^j = 0.$$

Multiplicando por g^{lk} se obtiene la ecuación del enunciado. \square

⁶Una teoría de la gravitación debida a Cartan, consideraba conexiones sin esta propiedad. Más tarde Einstein intentó unificar la gravitación y el campo electromagnético mediante conexiones con $\Gamma_{ij}^k \neq \Gamma_{ji}^k$. Tal intento, incluso en tiempos de Einstein, llegó a ser poco valorado porque no tomaba en cuenta los avances realizados en la física cuántica.

Corolario 3.3.4 Si $c : I \rightarrow M$ es una geodésica y $V = V(t)$ es su campo de vectores tangentes, entonces $G(V, V)$ es constante en t donde G es la métrica de la variedad.

Demostración: Basta usar $\frac{d}{dt}G(V, V) = G\left(\frac{DV}{dt}, V\right) + G\left(V, \frac{DV}{dt}\right)$. \square

Extenderemos a partir de ahora un poco la notación denotando mediante \dot{c} el campo de vectores tangentes de una curva c . Entonces el resultado anterior implica que una geodésica c es nula si $G(\dot{c}(0), \dot{c}(0)) = 0$, temporal si $G(\dot{c}(0), \dot{c}(0)) < 0$ y parametrizada por el tiempo propio si $G(\dot{c}(0), \dot{c}(0)) = -1$.

Comparando las ecuaciones de las geodésicas de la Proposición 3.3.3 y las que sabíamos obtener con el lagrangiano, no es necesario utilizar la fórmula del Teorema 3.3.2 para calcular los símbolos de Christoffel Γ_{ij}^k . En muchas ocasiones esto reduce drásticamente los cálculos.

Ejemplo: El lagrangiano $L = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$ de la métrica usual en \mathbb{R}^3 pasa a ser $L = \dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta$ cuando se incluye la esfera en \mathbb{R}^3 con la carta en esféricas $\phi = (\theta, \varphi)$ con el significado habitual ($x = \cos \varphi \sin \theta$, $y = \sin \varphi \sin \theta$, $z = \cos \theta$). En este sentido,

$$G = d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi$$

es la métrica usual en S^2 (más propiamente, la inducida por la usual). Calculemos sus símbolos de Christoffel. Las ecuaciones de Euler-Lagrange son

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) = 2\ddot{\theta}, \quad \frac{\partial L}{\partial \theta} = 2\dot{\varphi}^2 \sin \theta \cos \theta, \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\varphi}} \right) = 2\ddot{\varphi} \sin^2 \theta + 4\dot{\varphi}\dot{\theta} \sin \theta \cos \theta, \quad \frac{\partial L}{\partial \varphi} = 0.$$

Por consiguiente las ecuaciones diferenciales de las geodésicas son

$$\begin{cases} \ddot{\theta} - \sin \theta \cos \theta \dot{\varphi}^2 = 0 \\ \ddot{\varphi} + 2 \frac{\cos \theta}{\sin \theta} \dot{\theta} \dot{\varphi} = 0 \end{cases}$$

que comparadas con la definición implican $\Gamma_{12}^2 = \Gamma_{21}^2 = \cos \theta / \sin \theta$, $\Gamma_{22}^1 = -\sin \theta \cos \theta$ y que el resto de los símbolos de Christoffel son cero.

Un cálculo directo con la fórmula del Teorema 3.3.2 habría sido trabajoso.

Como ya hemos mencionado, en subvariedades de \mathbb{R}^n con la métrica inducida por la usual, DV/dt puede interpretarse como la proyección ortogonal de la derivada usual del campo de vectores (evaluado en la curva) sobre el hiperplano tangente, y según la Proposición 3.3.3, las geodésicas son las curvas tales que la derivada de sus vectores tangentes (la aceleración) es siempre un vector normal (la partícula no sufre una fuerza real si está ligada a la subvariedad). Esto permite identificar sin cálculos todas las geodésicas en una esfera y alguna de ellas en un toro usual.

Más concretamente, si tenemos dos superficies de \mathbb{R}^3 que son tangentes a lo largo de una geodésica de una de ellas, entonces también será geodésica en la otra.

Nuestra idea intuitiva sobre el espacio tangente $T_p(M)$ nos hace pensar en él como en una aproximación de M en las cercanías de p , sin embargo las definiciones presentan el espacio tangente y la variedad como objetos bien distintos. Una métrica riemanniana nos provee con una estructura suficientemente fuerte que permite reflejar nuestra intuición a través de cierta aplicación del espacio tangente en la variedad.

La positividad será importante en algunos pasos, por tanto en lo que resta de sección supondremos que M es una variedad riemanniana (no sólo semiriemanniana).

Fijado un punto p en M de dimensión n , el espacio tangente $T_p(M)$ es un espacio vectorial isomorfo a \mathbb{R}^n en el que podemos definir una bola centrada difeomorfa a la bola usual en \mathbb{R}^n

$$B_\epsilon = \{X|_p \in T_p(M) : G(X|_p, X|_p) < \epsilon^2\}$$

con $\epsilon > 0$ y G la métrica de M .

Definición: Para $B_\epsilon \subset T_p(M)$ como antes, se define la *aplicación exponencial* $\exp_p : B_\epsilon \rightarrow M$ como $\exp_p(\vec{v}) = \gamma_{\vec{v}}(1)$ donde $\gamma_{\vec{v}}$ es la geodésica con $\gamma_{\vec{v}}(0) = p$ y $\dot{\gamma}_{\vec{v}}(0) = \vec{v}$.

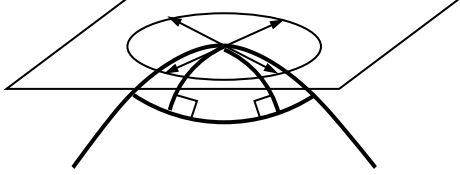
En términos mecánicos, la aplicación entre el espacio tangente en p y la variedad la llevamos a cabo lanzando una partícula con velocidad \vec{v} desde p y viendo dónde está al cabo de un segundo. La topología de la variedad puede estropear la inyectividad e incluso la buena definición de la aplicación exponencial para velocidades grandes (basta poner un agujero que impida avanzar durante un segundo). Todo lo que esperamos es que la variedad y el espacio tangente se parezcan cerca de p .

Proposición 3.3.5 *Existe $\epsilon > 0$ tal que \exp_p establece un difeomorfismo entre B_ϵ y un entorno de p en M . Al supremo de tales valores de ϵ se le llama radio de inyectividad en p .*

Demostración: Obviamente $\gamma_{\vec{0}}$ es la función constante, y por la dependencia continua en los parámetros de las ecuaciones diferenciales ordinarias, existen ϵ_1 y ϵ_2 tales que $\gamma_{\vec{v}}(t)$ está bien definida para $\vec{v} \in B_{\epsilon_1}$ y $|t| < \epsilon_2$. Por otro lado, la homogeneidad de la ecuación de las geodésicas (Proposición 3.3.3) implica que si $\gamma_{\vec{v}}(t)$ es geodésica, $\gamma_{\delta\vec{v}}(\delta^{-1}t)$ también lo es, y ambas coinciden por cumplir las mismas condiciones iniciales. Tomando $\delta = \epsilon_2/2$, el valor $t = 1$ estará en el dominio de definición de $\gamma_{\delta\vec{v}} = \gamma_{\delta\vec{v}}(t)$ y por tanto \exp_p estará bien definida en $B_{\epsilon_1\epsilon_2/2}$.

Si consideramos el espacio vectorial $T_p(M)$ como una variedad, su espacio tangente en el origen O se identifica con él mismo (cada vector se considera como vector tangente en el origen). Ello permite dar sentido a $d\exp_p(\vec{v})$ para $\vec{v} \in T_p(M)$, que por la regla de la cadena es $\dot{c}(0)$ para $c(t) = \exp_p(t\vec{v})$. Por definición $c(t) = \gamma_{t\vec{v}}(1) = \gamma_{\vec{v}}(t)$, así pues

$\dot{c}(0) = \vec{v}$ y $d\exp_p$ es la identidad en el origen de $T_p(M)$. El teorema de la función inversa asegura entonces que \exp_p es difeomorfismo local. \square



Intuitivamente la imagen de una esfera por \exp_p debería ser perpendicular a las geodésicas que parten de p . Este hecho, habitualmente en una versión generalizada (véase [dC92, 3.3.5]), se conoce bajo el nombre de *Lema de Gauss*. Esto permite establecer una especie de coordenadas polares generalizadas heredadas de las del espacio vectorial $T_p(M)$ con la métrica riemanniana particularizada en p .

Proposición 3.3.6 (Lema de Gauss) *Sea una aplicación $\theta \mapsto \vec{v}(\theta) \in T_p(M)$ tal que $\vec{v}(\theta)$ es unitario con la métrica riemanniana G de M en p y definamos $\alpha_\theta(r) = \beta_r(\theta) = \exp_p(r\vec{v}(\theta))$. Entonces para cada r y θ las curvas $\alpha_\theta(t)$ y $\beta_r(t)$ son ortogonales, esto es, $G(\dot{\alpha}_\theta(r), \dot{\beta}_r(\theta)) = 0$.*

Demostración: Comenzamos comprobando que $\frac{D\dot{\alpha}_\theta}{d\theta} = \frac{D\dot{\beta}_r}{dr}$, o escrito de manera más atrayente, $\frac{D}{d\theta} \frac{\partial f}{\partial r} = \frac{D}{dr} \frac{\partial f}{\partial \theta}$, donde $f(r, \theta) = \exp_p(r\vec{v}(\theta))$, lo que se traduce gracias a (3.11) en

$$\frac{\partial^2 f^k}{\partial \theta \partial r} + \Gamma_{ij}^k \frac{\partial f^j}{\partial r} \frac{\partial f^i}{\partial \theta} = \frac{\partial^2 f^k}{\partial r \partial \theta} + \Gamma_{ij}^k \frac{\partial f^j}{\partial \theta} \frac{\partial f^i}{\partial r}$$

y se sigue de $\Gamma_{ij}^k = \Gamma_{ji}^k$ (la propiedad de simetría de la conexión de Levi-Civita).

Por otro lado, la segunda propiedad de la conexión de Levi-Civita asegura

$$\frac{d}{dr} G(\dot{\alpha}_\theta(r), \dot{\beta}_r(\theta)) = G\left(\frac{D\dot{\alpha}_\theta(r)}{dr}, \dot{\beta}_r(\theta)\right) + G\left(\dot{\alpha}_\theta(r), \frac{D\dot{\beta}_r(\theta)}{dr}\right).$$

Por ser $\dot{\alpha}_\theta$ una geodésica, la primera derivada covariante se anula, y por la identidad comprobada al principio, el segundo sumando es

$$G\left(\dot{\alpha}_\theta(r), \frac{D\dot{\beta}_r(\theta)}{dr}\right) = \frac{1}{2} \frac{d}{d\theta} G(\dot{\alpha}_\theta(r), \dot{\alpha}_\theta(r)) = \frac{1}{2} \frac{d}{d\theta} G(\dot{\alpha}_\theta(0), \dot{\alpha}_\theta(0)) = 0$$

donde se ha empleado el Corolario 3.3.4 y que $\dot{\alpha}_\theta(0) = \vec{v}(\theta)$ es unitario.

En definitiva, se concluye que $G(\dot{\alpha}_\theta(r), \dot{\beta}_r(\theta))$ es constante como función de r . ahora bien, si $r \rightarrow 0^+$, $\dot{\alpha}_\theta(r)$ tiende a ser idénticamente p y $\dot{\alpha}_\theta(r) \rightarrow 0$ (si se necesitan detalles sobre este punto, obsérvese que $d\exp_p(r\vec{v}|_p) \rightarrow 0$ cuando $r \rightarrow 0^+$). \square

Para ilustrar la utilidad de la aplicación exponencial y del Lema de Gauss, probaremos una propiedad minimizante de las geodésicas. Empleando la Proposición 3.1.1 se puede obtener que si hay una curva $c \in C^\infty$ en una variedad M con métrica G que conecta dos puntos p y q en el sentido de que $p = c(0)$ y $q = c(1)$, y que minimiza el funcional de longitud

$$(3.15) \quad \ell(c) = \int_0^1 \sqrt{G(\dot{c}(t), \dot{c}(t))} dt$$

entonces necesariamente es una geodésica (véase la Proposición 3.1.4 en [Cha08]). sin embargo, la existencia de tal minimizante no está garantizada y, de hecho, puede no existir una geodésica conectando dos puntos. Todo funciona bien cuando fijamos el punto de partida y sólo consideramos puntos de llegada cercanos.

Teorema 3.3.7 *Sea B_ϵ como en la Proposición 3.3.5 y $\mathcal{U} = \exp_p(B_\epsilon)$. Sea $q \in \mathcal{U}$ y γ una geodésica con $\gamma(0) = p$, $\gamma(1) = q$; entonces $\ell(\gamma) \leq \ell(c)$ para toda curva diferenciable a trozos $c : I \rightarrow \mathcal{U}$ con $c(0) = p$, $c(1) = q$. Además la igualdad sólo se alcanza cuando c es una reparametrización de γ .*

Demostración: Por medio de \exp_p^{-1} la curva c se pasa a una curva en $B_\epsilon \subset T_p(M)$, que se puede escribir como $r(t)\vec{v}(t)$ con $r(t) > 0$ y $\vec{v}(t)$ unitario (con la métrica G de la variedad). Esto es, $c(t) = f(r(t), t)$ con $f(r, t) = \exp_p(r\vec{v}(t))$. De $c(0) = p$, $c(1) = q$ se deduce $r(0) = 0$ y $\gamma(t) = f(t, r(1)\vec{v}(1))$ es una geodésica que conecta p y q , en particular, por el Corolario 3.3.4,

$$\ell(\gamma) = \int_0^1 \sqrt{G(\dot{\gamma}(t), \dot{\gamma}(t))} dt = \int_0^1 \sqrt{G(\dot{\gamma}(0), \dot{\gamma}(0))} dt = r(1).$$

El Lema de Gauss asegura que $\partial f / \partial r$ y $\partial f / \partial t$ son ortogonales, así pues

$$G(\dot{c}, \dot{c}) = G\left(\frac{\partial f}{\partial r} r' + \frac{\partial f}{\partial t}, \frac{\partial f}{\partial r} r' + \frac{\partial f}{\partial t}\right) = (r')^2 G\left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial r}\right) + G\left(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial t}\right).$$

Ahora bien, con la notación empleada en la definición de aplicación exponencial, $f(r, t) = \gamma_{r\vec{v}(t)}(1) = \gamma_{\vec{v}(t)}(r)$ (recuérdese que $\gamma_{\vec{v}}(u) = \gamma_{\delta\vec{v}}(\delta^{-1}u)$ porque ambas geodésicas cumplen las mismas condiciones iniciales), y de nuevo por el Corolario 3.3.4, $G\left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial r}\right) = G(\dot{\gamma}_{\vec{v}(t)}(0), \dot{\gamma}_{\vec{v}(t)}(0)) = 1$. Entonces

$$\ell(c) = \int_0^1 \sqrt{G(\dot{c}, \dot{c})} \geq \int_0^1 \sqrt{(r')^2} = r(1) = \ell(\gamma).$$

Claramente la igualdad sólo se da si $\partial f / \partial t$ es nulo, es decir, cuando $\vec{v}(t)$ es constante y se tiene $c(t) = \exp_p(r(t)\vec{v}_0)$ y $\gamma(t) = \exp_p(r(1)\vec{v}_0)$, lo cual responde al cambio de parámetro $r(1)t \mapsto r(t)$. \square

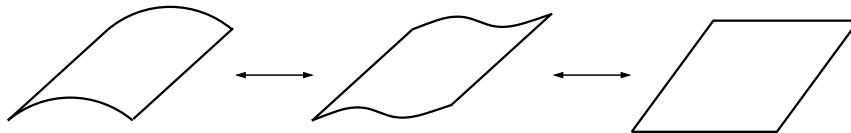
Capítulo 4

Curvatura en variedades

Una curva en \mathbb{R}^2 está intuitivamente muy curvada si su vector normal o tangente (uno viene de girar el otro) presenta grandes variaciones. La definición dada en los cursos básicos de geometría es fiel a esta idea, pues una vez normalizada la parametrización de la curva por longitud de arco, las fórmulas de Frenet-Serret nos dicen que la curvatura es el módulo de la derivada del vector tangente o normal unitario.

