



UNIVERSIDAD
POLITÉCNICA
DE MADRID



EPES
Escuela Politécnica
de Enseñanza Superior



UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA POLITÉCNICA DE ENSEÑANZA SUPERIOR (EPES)

TRABAJO FIN DE GRADO

GRADO EN MATEMÁTICAS

Fundamentos geométricos de la Relatividad General

Autor: Carlos Gómez Redondo

Dirigido por: M.^a Eugenia Rosado María
Alberto Navarro Garmendia

Madrid, 4 de julio de 2025



Fundamentos geométricos de la Relatividad General

Autor: Carlos Gómez Redondo

Dirigido por: M.^a Eugenia Rosado María
Alberto Navarro Garmendia

Trabajo Fin de Grado, 4 de julio de 2025

Escuela Politécnica de Enseñanza Superior (EPES)

Edificio C, Paseo Juan XXIII, II. Madrid

Si deseas citar este trabajo, la entrada completa en $\text{BIB}\text{T}\text{E}\text{X}$ es la siguiente:

```
@mastersthesis{citekey,  
title = {Fundamentos geométricos de la Relatividad General},  
author = {Gómez Redondo, Carlos y Rosado, M.a Eugenia y Navarro, Alberto},  
school = {Escuela Politécnica de Enseñanza Superior (EPES)},  
year = {2025},  
type = {Trabajo Fin de Grado}  
}
```

Esta obra está bajo una licencia [Creative Commons](https://creativecommons.org/licenses/by-nc-sa/4.0/) «Atribución-NoComercial-CompartirIgual 4.0 Internacional».

Plantilla derivada de <https://github.com/blazaid/UPM-Report-Template>.



Agradecimientos

Quiero empezar agradeciendo a mi familia y amigos, en particular a mis padres, por acompañarme y apoyarme siempre. Muy especialmente en los momentos más difíciles.

Quiero destacar mi profundo gratitud a mis tutores, Alberto Navarro y María Eugenia Rosado, por su dedicación y su guía constante durante la realización de este trabajo. Extiendo mi agradecimiento a todos los profesores del grado, por haber contribuido con sus enseñanzas a mi formación a lo largo de estos años. Sin el apoyo y compromiso de todos ellos, este trabajo no habría sido posible.

Resumen

Este trabajo explora los fundamentos geométricos sobre los que se construye la Teoría de la Relatividad. En primer lugar, comenzamos con un bloque en el que se presenta toda la geometría diferencial necesaria. Empezamos explorando el concepto de conexión y su relación con el transporte paralelo. Posteriormente se construye la generalización de la conexión para formas diferenciables y campos tensoriales, demostrando a su vez la unicidad de dicha extensión. Visto esto, obtenemos el tensor de curvatura a partir del concepto de holonomía, un camino diferente al de la literatura clásica, estudiando, además, su interpretación geométrica. A continuación se encuentra el bloque de la Teoría de la Relatividad, donde se proponen los postulados de la relatividad desde un enfoque matemático. A partir de esto, se demuestra cómo las únicas estructuras del espacio-tiempo compatibles con dichos postulados es la newtoniana y la lorentziana. Abordamos también los aspectos cinemáticos de la relatividad especial, así como algunos principios de conservación. Finalmente, se integran todos estos elementos para obtener la ecuación de campo de Einstein. Otros desarrollos a los que prestamos especial atención y que completan la literatura son la interpretación geométrica de la torsión, la caracterización de estructuras métricas compatibles con los postulados relativistas, la definición de la conexión a partir del transporte paralelo y la ley de conservación infinitesimal del impulso para el caso lorentziano.

Palabras clave: Espacio-tiempo, impulso, curvatura, tensor de Riemann, tensor de Einstein, conexión, y ecuación de Einstein.

Abstract

This project explores the geometric foundations underlying the Theory of Relativity. It begins with a section dedicated to developing the necessary tools from differential geometry. The concept of connection is introduced first, along with its relationship to parallel transport. The connection is then extended to differential forms and tensor fields, and the uniqueness of this extension is established. With these tools in hand, the curvature tensor is derived from the notion of holonomy, following a different approach from classical literature, and its geometric meaning is analyzed.

The second part focuses on the Theory of Relativity. The postulates of relativity are formulated from a mathematical perspective, and it is shown that the only space-time structures compatible with these postulates are the Newtonian and Lorentzian models. The discussion then turns to the mechanics of special relativity and several conservation principles. All these elements are eventually brought together to derive Einstein's field equation.

Additional developments that receive special attention and complement the existing literature include the geometric interpretation of torsion, the classification of metric structures compatible with the relativistic postulates, the definition of connection via parallel transport, and the infinitesimal conservation law of impulse in the Lorentzian case.

Keywords: space-time, impulse, curvature, Riemann tensor, Einstein tensor, connection, and Einstein's equation.

Índice general

1	Introducción	I
2	Preliminares	3
3	Conexión, transporte paralelo y geodésicas	5
3.1	Conexión en variedades abstractas	6
3.2	Equivalencia entre transporte paralelo y conexión	16
3.3	Conexión sobre formas y tensores	21
3.4	Geodésicas	27
4	Tensor de Curvatura	37
4.1	Curvatura como "holonomía puntual"	37
4.2	Curvatura a partir de las ecuaciones de estructura de Cartan	44
4.3	Tensor de Ricci	47
5	Postulados de la Relatividad	51
5.1	Caracterización de las estructuras sobre el espacio-tiempo	56
5.2	Variedad de métricas del espacio-tiempo	62
6	Espacio-tiempo de Minkowski	71
6.1	Cinemática	74
6.2	Masa y energía: Dos caras de la misma moneda	76
6.3	Fenómenos relativistas	81
7	Tensor de Materia	89
7.1	3-forma de masa	90
7.2	3-forma de impulso	91
7.3	Tensor de materia	93
7.4	Conservación del impulso	99
7.5	Formas de Killing y leyes de conservación	103
7.6	Movimiento con fuerzas externas	109
8	Relatividad General	113

8.1	Mecánica	114
8.2	Ecuación de Einstein	115
A	Apéndice	123
A.1	Conexión transporte paralelo y geodésicas	123
A.2	Apéndice: Tensor de Curvatura	131
A.3	Espacio-tiempo de Minkowski	137
A.4	Relatividad General	140
	Bibliografía	143
	Índice alfabético	145

1

Introducción

Una característica fundamental de la naturaleza humana es su obsesión por entender el mundo que lo rodea, desarrollando para ello teorías que explican todo tipo de fenómenos físicos. Todas estas construcciones se basan implícitamente en una idea tan simple como que dichos sucesos ocurren en un lugar y tiempo determinados. Sin embargo, pese a ser capaces de explicar muchos de estos fenómenos, preguntas como ¿qué es el espacio? o ¿qué es el tiempo? aún no tienen respuesta; lo más cercano que tenemos hasta el momento es precisamente la Teoría de la Relatividad. Este marco teórico nos permitió pasar de la descripción newtoniana en la que espacio y tiempo son el escenario donde ocurren todo lo que observamos, a otra en la que forman parte de los propios sucesos. En palabras de Wheeler: “la materia le dice al espacio-tiempo cómo curvarse, y el espacio-tiempo le dice a la materia cómo moverse”. En este contexto, el objetivo de este trabajo es estudiar la geometría necesaria para poder explorar la Teoría de la Relatividad.

El trabajo está organizado en dos bloques diferenciados por su naturaleza. El primero está dedicado al desarrollo de los aspectos puramente geométricos, que constituyen la base matemática necesaria. El segundo bloque aborda los elementos propios de la relatividad, aplicando los conceptos geométricos previamente establecidos al marco físico de la teoría.

En el bloque de la geometría comenzaremos presentando el concepto de conexión. A partir de esta nueva herramienta definiremos el transporte paralelo y extenderemos la idea de derivación para formas y tensores. En el segundo capítulo del bloque continuaremos con el estudio de la curvatura, construyendo el famoso tensor de curvatura de dos maneras diferentes y entendiendo el porqué de que esta es una extensión de la curvatura para superficies. Veremos también el tensor de Ricci y su interpretación geométrica.

En cuanto al bloque de la Relatividad, comenzaremos formulando sus postulados desde una perspectiva matemática y demostrando que sólo existen dos estructuras compatibles con estos: la newtoniana y la relativista. Posteriormente, en el capítulo 6 pasaremos al estudio de la relatividad especial, en la que formularemos la cinemática y veremos las consecuencias fundamentales de esta nueva perspectiva; desde la contracción espacial y dilatación temporal, hasta la equivalencia de masa y energía. En

el capítulo 7 desarrollaremos el lenguaje necesario para abordar problemas con una distribución de materia continua, presentando objetos fundamentales como la forma de impulso y el tensor de materia. Concluiremos este capítulo obteniendo algunas leyes de conservación e introduciendo para el caso de un fluido perfecto las ecuaciones de la dinámica. Acabaremos el trabajo empleando todo lo estudiado para demostrar la ecuación de campo de Einstein.

A lo largo del documento encontramos ciertos desarrollos que merecen ser destacados. En primer lugar, la sección 3.2 en la que exploramos de manera diferente a la habitual la relación entre el transporte paralelo y la conexión. Trataremos de establecer las propiedades básicas que debe verificar el transporte paralelo para inducir una conexión. Veremos que será necesario imponer que el transporte a lo largo de lazos infinitesimales debe converger a la identidad, una propiedad que no es inmediata, y que podría ser objeto de estudio en otro trabajo. En esa sección, encontramos más contenido que complementa a la literatura clásica en la interpretación geométrica de la torsión nula en una conexión. En la sección del 3.3 encontramos la extensión del concepto de conexión para formas y campos tensoriales tratando de justificar el porqué de las definiciones. Para el caso de las formas, veremos que podemos entenderlo como la aplicación dual de la conexión para campos modificada con un término para que verifique la regla de Leibniz. En el caso de los tensores, proponemos dos definiciones: la primera imponiendo la regla de Leibniz y en la segunda la del producto. Concluiremos que ambas son equivalentes y además demostraremos que sólo existe una extensión de la conexión para tensores que cumple la regla de Leibniz. Otra parte a destacar es la construcción del tensor de curvatura a partir del concepto de holonomía. Veremos cómo el tensor de Riemann se puede obtener tomando la diferencia entre un vector y el transportado alrededor de un lazo infinitesimalmente pequeño. Esto nos revelará el papel fundamental que juega el álgebra de los campos tangentes; viendo cómo surge la curvatura a partir de fundamentos algebraicos.

En cuanto a la parte de relatividad, empezamos destacando la sección 5.2. En ella demostramos que las estructuras del espacio-tiempo compatibles con los postulados definen una variedad con borde cuyo interior está formado por las métricas lorentzianas y el borde por las newtonianas. Demostrando así, no sólo que lo más probable es vivir en un universo relativista, sino que en caso de vivir en uno newtoniano no podríamos demostrarlo por medio del experimento de Michelson-Morley. En el capítulo 6 nos gustaría destacar el teorema 6.21, en el que se extiende la Segunda Ley de Newton para el caso relativista. En el capítulo 7 encontramos especialmente relevante la sección 7.4.2 en la que estudiamos la conservación del impulso para el caso lorentziano, complementando lo presentado en [NS04] que lo hace sólo para el newtoniano. Finalmente, hay que resaltar el resultado más importante del trabajo, el teorema 8.6 correspondiente a la ecuación de campo de Einstein.

2

Preliminares

En esta sección recordamos las definiciones básicas que utilizamos a lo largo del trabajo. Las referencias básicas empleadas para los contenidos de geometría diferencial son [Küh06, BG80, KoNo63]. En cuanto a la parte de relatividad, hemos seguido [NS17, Caro4, HE94, SW77]. Como [NS17] no está publicado, se puede consultar la versión antigua [NS04]

A lo largo del trabajo consideramos M *variedad diferenciable* C^∞ de dimensión finita m . Diremos que M es una variedad euclídea si $M \subset \mathbb{R}^n$. En caso de ser M una variedad con borde, a la variedad borde la denotaremos ∂M .

Sea $T_p M$ el *espacio tangente* a M en un punto $p \in M$ y denotaremos indistintamente $\partial_i, \frac{\partial}{\partial x_i}$ o $e_i, 1 \leq i \leq m$, los elementos de la base coordenada en función del contexto. Sea $T_p^* M$ el *espacio cotangente* a M en un punto $p \in M$ y $(\theta^i)_{i=1, \dots, m}$, la base dual de $(e_i)_{i=1, \dots, m}$. Sean $TM = \bigcup_{p \in M} T_p M$ y $T^* M = \bigcup_{p \in M} T_p^* M$ los *fibrados tangente* y *cotangente* de M . Sea $\mathfrak{X}(M)$ al álgebra de campos vectoriales en una variedad. Denotaremos Φ_t^X al flujo de un campo $X \in \mathfrak{X}(M)$. Diremos que M es *orientable* si admite una n -forma $\omega \in \Lambda^n M$ no nula, llamada *forma de volumen*.

Sea $T_p^p(M) = \bigotimes^p TM \otimes \bigotimes^q T^* M$ el fibrado de *campos tensoriales* de tipo (p, q) . Denotaremos la *contracción* del i -ésimo índice contravariante con el j -ésimo índice covariante como C_j^i . Llamamos *formas diferenciales* ω a las secciones alternadas de tensores de tipo $(0, p)$; el conjunto de p -formas diferenciales se escribirá como $\Lambda^p M$ y el álgebra de formas diferenciales como ΛM . Consideramos el *producto exterior* \wedge definido a partir de la operación de *antisimetrización* $\text{Alt}(\alpha)(v_1, \dots, v_r) = \frac{1}{r!} \sum_{\sigma \in S_r} \text{sign}(\sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(r)})$. Además, tenemos la *diferencial exterior* $d: \Lambda^k M \rightarrow \Lambda^{k+1} M$, y la *contracción interior* ι_X respecto a un campo vectorial $X \in \mathfrak{X}(M)$.

Denotaremos por G a un grupo de Lie y \mathfrak{g} a su álgebra de Lie. Escribiremos el *conmutador* y la *aplicación exponencial* como $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ y $\exp: \mathfrak{g} \rightarrow G$, respectivamente.

Trabajaremos también con *variedades métricas* (M, g) . Diremos que es *riemanniana* si g es definida positiva en cada $T_p M$, y *lorentziana* si g tiene signatura $(+, -, \dots, -)$. Si g es no degenerada, permite definir un tensor 2-contravariante simétrico llamado *métrica contravariante* h asociada a g que satisface $g(X, Y) = h(\alpha_X, \alpha_Y)$, $X, Y \in \mathfrak{X}(M)$, donde $\alpha_X(Z) = g(X, Z)$.

3

Conexión, transporte paralelo y geodésicas

En este capítulo analizamos la idea de conexión, concepto que surge de la necesidad de transportar vectores tangentes sobre la variedad de manera coherente. A diferencia de la geometría afín, en el caso de variedades, el espacio tangente en cada punto no es difeomorfo a la propia variedad. Por lo que no podemos identificar los espacios tangentes en cada punto como sí sucede en \mathbb{R}^m y, por tanto, no podemos definir operaciones como la derivación de manera inmediata. Es decir, en variedades necesitamos definir lo que se conoce como transporte paralelo mediante la conexión.

Con esta idea fundamental en mente, empezaremos presentando la conexión con su definición algebraica y, mediante ejemplos, veremos que se trata de un concepto intrínseco que generaliza la derivada direccional. También veremos las condiciones que nos permiten definir la conexión de manera única: la conocida como conexión de Levi-Civita. Al final del capítulo, estudiaremos las curvas geodésicas: unas curvas naturales dentro de la variedad que juegan un papel central en el estudio de la geometría intrínseca de la variedad y que son fundamentales a la hora de entender la relatividad.

De este capítulo caben resaltar ciertas partes que divergen del planteamiento usual. En primer lugar, el análisis de la relación entre el transporte paralelo y la conexión. En la sección 3.2 se estudia cómo construir una conexión a partir del transporte paralelo; al contrario de lo que suele ser habitual. En segundo lugar, se presenta la conexión para formas a partir de la aplicación dual. En la generalización para el caso de tensores en la sección 3.3.2, se presentan de dos maneras diferentes en función de si imponemos la regla de Leibniz o la del producto. Además, demostramos que ambas son equivalentes y que conmutan con todas las contracciones. A partir de esa idea, vemos que existe una única generalización para tensores verificando las propiedades anteriores. Finalmente, es importante mencionar que esta visión nos permitirá dar una interpretación clara tanto de la torsión como de la compatibilidad con la métrica que caracterizan la conexión de Levi-Civita.

3.1. Conexión en variedades abstractas

Empecemos con la definición más general de conexión.

Definición 3.1. Una *conexión* en una variedad M es una aplicación

$$\begin{aligned}\nabla: \mathfrak{X}(M) \times \mathfrak{X}(M) &\longrightarrow \mathfrak{X}(M), \\ (X, Y) &\longmapsto \nabla_X Y,\end{aligned}$$

que satisface las siguientes condiciones:

- 1) $\nabla_{(X_1+X_2)} Y = \nabla_{X_1} Y + \nabla_{X_2} Y, \quad \forall X_1, X_2, Y \in \mathfrak{X}(M),$
- 2) $\nabla_{fX} Y = f(\nabla_X Y), \quad \forall f \in C^\infty(M), \forall X, Y \in \mathfrak{X}(M),$
- 3) $\nabla_X (Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2, \quad \forall X, Y_1, Y_2 \in \mathfrak{X}(M),$
- 4) $\nabla_X (fY) = X(f)Y + f(\nabla_X Y), \quad \forall f \in C^\infty(M), \forall X, Y \in \mathfrak{X}(M).$

Nótese que las condiciones 1) y 2) en la definición 3.1 garantizan la linealidad en el primer argumento y la condición 4) es análoga a la regla de Leibniz. Con esto podemos empezar a ver que efectivamente se trata de una generalización de la derivada sobre variedades abstractas como ya adelantamos. Además, se pueden imponer condiciones extra como

$$\nabla_X Y - \nabla_Y X - [X, Y] = 0, \quad \forall X, Y \in \mathfrak{X}(M), \quad (3.1)$$

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \quad \forall X, Y, Z \in \mathfrak{X}(M). \quad (3.2)$$

Si verifica la condición (3.1) se dice que es una conexión sin torsión; al final de la sección ?? veremos la interpretación geométrica de esta condición y entenderemos el porqué de su nombre.

Definición 3.2. Llamamos *torsión* de la conexión ∇ al tensor Θ de tipo (1, 2) definido por

$$\Theta(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y], \quad \forall X, Y \in \mathfrak{X}(M). \quad (3.3)$$

La condición (3.2) se aplica para variedades equipadas con una métrica g . En los casos concretos en los que g es una métrica riemanniana o lorentziana decimos que es una conexión riemanniana en el primer caso y lorentziana en el segundo. Veremos en el teorema 3.10 que estas condiciones son las que nos garantizan que, dada una variedad, con una métrica no degenerada, induce una única conexión.

La definición que acabamos de ver es puramente algebraica y resulta muy útil a la hora de operar; sin embargo, no nos permite tener algo de intuición. Por ello, con el objetivo de tener una idea geométrica más allá de la algebraica empecemos por los casos sencillos del propio \mathbb{R}^m y variedades euclídeas donde podemos definir la que se conoce como conexión estándar en \mathbb{R}^m .

Ejemplo 3.3 (Conexión en \mathbb{R}^m). Sea $Y = \sum_{j=1}^m f_j \frac{\partial}{\partial x_j}$ un campo tangente en \mathbb{R}^m . Llamamos *conexión estándar* en \mathbb{R}^m , a la aplicación $\nabla: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ tal que

$$\nabla_X Y = \sum_{j=1}^m X(f_j) \frac{\partial}{\partial x^j}, \quad \forall X \in \mathfrak{X}(\mathbb{R}^m).$$

Es fácil verificar que efectivamente se trata de una conexión. Así, podemos ver que la conexión estándar en realidad no es otra cosa que una derivada direccional del campo Y siguiendo las curvas integrales del campo X .

Ejemplo 3.4 (Conexión en variedades euclídeas). Sea N una variedad euclídea \mathbb{R}^m y sea ∇ la conexión estándar en \mathbb{R}^m . Dados $X, Y \in \mathfrak{X}(N)$ dos campos tangentes a N , $(\nabla_X Y)_p$ está definido para cada $p \in N$. Denotamos $(\bar{\nabla}_X Y)_p$ su componente tangencial y $\alpha_p(X, Y)$ su componente normal, de manera que se tiene:

$$(\nabla_X Y)_p = (\bar{\nabla}_X Y)_p + \alpha_p(X, Y), \quad X, Y \in \mathfrak{X}(N), \quad p \in S, \quad (3.4)$$

con $(\bar{\nabla}_X Y)_p \in T_p N$ y $\alpha_p(X, Y) \in (T_p N)^\perp$, donde el símbolo $^\perp$ denota la proyección sobre el subespacio ortogonal a $T_p S$.

Expresándolo de esta forma se ve que $(\nabla_X Y)_p$ no tiene por qué ser un vector tangente a la variedad N . Por tanto, para definir una conexión en N es necesario tomar la parte tangente de dicho vector. La métrica en \mathbb{R}^m permite descomponer los vectores de \mathbb{R}^m en la componente tangencial y en la normal¹.

Proposición 3.5. *La aplicación $\bar{\nabla}$, dada en (3.4) es una conexión en N .*

Demostración. Las condiciones 1), 2) y 3) en la definición (3.1) son inmediatas teniendo en cuenta que ∇ es una conexión y la linealidad de la proyección $T_p M \rightarrow T_p N$. Veamos que la condición 4) en (3.1) también se satisface. Sean $f \in C^\infty(N)$, $X, Y \in \mathfrak{X}(M)$, entonces

$$\nabla_X(fY) = X(f)Y + f\nabla_X Y,$$

donde $X(f)Y \in \mathfrak{X}(N)$. Por tanto, tomando la componente tangencial en ambos lados de la igualdad anterior tenemos:

$$\bar{\nabla}_X(fY) = X(f)Y + f\bar{\nabla}_X Y,$$

con lo que se concluye. □

¹Nótese por tanto que dada una métrica sobre el espacio ambiente o sobre la propia variedad euclídea induce una única conexión, salvo que la métrica sea degenerada. En el caso de una variedad abstracta es necesario añadir la condición de torsión nula como comentamos.

Observación. Reescribiendo la definición de $\bar{\nabla}$ en términos de un límite de nuevo vemos que se trata de una generalización de la derivada direccional en N . Sea $c: I \rightarrow S$ una curva integral de $X \in \mathfrak{X}(S)$, con $c(t_0) = p$. Se tiene

$$(\bar{\nabla}_X Y)(p) = (Y \circ c)'(t_0)^T = \left[\lim_{t \rightarrow 0} \frac{Y(c(t_0 + t)) - Y(c(t_0))}{t} \right]^T \in T_p S,$$

donde el superíndice T indica la componente tangencial.