En una superficie S inmersa en \mathbb{R}^3 hay muchas direcciones en las que estudiar la variación de la normal alrededor de un punto. Siguiendo un importante y famoso trabajo de C.F. Gauss, un balance global de ellas consiste en medir la variación del área de la región descrita por el extremo del vector normal unitario. Precisamente, se introduce la *aplicación de Gauss* $N : S \rightarrow S^2$ que asigna a cada punto su normal unitaria e identificando los planos tangentes $T_p(S)$ y $T_{N(p)}(S^2)$, que son paralelos, dN se puede considerar un endomorfismo en $T_p(S)$ cuyo determinante mide la variación de área. Se define *curvatura de Gauss* en p , como el valor $K(p)$ de este determinante. Expresiones alternativas para $K(p)$ son el producto de las curvaturas principales o el cociente de los determinantes de la segunda y la primera formas fundamentales.

Gauss, tras introducir la curvatura K , probó el sorprendente resultado de que es intrínseca: depende sólo de la métrica de la superficie (inducida por la usual). Es decir, en contra de lo que sugiere la definición, es posible medir $K(p)$, desde dentro, sin referencia al espacio en el que está inmersa la superficie. A este resultado se le conoce como *Teorema Egregio*. En particular, si $f : S \rightarrow \tilde{S}$ es una *isometría*, una aplicación que preserva la métrica en el sentido de que $G(\vec{v}, \vec{v}) = \tilde{G}(df(\vec{v}), df(\vec{v}))$, entonces $K(p) = \tilde{K}(f(p))$.



Ello explica por qué $K = 0$ para un cilindro, un cono o cualquier cosa que localmente se pueda aplanar en \mathbb{R}^2 sin cambiar las distancias. La aparente curvatura de una porción de cilindro es sólo un efecto de mirarlo desde fuera y desaparecería si viviéramos dentro.

Originariamente, aunque no en su redacción final (véase [Spi79]), Gauss dedujo el Teorema Egregio a partir de lo que hoy en día se conoce como *Teorema de Gauss-Bonnet*:

$$\int_T K dS = \alpha + \beta + \gamma - \pi$$

donde T es un *triángulo geodésico* en la superficie (un triángulo formado por geodésicas) y α , β y γ son sus ángulos. Un habitante de una esfera podría percatarse de que no vive en un plano ($K = 0$) sin salirse de su mundo, comprobando que $\alpha + \beta + \gamma \neq \pi$ y podría estimar K dividiendo $\alpha + \beta + \gamma - \pi$ entre el área de T para triángulos muy pequeños.

Cuando intentamos generalizar la curvatura de Gauss a variedades, nos debemos olvidar de la definición original de Gauss porque no tiene sentido hablar de la normal en cada punto. La cantidad $\alpha + \beta + \gamma - \pi$ en el teorema de Gauss Bonnet es la variación del ángulo en un vector transportado paralelamente a lo largo de las geodésicas que conforman la frontera de T . Por ello, la nueva definición de curvatura interpreta que es lo que indica la variación por unidad de área de un vector por un transporte paralelo “cíclico”. La curvatura será en este sentido un tensor de tipo $(1, 3)$ porque se necesita un vector de partida para ser transportado, dos direcciones que en cierto modo indican dos lados del triángulo geodésico, y el resultado es una variación de carácter vectorial.

Hay también otra manera de entender la curvatura que no desarrollaremos aquí, y es como una especie de coeficientes de Taylor cuando se trata de aproximar una métrica Riemanniana por una euclídea [Gon05, §4.11], [Spi79]. Algo que se cancela si y sólo si estamos en el caso euclídeo, lo cual está más cerca del acercamiento del propio Riemann, quien estaba interesado en dar condiciones necesarias y suficientes para que una variedad fuera localmente isométrica a \mathbb{R}^n .

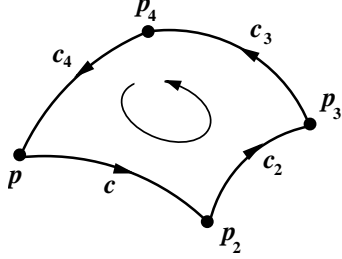
La curvatura tiene una importancia capital en la relatividad general, ya que de postula que la métrica del espacio-tiempo debe satisfacer que la curvatura es proporcional a la densidad de masa y energía. Por otro lado la curvatura, a pesar de su definición local, interviene en resultados “globales” de topología diferencial. Por ejemplo, es conocido que una superficie compacta $S \subset \mathbb{R}^3$ es homeomorfa a una esfera con g asas (la suma conexas de g toros). Esta cantidad g , puramente topológica, se expresa en términos de la curvatura de Gauss mediante la sencilla fórmula

$$\int_S K dS = 4\pi(1 - g).$$

4.1. Definición y significado

Intentamos adaptar a variedades la definición de curvatura de Gauss como una medida de la variación de un vector por transportes paralelos a lo largo de caminos cerrados infinitesimales.

En vez de considerar un triángulo geodésico en la variedad semiriemanniana, consideramos un “paralelogramo” cuyos lados están en las direcciones ∂_1 y ∂_2 , dado por las curvas $c_1(t) = \phi^{-1}(t, 0, 0, \dots)$, $c_2(s) = \phi^{-1}(\delta, s, 0, \dots)$, $c_3(t) = \phi^{-1}(\delta - t, \epsilon, 0, \dots)$ y $c_4(s) = \phi^{-1}(0, \epsilon - s, 0, \dots)$.



Sea $p_i = c_i(0)$, $p = p_1$ y sea P_i el transporte paralelo a lo largo de c_i , entonces la curvatura correspondiente a las direcciones ∂_1 y ∂_2 es la aplicación lineal

$$\vec{v} \in T_p(M) \mapsto - \lim_{\delta, \epsilon \rightarrow 0^+} \frac{P_4 P_3 P_2 P_1(\vec{v}) - \vec{v}}{\delta \epsilon}.$$

El signo menos es convencional para conservar cierto paralelismo con la definición original de la curvatura gaussiana.

Esta definición general es demasiado fea y poco operativa como para asignarle un lugar prominente. Se prefiere otra más sintética, que motivamos a continuación, utilizando la conexión de Levi-Civita.

Recordemos que la derivada covariante medía la variación absoluta de un campo de vectores, teniendo en cuenta no sólo el cambio en sus coordenadas sino también el sistema de referencia en el que se expresan. El transporte paralelo corresponde al caso en el que no hay variaciones a lo largo de una curva c , por ello es natural esperar

$$(4.1) \quad \nabla_{\dot{c}(0)} W = \lim_{t \rightarrow 0} \frac{W_t|_{c(0)} - W|_{c(0)}}{-t}$$

donde $W_t|_{c(0)}$ es el transporte paralelo de $W|_{c(t)}$ a $c(0)$ (véase [Wal04, §4] para una prueba). El denominador viene de que al pasar de $c(t)$ a $c(0)$ el incremento del parámetro es $-t$.

Dado $\vec{v} \in T_p(M)$, consideremos el transporte paralelo a lo largo de c_1 y c_2 (por este orden), lo que dará un vector en $p_3 = \phi^{-1}(\delta, \epsilon, 0, \dots)$. Como δ y ϵ son arbitrarios, podemos definir de esta forma un campo V con $\vec{v} = V|_p$ en un entorno de p , y en cualquier caso dentro del paralelogramo inicial.

Consideremos ahora en p el operador derivada cruzada $R(\partial_1, \partial_2)|_p = \nabla_{\partial_1|_p} \nabla_{\partial_2} - \nabla_{\partial_2|_p} \nabla_{\partial_1}$. Por la definición de V , se tiene $\nabla_{\partial_2} V = 0$ (porque el último transporte paralelo fue por c_2) y (4.1) implica

$$R(\partial_1, \partial_2)|_p V = -\nabla_{\partial_2|_p} \nabla_{\partial_1} V = - \lim_{\epsilon \rightarrow 0^+} \frac{P_4(\nabla_{\partial_1|_{p_4}} V) - \partial_1|_p V}{\epsilon}.$$

Como en c_1 el campo V viene del transporte paralelo de \vec{v} , el segundo término en el numerador se anula y una nueva aplicación de (4.1) lleva a

$$R(\partial_1, \partial_2)|_p V = - \lim_{\delta, \epsilon \rightarrow 0^+} \frac{P_4(P_3(V|_{p_3}) - V|_{p_4})}{\epsilon \delta}.$$

De nuevo por la definición de V se tiene V , se tiene $V|_{p_3} = P_2P_1(\vec{v})$ y $\vec{v} = P_4(V|_{p_4})$, así pues

$$R(\partial_1, \partial_2)|_p V = - \lim_{\delta, \epsilon \rightarrow 0^+} \frac{P_4P_3P_2P_1(\vec{v}) - \vec{v}}{\delta\epsilon}.$$

Entonces podemos medir la curvatura de una variedad semiriemanniana M a través del operador diferencial $R(\partial_1, \partial_2)$. Por supuesto, seleccionando otras direcciones obtendremos otras curvaturas $R(\partial_k, \partial_l)$. Es importante notar que gracias a las propiedades de las conexiones, se verifica $R(\partial_k, \partial_l)(fV) = fR(\partial_k, \partial_l)(V)$ para cualquier función $f : M \rightarrow \mathbb{R}$, entonces la construcción de V es irrelevante, cualquier campo de vectores con $V|_p = \vec{v}$ habría dado el mismo resultado. En otras palabras, $R(\partial_k, \partial_l)$ define una aplicación lineal $T_p(M) \rightarrow T_p(M)$ y por tanto puede considerarse un tensor de tipo $(1, 1)$. Todavía más, generalizando $R(\partial_k, \partial_l)$ a

$$(4.2) \quad R(Y, Z) = \nabla_Y \nabla_Z - \nabla_Z \nabla_Y - \nabla_{[Y, Z]}$$

se tiene, de nuevo por la propiedades de las conexiones, que $(X, Y, Z) \mapsto R(Y, Z)X$ es lineal en sus tres argumentos bajo multiplicación de funciones y así da lugar a una aplicación lineal bien definida $T_p(M) \times T_p(M) \times T_p(M) \rightarrow T_p(M)$. Con un elemento del dual, transformamos la imagen en \mathbb{R} y así obtenemos un tensor de tipo $(1, 3)$. De los tres argumentos vectoriales, el primero indica el vector que sufre el transporte paralelo y los otros dos indican el paralelogramo por el que se lleva a cabo¹.

Definición: En una variedad semiriemanniana se llama *tensor de curvatura* o *tensor de Riemann* al tensor de tipo $(1, 3)$ que corresponde a $(X, Y, Z) \mapsto R(Y, Z)X$ y se denota a sus componentes en una carta mediante R^i_{jkl} . Es decir, R^i_{jkl} es la componente i -ésima de $R(\partial_k, \partial_l)\partial_j$.

Observación: La definición se puede reinterpretar diciendo que para cualquier campo de vectores V se verifica $V^i_{;lk} - V^i_{;kl} = R^i_{jkl}V^j$. A primera vista es sorprendente que las derivadas covariantes cruzadas se cancelen de tal manera que sólo queden términos lineales en las componentes del campo de vectores. En principio (4.2) permite asignar un tensor de curvatura a una conexión sin referencia a una métrica. Sin embargo algunas de las propiedades enunciadas más adelante serían falsas sin una estructura semiriemanniana y supondremos en todo el capítulo que trabajamos en este ámbito con la conexión de Levi-Civita.

Si uno desea una fórmula explícita para R^i_{jkl} en términos de la métrica, debe enfrentarse a una expresión poco manejable.

¹Ésta es la notación habitual pero no es universal, algunos autores (por ejemplo [dC92, p.89]) definen $R(\cdot, \cdot)$ con el signo contrario y admiten otra ordenación de los argumentos. Las simetrías que veremos a continuación, hacen poco importantes estos cambios notacionales

Proposición 4.1.1 *Fijada una carta, se tiene*

$$(4.3) \quad R_{jkl}^i = \Gamma_{jl,k}^i - \Gamma_{jk,l}^i + \Gamma_{nk}^i \Gamma_{jl}^n - \Gamma_{nl}^i \Gamma_{jk}^n.$$

Demostración: Sabemos que $R(\partial_k, \partial_l)\partial_j = R_{jkl}^i \partial_i$, así pues debemos calcular las componentes de $\nabla_{\partial_k} \nabla_{\partial_l} \partial_j - \nabla_{\partial_l} \nabla_{\partial_k} \partial_j$. Por definición, de los símbolos de Christoffel, $\nabla_{\partial_l} \partial_j = \Gamma_{jl}^i \partial_i$, entonces

$$\nabla_{\partial_k} \nabla_{\partial_l} \partial_j = \nabla_{\partial_k} (\Gamma_{jl}^i \partial_i) = (\Gamma_{jl,k}^i + \Gamma_{nk}^i \Gamma_{jl}^n) \partial_i.$$

Intercambiando k y l y restando los resultados, se obtiene la fórmula buscada. \square

En las componentes del tensor de Riemann hay mucha información redundante debido a ciertas simetrías que se hacen más claras si damos a todos sus argumentos el mismo carácter covariante.

Proposición 4.1.2 *Sea $R_{ijkl} = g_{in} R_{jkl}^n$, entonces se cumplen las identidades*

$$a) \quad R_{ijkl} = -R_{ijlk} = -R_{jikl} = R_{klij}, \quad b) \quad R_{ijkl} + R_{iljk} + R_{iklj} = 0.$$

Demostración: La definición $R_{jkl}^i \partial_i = R(\partial_k, \partial_l)\partial_j$ implica $R_{ijkl} = G(R(\partial_k, \partial_l)\partial_j, \partial_i)$.

La primera igualdad de $a)$ se sigue inmediatamente de (4.2). Por otro lado $b)$ también se obtiene a partir de (4.2) con un cálculo más elaborado usando la simetría de los símbolos de Christoffel $\nabla_{\partial_k} \partial_j = \nabla_{\partial_j} \partial_k$.

Una forma bilineal B en un espacio vectorial verifica la identidad de polarización $B(\vec{x}, \vec{y}) + B(\vec{y}, \vec{x}) = B(\vec{x} + \vec{y}, \vec{x} + \vec{y}) - B(\vec{x}, \vec{x}) - B(\vec{y}, \vec{y})$. Entonces es antisimétrica si y sólo si $B(\vec{v}, \vec{v}) = 0$ para todo \vec{v} . Por ello, para obtener $R_{ijkl} + R_{jikl} = 0$ basta probar $G(R(\partial_k, \partial_l)X, X) = 0$. Por las propiedades de la conexión de Levi-Civita

$$\frac{1}{2} \partial_k \partial_l G(X; X) = \partial_k G(\nabla_{\partial_l} X, X) = G(\nabla_{\partial_k} \nabla_{\partial_l} X, X) + G(\nabla_{\partial_l} X, \nabla_{\partial_k} X).$$

Intercambiando k y l y restando, se sigue $G(R(\partial_k, \partial_l)X, X) = 0$.