Una vez tenemos ya una cierta intuición geométrica de qué representa una conexión, volvamos al caso general. Calculemos en la expresión en coordenadas de la conexión. Sean dos campos vectoriales $X = \sum_i c^i \frac{\partial}{\partial x^i}$ y $Y = \sum_j \eta^j \frac{\partial}{\partial x^j}$ en M , la conexión en coordenadas locales viene dada por

$$\nabla_X Y = \nabla_{\sum_i c^i \frac{\partial}{\partial x^i}} \sum_j \eta^j \frac{\partial}{\partial x^j} = \sum_i c^i \sum_j \nabla_{\frac{\partial}{\partial x^i}} (\eta^j \frac{\partial}{\partial x^j}) = \sum_{i,j} (c^i \frac{\partial \eta^j}{\partial x^i} \frac{\partial}{\partial x^j} + c^i \eta^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}).$$

Ahora por la definición de conexión 3.1, sabemos que su imagen debe ser un nuevo campo tangente por tanto podemos escribir que

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} := \sum_k \Gamma_{ij}^k \frac{\partial}{\partial x^k},$$

donde $\Gamma_{ij}^k \in C^\infty(M)$. Así, tenemos que

$$\nabla_X Y = \sum_k \left(\sum_i \xi^i \frac{\partial \eta^k}{\partial x^i} + \sum_{i,j} \xi^i \eta^j \Gamma_{ij}^k \right) \frac{\partial}{\partial x^k}. \quad (3.5)$$

Nótese que por la propiedad de linealidad en el primer argumento en realidad no hace falta escribir la fórmula anterior para campos X generales. Basta con escribirlo para campos de la forma $X = \frac{\partial}{\partial x^i}$. Así

$$\nabla_X Y = \nabla_{\frac{\partial}{\partial x^i}} \left(\sum_j \eta^j \frac{\partial}{\partial x^j} \right) = \sum_k \left(\frac{\partial \eta^k}{\partial x^i} + \sum_j \Gamma_{ij}^k \eta^j \right) \frac{\partial}{\partial x^k}.$$

De esta forma se puede ver que definir una conexión sobre una variedad consiste en dar las funciones diferenciables $\{\Gamma_{ij}^k\}_{i,j,k=1}^n$ correspondientes a los campos $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$.

Definición 3.6. Llamamos *símbolos de Christoffel* a las funciones diferenciables sobre M

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \sum_k \Gamma_{ij}^k \frac{\partial}{\partial x^k}. \quad (3.6)$$

Observación. Nótese que todo este razonamiento es local, por tanto los símbolos de Christoffel se definen para cada entorno coordenado. Debido a que son funciones diferenciables y los cambios de carta también, los símbolos definidos en dos entornos con intersección no nula son compatibles.

Proposición 3.7. Sea (M, ∇) una variedad diferenciable equipada con una conexión sin torsión, entonces $\Gamma_{ij}^k = \Gamma_{ji}^k$.

Demostración. Aplicando la definición de torsión nula para campos coordenados:

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} = - \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right] = 0 \implies \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i}.$$

□

Si nos restringimos a una única curva en el primer argumento de la conexión obtenemos lo que se conoce como derivada covariante.

Definición 3.8. Sea una curva $c : I \rightarrow M$ diferenciable. Sabemos que existe $X \in \mathfrak{X}(M)$ un campo tangente a la variedad tal que c es una curva integral. Entonces llamamos *derivada covariante* a lo largo de c a

$$\nabla_{\dot{c}(t)} Y := (\nabla_X Y)(c(t)).$$

En coordenadas locales tenemos que $\dot{c}(t) = \sum_i \dot{c}^i(t) \frac{\partial}{\partial x^i}$, y sustituyendo en la fórmula (3.5) tenemos que:

$$\nabla_{\dot{c}(t)} Y = \sum_k \left(\sum_i \dot{c}^i(t) \frac{\partial \eta^k(c(t))}{\partial x^i} + \sum_{i,j} \dot{c}^i(t) \eta^j(c(t)) \Gamma_{ij}^k(c(t)) \right) \frac{\partial}{\partial x^k}, \quad (3.7)$$

donde hemos escrito $c(t)$ en vez de $\varphi(c(t))$ con φ una aplicación coordenada, para aligerar la notación.

Observación. La derivada covariante es una restricción de la conexión a una única curva integral en el primer argumento. Por tanto, la derivada covariante también es una noción intrínseca a la variedad.

3.1.1. Unicidad de la conexión de Levi-Civita

Hasta ahora hemos visto que para definir esta generalización de la derivada direccional basta con escoger los símbolos de Christoffel. Al quedar esto a nuestra elección, podría parecer entonces que la conexión no nos dice nada acerca de la variedad. Sin embargo, veremos que existe una única conexión que además verifica las condiciones (3.1) y (3.2) de la definición 3.1. Para ello empecemos demostrando la Fórmula de Koszul.

Proposición 3.9 (Fórmula de Koszul). Sea (M, g) una variedad métrica. Toda conexión en M que cumple las condiciones (3.1) y (3.2), debe verificar la siguiente fórmula:

$$g(\nabla_X Y, Z) = \frac{1}{2} (X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) - g(Z, [Y, X]) - g(Y, [X, Z]) - g(X, [Y, Z])), \quad (3.8)$$

para todo $X, Y, Z \in \mathfrak{X}(M)$.

Demostración. A partir de la condición de compatibilidad métrica, para todo $X, Y, Z \in \mathfrak{X}(M)$, se satisfacen las siguientes identidades:

$$\begin{aligned} X(g(Y, Z)) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \\ Y(g(X, Z)) &= g(\nabla_Y X, Z) + g(X, \nabla_Y Z), \\ Z(g(X, Y)) &= g(\nabla_Z X, Y) + g(X, \nabla_Z Y). \end{aligned}$$

Sumando las dos primeras y restando la tercera, y reorganizando, obtenemos:

$$\begin{aligned} X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) &= g(\nabla_X Y, Z) + g(\nabla_Y X, Z) + g(\nabla_X Z, Y) \\ &\quad - g(\nabla_Z X, Y) + g(\nabla_Y Z, X) - g(\nabla_Z Y, X), \end{aligned}$$

con lo que se concluye, teniendo en cuenta la condición de torsión nula. \square

Teorema 3.10. *Sea (M, g) una variedad métrica con g no degenerada. entonces existe una única conexión ∇ sin torsión y compatible con la métrica denominada conexión de Levi-Civita.*

Demostración. A partir de la fórmula de Koszul (3.8), la unicidad es inmediata. Ahora debemos verificar que $\nabla_X Y$ definida a partir de la fórmula de Koszul siempre existe y que además es una conexión. Para la existencia basta darse cuenta de que la aplicación $g(X, \cdot)$ define un isomorfismo entre TM y T^*M ya que g es no degenerada. De esta forma, verificar que $g(\nabla_X Y, \cdot)$ es una 1-forma no nula nos garantiza la existencia del campo $\nabla_X Y$. Es inmediato comprobar que fijados $X, Y \in \mathfrak{X}(M)$, se cumple:

$$g(\nabla_X Y, Z_1 + Z_2) = g(\nabla_X Y, Z_1) + g(\nabla_X Y, Z_2), \quad \forall Z_1, Z_2 \in \mathfrak{X}(M),$$

pues los términos del lado derecho de la igualdad en la fórmula de Koszul, (3.8), se comportan bien con la suma de campos en el término Z .

Veamos ahora que $g(\nabla_X Y, fZ) = fg(\nabla_X Y, Z)$ para toda $f \in C^\infty(M)$. Sustituyendo fZ en la fórmula de Koszul (3.8), tenemos:

$$\begin{aligned} 2g(\nabla_X Y, fZ) &= X(g(Y, fZ)) + Y(g(X, fZ)) - fZ(g(X, Y)) \\ &\quad - g(fZ, [Y, X]) - g(Y, [X, fZ]) - g(X, [Y, fZ]). \end{aligned}$$

Sabiendo que $[X, fZ] = f[X, Z] + (Xf)Z$ para todo par de campos y para toda función diferenciable, sustituimos en la ecuación anterior obteniendo:

$$\begin{aligned} 2g(\nabla_X Y, fZ) &= X(fg(Y, Z)) + Y(fg(X, Z)) - fZg(X, Y) \\ &\quad + fg([X, Y], Z) - g(f[X, Z] + (Xf)Z, Y) - g(f[Y, Z] + (Yf)Z, X) \\ &= fXg(Y, Z) + (Xf)g(Y, Z) + fYg(X, Z) + (Yf)g(X, Z) \\ &\quad - fZg(X, Y) + fg([X, Y], Z) - fg([X, Z], Y) - (Xf)g(Z, Y) \\ &\quad - fg([Y, Z], X) - (Yf)g(Z, X). \end{aligned}$$

Observamos que los términos con Xf e Yf se cancelan y los multiplicados por f es precisamente $2g(\nabla_X Y, Z)$. Por tanto, la aplicación $Z \mapsto g(\nabla_X Y, Z)$ es una 1-forma y el campo $\nabla_X Y$ existe y es único. La comprobación de que ∇ así definida es una conexión se ve en el lema A.1. \square

Hemos dicho que una conexión queda definida mediante la elección de los símbolos de Christoffel. Sin embargo, el resultado anterior nos indica que la conexión bajo ciertas condiciones es única y por tanto en ese caso los símbolos de Christoffel también son únicos. A continuación vamos a ver cómo determinarlos a partir de la métrica.

Proposición 3.11. *Los símbolos de Christoffel de la única conexión métrica de una variedad métrica (M, g) verifican:*

$$\Gamma_{ij}^m = \frac{1}{2} g^{mk} \left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ki} - \frac{\partial}{\partial x^k} g_{ij} \right), \quad (3.9)$$

donde $(g^{mk}) = (g_{ij})^{-1}$.

Demostración. Sea $(\partial_1, \dots, \partial_n)$ la base de campos coordenados y sean $g_{ij} = g(\partial_i, \partial_j)$ las componentes de la métrica. Consideremos los campos $X = \partial_i, Y = \partial_j, Z = \partial_k$ y sustituyamos en la fórmula de Koszul (3.8). Teniendo en cuenta que $[\partial_i, \partial_j] = 0$, tenemos:

$$2g(\nabla_{\partial_i} \partial_j, \partial_k) = \partial_i g(\partial_j, \partial_k) + \partial_j g(\partial_k, \partial_i) - \partial_k g(\partial_i, \partial_j) = \partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}.$$

Y sustituyendo (3.6) en la expresión anterior se tiene:

$$2\Gamma_{ij}^l g_{lk} = \partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}.$$

Multiplicando por la inversa (g^{mk}) y sumando en k , se concluye. \square

Observación. Al estar unívocamente determinados por la métrica si verifican todas las condiciones de la definición 3.1, los símbolos de Christoffel encapsulan las propiedades de la geometría intrínseca de la variedad.

Observación. En el caso de tener una variedad con una métrica constante, un caso particular de una variedad plana, los símbolos de Christoffel son nulos. En el caso del espacio euclídeo de \mathbb{R}^m , vemos de nuevo cómo la conexión coincide con la derivada direccional. Sean $X = \sum_i \xi^i \frac{\partial}{\partial x^i}, Y = \sum_j \eta^j \frac{\partial}{\partial x^j}$ sustituyendo en la expresión de la conexión en coordenadas 3.5 se tiene,

$$\nabla_X Y = \sum_k \left(\sum_i \xi^i \frac{\partial \eta^k}{\partial x^i} \right) \frac{\partial}{\partial x^k}.$$

Para una interpretación de los símbolos de Christoffel ver ??.