Finalmente, para probar la igualdad restante, $R_{ijkl} = R_{klij}$, partimos de $b)$ y consideramos esta identidad permutando cíclicamente los nombres de los índices en R_{ijkl} . Por ejemplo, una de las cuatro identidades resultantes sería $R_{jkli} + R_{jikl} + R_{jlik} = 0$. Usando las antisimetrías ya probadas, la suma de todas es $R_{ijkl} - R_{klij} = 0$. \square

Incluso tomando en cuenta las simetrías, el tensor de Riemann con sus n^4 componentes es bastante intrincado y se echa de menos una simple función invariante por isometrías como la curvatura de Gauss en la teoría de superficies. El siguiente resultado, prácticamente de álgebra lineal, marca una dirección para dar una definición en este sentido.

Lema 4.1.3 *Sea M una variedad semiriemanniana con métrica G . Para vectores cualesquiera $\vec{a}, \vec{b}, \vec{x}, \vec{y} \in T_p(M)$ tales que $\{\vec{a}, \vec{b}\}$ y $\{\vec{x}, \vec{y}\}$ son bases de un mismo espacio bidimensional, se tiene*

$$\frac{G(R(\vec{a}, \vec{b})\vec{b}, \vec{a})}{G(\vec{a}, \vec{a})G(\vec{b}, \vec{b}) - G^2(\vec{a}, \vec{b})} = \frac{G(R(\vec{x}, \vec{y})\vec{y}, \vec{x})}{G(\vec{x}, \vec{x})G(\vec{y}, \vec{y}) - G^2(\vec{x}, \vec{y})}.$$

Demostración: Cada matriz real 2×2 se puede escribir como producto de matrices elementales de la forma:

$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \quad y \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Escribiendo la matriz de cambio de base de $\{\vec{a}, \vec{b}\}$ a $\{\vec{x}, \vec{y}\}$ de esta forma, se sigue el resultado, pues cada uno de estos cambios preservan la expresión del enunciado. Para el primer tipo de transformaciones, usando que G es bilineal, $G(\vec{a}, \vec{a})G(\lambda\vec{a} + \vec{b}, \lambda\vec{a} + \vec{b}) - G^2(\vec{a}, \lambda\vec{a} + \vec{b})$ no depende de λ y se puede tomar $\lambda = 0$. Por otro lado, se cumple que $G(R(\vec{a}, \lambda\vec{a} + \vec{b})(\lambda\vec{a} + \vec{b}), \vec{a})$ es igual a $G(R(\vec{a}, \vec{b})(\lambda\vec{a} + \vec{b}), \vec{a})$ ya que $R(\vec{a}, \vec{a}) = 0$ y la antisimetría $R_{ijkl} = -R_{jikl}$ de la Proposición 4.1.2 asegura que $G(R(\vec{a}, \vec{b})\vec{a}, \vec{a}) = 0$.

La invariancia por los otros dos tipos de transformaciones se sigue con cálculos más elementales. \square

Definición: Fijado un punto p en una variedad semiriemanniana M y un supespacio bidimensional σ_p de $T_p(M)$ se llama *curvatura seccional* de σ_p a

$$K(\sigma_p) = \frac{G(R(\vec{a}, \vec{b})\vec{b}, \vec{a})}{G(\vec{a}, \vec{a})G(\vec{b}, \vec{b}) - G^2(\vec{a}, \vec{b})}$$

con $\{\vec{a}, \vec{b}\}$ una base de σ_p .

Las curvaturas seccionales están bien definidas por el Lema 4.1.3 y dan tanta información como el propio tensor de Riemann [dC92, 4§3], que puede recuperarse por polarización [Wal04, p.143]. Además aparecen más naturalmente en algunos teoremas que relacionan curvatura y topología [Wal04, 5].

Para superficies $M \subset \mathbb{R}^3$, con la métrica inducida, $K(\sigma_p)$ es exactamente la curvatura de Gauss en p de $\exp_p(\mathcal{U})$ donde \mathcal{U} es un entorno del origen de σ_p en el que la aplicación exponencial \exp_p sea un difeomorfismo sobre su imagen [dC92, 6].

En dimensión dos, $\sigma_p = T_p(M)$ y sólo hay una curvatura seccional. Entonces podemos redefinir la curvatura de Gauss de una superficie (variedad bidimensional). Aunque no sea una subvariedad de \mathbb{R}^3 y sin referencia a ningún espacio en el que esté inmersa, como la función

$$(4.4) \quad K(p) = \frac{G(R(\partial_1, \partial_2)\partial_2, \partial_1)}{G(\partial_1, \partial_1)G(\partial_2, \partial_2) - G^2(\partial_1, \partial_2)} = \frac{R_{1212}}{g_{11}g_{22} - g_{12}^2}.$$

Gracias al Lema 4.1.3, $K(p)$ es realmente una función en la variedad: no depende del sistema de coordenadas empleado.

En cualquier dimensión, una forma de reducir la aparatosidad del tensor de Riemann sacrificando parte de la información es considerar contracciones, es decir, igualar un índice covariante y otro contravariante y sumar. Debido a las simetrías, básicamente sólo hay una posibilidad, la cual desempeña un papel fundamental en relatividad general y se asocia con el nombre abreviado de G. Ricci-Curbastro. Él y Levi-Civita, que era su estudiante, desarrollaron el cálculo tensorial antes del nacimiento de la relatividad.

Definición: En una variedad semiriemanniana se llama *tensor de Ricci* al tensor de tipo $(0, 2)$ cuyas componentes son $R_{ij} = R_{ikj}^k$. Equivalentemente, el tensor de Ricci asigna a cada par de campos de vectores (X, Y) la traza de la aplicación lineal $R(\cdot, Y)X$.

Hay una fórmula para las componentes del tensor de Ricci que da lugar a cálculos razonables para métricas sencillas en dimensión baja.

Proposición 4.1.4 *Si g es el determinante de la matriz de componentes de la métrica, se cumple*

$$R_{ij} = \frac{1}{\sqrt{|g|}} (\sqrt{|g|} \Gamma_{ij}^k)_{,k} - (\log \sqrt{|g|})_{,ij} - \Gamma_{li}^k \Gamma_{jk}^l.$$

Demostración: Consideramos sólo el caso $g > 0$, el otro es completamente similar.

Sea \vec{v}_i el vector cuya componente j -ésima es g_{ij} , entonces al ser el determinante una función multilineal

$$g_{,i} = \sum_{k=1}^n \det(\vec{v}_1, \dots, \vec{v}_{k,i}, \dots, \vec{v}_n).$$

Desarrollando el determinante por la k -ésima columna se tiene $g_{,i} = g_{jk,i} G^{jk}$ con G^{jk} el cofactor del elemento g_{jk} que por álgebra lineal elemental es $g g^{kj}$. De aquí, por la simetría de g^{kj} ,

$$g_{,i} = g g^{kj} g_{jk,i} = g g^{kj} (g_{jk,i} + g_{ki,j} - g_{ij,k}) = 2g \Gamma_{ik}^k$$

que implica $\Gamma_{ik}^k = (\log \sqrt{|g|})_{,i}$ y por tanto

$$\frac{1}{\sqrt{|g|}} (\sqrt{|g|} \Gamma_{ij}^k)_{,k} = \Gamma_{ij}^k \Gamma_{kl}^l + \Gamma_{ij,k}^k \quad \text{y} \quad (\log \sqrt{|g|})_{,ij} = \Gamma_{ik,j}^k.$$

Entonces la fórmula del enunciado equivale a

$$R_{ij} = \Gamma_{ij,k}^k - \Gamma_{ik,j}^k + \Gamma_{kl}^l \Gamma_{ij}^k - \Gamma_{li}^k \Gamma_{jk}^l$$

y esto se deduce de (4.3) y la definición $R_{ij} = R_{ikj}^k$. \square

O bien usando la Proposición 4.1.2 o directamente con la expresión en coordenadas de la Proposición 4.1.4, se tiene que el tensor de Ricci es simétrico, esto es, $R_{ij} = R_{ji}$.

Ejemplo: Calculemos el tensor de Ricci correspondiente a la métrica de Schwarzschild bidimensional con $r_0 = 1$, dada por $-(1 - r^{-1})dt^2 + (1 - r^{-1})^{-1}dr^2$.

Las ecuaciones de Euler-Lagrange con $L = -(1 - r^{-1})\dot{t}^2 + (1 - r^{-1})^{-1}\dot{r}^2$ dan lugar a las ecuaciones diferenciales de las geodésicas

$$\ddot{t} + \frac{\dot{t}\dot{r}}{r^2 - r} = 0 \quad \text{y} \quad \ddot{r} + \frac{r-1}{2r^3}\dot{t}^2 - \frac{\dot{r}^2}{2(r^2 - r)} = 0.$$

Empleando una notación común en relatividad, asignamos el índice cero al tiempo. Entonces los símbolos de Christoffel no nulos son

$$\Gamma_{01}^0 = \Gamma_{10}^0 = \frac{1}{2(r^2 - r)}, \quad \Gamma_{00}^1 = \frac{r-1}{2r^3} \quad \text{y} \quad \Gamma_{11}^1 = \frac{-1}{2(r^2 - r)}.$$

Según la Proposición 4.1.4 y estos cálculos, $R_{01} = \Gamma_{01,k}^k - \Gamma_{10}^k \Gamma_{1k}^l = 0 - \Gamma_{00}^k \Gamma_{1k}^0 - \Gamma_{10}^k \Gamma_{1k}^1 = 0$. Así pues $R_{01} = R_{10} = 0$, de hecho esto es general en el caso de dimensión dos para métricas diagonales. Por otro lado,

$$R_{00} = \Gamma_{00,1}^1 - \Gamma_{00}^k \Gamma_{0k}^0 - \Gamma_{10}^k \Gamma_{0k}^1 = (-r^{-3} + \frac{3}{2}r^{-4}) - \frac{1}{4}r^{-4} - \frac{1}{4}r^{-4} = \frac{1-r}{r^4}.$$

Análogamente

$$R_{11} = \Gamma_{11,1}^1 - \Gamma_{01}^k \Gamma_{1k}^0 - \Gamma_{11}^k \Gamma_{1k}^1 = \frac{2r-1}{2(r^2 - r)^2} - \frac{1}{4(r^2 - r)^2} - \frac{1}{4(r^2 - r)^2} = \frac{1}{r^2(r-1)}.$$

Con $g^{ik}R_{jk}$ se consigue elevar uno de los índices del tensor de Ricci y así es posible llevar a cabo una contracción más para conseguir una función.

Definición: Se llama *curvatura escalar* de una variedad semiriemanniana a la función $R = g^{ij}R_{ij}$.

Como era de esperar, en el caso bidimensional no hay nada nuevo respecto a la curvatura de Gauss de (4.4).

Proposición 4.1.5 *En una variedad semiriemanniana bidimensional, se cumple $R = 2K$ donde R es la curvatura escalar y K es la curvatura de Gauss.*

Demostración: Con la notación de la Proposición 4.1.2, $R_{jl} = g^{ik}R_{ijkl}$ y las simetrías allí indicadas permiten escribir $R_{11} = g^{22}R_{1212}$, $R_{12} = R_{21} = -g^{12}R_{1212}$, $R_{22} = g^{11}R_{1212}$. Entonces $R = g^{ij}R_{ij}$ y (4.4) implican $R = 2(g^{11}g^{22} - (g^{12})^2)(g_{11}g_{22} - (g_{12})^2)K$. El producto de ambos paréntesis vale uno porque es el determinante de una matriz por el de su inversa, por tanto la curvatura escalar es el doble de la curvatura de Gauss. \square

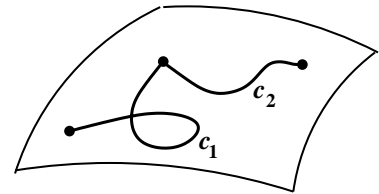
4.2. Curvatura y topología

Una métrica riemanniana en M permite medir vectores en $T_p(M)$ y con ello se pueden calcular (definir) longitudes de curvas simplemente utilizando que los vectores tangentes son velocidades y que el espacio es la velocidad por el tiempo. Con esta idea, ya habíamos definido en (3.15) la longitud de una curva $c : [0, 1] \rightarrow M$ que conecta $p = c(0)$ y $q = c(1)$. Si el dominio de definición es $[t_1, t_2]$ o si queremos medir sólo el arco correspondiente a este intervalo, la definición sería

$$\ell(c) = \int_{t_1}^{t_2} \sqrt{G(\dot{c}(t), \dot{c}(t))} dt$$

que sigue siendo invariante por reparametrizaciones. Cuando c es una geodésica el integrando es constante por el Corolario 3.3.4 y se tiene $\ell(c) = (t_2 - t_1)\sqrt{G(\dot{c}(0), \dot{c}(0))}$.

Claramente no causa ningún problema apreciable que la curva tenga algún punto aislado en el que no sea suave (donde suave significa C^∞) y en contra de la política seguida en el resto del curso, aquí convendrá considerar también estos objetos con fallos en la regularidad en puntos aislados, con el único propósito de conectar curvas distintas. Concretamente, si $c_1 : [t_1, t_2] \rightarrow M$ y $c_2 : [t_2, t_3] \rightarrow M$ con $c_1(t_2) = c_2(t_2)$, la curva $c : [t_1, t_3] \rightarrow M$ dada por $c(t) = c_1(t)$ si $t \in [t_1, t_2]$ y $c(t) = c_2(t)$ si $t \in [t_2, t_3]$ es continua pero no regular en general en $t = t_2$, sin embargo eso no cambia la relación $\ell(c) = \ell(c_1) + \ell(c_2)$.



Si M es una variedad riemanniana conexa, existe siempre una curva que conecta dos puntos (conexo y localmente conexo por arcos implica conexo por arcos) y es natural definir la longitud del camino más corto.

Definición: Sea M una variedad riemanniana conexa. Se define la *distancia riemanniana* en M como la función $d : M \times M \rightarrow \mathbb{R}$ dada por

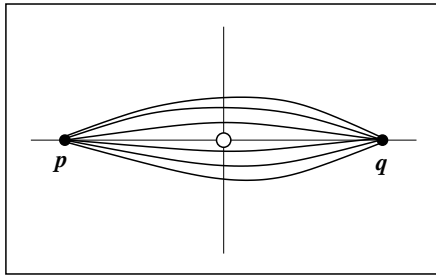
$$d(p, q) = \inf \{ \ell(c) : c(0) = p, c(1) = q \text{ para } c \text{ suave a trozos} \}.$$

Para nuestra tranquilidad notacional, deberíamos convencernos de que la distancia riemanniana es verdaderamente una distancia.