Ejemplo 3.12 (Conexión en la esfera.) Consideremos la esfera unidad \mathbb{S}^2 inmersa en \mathbb{R}^3 . La métrica inducida por la usual en coordenadas esféricas queda $g = d\varphi^2 + \sin^2 \varphi d\theta^2$, donde θ es el ángulo en el plano del ecuador y φ es el ángulo polar. Calculemos los símbolos de Christoffel empleando la fórmula 3.9.

$$\begin{aligned}\Gamma_{\varphi\varphi}^{\varphi} &= \frac{1}{2}g^{\varphi\varphi} (g_{\varphi\varphi,\varphi} + g_{\varphi\varphi,\varphi} - g_{\varphi\varphi,\varphi}) + \frac{1}{2}g^{\varphi\theta} (g_{\varphi\theta,\varphi} + g_{\theta\varphi,\varphi} - g_{\varphi\varphi,\theta}) = 0, \\ \Gamma_{\varphi\theta}^{\theta} &= \frac{1}{2}g^{\theta\varphi} (g_{\theta\varphi,\varphi} + g_{\varphi\varphi,\theta} - g_{\varphi\theta,\varphi}) + \frac{1}{2}g^{\theta\theta} (g_{\theta\theta,\varphi} + g_{\theta\varphi,\theta} - g_{\varphi\theta,\theta}) = \frac{1}{2} \frac{1}{\sin^2 \varphi} (2 \sin \varphi \cos \varphi) = \frac{\cos \varphi}{\sin \varphi}, \\ \Gamma_{\theta\theta}^{\varphi} &= \frac{1}{2}g^{\varphi\varphi} (g_{\theta\varphi,\theta} + g_{\varphi\theta,\theta} - g_{\theta\theta,\varphi}) + \frac{1}{2}g^{\varphi\theta} (g_{\theta\theta,\theta} + g_{\theta\theta,\theta} - g_{\theta\theta,\theta}) = -\sin \varphi \cos \varphi.\end{aligned}$$

Análogamente, se obtienen todos los símbolos de Christoffel:

$$\begin{aligned}\Gamma_{\theta\theta}^{\theta} &= 0, & \Gamma_{\varphi\theta}^{\theta} &= \frac{\cos \varphi}{\sin \varphi}, \\ \Gamma_{\varphi\varphi}^{\theta} &= 0, & \Gamma_{\theta\theta}^{\varphi} &= -\sin \varphi \cos \varphi, \\ \Gamma_{\theta\varphi}^{\varphi} &= 0, & \Gamma_{\varphi\varphi}^{\varphi} &= 0.\end{aligned}$$

Consideremos los campos $X = \frac{\partial}{\partial \varphi}$ y $Y = \frac{\partial}{\partial \theta}$. Finalmente, sustituyendo en la fórmula 3.5 obtenemos

$$\nabla_X Y = \sum_{k=\varphi,\theta} \Gamma_{\varphi\theta}^k \partial_k = \Gamma_{\varphi\theta}^{\theta} \partial_{\theta} = \frac{\cos \varphi}{\sin \varphi} \partial_{\theta},$$

y del mismo modo

$$\nabla_Y X = \sum_{k=\varphi,\theta} \Gamma_{\theta\varphi}^k \partial_k = \Gamma_{\theta\varphi}^{\varphi} \partial_{\varphi} = \frac{\cos \varphi}{\sin \varphi} \partial_{\varphi}.$$

Siguiendo con la idea de que la conexión generaliza la derivada direccional, este resultado se puede interpretar como que el desplazamiento en la dirección de los meridianos², ∂_{φ} , un vector en la dirección del paralelo ∂_{θ} experimenta un giro proporcional a $\frac{\cos \varphi}{\sin \varphi}$. Es decir, este ejemplo ilustra que la conexión indica cómo se transportan los vectores tangentes de un espacio tangente a otro.

3.1.2. Transporte paralelo

Como comentamos al principio de la sección, cuando trabajamos con \mathbb{R}^m que podemos comparar vectores de diferentes espacios tangentes de manera natural, mientras que en el resto de variedades no. Sin embargo, con el ejemplo anterior se puede de manera intuitiva como la conexión nos permite definir esta correspondencia entre espacios tangentes. Esta es aplicación inducida por la conexión se conoce como transporte paralelo. Veamos cómo construirla.

²Se razona de manera análoga para $\nabla_Y X$.

Definición 3.13. Un campo $Y \in \mathfrak{X}(M)$ se dice *paralelo* si $\nabla_X Y = 0$, $\forall X \in \mathfrak{X}(M)$. Análogamente si $c: I \rightarrow M$ es una curva diferenciable y $V \in \mathfrak{X}(c)$ un campo con soporte en dicha curva. Decimos que V es *paralelo a lo largo de c* si $\nabla_{\dot{c}} V = 0$.

Observación. El conjunto de campos paralelos forman un \mathbb{R} -submódulo de $\mathfrak{X}(M)$. Sean $Y, Z \in \mathfrak{X}(M)$ campos paralelos. Entonces se verifica

$$\nabla_X(aY + bZ) = \nabla_X(aY) + \nabla_X(bZ) = a'Y + a\nabla_X Y + b'Z + b\nabla_X Z = 0, \quad \forall X \in \mathfrak{X}(M) \quad \forall a, b \in \mathbb{R}.$$

De manera análoga se ve que los campos paralelos a lo largo de una curva forman un \mathbb{R} -submódulo de $\mathfrak{X}(c)$.

Teorema 3.14. Sea $c: I \rightarrow M$ una curva diferenciable no constante. Sean $t_0 \in I$, $p_0 = c(t_0) \in M$ y $v \in T_{p_0}M$. Entonces existe un único campo $Y \in \mathfrak{X}(c)$ con $Y(t_0) = v$ y que $\nabla_{\dot{c}} Y = 0$.

Demostración. En coordenadas tenemos $c = \sum_i c^i \frac{\partial}{\partial x^i}$, $Y = \sum_j \eta^j \frac{\partial}{\partial x^j}$. Como buscamos que Y sea paralelo a lo largo de c , verifica que $\nabla_{\dot{c}} Y = 0$. Sustituyendo en la expresión de la derivada covariante en coordenadas 3.7:

$$\frac{\partial \eta^k}{\partial x^i} + \sum_{i,j} \eta^j \Gamma_{ij}^k(c(t)) = 0, \quad \forall k = 1, \dots, m. \quad (3.10)$$

Por tanto, para que el campo sea paralelo debe verificar este sistema de ecuaciones diferenciales. Por el teorema de existencia y unicidad de soluciones ecuaciones diferenciales ordinarias, tenemos que para la condición inicial $(\eta^1, \dots, \eta^m)(\tilde{c}(t_0)) = (v_1, \dots, v_m)$ con $\tilde{c}(t_0)$ el campo c es único. Debemos ver que esto es independiente de la carta empleada. Pare ello, hacemos en cartas cuya intersección contenga a la curva. De nuevo, por el teorema de existencia y unicidad tenemos que la solución obtenida en la intersección de dos abiertos debe de ser única. \square

Veamos ahora si los siguientes campos sobre la esfera son paralelos o no a modo de ejemplo.

Ejemplo 3.15. ■ Consideremos el campo $Y = \frac{\partial}{\partial x^\varphi}$ sustituyendo en el sistema de ecuaciones diferenciales 3.10 tenemos

$$\begin{cases} \Gamma_{\theta\varphi}^\theta + \Gamma_{\varphi\varphi}^\theta = 0, \\ \Gamma_{\theta\varphi}^\varphi + \Gamma_{\varphi\varphi}^\varphi = 0. \end{cases}$$

En el ejemplo 3.12 ya calculamos los símbolos de Christoffel. Así, llegamos a que Y es paralelo a la curva tal que

$$\frac{\cos(\varphi)}{\sin(\varphi)} = 0;$$

es decir, $c(t) = (\theta(t), \frac{\pi}{2})$. Por tanto, el campo Y es paralelo a lo largo del ecuador. De manera similar podemos ver que $Y = \frac{\partial}{\partial x^\theta}$ también es paralelo a lo largo del ecuador.

- Consideremos ahora la curva $c(t) = (\omega t, \frac{\pi}{3})$ contenida en un meridiano. y el campo $Y = \frac{\partial}{\partial x^\theta}$. Veamos que Y no es paralelo a lo largo de c . Por las propiedades de la conexión y con los resultados del ejemplo 3.12 tenemos

$$\nabla_{\dot{c}} Y = \nabla_{(\omega, 0)}(1, 0) = \omega \nabla_{(1, 0)}(1, 0) = \omega \left(\Gamma_{\theta\theta}^\varphi \frac{\partial}{\partial x^\varphi} + \Gamma_{\theta\theta}^\theta \frac{\partial}{\partial x^\theta} \right) = -\omega \sin(\theta) \cos(\theta) \neq 0$$

La siguiente propiedad de los campos paralelos nos permitirá definir el transporte paralelo.

Proposición 3.16. *Sea (M, g) una variedad diferenciable junto con una métrica no degenerada. Sea ∇ la conexión inducida por g . Sea $X \in \mathfrak{X}(M)$ un campo y c una de sus curvas integrales. Sean $Y, Z \in \mathfrak{X}(c)$ dos campos paralelos a lo largo de c , entonces $g(Y(t), Z(t))$ es constante a lo largo de c .*

Demostración. Como ∇ es compatible con la métrica y Y y Z son campos paralelos a lo largo de c tenemos que

$$X(c(t))(g(Y, Z)) = g(\nabla_{\dot{c}} Y, Z) + g(Y, \nabla_{\dot{c}} Z) = g(0, Z) + g(Y, 0) = 0.$$

Por lo tanto, $g(Y, Z)$ es constante a lo largo de c . □

Definición 3.17. Dada una curva diferenciable $c : I \rightarrow M$ no constante, con $c(t_0) = p$ y $c(t_1) = q$. Se define el *transporte paralelo* de p a q a lo largo de c como:

$$\begin{aligned} P_{t_0}^{t_1}(c) : T_p M &\longrightarrow T_q M \\ v &\longmapsto V(t_1) \end{aligned}$$

donde $V \in \mathfrak{X}(c)$ es el único campo paralelo a lo largo de c con $V(t_0) = v$.

Observación. Nótese que esta definición del transporte paralelo falla para curvas constantes ya que todos los campos son paralelos a lo largo de la curva. Esto no es un problema cuando lo que tratemos sea comparar elementos de diferentes espacios tangentes. Sin embargo, cuando definamos conexiones a partir del transporte paralelo y para ello deberemos imponer que para curvas constantes P sea la identidad.

Antes de seguir adelante debemos ver que la aplicación está bien definida; que no depende de la parametrización de la curva.

Proposición 3.18. *El transporte paralelo $P_{t_0}^{t_1}(c)$ no depende de la parametrización de la curva c .*

Demostración. Sea $\gamma : J \rightarrow I$ una función diferenciable y biyectiva tal que $\gamma(s_0) = t_0$ y $\gamma(s_1) = t_1$. De esta forma podemos considerar la curva reparametrizada como $\tilde{c} = c \circ \gamma : J \rightarrow M$.

Sea $Y \in \mathfrak{X}(c)$ el campo paralelo a lo largo de c con $Y(t_0) = y$. Definimos $\tilde{Y} \in \mathfrak{X}(\tilde{c})$ como la reparametrización de Y , es decir que $\tilde{Y}(s) = Y(\gamma(s))$. Por la regla de la cadena y las propiedades de una conexión tenemos que

$$\nabla_{\dot{\tilde{c}}(s)} \tilde{Y} = \nabla_{\gamma'(s)\dot{c}(\gamma(s))} Y = \gamma'(s) \nabla_{\dot{c}(\gamma(s))} Y = \gamma'(s) \cdot 0 = 0$$

Por tanto, \tilde{Y} es paralelo a lo largo de \tilde{c} . Así, podemos definir la aplicación del transporte paralelo sobre \tilde{c} y como se verifica que $\tilde{Y}(s_1) = Y(\gamma(s_1)) = Y(t_1)$ y $\tilde{Y}(s_0) = Y(\gamma(s_0)) = Y(t_0) = y$. Podemos concluir que, $P_{s_0}^{s_1}(\tilde{c})(y) = \tilde{Y}(s_1) = Y(\gamma(s_1)) = Y(t_1) = P_{t_0}^{t_1}(c)(y)$. Además, por la unicidad del campo paralelo demostrada anteriormente podemos afirmar que no existe ningún otro campo \tilde{Y} que lo verifique. Así, concluimos que $P_{s_0}^{s_1}(\tilde{c}) = P_{t_0}^{t_1}(c)$, y por tanto el transporte paralelo no depende de la parametrización escogida. \square

Veamos ahora alguna propiedad del transporte paralelo 3.2. Para demostración, véase A.2.

Proposición 3.19. *El transporte paralelo $P_{t_0}^{t_2}(c): T_{c(t_0)}M \rightarrow T_{c(t_2)}M$ satisface*

$$P_{t_0}^{t_2}(c) = P_{t_1}^{t_2}(c_2) \circ P_{t_0}^{t_1}(c_1),$$

$c: [t_0, t_2] \rightarrow M$ curva concatenación diferenciable de las curvas $c_1: [t_0, t_1] \rightarrow M$ y $c_2: [t_1, t_2] \rightarrow M$. Además, el transporte paralelo $P_{t_0}^t(c): T_{c(t_0)}M \rightarrow T_{c(t)}M$ es una isometría lineal.

Para la demostración del siguiente teorema, véase [I].

Teorema 3.20 (Teorema de Taylor en variedades). *Sea (M, ∇) una variedad de dimensión m equipada con una conexión. Sea $c: [0, t] \rightarrow M$ una curva diferenciable con $c(0) = p$, $c(t) = q$. Supongamos que $V(t)$ un campo paralelo a lo largo de c . Sea $V^*(t)$ el campo trasladado por c desde q hasta p , es decir*

$$V^*(t) = P_t^0(c)(V(t)) \in T_pM,$$

donde $P_0^t(c)$ es el transporte paralelo de T_pM a $T_{c(t)}M$ y $P_t^0(c)$ su inversa. Entonces, alrededor de $t = 0$,

$$V^*(t) = V(0) + t \nabla_{\dot{c}} V(0) + \frac{t^2}{2} \nabla_{\dot{c}}^2 V(0) + o(t^2).$$

Una vez tenemos el transporte paralelo podemos comparar vectores de diferentes espacios tangentes y ver una conexión como un cociente muy similar al de la derivada usual. Sea ∇ una conexión en M y P la aplicación de transporte paralelo que define. Sean $X, Y \in \mathfrak{X}(M)$ campos tangentes y sea c la curva integral de X tal que $c(0) = p$ entonces

$$\nabla_X Y|_p = \lim_{t \rightarrow 0} \frac{(P_0^t(c(t)))^{-1} - c(0)}{t} = \lim_{t \rightarrow 0} \frac{P_t^0(c(t)) - c(0)}{t}.$$

Es decir, el único cambio con la derivada usual es que antes de hacer la diferencia entre dos vectores es que tenemos que trasladar el vector mediante la aplicación $P_0^t(c(t))$.

Por otro lado, si observamos la expresión anterior podemos ver como el miembro derecho depende únicamente del transporte paralelo sobre una curva. Esto hace que de manera natural surja la pregunta de si podemos construir una conexión a partir de un transporte paralelo. En la siguiente sección trataremos de definir la conexión a partir de las propiedades vistas.

3.2. Equivalencia entre transporte paralelo y conexión

Definición 3.21. Un *pseudotransporte paralelo* en M es una asignación, en la que a cada curva diferenciable $c: [a, b] \rightarrow M$ le hace corresponder una familia de aplicaciones $P_a^t(c) : T_{c(a)}M \rightarrow T_{c(t)}M$, $t \in [a, b]$, que satisfacen las siguientes propiedades:

1. **Linealidad.** Para cada curva diferenciable $c: [a, b] \rightarrow M$ y cada $t \in [a, b]$, el transporte paralelo

$$P_a^t(c) : T_{c(a)}M \rightarrow T_{c(t)}M,$$

es una aplicación lineal.

2. **Condición inicial.** Para cualquier curva $c: [a, b] \rightarrow M$ y cualquier $t \in [a, b]$, se tiene

$$P_t^t(c) = \text{id}_{T_{c(t)}M}.$$

3. **Composición.** Sea $c: [a, b] \rightarrow M$ una curva diferenciable y sea $s \in (a, b)$ entonces

$$P_s^b(c) \circ P_a^s(c) = P_a^b(c).$$

4. **Invarianza bajo reparametrizaciones.** Sea $\phi : [a', b'] \rightarrow [a, b]$ un difeomorfismo con derivada no nula en todo punto tal que $\phi(a') = a$ y $\phi(b') = b$, entonces

$$P_{a'}^{b'}(c \circ \phi) = P_a^b(c).$$

Observación. Nótese que a partir de la propiedad 3 y por inducción podemos extenderlo a una composición finita. Sea una partición $a = t_0 < t_1 < \dots < t_n = b$ del intervalo $[a, b]$. Entonces

$$P_a^b(c) = P_{t_{n-1}}^b(c) \circ P_{t_{n-2}}^{t_{n-1}}(c) \circ \dots \circ P_{t_1}^{t_2}(c) \circ P_{t_0}^{t_1}(c)$$

Observación (Intento de definición de la conexión a partir del pseudotransporte paralelo). Sea M una variedad diferenciable dotada de un pseudotransporte paralelo 3.21. La aplicación dada por

$$\nabla_X Y|_p = \lim_{t \rightarrow 0} \frac{P_0^t(c)(Y_{c(t)}) - Y_p}{t}, \quad X, Y \in \mathfrak{X}(M),$$

donde c es la curva integral de X que pasa por $p \in M$, verifica todas las propiedades de las conexiones menos la aditividad.

Demostración. Dados $X, Y \in \mathfrak{X}(M)$ y sea c la curva integral de X que pasa por $p \in M$, veamos que el siguiente límite efectivamente define una conexión:

$$\nabla_X Y|_p = \lim_{t \rightarrow 0} \frac{P_0^t(c)(Y_{c(t)}) - Y_p}{t},$$

donde $c(t) = \Phi_t^X(p)$ es el flujo de X con $c(0) = p$. Veamos que este límite satisface todas las propiedades de una conexión.

1. Producto por una función diferenciable.

$$\nabla_{fX} Y(p) = f(p)(\nabla_X Y)(p), \quad \forall f \in C^\infty(M).$$

Sea $c(t) = \Phi_t^X(p)$ la curva integral de X en p . Es fácil ver entonces que la curva integral de fX con $f \in C^\infty(M)$ es una reparametrización de la anterior. De esta forma podemos considerar $c_f(t) = c(b(t))$, con $b(t)$ una función diferenciable tal que $b(0) = 0$ y $b'(0) = f(p)$. Entonces por la propiedad 4 de invarianza bajo reparametrizaciones del transporte paralelo,

$$\nabla_{fX} Y(p) = \lim_{t \rightarrow 0} \frac{P_0^t(c_f)(Y_{c_f(t)}) - Y(p)}{t} = \lim_{t \rightarrow 0} \frac{P_0^{b(t)}(c)(Y_{c(b(t))}) - Y(p)}{t}.$$

Esto lo podemos reescribir como

$$\nabla_{fX} Y(p) = \lim_{t \rightarrow 0} \frac{P_0^{b(t)}(c)(Y_{c(b(t))}) - Y(p)}{b(t)} \cdot \frac{b(t)}{t}.$$

Por definición, el primer factor tiende a $(\nabla_X Y)(p)$. Como $f(p) = b'(0)$, entonces

$$\nabla_{fX} Y(p) = f(p)(\nabla_X Y)(p).$$

2. Aditividad en el segundo argumento

$$\nabla_X (Y_1 + Y_2)(p) = \nabla_X Y_1(p) + \nabla_X Y_2(p).$$

De nuevo consideramos $c(t) = \Phi_t^X(p)$. Para demostrar esta propiedad basta con aplicar la linealidad del transporte paralelo en el segundo argumento:

$$\begin{aligned}\nabla_X(Y_1 + Y_2)(p) &= \lim_{t \rightarrow 0} \frac{P_0^t(c)((Y_1 + Y_2)_{c(t)}) - (Y_1 + Y_2)(p)}{t} \\ &= \lim_{t \rightarrow 0} \frac{P_0^t(c)(Y_{1,c(t)}) - Y_1(p)}{t} + \lim_{t \rightarrow 0} \frac{P_0^t(c)(Y_{2,c(t)}) - Y_2(p)}{t} \\ &= \nabla_X Y_1(p) + \nabla_X Y_2(p).\end{aligned}$$

3. Regla de Leibniz

$$\nabla_X(fY)(p) = X(f)(p)Y(p) + f(p)(\nabla_X Y)(p), \quad \forall f \in C^\infty(M).$$

Sea $c(t) = \Phi_t^X(p)$ y sea $f \in C^\infty(M)$. Por definición y como $P_0^t(c)$ es lineal tenemos que

$$\nabla_X(fY)(p) = \lim_{t \rightarrow 0} \frac{f(c(t))P_0^t(c)(Y_{c(t)}) - f(p)Y(p)}{t}.$$

Si sumamos y restamos $f(c(t))Y(p)$, llegamos a que:

$$\nabla_X(fY)(p) = \lim_{t \rightarrow 0} \frac{f(c(t)) \left[P_0^t(c)(Y_{c(t)}) - Y(p) \right]}{t} + \lim_{t \rightarrow 0} \frac{\left[f(c(t)) - f(p) \right] Y(p)}{t}.$$

Así, es fácil ver que

$$\nabla_X(fY)(p) = X(f)(p)Y(p) + f(p)(\nabla_X Y)(p).$$

□

Observación. Para garantizar que se verifica la aditividad es necesario imponer que el transporte a lo largo de lazos infinitesimales debe converger a la identidad. Esto se puede ver en [KN51].

Así, si añadimos esta última propiedad podemos demostrar que dar un transporte paralelo es equivalente a dar una conexión.

Definición 3.22. Un *transporte paralelo* P en M es una asignación, en la que a cada curva diferenciable sobre M , le corresponde una familia de aplicaciones que satisfacen las condiciones enunciadas en 3.21. Además, se requiere que esta asignación verifique una condición de continuidad: al considerar lazos infinitesimales, el transporte debe converger a la identidad (ver [KN51]).

Veamos ahora que no sólo podemos definir una conexión, si no que podemos imponiendo dos condiciones más podemos obtener la conexión de Levi-Civita.

Proposición 3.23. *Sea (M, g, P) una variedad diferenciable con una métrica no degenerada y un transporte paralelo P que verifica las propiedades de la definición 3.22. Si además $P_a^t(c)$ es una isometría para toda curva c y todo $a, t \in [a, b]$, es decir,*

$$g_{c(t)}(P_a^t(c)(v), P_a^t(c)(w)) = g_{c(a)}(v, w), \quad \forall v, w \in T_{c(a)}M,$$

entonces la conexión ∇ inducida por P es compatible con la métrica g .