Proposición 4.2.1 *Con la notación de la definición anterior se tiene que (M, d) es un espacio métrico.*

Demostración: Verificar las propiedades es una tarea más o menos rutinaria (que se deja al lector) salvo probar que $d(p, q) = 0$ implica $p = q$. Para ello, recordemos que la

Proposición 3.3.5 asegura que siempre existe una bola centrada B_ϵ en $T_p(M)$ tal que \exp_p define un difeomorfismo en $B_{2\epsilon}$. Si $p \neq q$, escogiendo ϵ suficientemente pequeño, se tiene $q \notin \exp_p(B_\epsilon)$. Si fuera $d(p, q) < \epsilon$ entonces existiría una curva $c(0) = p$, $c(1) = q$ con $\ell(c) < \epsilon$. Sea $t_0 = \inf\{t : c(t) \notin \exp_p(B_\epsilon)\}$, entonces $\tilde{c}(t) = c(t_0 t)$ conectaría $p = \tilde{c}(0)$ con un punto $\tilde{c}(1)$ de la frontera de B_ϵ y $\ell(\tilde{c}) \leq \ell(c) < \epsilon$ porque $0 < t_0 \leq 1$ pero esto contradice el Teorema 3.3.7 (la frontera de B_ϵ son los puntos alcanzables con geodésicas de longitud ϵ). \square



Sabemos que cada par de puntos se conecta con una curva y también, por el Teorema 3.3.7 (ver también la demostración anterior) que localmente se puede reemplazar un arco de curva por uno de geodésica, pero ello no implica que para cada par de puntos exista una geodésica que los une dando un minimizante de la longitud. De hecho tal minimizante puede no existir. Por ejemplo, en \mathbb{R}^2 claramente $d(p, q) = 2$ para $p = (-1, 0)$, $q = (1, 0)$ con el minimizante dado por el segmento de recta que une p y q . En $\mathbb{R}^2 - \{(0, 0)\}$ todavía se tiene $d(p, q) = 2$ pero sólo podemos construir una familia de curvas para la que $\inf \ell(c) = 2$ porque excluir el origen impide tomar el segmento de recta anterior.

El teorema de Hopf-Rinow afirma, entre otras cosas, que la única forma de que no haya *geodésicas minimizantes* (geodésicas c con $c(0) = p$, $c(1) = q$ y $d(p, q) = \ell(c)$ para p y q dados) es que topológicamente falte algún punto. Aquí seguimos esencialmente la exposición de [O'N83].

Teorema 4.2.2 (de Hopf-Rinow) *Sea M una variedad riemanniana conexa y sea d su distancia riemanniana. Las siguientes afirmaciones son equivalentes:*

- El espacio métrico (M, d) es completo.*
- La aplicación \exp_p está bien definida en todo $T_p(M)$ cualquiera que sea $p \in M$.*
- Un subconjunto de M es compacto si y sólo si es cerrado y acotado.*

La mayor dificultad de la prueba viene del siguiente resultado que tiene interés independiente.

Proposición 4.2.3 *Sea M una variedad riemanniana conexa y $p \in M$. Si la aplicación \exp_p está definida en todo $T_p(M)$, entonces para cualquier $q \in M$ existe una geodésica minimizante que conecta p y q .*

Veamos primero cómo a partir de este resultado se deduce el teorema de Hopf-Rinow y la afirmación con la que lo hemos presentado.

Demostración del Teorema 4.2.2: Probaremos $a) \Rightarrow b) \Rightarrow c) \Rightarrow a)$.

a) \Rightarrow b). Si para cierto $p \in M$, el dominio de \exp_p no fuera $T_p(M)$, habría alguna geodésica c con $c(0) = p$ y $G(\dot{c}, \dot{c}) = 1$ tal que $c(t)$ no está definida para todo $t \in \mathbb{R}^+$. Sea $t_0 \in \mathbb{R}^+$ el máximo valor tal que c está definida en $I = [0, t_0)$. Para cualquier $\{t_n\}_{n=1}^\infty \subset I$ con $t_n \rightarrow t_0$ se tiene $d(c(t_n), c(t_m)) \leq |t_n - t_m| \rightarrow 0$, así pues $\{c(t_n)\}_{n=1}^\infty$ es una sucesión de Cauchy y por tanto converge a cierto punto de M . Por consiguiente, c se extiende por continuidad a $[0, t_0]$, lo que contradice la teoría de existencia y unicidad de las ecuaciones diferenciales ordinarias² por la elección de t_0 .

b) \Rightarrow c). Dado $K \subset M$ cerrado y acotado, elijamos $p \in K$ y sea δ el supremo de las distancias de p a los elementos de K . Por la Proposición 4.2.3, para cada $q \in K$ existe una geodésica c_q con $c_q(0) = p$, $c_q(1) = q$ y $\ell(c_q) = d(p, q)$. Entonces $\sqrt{G(\dot{c}_q(0), \dot{c}_q(0))} \leq \delta$ y esto significa que $K \subset \exp_p(\overline{B_\delta})$. El conjunto $\overline{B_\delta}$ es compacto (una bola en un espacio vectorial isomorfo a \mathbb{R}^n) y \exp_p es continua. La topología general básica implica que $\exp_p(\overline{B_\delta})$ es compacto y K también lo es (por ser cerrado dentro de compacto).

c) \Rightarrow a). Una sucesión de Cauchy conforma siempre un conjunto acotado, su cierre será entonces compacto y la sucesión tendrá una subsucesión convergente pero en una sucesión de Cauchy $\lim a_{n_k} = \lim a_n$ siempre que el primer límite exista. \square

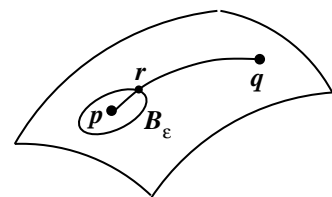
Corolario 4.2.4 *En una variedad riemanniana conexa y completa, cualquier par de puntos están conectados por una geodésica minimizante. En particular $d(p, q)$ es la longitud de la geodésica más corta que conecta p y q .*

Demostración: Basta aplicar a) \Rightarrow b) del teorema de Hopf-Rinow y la Proposición 4.2.3. \square

Demostración de la Proposición 4.2.3: La idea intuitiva básica es sencilla. En un entorno de p las geodésicas son minimizantes por el Teorema 3.3.7, por ello en el camino de p a q se puede comenzar usando una geodésica. Si en algún punto debiéramos cambiar de curva, llegaríamos a una contradicción con el Teorema 3.3.7 en un entorno suyo.

Los detalles topológicos complican el desarrollo de esta idea intuitiva.

Consideremos un entorno de p de la forma $\exp_p(B_\epsilon)$ que no contenga a q y sea $r = \exp_p(\epsilon \vec{v})$, $G(\vec{v}, \vec{v}) = 1$, en su frontera tal que $d(r, q)$ sea mínimo (tal frontera existe y es compacta para ϵ pequeño porque \exp_p es difeomorfismo local). Por el Teorema 3.3.7, $d(p, r) = \epsilon$ y esta distancia se realiza por medio del arco de la geodésica $\gamma(t) = \exp_p(t\vec{v})$ con $t \in [0, \epsilon]$. Para cualquier curva con $c(0) = p$ y $c(1) = q$, sea c_1 el arco de c con $t \in [0, t_0]$ donde $t_0 = \inf\{t \in [0, 1] : c(t) \notin \exp_p(B_\epsilon)\}$ y c_2 el arco $t \in [t_0, 1]$. De



²La única manera de que la solución de una ecuación diferencial ordinaria bien planteada deje de existir, es que explote (como ocurre con la solución de $x' = x^2$, $x(0) = 1$ en $t = 1$). En nuestro caso habría existencia y unicidad en un entorno de t_0 y conseguiríamos una solución en $[0, t_0 + \epsilon)$ en contra de la definición de t_0 .

$d(p, r) = \epsilon$ y la elección de r , $\ell(c) = \ell(c_1) + \ell(c_2) \geq d(p, r) + d(r, q)$ y tomando ínfimos en c ,

$$d(p, q) = d(p, r) + d(r, q) = d(p, \gamma(\epsilon)) + d(\gamma(\epsilon), q).$$

De hecho $d(p, \gamma(t)) = t$ para $\gamma(t) \in \exp_p(B_\epsilon)$ porque $G(\vec{v}, \vec{v}) = 1$ y entonces se tiene

$$(4.5) \quad d(p, q) = t + d(\gamma(t), q) \quad \text{para } t \in [0, \epsilon].$$

Por hipótesis, la geodésica $\gamma(t)$ está definida para todo $t \in \mathbb{R}$, en particular $[0, d(p, q)]$.

Repitiendo el argumento que ha llevado a (4.5) pero con $r = \gamma(\epsilon)$ en lugar de p , se obtiene

$$(4.6) \quad d(r, q) = t + d(\tilde{\gamma}(t), q) \quad \text{para } t \in [0, \epsilon']$$

con $\tilde{\gamma}(t) = \exp_r(t\vec{w})$, $G(\vec{w}, \vec{w}) = 1$ y $\epsilon' > 0$.

Entonces, recordando que $d(p, q) = d(p, r) + d(r, q) = \epsilon + d(r, q)$, la igualdad de (4.5) se cumple para $t \in [0, \epsilon + \epsilon']$ reemplazando $\gamma(t)$ por $c(t) = \gamma(t)$ si $t \in [0, \epsilon]$ y $c(t) = \tilde{\gamma}(t - \epsilon)$ si $t \in [\epsilon, \epsilon + \epsilon']$. Por (4.5) y (4.6), escribiendo $p_1 = c(\epsilon - \delta)$, $q_1 = c(\epsilon + \delta)$ con $0 < \delta < \min(\epsilon, \epsilon')$,

$$d(p, q) = d(p, r) + d(r, q) = d(p, p_1) + d(p_1, r) + d(r, q_1) + d(q_1, q)$$

de donde se deduce, con la ayuda de la triangular, $d(p_1, q_1) = d(p_1, r) + d(r, q_1) = 2\delta$. Si δ es suficientemente pequeño, entonces q_1 está en $\exp_{p_1}(B_\lambda)$ para algún $\lambda > 0$ y por el Teorema 3.3.7, la curva $c(t)$, que conecta $p_1 = c(\epsilon - \delta)$ y $q_1 = c(\epsilon + \delta)$, es una geodésica. En particular, $\tilde{\gamma}$ es la continuación de la geodésica γ y (4.5) se cumple para $t \in [0, \epsilon + \epsilon']$.

La condición (4.5) es cerrada (una distancia es continua por topología general) así pues existe un intervalo maximal $[0, t^*] \subset [0, d(p, q)]$ en el que se satisface, pero acabamos de probar que cualquier intervalo propio se puede extender. La única posibilidad es que (4.5) sea válido para todo $t \in [0, d(p, q)]$ y entonces $\gamma : [0, d(p, q)] \rightarrow M$ es una geodésica de longitud $d(p, q)$ que conecta p y q . \square

Después de que hemos visto propiedades minimizantes de las geodésicas vamos a dar una nueva orientación a este tema relacionando la separación entre geodésicas próximas con la curvatura. Los conceptos iniciales tiene sentido en el ámbito general semiriemanniano.

Definición: Dada una curva $c : [0, 1] \rightarrow M$ en una variedad semiriemanniana M se dice que una función $\alpha : [0, 1] \times (-\delta, \delta) \rightarrow M$ es una *variación* de c si $c(t) = \alpha(t, 0)$. Si además para cada s fijado $t \mapsto \alpha(t, s)$ es una geodésica, diremos que α es una *variación geodésica*. Denotaremos con $\frac{\partial \alpha}{\partial t}$ y $\frac{\partial \alpha}{\partial s}$ a los vectores $d\alpha(\partial_1)$ y $d\alpha(\partial_2)$, respectivamente.

Intuitivamente, $\alpha(t, \epsilon) - \alpha(t, 0)$ mide la variación relativa de la curva central con respecto a otra cercana y la dirección en la que habría que moverse desde $\alpha(t, 0)$ para llegar a la “curva contigua” es $\frac{\partial}{\partial s}\alpha(t, 0)$. Nótese que éste es el vector tangente $\dot{c}(0)$ de cada curva transversal $c(s) = \alpha(t_0, s)$. Análogamente $\frac{\partial}{\partial t}\alpha$ es el vector tangente $\dot{c}(t)$ de cada curva $c(t) = \alpha(t, s_0)$.

Antes de seguir recordemos un resultado que apareció al comienzo de la prueba del Lema de Gauss (Lema 3.3.6). Con nuestra notación para variaciones es

$$(4.7) \quad \frac{D}{dt} \frac{\partial \alpha}{\partial s} = \frac{D}{ds} \frac{\partial \alpha}{\partial t}.$$

La prueba se reduce a la simetría de los símbolos de Christoffel después de escribir la definición de la derivada covariante.

En \mathbb{R}^n hay familias de geodésicas paralelas, como $\alpha(t, s) = (t, s)$ en \mathbb{R}^2 . El siguiente resultado explica esto como consecuencia de que la curvatura sea nula, relacionándola con cierta convexidad o concavidad (una derivada segunda) de la separación de las geodésicas contiguas.

Proposición 4.2.5 *Sea $\alpha(t, s)$ una variación geodésica en una variedad semiriemanniana y sea $V(t)$ el campo de vectores $\frac{\partial}{\partial s}\alpha(t, 0)$ a lo largo de la geodésica $c(t) = \alpha(t, 0)$, entonces se cumple la relación*

$$(4.8) \quad \frac{D^2 V}{dt^2} + R(V, \dot{c})\dot{c} = 0$$

donde $\frac{D^2 V}{dt^2} = \frac{D}{dt} \left(\frac{DV}{dt} \right)$ y $\frac{D}{dt}$ representa la derivada covariante a lo largo de c .

Demostración: Escribamos c_s para indicar la curva $t \mapsto \alpha(t, s)$ con s fijado. Recordando la definición de derivada covariante y la del operador R , se tiene

$$R(Y, \dot{c}_s) = \frac{D}{ds} \frac{D}{dt} X - \frac{D}{dt} \frac{D}{ds} X$$

donde X es el campo $\dot{c}_s(t) = \frac{\partial}{\partial t}\alpha(t, s)$ que podemos considerar indistintamente como función de t o de s y, análogamente, $Y = \frac{\partial}{\partial s}\alpha(t, s)$. Si s está fijado, X es el campo de vectores tangentes de una geodésica y por la Proposición 3.3.3, $\frac{DX}{dt} = 0$. Entonces

$$R(Y, \dot{c}_s) = -\frac{D}{dt} \frac{D}{ds} \frac{\partial}{\partial t} \alpha(t, s) = -\frac{D}{dt} \frac{D}{dt} \frac{\partial}{\partial s} \alpha(t, s).$$

Donde se ha empleado (4.7) para la última igualdad. Particularizando en $s = 0$ se obtiene la igualdad buscada. \square

La ecuación (4.8) es crucial en la relación entre curvatura y topología. Para hacer hincapié en ello, se nombra la propia ecuación y los campos que la verifican.

Definición: Dada una geodésica c , se dice que un campo de vectores a lo largo de ella es un *campo de Jacobi* si se verifica la ecuación (4.8), llamada *ecuación de Jacobi*.

Observación: En realidad separar el concepto de campo de Jacobi en abstracto es un poco gratuito, pues se puede probar (Prop.6.2 [Wal04]) que todos ellos provienen de una variación geodésica.

Si podemos conseguir una variación geodésica que conserve fijos los puntos inicial y final (una homotopía de caminos), estaremos diciendo que hay geodésicas cercanas que pasan por los mismos puntos, como ocurre en S^2 con meridianos próximos uniendo el polo norte y el polo sur. Ésta es la idea que radica en la siguiente definición.

Definición: Sea $c : [0, 1] \rightarrow M$ conectando $p = c(0)$ y $q = c(1)$. Se dice que p y q son *puntos conjugados* si existe un campo de Jacobi J no idénticamente nulo tal que $J(0) = J(1) = \vec{0}$.

Lo que probaremos a continuación, es que la curvatura seccional negativa fuerza a que las geodésicas se separen y por tanto a que no haya puntos conjugados, justo al revés de lo que ocurre en S^2 que tiene curvatura seccional positiva.

Proposición 4.2.6 *Sea M una variedad riemanniana tal que todas sus curvaturas seccionales son menores o iguales que cero, entonces no hay puntos conjugados.*

Demostración: Si p y q fueran conjugados, por definición, existiría un campo de Jacobi J a lo largo de una geodésica c con $c(0) = p$, $c(1) = q$ y $J(0) = J(1) = \vec{0}$.

Consideremos $f : [0, 1] \rightarrow \mathbb{R}$ dada por $f(t) = G(J(t), J(t))$ donde G es la métrica. Por las propiedades de la conexión de Levi-Civita,

$$\frac{1}{2}f''(t) = \frac{1}{2}\frac{d}{dt}f'(t) = \frac{d}{dt}G\left(J, \frac{DJ}{dt}\right) = G\left(\frac{DJ}{dt}, \frac{DJ}{dt}\right) + G\left(\frac{D^2J}{dt^2}, J\right).$$

Utilizando la ecuación de Jacobi y la definición de la curvatura seccional,

$$G\left(\frac{D^2J}{dt^2}, J\right) = -G(R(J, \dot{c})\dot{c}, J) = -K(\sigma)(G(J, J)G(\dot{c}, \dot{c}) - G^2(J, \dot{c}))$$

con σ el subespacio de $T_{c(t)}(M)$ generado por $J(t)$ y $\dot{c}(t)$. El resultado es no negativo por nuestra hipótesis y la desigualdad de Cauchy-Schwarz.

Entonces f es convexa, no negativa y $f(0) = f(1) = 0$, lo cual conduce a que f es idénticamente nula y que J también lo es. \square

Recordemos que definimos geodésica en primer lugar como una curva estacionaria para cierto funcional (a través del principio de mínima acción). Concretamente, si una geodésica conecta dos puntos, entonces $\int G(\dot{c}, \dot{c})$ es estacionario cuando se considera

sobre todas las curvas suaves que conectan esos puntos. Para deducir las ecuaciones de Euler-Lagrange en la Proposición 3.1.1, esencialmente consideramos una variación de una curva preservando los extremos y después integramos por partes. Esta variación no es geodésica en general, de hecho por la Proposición 4.2.6 no puede serlo si la curvatura seccional no es positiva. Lo que estamos diciendo es que las geodésicas que unen dos puntos son estacionarias frente a todas las otras curvas cercanas que los unen. Por conveniencia notacional, demos un nombre a estas variaciones no geodésicas con extremos fijos y al funcional al que se aplican.

Definición: Una variación α de una curva $c : [0, 1] \rightarrow M$ que conecta $p = c(0)$ y $q = c(1)$ se dice que es una *variación propia* de c si $\alpha(0, s) = p$ y $\alpha(1, s) = q$. Se definen la *energía* de α y su *campo variacional*, respectivamente como

$$E(s) = \int_0^1 G(\dot{c}_s, \dot{c}_s) \quad \text{y} \quad V(t) = \frac{\partial}{\partial s} \alpha(t, 0),$$

donde para cada s fijado, c_s es la curva $c_s(t) = \alpha(t, s)$.

Nótese que E es una función $E : (-\delta, \delta) \rightarrow \mathbb{R}$ y V es un campo a lo largo de c con $V(0) = V(1) = \vec{0}$.

Con este lenguaje, la prueba de que las curvas que satisfacen $D\dot{c}/dt = \vec{0}$ (las geodésicas) son estacionarias para la energía, es bastante directa.

Proposición 4.2.7 (primera variación de la energía) *Si V es el campo variacional de una variación propia de una curva $c : [0, 1] \rightarrow M$ en una variedad semiriemanniana,*

$$E'(0) = -2 \int_0^1 G(V, \frac{D\dot{c}}{dt}) dt.$$

Demostración: Vamos a probar la fórmula más general:

$$(4.9) \quad E'(s) = -2 \int_0^1 G\left(\frac{\partial \alpha}{\partial s}, \frac{D\dot{c}_s}{dt}\right) dt.$$

Empleando (4.7),

$$\frac{d}{ds} E(s) = 2 \int_0^1 G\left(\frac{D}{ds} \frac{\partial \alpha}{\partial t}, \dot{c}_s\right) dt = 2 \int_0^1 G\left(\frac{D}{dt} \frac{\partial \alpha}{\partial s}, \dot{c}_s\right) dt.$$

Por otro lado,

$$\int_0^1 \frac{d}{dt} G\left(\frac{\partial \alpha}{\partial s}, \dot{c}_s\right) dt = \int_0^1 G\left(\frac{D}{dt} \frac{\partial \alpha}{\partial s}, \dot{c}_s\right) dt + \int_0^1 G\left(\frac{\partial \alpha}{\partial s}, \frac{D\dot{c}_s}{dt}\right) dt.$$

Al ser una variación propia, $\frac{\partial}{\partial s} \alpha(0, s) = \frac{\partial}{\partial s} \alpha(1, s) = \vec{0}$ y el primer miembro se anula. Despejando y sustituyendo en la fórmula anterior se obtiene (4.9). \square

Proposición 4.2.8 (segunda variación de la energía) *Con la notación de la proposición anterior, si c es una geodésica*

$$E''(0) = -2 \int_0^1 G\left(V, \frac{D^2V}{dt^2} + R(V, \dot{c})\dot{c}\right) dt.$$

Demostración: Partimos de (4.9) en la prueba de la Proposición 4.2.7. Derivando,

$$E''(s) = -2 \int_0^1 G\left(\frac{D}{ds} \frac{\partial \alpha}{\partial s}, \frac{D\dot{c}_s}{dt}\right) dt - 2 \int_0^1 G\left(\frac{\partial \alpha}{\partial s}, \frac{D}{ds} \frac{D\dot{c}_s}{dt}\right) dt.$$

La primera integral se anula para $s = 0$ porque $c_0 = c$ y $D\dot{c}/dt = \vec{0}$ por ser una geodésica. Usando las definiciones y (4.7),

$$\frac{D}{ds} \frac{D\dot{c}_s}{dt} = \frac{D}{ds} \frac{D}{dt} \frac{\partial \alpha}{\partial s} = \frac{D}{dt} \frac{D}{ds} \frac{\partial \alpha}{\partial s} + R\left(\frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t}\right) \frac{\partial \alpha}{\partial t} = \frac{D}{dt} \frac{D}{dt} \frac{\partial \alpha}{\partial s} + R\left(\frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t}\right) \frac{\partial \alpha}{\partial t}.$$

Sustituyendo $s = 0$ se consigue la fórmula buscada. \square

Utilizando estos resultados, vamos a ver que la curvatura condiciona de algún modo lo lejos que pueden estar los puntos.

Teorema 4.2.9 (Bonnet-Myers) *Sea M una variedad riemanniana completa y conexa de dimensión n y supongamos que existe $r > 0$ tal que*

$$R_{ij}v^i v^j \geq (n-1)r^{-2}$$

para todo $p \in M$ y todo vector $\vec{v} = v^i \partial_i \in T_p(M)$ con $G(\vec{v}, \vec{v}) = 1$. Entonces $d(p, q) \leq \pi r$ cualesquiera que sean $p, q \in M$.

Observación: Para la esfera usual de radio r se verifica la condición del teorema y la conclusión es óptima porque para cualquier par de puntos antipodales p y q se tiene $d(p, q) = \pi r$.

En la demostración emplearemos un resultado auxiliar que es básicamente de álgebra lineal.

Lema 4.2.10 *Sean $Z_m = Z_m^i \partial_i$, $m = 1, 2, \dots, n$ campos de vectores que en p dan una base ortonormal de $T_p(M)$ con la métrica G de una variedad semiriemanniana M y sea $X = X^i \partial_i$ otro campo de vectores. En p se verifica la igualdad*

$$R_{ij}X^i X^j = \sum_{m=1}^n G(Z_m, R(Z_m, X)X)$$

donde R_{ij} son las componentes del tensor de Ricci.

Demostración: Es un hecho conocido de álgebra lineal que si F es una aplicación lineal autoadjunta en un espacio de dimensión finita, su traza es $\sum_i \langle \vec{e}_i, F(\vec{e}_i) \rangle$ para cualquier base ortonormal $\{\vec{e}_1, \dots, \vec{e}_n\}$.

En nuestro caso tomamos $F(Z) = R(Z, X)X$ y $\langle X, Y \rangle = G(X, Y)$ en el espacio vectorial $T_p(M)$. La aplicación F es autoadjunta por la Proposición 4.1.2 ya que para cada $Y = Y^i \partial_i$

$$G(Y, F(Z)) = Y^i R_{ijkl} Z^k X^j X^l = Z^k R_{klj} Y^i X^j X^l = G(Z, F(Y))$$

y el resultado se sigue de la definición del tensor de Ricci como una traza. \square

Demostración del Teorema 4.2.9: Por el Corolario 4.2.4, existe una geodésica minimizante $c : [0, 1] \rightarrow M$ que conecta dos puntos dados $p = c(0)$ y $q = c(1)$. Sea $\{\vec{w}_1, \dots, \vec{w}_n\}$ una base ortonormal de $T_p(M)$ con \vec{w}_1 proporcional a $\dot{c}(0)$. Consideramos los campos de vectores en c dados por $V_i(t) = \text{sen}(\pi t)W_i(t)$, $1 \leq i \leq n$, donde $W_i(t)$ es el transporte paralelo de \vec{w}_i desde $c(0)$ a $c(t)$ a lo largo de c . Por la definición del transporte paralelo,

$$\frac{DW_i}{dt} = 0, \quad \frac{DV_i}{dt} = \pi \cos(\pi t)W_i(t) \quad \text{y} \quad \frac{D^2V_i}{dt^2} = -\pi^2 V_i(t).$$

Cada uno de los campos V_i es campo variacional de alguna variación propia que se hallaría resolviendo una ecuación diferencial (nótese que $V_i(0) = V_i(1) = \vec{0}$ permite que sea propia). La Proposición 4.2.8 con $V = V_i$ y el hecho de que c sea minimizante, aseguran

$$0 \leq -2 \int_0^1 G(V_i, -\pi^2 V_i + R(V_i, \dot{c})\dot{c}) dt = \pi^2 - 2 \int_0^1 \text{sen}^2(\pi t) G(W_i, R(W_i, \dot{c})\dot{c}) dt.$$

Sumando esta desigualdad para $2 \leq i \leq n$, se tiene

$$\pi^2(n-1) \geq 2 \int_0^1 \text{sen}^2(\pi t) \sum_{i=1}^n G(W_i, R(W_i, \dot{c})\dot{c}) dt.$$

El sumando correspondiente a $i = 1$ puede añadirse porque $R(W_1, \dot{c})\dot{c}$ es nulo ya que $W_1(t)$ y $\dot{c}(t)$ son proporcionales para cada t , por ser $\frac{d}{dt}G(W_1, \dot{c}) = 0$, y R es antisimétrico.

Por el Lema 4.2.10, el sumatorio se escribe en términos del tensor de Ricci y de nuestras hipótesis se sigue

$$\pi^2(n-1) \geq 2 \int_0^1 \text{sen}^2(\pi t) (n-1) r^{-2} G(\dot{c}(0), \dot{c}(0)) dt = (n-1) r^{-2} (\ell(c))^2.$$

Recordando que $\ell(c) = d(p, q)$, se deduce el resultado. \square

Sabiendo un poco de topología algebraica se deduce del teorema anterior una propiedad del grupo fundamental.

Corolario 4.2.11 *Bajo las hipótesis del Teorema 4.2.9, la variedad M es compacta y tiene grupo fundamental finito.*

Demostración: Según el Teorema 4.2.9 la propia variedad en su totalidad es un conjunto acotado con la distancia riemanniana, entonces el teorema de Hopf-Rinow (Teorema 4.2.2) implica que M es compacta.

Sea \widetilde{M} la variedad que constituye el recubridor universal de M y $\pi : \widetilde{M} \rightarrow M$ la aplicación recubridora (véase el apéndice A de [O’N83] para las definiciones). Se puede dotar a \widetilde{M} de una métrica riemanniana dada por $\widetilde{G}(\vec{v}, \vec{w}) = G(d\pi(\vec{v}), d\pi(\vec{w}))$ donde G es la métrica en M . Esto hace que M y \widetilde{M} sean localmente isométricas (π define siempre un difeomorfismo local por definición de aplicación recubridora) y por tanto que \widetilde{M} también satisfaga las hipótesis del Teorema 4.2.9, lo cual implica, como antes, que \widetilde{M} es compacta y de aquí, que $\text{card}(\pi^{-1}(\{p\})) < \infty$ para cada $p \in M$ (si fuera infinito habría un punto de acumulación y no sería aplicación recubridora). Una de las propiedades del recubridor universal es que este cardinal coincide con el del grupo fundamental. \square

Terminamos enunciando otros dos teoremas que relacionan la curvatura seccional y la topología de una variedad (véanse [dC92] y [Wal04] para las demostraciones y [Ber03] para otros enunciados).

Una variedad completa simplemente conexa puede ser bien diferente de \mathbb{R}^n (por ejemplo S^n), sin embargo una condición adicional sobre la curvatura es suficiente para caracterizar \mathbb{R}^n .

Teorema 4.2.12 (Hadamard) *Sea M una variedad riemanniana completa y simplemente conexa tal que todas sus curvaturas seccionales son menores o iguales que cero, entonces. $\exp_p : T_p(M) \rightarrow M$ define un isomorfismo, en particular es difeomorfa a \mathbb{R}^n .*

En la misma línea (pero mucho más profundo en su demostración), hay un resultado que caracteriza las esferas.

Teorema 4.2.13 (teorema de la esfera) *Sea M una variedad riemanniana compacta y simplemente conexa de dimensión n . Si sus curvaturas seccionales máxima y mínima, K_{max} y K_{min} , satisfacen $K_{min} > \frac{1}{4}K_{max}$ entonces M es homeomorfa a la esfera S^n .*

El teorema es óptimo, en el sentido de que existiría un contraejemplo si permitimos $K_{min} = \frac{1}{4}K_{max}$. Sólo muy recientemente, con técnicas muy avanzadas, se ha conseguido reemplazar “homeomorfa” por “difeomorfa” en la conclusión del teorema.

4.3. Las ecuaciones de campo

Según la teoría clásica de la gravitación, una masa puntual en el origen genera un campo de potencial $\Phi = -GM/r$ donde $r = \sqrt{x^2 + y^2 + z^2}$, $G = 6.670 \cdot 10^{-11} Nm^2/kg^2$ es la constante de gravitación universal y M es el valor de la masa. Esto significa que la trayectoria de una partícula de prueba en este campo, viene determinada por

$$\frac{d^2x}{dt^2} = -\frac{\partial\Phi}{\partial x}, \quad \frac{d^2y}{dt^2} = -\frac{\partial\Phi}{\partial y}, \quad \frac{d^2z}{dt^2} = -\frac{\partial\Phi}{\partial z}.$$

Para utilizar el convenio de sumación, introducimos $\delta^{ij} = 1$ si $i = j$ y $\delta^{ij} = 0$ si $i \neq j$ (que corresponde a los g^{ij} para la métrica euclídea) y escribimos $(x^1, x^2, x^3) = (x, y, z)$. Entonces

$$(4.10) \quad \frac{d^2x^i}{dt^2} = -\delta^{ij}\partial_j\Phi.$$

Si el campo gravitatorio no estuviera generado por una masa puntual, habría que sumar o integrar la contribución de todas las masas (quizá infinitesimales) que generan el campo para obtener Φ . Cuando la partícula de prueba se mueve fuera de la región donde están estas masas, todas las posibilidades quedan representadas por la *ecuación de Poisson*

$$(4.11) \quad \delta^{ij}\partial_i\partial_j\Phi = 0, \quad \text{es decir} \quad \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 0.$$

Concretamente, la función $\Phi(\vec{x}) = -G \int \rho(\vec{y}) \|\vec{x} - \vec{y}\|^{-1} d\vec{y}$ resuelve esta ecuación para \vec{x} fuera del soporte de la función de densidad ρ . Físicamente es natural preguntarse también por el caso en que \vec{x} está en el soporte de ρ , por ejemplo si queremos estudiar la fuerza gravitatoria en el interior de la Tierra. Entonces la solución anterior corresponde a la ecuación de Poisson general

$$(4.12) \quad \delta^{ij}\partial_i\partial_j\Phi = 4\pi G\rho \quad \text{o más claramente,} \quad \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 4\pi G\rho.$$

Parametricemos las trayectorias de partículas cercanas bajo la acción de un potencial Φ mediante una función $\alpha(t, s) = (x^1(t, s), x^2(t, s), x^3(t, s))$ de forma que para cada $s = s_0$ fijo $t \mapsto \alpha(t, s_0)$ sea la ecuación de movimiento de una partícula. En otras palabras, estamos considerando variaciones de una curva como las de la sección anterior donde t representa ahora el tiempo y s cambia de partícula (de curva).