Demostración. Sea $X \in \mathfrak{X}(M)$ un campo y $c(t) = \Phi_t^X(p)$ una de sus curvas integrales. Dado que $P_0^t(c)$ es una isometría:

$$g_p(P_t^0(c)(Y_{c(t)}), P_t^0(c)(Z_{c(t)})) = g_{c(t)}(Y_{c(t)}, Z_{c(t)}).$$

Entonces derivando y sustituyendo en la igualdad anterior obtenemos:

$$\frac{d}{dt}\Big|_{t=0} g_{c(t)}(Y_{c(t)}, Z_{c(t)}) = g_p\left(\frac{d}{dt}\Big|_{t=0} P_t^0(c)(Y_{c(t)}), P_t^0(c)(Z_{c(t)}) + g_p(P_t^0(c)(Y_{c(t)}), \frac{d}{dt}\Big|_{t=0} P_t^0(c)(Z_{c(t)}).\right.$$

Por definición tenemos que $\frac{d}{dt}\Big|_{t=0} P_t^0(c)(Y_{c(t)}) = \nabla_X Y$, y por tanto

$$\frac{d}{dt}\Big|_{t=0} g_{c(t)}(Y_{c(t)}, Z_{c(t)}) = g_p(\nabla_X Y_p, Z_p) + g_p(Y_p, \nabla_X Z_p).$$

Por último, como estamos derivando en la dirección del vector tangente a c en p ; podemos reescribirlo como la derivada direccional según el campo X :

$$X(g(Y, Z))(p) = g(\nabla_X Y, Z)_p + g(Y, \nabla_X Z)_p.$$

□

Para ver qué se tiene que dar para que la conexión inducida tenga torsión nula, primero debemos presentar el siguiente lema que nos da una interpretación geométrica del corchete de Lie.

Lema 3.24. *Sea $X \in \mathfrak{X}(M)$ y sea Φ_t^X su flujo. Entonces, para cada $p \in M$ se tiene*

$$[X, Y]_p = \frac{d}{dt}\Big|_{t=0} (\Phi_{-t}^X \cdot Y)_p, \quad \forall Y \in \mathfrak{X}(M),$$

donde

$$\Phi_{-t}^X \cdot Y = T\Phi_{-t}^X \circ Y \circ \Phi_t^X$$

es el campo pull-back del campo Y por el difeomorfismo Φ_t^X .

Demostración. Véase [KoNo63, Prop. 1.9].

□

Es decir, que el corchete de Lie $[X, Y]$ es la variación infinitesimal del campo Y a lo largo de las curvas integrales del campo X .

Proposición 3.25. *Sea M una variedad diferenciable. Entonces definir una conexión en M con torsión nula es equivalente a definir un transporte paralelo P que verifica las propiedades 1-4 de la definición 3.22 y, siguiendo la notación anterior, que para todo $X, Y \in \mathfrak{X}(M)$ y $p \in M$,*

$$\lim_{t \rightarrow 0} \frac{P_t^0(\Phi_t^X(p))(Y_{\Phi_t^X(p)(t)}) - T\phi_X^{-t}(Y_{\Phi_t^X(p)})}{t} = \lim_{t \rightarrow 0} \frac{P_t^0(c_Y)(X_{\phi_Y^t(p)(t)}) - T\phi_Y^{-t}(X_{\phi_Y^t(p)})}{t} = -[X, Y]_p.$$

Demostración. Sean $X, Y \in \mathfrak{X}(M)$. Hemos visto que la relación entre conexión y transporte paralelo es

$$\nabla_X Y = \lim_{t \rightarrow 0} \frac{P_0^t(\phi_t^X(p))(Y_{\phi_t^X(p)}) - Y_p}{t}.$$

Ahora, si sumamos y restamos $T\phi_{-t}^X(Y_{\phi_t^X(p)})$ en el numerador tenemos

$$\nabla_X Y = \lim_{t \rightarrow 0} \frac{P_0^t(\phi_t^X(p))(Y_{\phi_t^X(p)}) - T\phi_{-t}^X(Y_{\phi_t^X(p)}) + T\phi_{-t}^X(Y_{\phi_t^X(p)}) - Y_p}{t}.$$

Definimos

$$E_X(p) := \lim_{t \rightarrow 0} \frac{P_0^t(\phi_t^X(p))(Y_{\phi_t^X(p)}) - T\phi_{-t}^X(Y_{\phi_t^X(p)})}{t},$$

entonces sustituyéndolo y aplicando el lema anterior, tenemos que:

$$\nabla_X Y_p = [X, Y]_p + E_X(p).$$

Podemos hacer el mismo razonamiento para $\nabla_Y X$ definiendo $E_Y(p)$ y obtenemos

$$\nabla_Y X_p = [Y, X]_p + E_Y(p).$$

Considerando ahora la diferencia

$$\nabla_X Y_p - \nabla_Y X_p = [X, Y]_p + E_X(p) - [Y, X]_p - E_Y(p) = 2[X, Y]_p + (E_X(p) - E_Y(p)).$$

Finalmente igualando con la condición de torsión nula:

$$[X, Y] = 2[X, Y] + (E_X - E_Y) \iff E_X - E_Y = -[X, Y].$$

□

Teorema 3.26. *Sea (M, g, P) una variedad diferenciable equipada con una métrica no degenerada y un transporte paralelo que además verifica*

1. $P_a^t(c)$ es una isometría para toda curva c y todo $a, t \in [a, b]$, es decir,

$$g_{c(t)}(P_a^t(c)(v), P_a^t(c)(w)) = g_{c(a)}(v, w), \quad \forall v, w \in T_{c(a)}M.$$

2. Para todo $X, Y \in \mathfrak{X}(M)$ y $p \in M$,

$$\lim_{t \rightarrow 0} \frac{P_t^0(\Phi_t^X(p))(Y_{\Phi_t^X(p)}(t)) - T\phi_X^{-t}(Y_{\Phi_t^X(p)})}{t} - \lim_{t \rightarrow 0} \frac{P_t^0(c_Y)(X_{\phi_Y^t(p)}(t)) - T\phi_Y^{-t}(X_{\phi_Y^t(p)})}{t} = -[X, Y]_p.$$

entonces induce la conexión de Levi-Civita.

Demostración. Como P verifica las condiciones de 3.22 entonces induce una conexión. Aplicando las proposiciones 3.23 y 3.25 obtenemos que la conexión inducida ∇ , tiene torsión nula y es compatible con la métrica. Por tanto, podemos aplicar el teorema 3.10, que nos garantiza que dicha conexión es la de Levi-Civita. \square

La proposición 3.25 nos permite dar una interpretación geométrica a la condición de torsión nula que al principio puede parecer sólo un detalle técnico que permite obtener la unicidad en la conexión.

Sabemos que los campos definen una derivada direccional para funciones diferenciables. Además, si $[X, Y] = 0$ ambos campos conmutan y $X(Y(f)) = Y(X(f))$. Lo mismo sucede con la composición de sus flujos, en el límite, es igual ir primero por la curva integral de X y luego por la de Y que al revés. Por otro lado, si la torsión es nula y consideramos campos que conmutan tenemos que $E_X - E_Y = -[X, Y] = 0$, y por tanto la diferencia entre el flujo y transporte paralelo es igual hasta primer orden independientemente de la dirección. Por ello, si la torsión es nula y éstos conmutan, el transporte paralelo de vectores también. En el caso de que $E_X - E_Y \neq -[X, Y]$, aunque los campos conmuten, al transportar un vector a lo largo de un campo y luego a lo largo del otro, incluso en el límite es diferente. Por tanto, una conexión con torsión no nula introduce un término extra, ya que su comportamiento difiere en función de su dirección, induciendo un “giro” a lo largo del traslado y que tiene como consecuencia que el transporte paralelo no conmuta aunque los campos sí lo hagan.

3.3. Conexión sobre formas y tensores

En las secciones anteriores hemos trabajado con la conexión y transporte paralelo definido para campos. De esta forma, surge de manera natural la pregunta de si es posible extender esta idea para formas diferenciables a partir de la dualidad entre T_pM y T_p^*M . A priori, lo podemos hacer de dos

formas diferentes. La primera, partiendo de que el transporte paralelo es una aplicación lineal podríamos tomar su aplicación dual en los espacios cotangentes:

$$P_0^t(c)^* : T_q^*M \longrightarrow T_p^*M.$$

La segunda, viendo que fijando el primer argumento en la conexión obtenemos un endomorfismo en el espacio tangente, y mediante la dualidad, definir una conexión en el cotangente. Por la equivalencia que acabamos de ver independientemente de que camino escojamos podríamos construir una teoría análoga a la vista para campos³.

3.3.1. Conexión sobre formas

Sean (M, ∇) una variedad diferenciable equipada con una conexión y $X \in \mathfrak{X}(M)$ un campo sobre M . Consideremos el endomorfismo inducido por X :

$$\nabla_X : \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M).$$

De manera natural la aplicación candidata es la aplicación dual

$$\begin{aligned} \nabla_X^* : \Lambda^1 M &: \longrightarrow \Lambda^1 M \\ \omega &\longmapsto \nabla_X^* \omega := \omega \circ \nabla_X : \mathfrak{X}(M) \longrightarrow C^\infty(M). \end{aligned}$$

Es fácil ver que esta aplicación verifica las tres primeras propiedades de la definición 3.1. Sin embargo, no verifica la regla de Leibniz:

$$(\nabla_X^*(f\omega))(Y) = f \cdot (\nabla_X^*\omega)(Y) \quad X, Y \in \mathfrak{X}(M), \quad f \in C^\infty(M), \quad \omega \in \Lambda^1 M.$$

Es decir, esta aplicación así definida no puede ser una conexión para formas; intentemos pensarlo de otra forma. Como ω es una forma diferenciable, su evaluación sobre un campo cualquiera $\omega(Y)$ es una función diferenciable. Para esta función su derivada respecto de un campo $X(\omega(Y))$ sí que está definida. Sin embargo, $X(\omega(Y))$ captura todo su cambio sin distinguir si proviene de la propia ω respecto de X o del campo Y . Entonces, podemos escribir

$$X(\omega(Y)) = \nabla_X(\omega)(Y) + \omega(\nabla_X Y),$$

Si despejamos, tenemos que

$$\nabla_X(\omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y).$$

³Esta idea en realidad se trata de un caso particular del estudio conexiones en fibrados cuando nos restringimos a TM y T^*M . No lo haremos desde esta perspectiva pues el estudio de esta rama se sale de los objetivos del trabajo. Además, la construcción de la conexión definiendo primero una conexión para campos vectoriales y a partir de ella deducir la correspondiente para 1-formas puede proporcionar una mayor intuición

Esta es una nueva candidata a ser una conexión para formas. Al igual que antes, la verificación de las tres primeras condiciones es inmediata. Veamos si cumple la regla de Leibniz.

$$\begin{aligned} (\nabla_X(f\alpha))(Z) &= X((f\alpha)(Z)) - (f\alpha)(\nabla_X Z) = X(f\alpha(Z)) - f\alpha(\nabla_X Z) \\ &= X(f)\alpha(Z) + f[X(\alpha(Z)) - \alpha(\nabla_X Z)] \\ &= X(f)\alpha(Z) + f(\nabla_X \alpha)(Z). \end{aligned}$$

Así se tiene que

$$\nabla_X(f\alpha) = X(f)\alpha + f\nabla_X\alpha.$$

Proposición 3.27. *Sea (M, ∇) una variedad diferenciable equipada con una conexión. Entonces induce una única conexión para formas diferenciables:*

$$\nabla_X(\alpha)(Y) = X(\alpha(Y)) - \alpha(\nabla_X Y), \quad X, Y \in \mathfrak{X}(M). \quad (3.11)$$

Proposición 3.28. *Sea (M, g, ∇) una variedad diferenciable con una métrica no degenerada y una conexión compatible con la métrica. Entonces, la conexión inducida sobre 1-formas es compatible con la métrica contravariante h , es decir:*

$$X(h(\alpha, \beta)) = h(\nabla_X \alpha, \beta) + h(\alpha, \nabla_X \beta),$$

para todo $X \in \mathfrak{X}(M)$ y $\alpha, \beta \in \Lambda^1 M$.