Siguiendo la analogía, el campo asociado a partículas contiguas es $V = \frac{\partial\alpha}{\partial s}(t, 0)$ y la "aceleración relativa" es d^2V/dt^2 . En coordenadas, empleando (4.10), esta aceleración es

$$\frac{d^2V^i}{dt^2} = \frac{\partial^2}{\partial t^2} \frac{\partial x^i}{\partial s}(t, 0) = -\frac{\partial}{\partial x} \Big|_{s=0} \left(\delta^{ik} \partial_k \Phi(x^1(t, s), x^2(t, s), x^3(t, s)) \right) = -\delta^{ik} (\partial_j \partial_k \Phi) V^j.$$

Entonces $d^2V^i/dt^2 = A_j^iV^j$ y la ecuación de Poisson en ausencia de masas (4.11) es $A_i^i = 0$. En términos físicos ésta es una ley de conservación para el flujo de la intensidad de campo, algo así como que toda la fuerza gravitatoria que entra en una región cerrada debe también salir.

En el marco de la relatividad general, las trayectorias vienen dadas por geodésicas y, según la Proposición 4.2.5, las variaciones geodésicas satisfacen

$$\frac{D^2V^i}{dt^2} = -R_{jkl}^i \dot{c}^j \dot{c}^l V^k.$$

El análogo natural de $A_i^i = 0$ es en este contexto que la contracción $R_{jil}^i \dot{c}^j \dot{c}^l$ se anule. Como el vector velocidad \dot{c} es arbitrario, se obtienen las *ecuaciones de campo en el vacío*

$$(4.13) \quad \boxed{R_{ij} = 0} \quad 1 \leq i, j \leq 4.$$

Teniendo en cuenta la simetría $R_{ij} = R_{ji}$, constituyen 10 ecuaciones en derivadas parciales que son intrínsecamente muy complicadas debido a su no linealidad y la métrica de Schwarzschild es una de las pocas soluciones explícitas de interés físico que se conocen. Puede resultar chocante que al pasar de la gravedad newtoniana a la relativista tengamos que cambiar una ecuación (la de Poisson) por 10 ecuaciones. La razón es que en el primer caso tenemos que determinar una función potencial Φ , mientras que en el segundo debemos determinar toda la estructura riemanniana definida por la métrica que está dada por 10 funciones g_{ij} (de nuevo teniendo en cuenta las simetrías). Éste es en parte el precio a pagar por una teoría de la gravitación independiente de las coordenadas.

En el interior de una masa o, por la equivalencia masa-energía, cuando hay energía radiante, las ecuaciones en el vacío (4.13) deberían sustituirse por un análogo de (4.12) y consecuentemente R_{ij} debería ser proporcional al *tensor de energía-momento* T_{ij} que, sin entrar en detalles (véase [Sch85], [HT90]), da las densidades (más bien flujos) de energía y momento lineal. Sin embargo la ecuación natural $R_{ij} = KT_{ij}$ con K constante, o equivalentemente $R^{ij} = KT^{ij}$, no puede ser cierta ya que ciertas razones físicas (esencialmente la conservación de la energía-momento) imponen $T_{;i}^{ij} = 0$ mientras que se sabe que $R_{;i}^{ij} = \frac{1}{2}g^{ij}R_{;i}$ donde R es la curvatura escalar. La posibilidad más inmediata es reemplazar R^{ij} por el llamado *tensor de Einstein* $G^{ij} = R^{ij} - \frac{1}{2}g^{ij}R$ que ya cumple $G_{;i}^{ij} = 0$. Con algunos argumentos físico-matemáticos [Sch85], se deduce que para que en el límite para velocidades pequeños las geodésicas de la relatividad general tiendan a las trayectorias newtonianas, debe tenerse $K = 8\pi G$ (siempre bajo las unidades relativistas con las que $c = 1$). En definitiva, las ecuaciones de campo generales son

$$R_{ij} - \frac{1}{2}g_{ij}R = 8\pi GT_{ij}.$$

Multiplicando por g^{ij} para despejar R , es fácil obtener la formulación equivalente

$$(4.14) \quad \boxed{R_{ij} = 8\pi G\left(T_{ij} - \frac{1}{2}g_{ij}T\right)} \quad \text{con } T = g^{ij}T_{ij}.$$

De esta forma es más claro que cuando $T_{ij} = 0$ (ausencia de masas y energías en el punto de prueba), se obtienen las ecuaciones de campo en el vacío.

D. Hilbert dedujo las ecuaciones de campo a partir de un principio de mínima acción [Fra12], algo así como postular que la gravedad intenta curvar el espacio tiempo lo menos posible, lo cual es física y matemáticamente muy atractivo.

La complicación de las ecuaciones de campo sugiere que busquemos soluciones con simetrías especiales. El caso de simetría radial en el espacio es particularmente interesante y tiene interés tanto en la gravedad a escala local como en los modelos cosmológicos.

Aunque es posible dar bases teóricas rigurosas al significado de la simetría radial [MTW73, Box 32.3], procederemos aquí de modo intuitivo.

La métrica usual en una superficie esférica (en \mathbb{R}^3) de radio L es en coordenadas esféricas $L^2 d\theta^2 + L^2 \sin^2 \theta d\varphi^2$. Consideremos una curva, en \mathbb{R}^3 para fijar ideas, parametrizada por longitud de arco con parámetro u (y por tanto métrica inducida du^2). Si “pegamos” las superficies esféricas anteriores ortogonalmente en los puntos de la curva y les asignamos un radio en función de u , la métrica obtenida será:

$$du^2 + (L(u))^2 d\theta^2 + (L(u))^2 \sin^2 \theta d\varphi^2$$

que con el cambio de variable (carta) natural $r = L(u)$ se escribe como

$$(4.15) \quad B(r)dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2 \quad \text{con } B(r) > 0.$$

Al añadir a este espacio una función coordenada temporal, consideramos métricas de la forma

$$(4.16) \quad g_{11} dt^2 + 2g_{12} dt dr + 2g_{13} dt d\theta + 2g_{14} dt d\varphi + B(r)dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2$$

con g_{11} negativa.

Si estudiamos un fenómeno gravitatorio estático entonces las funciones g_{ij} no deben depender de la coordenada tiempo t que hemos introducido. Además $g_{12} = g_{13} = g_{14} = 0$ porque en otro caso la falta de invariancia con respecto a la inversión del tiempo $t \mapsto -t$ indicaría algún tipo de “sentido de movimiento”. Por otro lado, si extendemos la simetría radial hasta pedir que el tiempo avance al mismo ritmo en todos los puntos a la misma “distancia” r (en la misma esfera) entonces la hipótesis debería ser que g_{11} sólo depende de r .

En definitiva, bajo estas hipótesis se llega a que la métrica debería tener la forma

$$(4.17) \quad -A(r) dt^2 + B(r) dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2.$$

En términos físicos, la métrica correspondiente a la gravedad ejercida por una masa esférica estática y homogénea en el exterior de ella se ajusta a las hipótesis, pues la propia

simetría del problema sugiere que no hay direcciones distinguibles ni posibles cambios en el tiempo (con una Tierra estática y homogénea el norte o el sur son meramente convencionales).

Nuestra experiencia con la gravedad añade otra hipótesis no relacionada con la simetría. La atracción gravitatoria disminuye con la distancia y cuando r es grande es despreciable, así pues (4.17) se debería parecer en esas condiciones a la métrica de Minkowski que en esféricas es $-dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2$ y surge la condición natural $A(+\infty) = B(+\infty) = 1$. Dicho de otra forma, cuando la gravedad es muy pequeña, debemos recuperar la métrica de la relatividad especial.

Teorema 4.3.1 *Si una métrica de la forma (4.17) satisface (4.13) con $A, B \in C^2(\mathbb{R}^+)$ tales que $A(r), B(r) \rightarrow 1$ cuando $r \rightarrow +\infty$, entonces es la métrica de Schwarzschild*

$$-(1 - \frac{r_0}{r})dt^2 + (1 - \frac{r_0}{r})^{-1}dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2$$

donde r_0 es una constante.

Demostración: Consideramos la ordenación natural $(x^1, x^2, x^3, x^4) = (t, r, \theta, \varphi)$ de las funciones coordenadas que ya ha aparecido antes implícitamente (en relatividad es también muy común utilizar la numeración de 0 a 3).

El lagrangiano correspondiente a las métricas (4.17) es $L = -At^2 + Br^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2$ y las ecuaciones de Euler-Lagrange correspondientes son

$$\begin{cases} \ddot{t} + \frac{A'}{A}\dot{t} = 0, & \ddot{r} + \frac{A'}{2B}\dot{t}^2 - \frac{r}{B}\dot{\theta}^2 + \frac{B'}{2B}\dot{r}^2 - \frac{r}{B}\sin^2 \theta \dot{\varphi}^2 = 0 \\ \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin \theta \cos \theta \dot{\varphi}^2 = 0, & \ddot{\varphi} + \frac{2}{r}\dot{r}\dot{\varphi} + \frac{2 \cos \theta}{\sin \theta}\dot{\theta}\dot{\varphi} = 0. \end{cases}$$

De aquí se obtiene que los símbolos de Christoffel no nulos son

$$\begin{aligned} \Gamma_{21}^1 &= \Gamma_{12}^1 = \frac{A'}{2A}, & \Gamma_{11}^2 &= \frac{A'}{2B}, & \Gamma_{22}^2 &= \frac{B'}{2B}, & \Gamma_{33}^2 &= -\frac{r}{B}, & \Gamma_{44}^2 &= -\frac{r}{B}\sin^2 \theta, \\ \Gamma_{23}^3 &= \Gamma_{32}^3 = \frac{1}{r}, & \Gamma_{44}^3 &= -\sin \theta \cos \theta, & \Gamma_{24}^4 &= \Gamma_{42}^4 = \frac{1}{r}, & \Gamma_{34}^4 &= \Gamma_{43}^4 = \frac{\cos \theta}{\sin \theta}. \end{aligned}$$

Ahora por la Proposición 4.1.4 con $g = -r^4 AB \sin^2 \theta$, se deduce tras algunos cálculos tediosos que $R_{ij} = 0$ para $i \neq j$ mientras que

$$(4.18) \quad \begin{cases} R_{11} = \frac{A''}{2B} - \frac{(A')^2}{4AB} - \frac{A'B'}{4B^2} + \frac{A'}{rB}, & R_{22} = -\frac{A''}{2A} + \frac{(A')^2}{4A^2} + \frac{A'B'}{4AB} + \frac{B'}{rB}, \\ R_{33} = \frac{rB'}{2B^2} - \frac{1}{B} - \frac{rA'}{2AB} + 1, & R_{44} = R_{33} \sin^2 \theta. \end{cases}$$

Las ecuaciones de campo (4.13) se traducen en $R_{11} = R_{22} = R_{33} = 0$ que dan lugar a un sistema de ecuaciones diferenciales para A y B . Eliminando A' de R_{11} y R_{22} por medio de una combinación lineal, se llega a una ecuación particularmente sencilla:

$$0 = \frac{B}{A}R_{11} + R_{22} = \frac{A'}{rA} + \frac{B'}{rB} = \frac{(AB)'}{rAB}$$

que implica que AB es constante. De la condición $A, B \rightarrow 1$ cuando $r \rightarrow +\infty$ se deduce $B = A^{-1}$. Sustituyendo en la ecuación $R_{22} = 0$ se sigue $-(rA)' + 1 = 0$ y se obtiene finalmente $A(r) = 1 + r_0/r$ donde r_0 es una constante arbitraria.

Es fácil comprobar que las funciones halladas realmente verifican las ecuaciones y que por tanto no hemos introducido soluciones extrañas. \square

En la gravitación clásica el problema de los dos cuerpos (dos cuerpos que interactúan gravitatoriamente) admite soluciones explícitas bien conocidas, mientras que en el contexto de la relatividad general este problema es ya demasiado difícil y hay que conformarse con el caso de un cuerpo representado por la métrica de Schwarzschild. Por ello parece pretencioso (y seguramente lo es) obtener una solución de las ecuaciones de campo que represente la gravedad a escala cosmológica: la geometría del Universo. La suposición bajo la cual se aborda este problema es la llamada *hipótesis cosmológica* consistente en suponer que el espacio es homogéneo e isótropo, es decir, que los puntos son indistinguibles entre sí y lo mismo se aplica a las direcciones. Basta mirar al cielo nocturno para ver estrellas y planetas que dan al traste con un ajuste perfecto a la hipótesis cosmológica pero datos astronómicos (como la radiación de fondo) sugieren que es correcta a gran escala.

La isotropía es la simetría radial y por tanto nos centraremos en coordenadas espaciales que respondan a métricas de la forma (4.15). Por otro lado, la homogeneidad implica que la curvatura escalar debe ser constante (si variase, unos puntos serían distinguibles de otros). Resulta que esto es suficiente para determinar B .

Proposición 4.3.2 *Si una variedad tridimensional admite una métrica del tipo (4.15) y tiene curvatura escalar R constante, entonces $B(r) = (1 - Rr^2/6)^{-1}$.*

Demostración: En vez de realizar los cálculos partiendo de cero, vamos a aprovechar la prueba del Teorema 4.3.1. Tomemos $A = 1$ en (4.17), entonces $\Gamma_{ij}^k = 0$ si $i = 1$ ó $j = 1$ y la Proposición 4.1.4 implica $R_{ij} = 0$ para estos índices. Además el resto de las componentes del tensor de Ricci coinciden con las correspondientes del tensor de Ricci para la métrica (4.15) porque los g_{ij} no dependen de t y $|g|$ es igual para ambas métricas. Por tanto, teniendo en mente (4.18),

$$R = g^{ij}R_{ij} = \frac{1}{B} \cdot \frac{B'}{rB} + \frac{2}{r^2} \cdot \left(\frac{rB'}{2B^2} - \frac{1}{B} + 1 \right) = \frac{2B'}{rB^2} + \frac{2}{r^2} \left(1 - \frac{1}{B} \right).$$

Esta ecuación se reescribe como $(Rr^3/6 - r + r/B)' = 0$ que integrando implica el resultado. \square

De nuevo, como en (4.16), postulamos que podemos introducir una coordenada temporal t de forma que $g_{12} = g_{13} = g_{14} = 0$, ya que en otro caso la longitud de un vector dependería del sentido espacial en el que apuntase. También extendemos la homogeneidad suponiendo que el tiempo transcurre a la misma velocidad en cada punto, en términos matemáticos, que g_{11} no depende de r , θ y φ . En principio la curvatura escalar sí podría depender de t y consideramos métricas de la forma

$$-A(t)dt^2 + \frac{dr^2}{1 - R(t)r^2/6} + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2.$$

Un cambio $\int \sqrt{A} \mapsto t$ permite suponer $A = 1$ y un cambio posterior $r \mapsto C(t)r$ con $C(t) = \sqrt{6/|R(t)|}$ si $R \neq 0$, lleva finalmente a las *métricas de Robertson-Walker*

$$(4.19) \quad \boxed{-dt^2 + C^2(t) \left(\frac{dr^2}{1 - kr^2/6} + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2 \right)} \quad \text{con } k \in \{-1, 0, 1\}$$

como candidatas a ser la métrica del Universo.

Es fácil ver que cualquier recta $t = \tau$ con r , θ y φ constantes, define una geodésica temporal. Según la interpretación de la relatividad general, en estas coordenadas, cada punto espacial del Universo permanece inmóvil y sólo transcurre el tiempo. En estas condiciones y suponiendo que la influencia de la energía radiante en el universo actual es despreciable, la elección físicamente natural del tensor de energía-momento es que todas sus componentes sean nulas excepto $T_{00} = \rho(t)$ que representa cierta densidad [Sch85]. Con este tensor T_{ij} las ecuaciones de campo (4.14) para (4.19) dan lugar a dos ecuaciones diferenciales sencillas que admiten una solución explícita [FN79]. Esta solución tiene un valor en el pasado para el que la métrica colapsa porque $C(t) \rightarrow 0$ y $\rho(t) \rightarrow \infty$. Esto es lo que se llama el *Big Bang*. Si $k = 1$, además hay un valor en el futuro con la misma propiedad, lo que se llama el *Big Crunch*. En los casos $k = 0$ y $k = -1$, $C(t)$ es monótona creciente, lo cual se interpreta como la expansión del Universo ya que una geodésica con $t = t_0$ que conecta dos puntos tiene longitud proporcional a $C(t_0)$.

Los datos astronómicos de principios de los años 90 del siglo XX (y otros datos más precisos posteriores), sugerían que el Universo responde con gran aproximación al caso $k = 0$ de (4.19) pero durante esa misma década, mediciones realizadas a partir de la observación de supernovas muy lejanas, mostraban un desacuerdo con las ecuaciones de campo. Es como si la expansión del Universo tuviera una aceleración añadida que da lugar a un crecimiento de $C(t)$ incompatible con las ecuaciones. Hasta hoy no se ha encontrado una explicación satisfactoria de este fenómeno.

Apéndice A

Hojas del curso 2011-2012

Las cuatro hojas de problemas incluidas en las páginas siguientes se emplearon para la evaluación en el curso 2011-2012. Cada una de ellas está asociada a un capítulo de estos apuntes. Las “notas e indicaciones” acompañaban originalmente a los enunciados.

Algunos de los ejercicios son de dificultad media-alta. Se disponía de tiempo suficiente para cada hoja (en general más de dos semanas) y por supuesto se podía emplear la bibliografía que se considerase conveniente. También cabía la posibilidad de pedir algunas indicaciones más al profesor, que en caso de ser demasiado cercanas a la solución, descontaban sobre la puntuación máxima asignada al ejercicio correspondiente.

Se incluye también el examen final voluntario, una opción que eligieron pocos alumnos. No disminuía la calificación ya obtenida con las hojas de problemas.

Fecha límite de entrega: 3 de noviembre

Ejercicios

1) Sea $O(2)$ el conjunto de matrices ortogonales 2×2 (las que corresponden a movimientos en \mathbb{R}^2). Este conjunto viene determinado por ciertas ecuaciones en los cuatro elementos de las matrices y por tanto puede considerarse como una subvariedad de \mathbb{R}^4 . ¿Qué dimensión tiene? Encontrar una variedad sencilla a la que sea difeomorfa.

2) Consideremos el hemisferio superior de la esfera unidad $H = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, z > 0\}$ con la carta que da la proyección en las dos primeras coordenadas $(x, y, z) \mapsto (x, y)$. Sea $f : H \rightarrow H$ la función que asigna a cada punto p el punto medio del arco de meridiano que une p con el polo norte. Hallar las coordenadas de $Y|_{f(p)} = df(X|_p)$ en la base $\{\frac{\partial}{\partial x}|_{f(p)}, \frac{\partial}{\partial y}|_{f(p)}\}$ donde df es la aplicación tangente y $X|_p = \frac{x}{\sqrt{x^2+y^2}} \frac{\partial}{\partial x}|_p + \frac{y}{\sqrt{x^2+y^2}} \frac{\partial}{\partial y}|_p$.

3) Se dice que un tensor (campo tensorial) de tipo $(0, 2)$ es *simétrico* si $T_{ij} = T_{ji}$ donde T_{ij} son sus componentes. Demostrar que este concepto de simetría está bien definido, es decir, que no depende de la carta empleada para calcular las componentes. Comprobar que sin embargo no se puede extender a tensores de tipo $(1, 1)$, concretamente, construir un ejemplo para el que $T_j^i = T_i^j$ se cumpla usando una carta pero no otra.

4) Clásicamente el campo electromagnético se representa por dos campos vectoriales, \vec{E} y \vec{B} , en \mathbb{R}^3 que en ausencia de cargas y corrientes están gobernados por las *ecuaciones de Maxwell*

$$\operatorname{div} \vec{E} = 0, \quad \operatorname{div} \vec{B} = 0, \quad \operatorname{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \operatorname{rot} \vec{B} = \frac{\partial \vec{E}}{\partial t}.$$

Consideramos \mathbb{R}^4 con la carta identidad $(\mathbb{R}^4, \phi = (x^1, x^2, x^3, x^4))$. Sea $*$ el endomorfismo lineal en $\operatorname{Alt}^2(T_p(\mathbb{R}^4))$ tal que $*(dx^i \wedge dx^j) \wedge dx^i \wedge dx^j = \pm dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^4$ para $i \neq j$ donde se escoge el signo $-$ si y sólo si $i = 1$ o $j = 1$. Por ejemplo, $*(dx^3 \wedge dx^4) = dx^1 \wedge dx^2$ y $*(dx^1 \wedge dx^3) = dx^2 \wedge dx^4$.

Comprobar que definiendo el *tensor de Faraday* $F = F_{ij} dx^i \wedge dx^j$ con $(F_{21}, F_{31}, F_{41}) = \vec{E}$, $(F_{34}, F_{42}, F_{23}) = \vec{B}$ y el resto de los F_{ij} nulos, las ecuaciones de Maxwell equivalen a:

$$dF = d * F = 0 \quad \text{con la notación } (t, x, y, z) = (x^1, x^2, x^3, x^4).$$

Notas e indicaciones

1) De cursos anteriores debes saber que ecuaciones adecuadas en \mathbb{R}^n definen un subconjunto con una estructura natural de subvariedad. El resultado más conocido en este sentido es que $F(\mathbf{x}) = 0$ define una subvariedad $n - 1$ dimensional de \mathbb{R}^n si $\nabla F \neq \vec{0}$ en todos sus puntos. El resultado se extiende con modificaciones a submersiones. Si los conocimientos previos no permiten dar una solución teórica de este problema, se valorará una explicación con palabras.

2) Si uno quiere proceder a partir de la definición, quizá sea conveniente cambiar a la carta en esféricas, con la que f tiene una expresión sencilla, y reescribir el resultado en cartesianas. También es posible anticipar la solución geoméricamente.

3) En cada punto fijado, un tensor de tipo $(1, 1)$ corresponde a un endomorfismo lineal. Un buen complemento al problema es pensar qué matrices simétricas lo son en cualquier base.

4) A. Einstein fue el primero en sugerir que el campo electromagnético debería representarse por un tensor en lugar de por los dos vectores clásicos. El operador $*$ es un caso especial del *operador estrella de Hodge*. Aunque aquí no sea claro, admite una definición libre de coordenadas una vez especificado un producto escalar y nuestro caso es el que corresponde al producto escalar de la relatividad especial.

Fecha límite de entrega: 22 de noviembre

Ejercicios

1) En $(\mathbb{R}^+)^3$ se consideran los campos de vectores dados por $X_1 = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + h(z) \frac{\partial}{\partial z}$ y $X_2 = \left(\frac{y}{x} + \frac{x}{y}\right) \frac{\partial}{\partial y} + \frac{z}{x} \frac{\partial}{\partial z}$. Hallar una función con $h(1) = 1$ de manera que la distribución generada por X_1 e X_2 sea completamente integrable. Comprobar que en ese caso los conos $z = K\sqrt{x^2 + y^2}$ con $K \in \mathbb{R}^+$ son superficies integrales.

2) En $M = \mathbb{R}^2 - \{(0, 0)\}$ consideramos para cada $p = (a, b) \in M$ el difeomorfismo $f_p(x, y) = (ax - 2by, bx + (a + 2b)y)$ y definimos el campo de vectores $X|_p = df_p\left(\frac{\partial}{\partial y}\Big|_{p_0}\right)$ con $p_0 = (1, 0)$.

a) Demostrar que (M, \odot) tiene estructura de grupo abeliano con la operación $p \odot q = f_p(q)$ y que su elemento neutro es p_0 .

b) Probar que el flujo local de X es $\Phi_t(p) = p \odot c_{p_0}(t)$ donde $c_{p_0}(t)$ es la curva integral de X con $c_{p_0}(0) = p_0$.

c) Hallar una fórmula explícita para c_{p_0} y para Φ_t .

3) Sea $\omega = a dx + b dy + c dz$ una 1-forma en \mathbb{R}^3 (con la carta identidad) tal que a , b y c son funciones que no se anulan simultáneamente. Probar que

$$a\left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) + b\left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\right) + c\left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) = 0$$

es una condición necesaria y suficiente para que la distribución $\Delta = \{X : \omega(X) = 0\}$ sea completamente integrable. Intentar reescribir esta condición sólo en términos de ω y $d\omega$, sin referencia a las coordenadas usadas.

4) En mecánica clásica, el movimiento unidimensional de una partícula de masa m en un campo de fuerzas de potencial V está determinado por un punto (p, q) en el plano de fases \mathbb{R}^2 con p el momento y q la posición. La ley de Newton ($F = ma$) asegura que la trayectoria de una partícula corresponde a una curva integral de $X = -V'(q) \frac{\partial}{\partial p} + \frac{p}{m} \frac{\partial}{\partial q}$ en el plano de fases. Hallar explícitamente el flujo local Φ_t en el caso $m = 1$ con $V(q) = 2q^2$ (correspondiente al movimiento armónico simple) y comprobar que para cualquier región acotada R , la integral $\int_{\Phi_t(R)} dp \wedge dq$ no depende del tiempo.

Notas e indicaciones

1) La condición de ser completamente integrales no varía al multiplicar vectores por funciones no nulas, pues la distribución generada es la misma.

Estos conos son todas las superficies integrales (maximales), aunque esto no se pide en el ejercicio.

2) La prueba de la propiedad asociativa es larga sin emplear trucos ingeniosos. Se consigue cierta reducción en los cálculos directos notando que el resultado de $(p \odot q) \odot r$ es simétrico en p y r . De esta forma, $(p \odot q) \odot r = (r \odot q) \odot p$ y la asociativa se deduce de la conmutativa.

Éste es ejemplo de grupo de Lie, un grupo con estructura de variedad, y de un campo de vectores invariante, es decir, obtenido a partir de un vector en el elemento identidad al que se le aplican todas las transformaciones del grupo. La teoría que se atisba en este problema implica que todos los campos de vectores de este tipo son completos.

3) La segunda parte es un caso particular de la formulación del teorema de Frobenius en términos de formas diferenciales, que no hemos visto en el curso. Puede ayudar buscar en la literatura esta formulación (curiosamente más próxima a la original de Frobenius). Nótese que es bastante sorprendente que esta condición con derivadas parciales sea invariante por cambios de carta.

4) La notación p y q es común en mecánica. Si causa alguna confusión escríbase x e y en su lugar, que son los nombres matemáticos habituales para las coordenadas de \mathbb{R}^2 .

La invariancia de la integral indica que el flujo Hamiltoniano preserva el área y es un hecho general para sistemas cerrados, incluso con muchas partículas, llamado teorema de Liouville. Físicamente refleja de alguna forma la homogeneidad del tiempo. Si a lo largo del tiempo Φ_t cambiase el área de R , tendríamos una manera de distinguir si ha pasado mucho o poco tiempo desde el instante inicial.

Fecha límite de entrega: 12 de enero

Ejercicios

1) Calcular las componentes de la derivada covariante del campo de vectores $V = x \frac{\partial}{\partial x} + 2x \frac{\partial}{\partial y}$ cuando la métrica es $dx^2 + x^2 dy^2$.

2) El *semiplano de Poincaré* es el semiplano complejo $\mathbb{H} = \{x+iy : x \in \mathbb{R}, y \in \mathbb{R}^+\}$ dotado de la *métrica de Poincaré* $G = y^{-2}(dx^2 + dy^2)$. Para cada $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{R})$, el grupo de matrices reales 2×2 con determinante 1, se define $g(z) = (az + b)/(cz + d)$.

a) Probar que g es una isometría, es decir, que $g : \mathbb{H} \rightarrow \mathbb{H}$ es un difeomorfismo que preserva la métrica en el sentido de que se verifica $G(\vec{v}, \vec{w}) = G(dg(\vec{v}), dg(\vec{w}))$.

b) Explicar por qué las curvas $c(t) = g(ie^{Kt})$ son geodésicas.

3) Probar que con el modelo de gravitación dado por la métrica de Schwarzschild, a lo largo de la trayectoria de una partícula la cantidad $r^4(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta)$ permanece constante.

4) Leyes físicas básicas implican que en presencia de un monopolo magnético en el origen, el movimiento de una partícula cargada viene descrito por una curva $c(t) = (x(t), y(t), z(t))$ que satisface $\ddot{c} = \|c\|^{-3}(c \times \dot{c})$ donde \times indica el producto vectorial. Probar que si la partícula parte de $(1, 0, 1)$ con velocidad $(0, 1/\sqrt{2}, 0)$, entonces $c(t)$ es una geodésica del cono $z = \sqrt{x^2 + y^2}$.

Notas e indicaciones

1) Con la notación del curso, lo que se pide es $V_{;j}^i$ para $i, j \in \{1, 2\}$.

2) En a) los cálculos se reducen drásticamente escribiendo la métrica como $(\text{Im}(z))^{-2}|dz|^2$ con $z = x + iy$ y estudiando cómo se transforman $\text{Im}(z)$ y dz .

Se dará puntuación extra por decidir razonadamente si todas las geodésicas son de la forma indicada en el segundo apartado.

3) Recuérdese que llamamos la métrica de Schwarzschild a

$$-\left(1 - \frac{r_0}{r}\right)dt^2 + \left(1 - \frac{r_0}{r}\right)^{-1}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\varphi^2.$$

4) Se sobreentiende que el cono tiene la métrica usual, es decir, $G(\vec{v}, \vec{w}) = E(di(\vec{v}), di(\vec{w}))$ donde i es la inclusión y E es la métrica euclídea $dx^2 + dy^2 + dz^2$ en \mathbb{R}^3 . Con esto no se sugiere que sea una buena elección trabajar con coordenadas cartesianas.

Se dará puntuación extra por explicar qué sucede bajo condiciones iniciales arbitrarias.

Fecha límite de entrega: 26 de enero

Ejercicios

1) Sea $T(X, Y, Z) = \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X - \nabla_{[Y, Z]} X$ donde X, Y y Z son campos de vectores en una variedad semiriemanniana M . Comprobar que $fT(X, Y, Z) = T(fX, Y, Z) = T(X, fY, Z) = T(X, Y, fZ)$ para cualquier $f : M \rightarrow \mathbb{R}$.