Demostración. Siguiendo la notación del lema anterior, sean $\alpha_Y, \alpha_Z \in \Lambda^1 M$. Por la definición de la métrica contravariante y por la compatibilidad de la métrica con la conexión, tenemos que

$$X(h(\alpha_Y, \alpha_Z)) = X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = h(\alpha_{\nabla_X Y}, \alpha_Z) + h(\alpha_Y, \alpha_{\nabla_X Z}).$$

Por el lema A.3 se tiene que

$$X(h(\alpha_Y, \alpha_Z)) = h(\nabla_X \alpha_Y, \alpha_Z) + h(\alpha_Y, \nabla_X \alpha_Z).$$

Finalmente, como g es no degenerada, define un isomorfismo lineal entre campos y formas, con lo que se concluye. \square

3.3.2. Conexión sobre tensores

Una vez hemos definido una conexión para campos y para formas, entonces deberíamos ser capaces de definir una conexión para campos tensoriales. Para hacerlo, razonaremos como antes, imponiendo a

priori unas condiciones que nos pueden parecer razonables y que se asemejan al comportamiento de la derivación a la que estamos acostumbrados y veremos que satisfacen las propiedades de una conexión.

La primera vía es definir la derivada de tensores siguiendo la misma idea de antes. Si evaluamos el tensor, obtenemos una función diferenciable y podemos calcular su diferencial aplicándole un campo. Sin embargo, este diferencial engloba el cambio del tensor como de todas sus componentes.

Definición 3.29. Sean (M, ∇) una variedad diferenciable equipada con una conexión y $T \in T_p^r(M)$ un campo tensorial. Sea $X \in \mathfrak{X}(M)$ un campo. Entonces para toda colección de r 1-formas $\omega_1, \dots, \omega_q$ y s campos de vectores Y_1, \dots, Y_p , definimos la derivada tensorial en la dirección de X como

$$\begin{aligned} (\nabla_X T)(\omega_1, \dots, \omega_q, Y_1, \dots, Y_p) &:= X\left(T(\omega_1, \dots, \omega_q, Y_1, \dots, Y_p)\right) \\ &\quad - \sum_{i=1}^r T(\omega_1, \dots, \nabla_X \omega_i, \dots, \omega_q, Y_1, \dots, Y_p) \\ &\quad - \sum_{j=1}^s T(\omega_1, \dots, \omega_q, Y_1, \dots, \nabla_X Y_j, \dots, Y_p). \end{aligned}$$

En esta definición, las derivadas $\nabla_X \omega_i$ y $\nabla_X Y_j$ se obtienen de la conexión inducida en los fibrados cotangente y tangente, respectivamente. Para comprobar que la derivada tensorial definida como 3.29 verifica las propiedades 1-4 de la definición 3.1, véase la proposición A.4.

Observación. Esta definición es la extensión natural para tensores de tipo $(0, 1)$ y $(1, 0)$. Además, para el caso de funciones diferenciables se tiene que $\nabla_X f = X(f)$. Por tanto, la derivada tensorial es una extensión válida de la derivación que conocemos hasta ahora.

Al principio del capítulo presentamos las condiciones de torsión nula y la de compatibilidad con la métrica con la única motivación de que estas nos garantizaban la existencia y unicidad de la conexión. Sin embargo, sus nombres claramente derivaban de una interpretación geométrica. La condición de torsión nula quedó explicada con la proposición 3.25. Ahora que ya hemos definido la derivada tensorial podemos darle una interpretación a la propiedad de compatibilidad con la métrica.

Proposición 3.30. Sea (M, g, ∇) una variedad diferenciable con una métrica no degenerada y una conexión. La condición de compatibilidad con la métrica es equivalente a

$$\nabla g = 0.$$

Demostración. Sean $X, Y, Z \in \mathfrak{X}(M)$ tres campos sobre M . Entonces es fácil ver que

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \iff X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = 0 \iff \nabla_X g(Y, Z) = 0.$$

Como esto se verifica para cualesquiera tres campos, podemos escribirlo como $\nabla g = 0$. \square

Ya vimos que si una conexión es compatible con la métrica, tanto longitudes como ángulos se conservan al trasladar vectores mediante el transporte paralelo inducido. Es decir, los productos $g(X, X)$ y $g(X, Y)$ no cambian a lo largo del traslado. Como podemos expresar esta condición $\nabla g = 0$, podemos darle una interpretación geométrica análoga a la usual. Pues $\nabla g = 0$ significa que la métrica no cambia según nos movemos siguiendo el transporte paralelo inducido.

Antes de seguir avanzando, veamos la segunda forma de definir la derivada tensorial que adelantamos al principio de esta sección y veamos que es equivalente a la que acabamos de estudiar.

Definición 3.31. Sea (M, ∇) una variedad diferenciable equipada con una conexión. Sean un campo $X \in \mathfrak{X}(M)$ y $T = \omega_1 \otimes \cdots \otimes \omega_q \otimes X_1 \otimes \cdots \otimes X_p$ un tensor de tipo (p, q) . Entonces definimos la conexión para campos tensoriales como

$$\nabla_X T = \sum_{i=1}^q \omega_1 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_q \otimes X_1 \otimes \cdots \otimes X_p + \sum_{j=1}^p \omega_1 \otimes \cdots \otimes \omega_q \otimes X_1 \otimes \cdots \otimes \nabla_X X_j \otimes \cdots \otimes X_p$$

donde $\nabla_X \omega_i$ y $\nabla_X X_j$ denotan las conexiones para campos y para formas, respectivamente.

Observación. En el primer caso hemos definido la derivada a partir de la regla de Leibniz mientras que en el segundo con una análoga a la del producto.

Para ver que la derivada tensorial definida por la regla del producto es equivalente a la definida por componentes, véase lema A.5.

Corolario 3.32. Sea (M, ∇) una variedad diferenciable con una conexión. Entonces la conexión para tensores es única.

Corolario 3.33. Sea (M, ∇) una variedad diferenciable con una conexión. Entonces ∇ conmutará con cualquier contracción de índices.

Demostración. Sean dos campos $X, Y \in \mathfrak{X}(M)$ y una forma diferenciable $\omega \in \Lambda^1 M$. Entonces, por definición de la conexión sobre tensores y la linealidad de la conexión tenemos que:

$$C(\nabla_X(\omega \otimes Y)) = C((\nabla_X \omega) \otimes Y) + C(\omega \otimes \nabla_X Y) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y).$$

Por la regla de derivación de formas sabemos que $\nabla_X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y)$. Con lo que concluimos: $C(\nabla_X(\omega \otimes Y)) = \nabla_X(\omega(Y)) = \nabla_X(C(\omega \otimes Y))$. \square

El proceso de extensión del concepto de derivación a objetos matemáticos más abstractos se ha basado en imponer unas reglas que considerábamos análogas a las que ya conocíamos en la derivación habitual. Sin embargo, puede ser que exista alguna otra regla *a priori* menos intuitiva que también generalice la derivación y no sea equivalente a las anteriores. Debemos comprobar que esto no pasa y,

para hacerlo, resulta clave el resultado anterior, pues es una propiedad que ambas derivaciones conocidas comparten. De hecho, el resultado siguiente muestra que, aunque *a priori* no es una forma intuitiva de extender la definición, el que la derivada conmute con la contracción es el lenguaje natural para caracterizar las extensiones que buscamos. Es decir, toda generalización de la derivada que respete las reglas impuestas debe conmutar con la contracción.

Proposición 3.34. *Una extensión de la conexión ∇ para campos tensoriales que sea $C^\infty(M)$ -lineal en el primer argumento y conmute con toda contracción es única y coincide con la derivada tensorial definida por la regla de Leibniz.*

Demostración. Supongamos que existen dos extensiones, ∇^1 y ∇^2 , que cumplen estas propiedades y coinciden con ∇ cuando se aplican sobre campos vectoriales $\mathfrak{X}(M)$ y 1-formas $\Lambda^1 M$.

Definimos la diferencia $\Delta_X T := \nabla_X^1 T - \nabla_X^2 T$. Nótese que en el caso de campos o formas $\Delta_X T = 0$. Por otro lado, como conmutan con la contracción $C(\Delta_X T) = \Delta_X(C(T))$. Entonces, por inducción sobre el orden del tensor, es fácil ver que $C(\Delta_X T) = \Delta_X(C(T)) = 0$ para toda contracción y orden del tensor. Es decir, se trata del tensor nulo ya que todas sus contracciones son nulas. Por tanto, $\nabla^1 = \nabla^2$, la extensión es única. Por lo visto anteriormente podemos afirmar que dicha extensión es la definida anteriormente. Para ver la comprobación de la regla de Leibniz consultar el lema A.6 \square

Por tanto, con esto lo que podemos ver es que la única extensión posible de la idea de derivación es precisamente la que habíamos construido a lo largo de la sección. Además, nos da una tercera vía para definirla sin necesidad de imponer que se verifique la regla de Leibniz; basta con que conmute con todas las contracciones.

3.3.3. Conexión como tensor

Antes de concluir el estudio de la conexión es importante mencionar que podemos verla de una forma ligeramente diferente a como veníamos haciéndolo. Empecemos por el caso más sencillo. Vemos que fijando el segundo argumento de la conexión obtenemos una aplicación que en cada espacio tangente actúa como endomorfismo $C^\infty(M)$ -lineal

$$\begin{aligned} \nabla Y : \mathfrak{X}(M) &\longrightarrow \mathfrak{X}(M) \\ X &\longmapsto \nabla_X Y. \end{aligned}$$

Por tanto lo podemos ver como un campo tensorial de tipo $(1, 1)$.

Observación. Nótese que esto sólo es posible al fijar el segundo argumento de la conexión, pues por definición es $C^\infty(M)$ -lineal en el primero. Es decir, el endomorfismo ∇Y es un tensor $(1, 1)$ mientras que ∇_X o ∇ no ya que en la segunda componente verifica la regla de Leibniz.

Esta misma idea se puede generalizar al caso de tensores. Podemos considerar la conexión en el caso de tensores como

$$\nabla : \mathfrak{X}(M) \times T_q^p(M) \longrightarrow T_q^p(M).$$

Si ahora fijamos el segundo argumento tomando un tensor $T \in T_q^p(M)$ obtenemos una aplicación $C^\infty(M)$ -lineal

$$\begin{aligned} \nabla T : \mathfrak{X}(M) &\longrightarrow T_q^p(M) \\ X &\longmapsto \nabla_X T. \end{aligned}$$

Es decir, podemos ver ∇T como un tensor de tipo $(r, s+1)$ que al evaluar, o contraer, en su última entrada en los campos obtenemos la derivada en dirección de dicho campo.

3.4. Geodésicas

En este apartado dejamos a un lado el estudio de la conexión para centrarnos en un concepto clave en geometría diferencial como es el de las curvas geodésicas, cuyo desarrollo se fundamenta en la teoría de conexiones vista hasta ahora.