2) Sea $M = \mathbb{R}^+ \times (0, 1)$ con la métrica $ds^2 = dx^2 + (e^x - e^{-x})^2 dy^2$. Calcular la curvatura escalar de esta variedad riemanniana.

3) Probar que si en \mathbb{R}^2 tenemos una métrica riemanniana de la forma $A(x)dx^2 + B(y)dy^2$ entonces el tensor de Ricci es nulo.

4) Consideremos el semiplano de Poincaré $\mathbb{H} = \{x + iy : x \in \mathbb{R}, y \in \mathbb{R}^+\}$ dotado de la métrica $G = y^{-2}(dx^2 + dy^2)$. Recuerdese que la acción de $g \in \text{SL}_2(\mathbb{R})$ dada por $g(z) = (az + b)/(cz + d)$, define una isometría. En particular si d es la función distancia en \mathbb{H} se tiene $d(z, w) = d(g(z), g(w))$.

- Probar que la función $F(z, w) = 1 + \frac{|z-w|^2}{2\text{Im}(z)\text{Im}(w)}$ cumple $F(z, w) = F(g(z), g(w))$.
- Probar que si $\text{Re}(z) = \text{Re}(w) = 0$, la distancia entre z y w es $\text{arc cosh } F(z, w)$.
- Deducir que $d(z, w) = \text{arc cosh } F(z, w)$ para todo $z, w \in \mathbb{H}$.

Notas e indicaciones

1) No se puede apelar a que el tensor de curvatura es en realidad un tensor de tipo $(1, 3)$, porque en el curso hemos deducido esto último dando por supuesto este ejercicio. Lo que se espera es una demostración usando las propiedades de las conexiones.

2) Una manera de hacerlo, no necesariamente la más rápida, es calcular la curvatura de Gauss y emplear su relación con la curvatura escalar.

3) Hay una prueba brevísima usando la tensorialidad bajo un cambio de coordenadas adecuado.

4) Aquí $\operatorname{arc} \cosh x$ es la función inversa de $\cosh x = (e^x + e^{-x})/2$ en $\mathbb{R}^+ \cup \{0\}$.

Para el segundo apartado, nótese que la vertical $x = 0$, convenientemente parametrizada es una geodésica. La deducción del tercer apartado pasa por entender por qué podemos llevar siempre el problema a $x = 0$.

Éste es uno de los pocos ejemplos sencillos en que se puede dar una fórmula explícita para la distancia en una variedad riemanniana.

EXAMEN FINAL

Duración: dos horas

3 de febrero de 2012

Ejercicios

1) Sean $M = \mathbb{R}^+ \times \mathbb{R}^+$, $N = \mathbb{R}^2 - \{(0, 0)\}$ y $f : M \rightarrow N$ dada por $f(x, y) = (x + y, x - y)$. Consideremos $\omega = \frac{y}{x^2+y^2}dx - \frac{x}{x^2+y^2}dy \in \Omega^1(N)$. Hallar $\eta \in \Omega^1(M)$ tal que $\eta(\vec{v}) = \omega(df(\vec{v}))$ para cualquier vector \vec{v} del espacio tangente de M .

2) Sea Δ la distribución en $M = \mathbb{R}^3 - \{(0, 0, 0)\}$ generada por $X_1 = (y \cos z + z \sin z) \frac{\partial}{\partial x} + (z - x \cos z) \frac{\partial}{\partial y} + (-y - x \sin z) \frac{\partial}{\partial z}$ y $X_2 = (y \sin z - z \cos z) \frac{\partial}{\partial x} + (z - x \sin z) \frac{\partial}{\partial y} + (-y + x \cos z) \frac{\partial}{\partial z}$. Decidir si la esfera unidad $S^2 \subset \mathbb{R}^3$ es una subvariedad integral para Δ .

3) Hallar las ecuaciones diferenciales de las geodésicas para el paraboloido $z = x^2 + y^2$ entendido como subvariedad de \mathbb{R}^3 con la métrica inducida.

4) Recuérdese que el tensor de Riemann tiene las simetrías $R_{ijkl} = -R_{ijlk} = -R_{jikl} = R_{klij}$, donde $R_{ijkl} = g_{in} R_{jkl}^n$. Probar que el tensor de Ricci $R_{ij} = R_{ikj}^k$ de cualquier métrica en \mathbb{R}^2 de la forma $A(x, y)dx^2 + B(x, y)dy^2$ cumple $R_{12} = R_{21} = 0$.

Bibliografía

- [Arn78] V. I. Arnold. *Mathematical methods of classical mechanics*. Springer-Verlag, New York, 1978. Translated from the Russian by K. Vogtmann and A. Weinstein, Graduate Texts in Mathematics, 60.
- [Aub01] T. Aubin. *A course in differential geometry*, volume 27 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [AW01] G. B. Arfken and H. J. Weber. *Mathematical methods for physicists*. Harcourt/Academic Press, Burlington, MA, fifth edition, 2001.
- [BC70] F. Brickell and R. S. Clark. *Differentiable manifolds. An introduction*. Van Nostrand Reinhold Company, London-New York-Toronto, Ont., 1970. The New University Mathematics Series.
- [BE81] J. K. Beem and P. E. Ehrlich. *Global Lorentzian geometry*, volume 67 of *Monographs and Textbooks in Pure and Applied Math*. Marcel Dekker Inc., New York, 1981.
- [Ber03] M. Berger. *A panoramic view of Riemannian geometry*. Springer-Verlag, Berlin, 2003.
- [BG80] R. L. Bishop and S. I. Goldberg. *Tensor analysis on manifolds*. Dover Publications Inc., New York, 1980. Corrected reprint of the 1968 original.
- [Boo75] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, No. 63.
- [Cha08] F. Chamizo. Geometría IV (tensores, formas, curvatura, relatividad y todo eso). <http://www.uam.es/fernando.chamizo/libreria/fich/apgeomiv08.pdf>, 2008.
- [dC92] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

- [DNF85] B. Doubrovine, S. Novikov, and A. Fomenko. *Géométrie contemporaine. Méthodes et applications. 1^{re} partie*. Traduit du Russe: Mathématiques. [Translations of Russian Works: Mathematics]. “Mir”, Moscow, 1985. Géométrie des surfaces, des groupes de transformations et des champs. [Geometry of surfaces, transformation groups and fields], Translated from the Russian by Vladimir Kotliar, Reprint of the 1982 translation.
- [Ein39] A. Einstein. On a stationary system with spherical symmetry consisting of many gravitating masses. *Ann. of Math. (2)*, 40:922–936, 1939.
- [FLS64] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman lectures on physics. Vol. 2: Mainly electromagnetism and matter*. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1964.
- [FN79] J. Foster and J. D. Nightingale. *A short course in general relativity*. Longman, London, 1979. Longman Mathematical Texts.
- [Fra12] T. Frankel. *The geometry of physics*. Cambridge University Press, Cambridge, third edition, 2012. An introduction.
- [GHL87] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, 1987.
- [Gon05] J. Gonzalo. *Varietades y Geometría: un curso breve*, volume 64 of *Documentos de Trabajo*. Ediciones de la Universidad Autónoma de Madrid, Madrid, 2005.
- [Gua08] M. Gualtieri. Geometry and Topology (1300Y). <http://www.math.toronto.edu/mgualt/MAT1300/1300%20Lecture%20notes.pdf>, 2008.
- [HE73] S. W. Hawking and G. F. R. Ellis. *The large scale structure of space-time*. Cambridge University Press, London, 1973. Cambridge Monographs on Mathematical Physics, No. 1.
- [Hir76] M. W. Hirsch. *Differential topology*. Springer-Verlag, New York, 1976. Graduate Texts in Mathematics, No. 33.
- [HT90] L. P. Hughston and K. P. Tod. *An introduction to general relativity*, volume 5 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1990.
- [Jän01] K. Jänich. *Vector analysis*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2001. Translated from the second German (1993) edition by Leslie Kay.

- [Lee09] Jeffrey M. Lee. *Manifolds and differential geometry*, volume 107 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2009.
- [LEMWed] H. A. Lorentz, A. Einstein, H. Minkowski, and H. Weyl. *The principle of relativity*. A collection of original memoirs on the special and general theory of relativity. Dover Publications Inc., New York, N. Y., undated. With notes by A. Sommerfeld, Translated by W. Perrett and G. B. Jeffery.
- [LL76] L. D. Landau and E. M. Lifshitz. *Course of theoretical physics. Vol. 1*. Pergamon Press, Oxford, third edition, 1976. Mechanics, Translated from the Russian by J. B. Skyes and J. S. Bell.
- [Lun92] A. T. Lundell. A short proof of the Frobenius theorem. *Proc. Amer. Math. Soc.*, 116(4):1131–1133, 1992.
- [MTW73] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman and Co., San Francisco, Calif., 1973.
- [Mun75] J. R. Munkres. *Topology: a first course*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.
- [Nab88] G. L. Naber. *Spacetime and singularities*, volume 11 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1988. An introduction.
- [NT06] S. P. Novikov and I. A. Taimanov. *Modern geometric structures and fields*, volume 71 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006. Translated from the 2005 Russian original by Dimitry Chibisov.
- [O’N83] B. O’Neill. *Semi-Riemannian geometry*, volume 103 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983. With applications to relativity.
- [Sch85] B. F. Schutz. *A First Course in General Relativity*. Cambridge University Press, Cambridge, 1985.
- [Spi79] M. Spivak. *A comprehensive introduction to differential geometry. Vol. II*. Publish or Perish Inc., Wilmington, Del., second edition, 1979.
- [Spi82] M. Spivak. *Cálculo en variedades*. Editorial Reverté, 1982.
- [Wal84] R. M. Wald. *General relativity*. University of Chicago Press, Chicago, IL, 1984.

- [Wal04] G. Walschap. *Metric structures in differential geometry*, volume 224 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2004.
- [War83] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1983. Corrected reprint of the 1971 edition.
- [Whi36] H. Whitney. Differentiable manifolds. *Ann. of Math. (2)*, 37(3):645–680, 1936.

Índice alfabético

- acción, 38
- aceleración de Coriolis, 51
- afelio, 49
- agujero negro, 47
- álgebra de Lie, 25
- álgebra exterior, 16
- anticonmutativa, propiedad, 18
- aplicación de Gauss, 61
- aplicación exponencial, 57
- aplicación tangente, 6
- Arnold, V.I. (1937–2010), 29
- atlas maximal, 3

- base local, 31
- Big Bang, 84
- Big Crunch, 84
- bola de pelo, teorema de la, 31
- Bonnet-Myers, teorema de, 76

- campo conservativo, 24
- campo de Jacobi, 74
- campo de vectores, 12
 - completo, 24
- campo tensorial, 13
- campo variacional, 75
- carta, 2
- Cartan, É. (1869–1951), 15, 18, 55
- cartas compatibles, 3
- Christoffel, símbolos de, 54
- componentes de un tensor, 10
- conexión afín, 52
- conexión de Levi-Civita, 53
- conmutador, 27

- constante de gravitación universal, 46, 47, 79
- contracción, 12
- contravector, 10
- convenio de sumación de Einstein, 11
- coordenadas generalizadas, 38
- corchete de Lie, 27
- covector, 13
- curva integral, 24
 - maximal, 24
- curvatura
 - de Gauss, 61, 66
 - de una curva plana, 61
 - escalar, 68
 - seccional, 66
 - tensor de, 64

- derivación, 4
- derivada absoluta, 51, 53
- derivada covariante, 36, 53
- derivada de Lie, 29
- derivada exterior, 19
- derivada parcial, 3
- difeomorfismo, 4
- distancia riemanniana, 69
- distribución, 30
 - completamente integrable, 31
 - involutiva, 32

- ecuación de Jacobi, 74
- ecuación de ondas, 42
- ecuación de Poisson, 79
- ecuaciones de campo, 80
- ecuaciones de Maxwell, 42

Einstein, A. (1879–1955), 36, 41, 44
 Einstein, tensor de, 80
 energía, 75
 esferas exóticas, 3
 espacio cotangente, 13
 espacio de configuración, 38
 espacio de Minkowski, 44
 espacio proyectivo, 1
 espacio tangente, 4
 espacio-tiempo, 41
 estructura diferenciable, 3
 Euler-Lagrange, ecuaciones de, 37

 fibrado tangente, 12
 FitzGerald, G. (1851–1901), 44
 flujo local, 25
 forma alternada, 16
 forma diferencial, 16
 Frobenius, teorema de, 32
 funciones coordenadas, 3

 Gauss, C.F.. (1777–1855), 61
 Gauss-Bonnet, teorema de, 62
 geodésica, 36, 40
 minimizante, 70
 nula, 45
 temporal, 45
 gradiente, 15
 grados de libertad, 38
 Grassmann, H. (1809–1877), 16, 18
 Green, teorema de, 22
 grupo de Lie, 25
 grupo uniparamétrico, 25

 Hadamard, teorema de, 78
 Hilbert, D. (1862–1943), 81
 hipótesis cosmológica, 83
 Hopf-Rinow, teorema de, 70
 horizonte de sucesos, 47

 imagen recíproca, 21
 inmersión, 7

 isometría, 61

 Kepler, J. (1571–1630), 49

 lagrangiano, 36
 Laplace, P.S. (1749–1827), 48
 Lema de Gauss, 58
 Lema de Sard, 8
 Levi-Civita, T. (1873–1941), 52
 Lie, S. (1842–1899), 25
 longitud de una curva, 59
 Lorentz, H.A. (1853–1928), 44
 Lorentz, variedad de, 45

 Maxwell, J.C. (1831–1979), 42
 métrica, 36, 39
 de Minkowski, 43
 de Robertson-Walker, 84
 de Schwarzschild, 46, 82
 inducida, 54
 usual, 39, 54
 Milnor, J. (1931–), 3
 Minkowski, H. (1864–1909), 41
 multilineal, aplicación, 9

 Newton, I. (1642–1727), 45, 49

 orientación, 21
 inducida, 22

 particiones de la unidad, 6
 perihelio, 49
 Poincaré, H. (1854–1912), 44
 primera variación de la energía, 75
 principio de Hamilton, 38
 principio de inercia, 36, 39, 40, 55
 principio de mínima acción, 38
 producto exterior, 17, 18
 producto tensorial, 39
 propiedad de Hausdorff, 2
 pullback, 21
 puntos conjugados, 74

 radio de inyectividad, 57

radio de Schwarzschild, 46
Ricci, tensor de, 67
Ricci-Curbastro, G. (1853–1925), 67
Riemann, B. (1826–1866), 36
Riemann, tensor de, *véase* curvatura, tensor de

Schwarzschild, K. (1873–1916), 46
segunda variación de la energía, 76
segundo axioma de numerabilidad, 2
sistema normal de coordenadas, 55
Stokes, teorema de, 22
subvariedad, 31
subvariedad integral, 31
superconmutativa, propiedad, 18

tensor, 9, 13
tensor de energía-momento, 80, 84
tensor simétrico, 86
tensorialidad, 14
teorema de la esfera, 78
Teorema Egregio, 61
tiempo propio, 45
tipo de un tensor, 9
transformaciones de Lorentz, 44
transporte paralelo, 53
triángulo geodésico, 62

uno forma, 13, 16

variación, 72
 geodésica, 72
 propia, 75
variedad
 con borde, 21
 diferenciable, 3
 pseudoriemanniana, *véase* semiriemanniana
 riemanniana, 39
 semiriemanniana, 39
vector tangente, 4

Whitney, H. (1907–1989), 1

Whitney, teorema de, 7