Definición 3.35. Llamamos *geodésica* en una variedad diferenciable M a una curva diferenciable $c : I \longrightarrow M$ parametrizada por arco cuyo campo de velocidades es paralelo a lo largo de la propia curva, es decir, que si c

$$\nabla_{\dot{c}} \dot{c} = 0.$$

Veamos con un ejemplo la importancia de la condición sobre la parametrización de la curva.

Ejemplo 3.36. Sean las curvas $c_1(t) = (t, 0, 0)$ y $c_2(t) = (t^2, 0, 0)$ definidas para t positivo. Vemos que c_1 es una geodésica ya que $\nabla_{\dot{c}_1} \dot{c}_1 = \nabla_{(1,0,0)} (1, 0, 0) = 0$. En cambio, c_2 no lo es

$$\nabla_{\dot{c}_2} \dot{c}_2 = \nabla_{(2t,0,0)} (2t, 0, 0) = 2t \nabla_{(1,0,0)} (2t, 0, 0) = 2t((2, 0, 0) + (\Gamma_{11}^1, \Gamma_{11}^2, \Gamma_{11}^3)) = (4t, 0, 0).$$

Es decir, lo que hace que una curva sea o no una geodésica no es sólo la trayectoria, sino su velocidad. Mientras que la velocidad de la primera es constante la de la segunda aumenta linealmente con el tiempo; tiene aceleración.

Observación. Nótese que si $\nabla_{\dot{c}} \dot{c} = \lambda(t)\dot{c}$ entonces existe una reparametrización $\tilde{c} = c \circ h$ tal que \tilde{c} es una geodésica. Por tanto, podríamos haber añadido este caso a la definición como en [Küh06, p. 229]. Para evitarlo, basta tomar curvas que siempre estén parametrizadas por longitud de arco.

Veamos ahora algunas de las propiedades de las geodésicas.

Proposición 3.37. Sea $c : I \rightarrow M$ una geodésica en una variedad diferenciable (M, g) . Entonces:

1. La norma $\|\dot{c}(s)\|$ es constante.
2. O bien c es regular o bien es una curva constante.
3. Sea $\gamma = c \circ h$ una reparametrización de c donde $h : J \rightarrow I$. Entonces γ es una geodésica si y solo si $h(s) = as + b$ para constantes $a, b \in \mathbb{R}$.
4. En particular, si $v \neq 0$, la geodésica $\gamma(s) = c\left(\frac{s}{v}\right)$ reparametriza c por la longitud de arco.

Demostración. 1. Derivando obtenemos que $\frac{d}{dt}\|\dot{c}(t)\|^2 = \frac{d}{dt}g(\dot{c}(t), \dot{c}(t)) = 2g(\dot{c}(t), \nabla_{\dot{c}}\dot{c}(t))$. Como $\nabla_{\dot{c}}\dot{c}(t) = 0$, se concluye.

2. Si c no es constante, existe t_0 tal que $\|\dot{c}(t_0)\| > 0$ y como la norma $\|\dot{c}(t)\|$ es constante, se concluye.
3. Es fácil ver que la derivada segunda de γ es

$$\nabla_{\dot{\gamma}}\dot{\gamma}(s) = b'(s)\nabla_{\dot{c}}(\dot{c}(h(s))b'(s)) = b'(s)((b'(s))^2\nabla_{\dot{c}}\dot{c} + b''(s)\dot{c}(h(s)))$$

Teniendo en cuenta que c es geodésica tenemos que

$$\nabla_{\dot{\gamma}}\dot{\gamma} = b'(s)b''(s)\dot{c}(h(s)).$$

Para que γ sea geodésica, debe cumplirse $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$, es decir que $b'(s)b''(s)\dot{c}(h(s)) = 0$. Recordemos que c es una curva regular entonces se deduce que

$$b''(s) = 0 \iff b(s) = as + b, \quad a, b \in \mathbb{R}$$

4. Por lo anterior, la curva $\gamma(s) = c\left(\frac{s}{v}\right)$ sigue siendo una geodésica. Además, si calculamos su norma tenemos que $\|\dot{\gamma}(s)\| = \|\dot{c}\left(\frac{s}{v}\right)\frac{1}{v}\| = \frac{v}{v} = 1$.

□

Proposición 3.38. Sea $\varphi : (M, g) \rightarrow (N, h)$ una isometría. Entonces, para cualesquiera campos vectoriales $X, Y \in \mathfrak{X}(M)$, se tiene

$$d\varphi(\nabla_X^g Y) = \nabla_{d\varphi(X)}^h (d\varphi Y),$$

donde ∇^g, ∇^h son conexiones compatibles con las respectivas métricas.

Demostración. Por compatibilidad con la métrica, para todo $Z \in \mathfrak{X}(M)$ se cumple

$$X(g(Y, Z)) = g(\nabla_X^g Y, Z) + g(Y, \nabla_X^g Z).$$

De igual manera en la variedad N se tiene

$$d\varphi(X)\left(b(d\varphi Y, d\varphi Z)\right) = b\left(\nabla_{d\varphi(X)}^b(d\varphi Y), d\varphi Z\right) + b\left(d\varphi Y, \nabla_{d\varphi(X)}^b(d\varphi Z)\right).$$

Por el lema A.7, podemos igualar ambas dos expresiones:

$$g(\nabla_X^g Y, Z) + g(Y, \nabla_X^g Z) = b\left(\nabla_{d\varphi(X)}^b(d\varphi Y), d\varphi Z\right) + b\left(d\varphi Y, \nabla_{d\varphi(X)}^b(d\varphi Z)\right).$$

Como φ es una isometría tenemos que $g(Y, Z) = b(d\varphi Y, d\varphi Z)$ y por tanto:

$$b\left(d\varphi(\nabla_X^g Y), d\varphi Z\right) + b\left(d\varphi Y, d\varphi(\nabla_X^g Z)\right) = b\left(\nabla_{d\varphi(X)}^b(d\varphi Y), d\varphi Z\right) + b\left(d\varphi Y, \nabla_{d\varphi(X)}^b(d\varphi Z)\right).$$

Finalmente, sabiendo que b es no degenerada y que la igualdad anterior se verifica para todo $Z \in \mathfrak{X}(M)$, se concluye:

$$d\varphi(\nabla_X^g Y) = \nabla_{d\varphi(X)}^b(d\varphi Y).$$

□

Corolario 3.39. *Las geodésicas son invariantes bajo isometrías.*

Demostración. Particularizando la proposición anterior para $X = Y = \dot{\gamma}(t)$ con γ una geodésica en M obtenemos

$$d\varphi\left(\nabla_{\dot{\gamma}(t)}^g \dot{\gamma}(t)\right) = \nabla_{d\varphi(\dot{\gamma}(t))}^b\left(d\varphi(\dot{\gamma}(t))\right) \implies \nabla_{d\varphi(\dot{\gamma}(t))}^b\left(d\varphi(\dot{\gamma}(t))\right) = 0.$$

Como $(\varphi \circ \gamma)'(t) = d\varphi(\dot{\gamma}(t))$, se concluye que

$$\nabla_{(\varphi \circ \gamma)'(t)}^b(\varphi \circ \gamma)'(t) = 0.$$

□

Ahora que ya sabemos qué es una geodésica y algunas de sus propiedades más importantes, veamos cómo podemos calcularlas.

Proposición 3.40. *Sea (M, ∇) una variedad diferenciable equipada con una conexión. Entonces existe una única geodésica maximal $c(t) = (c^1(t), \dots, c^m(t))$ tal que $c(0) = p$ y $\dot{c}(0) = v$ satisface el siguiente sistema de ecuaciones diferenciales*

$$\frac{d^2 c^k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{dc^i}{dt} \frac{dc^j}{dt} = 0 \quad \forall k = 1, \dots, m. \quad (3.12)$$

Demostración. Sustituyendo en la expresión local de la derivada covariante 3.7 la condición $\nabla_{\dot{c}}\dot{c} = 0$, tenemos que

$$\nabla_{\dot{c}}\dot{c} = \sum_k \left(\sum_i \dot{c}^i \frac{\partial \dot{c}^k}{\partial x^i} + \sum_{ij} \dot{c}^i \dot{c}^j \Gamma_{ij}^k(c(t)) \right) \frac{\partial}{\partial x^k} = 0.$$

Viendo que por la regla de la cadena, tenemos que $\sum_i \dot{c}^i \frac{\partial \dot{c}^k}{\partial x^i} = \frac{d^2 c^k}{dt^2}$, concluimos

$$\frac{d^2 c^k}{dt^2} + \sum_{ij} \Gamma_{ij}^k \frac{dc^i}{dt} \frac{dc^j}{dt} = 0 \quad \forall k = 1, \dots, m.$$

Por el teorema de existencia y unicidad de ecuaciones diferenciales ordinarias tenemos que existe una única curva tal que $c(0) = p$ y $\dot{c}(t) = v$ que satisface el sistema anterior. Es decir, dados $c(0) = p$ y $\dot{c}(t) = v$ existe una única geodésica definida sobre I que las satisface.

Veamos cómo construir la geodésica maximal. Consideremos otra geodésica $\alpha : J \rightarrow M$ tal que $\alpha(0) = c(0)$ y $\alpha'(0) = \dot{c}(0)$. Definamos ahora el conjunto de tiempos en los que ambas geodésicas coinciden; es decir,

$$S = \{t \in I \cap J : c(t) = \alpha(t), \dot{c}(t) = \alpha'(t)\}.$$

En primer lugar, podemos afirmar que es no vacío ya que $0 \in S$. Además, es un conjunto abierto por el teorema de la unicidad de soluciones de los sistemas de ecuaciones diferenciales ordinarias. También es un conjunto cerrado. Sea la función diferenciable $f(t) := c(t) - \alpha(t)$, entonces podemos redefinir S como $S = f^{-1}(0) \cap (f')^{-1}(0)$. Es decir, por continuidad de f y f' , S es una intersección de dos cerrados; por tanto, cerrado. De esta forma tenemos que $S = I \cap J$. Para terminar, tomamos la siguiente unión de abiertos $\mathcal{I} := \bigcup_{\lambda \in \Lambda} \{t \in I_\lambda \mid c_\lambda : I_\lambda \rightarrow M \text{ es una geodésica tal que } c_\lambda(0) = p, c'_\lambda(0) = v\}$ y definimos la geodésica maximal como

$$\begin{aligned} \tilde{c} : \mathcal{I} &\longrightarrow M \\ t &\longmapsto \tilde{c}(t) := c_\lambda(t) \text{ tal que } t \in I_\lambda. \end{aligned}$$

□

Observación. La unicidad de la geodésica maximal también se puede ver directamente como un corolario del teorema 3.14.

Calculemos ahora las geodésicas en algunas variedades euclídeas conocidas.

Ejemplo 3.41 (\mathbb{R}^m). Consideremos el espacio \mathbb{R}^m con el producto escalar habitual. Como los símbolos de Christoffel son nulos, las geodésicas verifican $\frac{d^2 c^k}{dt^2} = 0 \quad \forall k = 1, \dots, m$. Es decir, las geodésicas en este espacio son rectas.

Ejemplo 3.42 (Cilindro). Es fácil ver que la métrica inducida por la usual de \mathbb{R}^3 sobre el cilindro es $ds^2 = r^2 d\theta^2 + dz^2$. De nuevo, tenemos que todos los símbolos de Christoffel son nulos y por tanto las geodésicas son rectas en estas coordenadas:

$$c(t) = (\theta(t), z(t)) = (a + bt, c + dt).$$

Esto nos da tres tipos de curvas: rectas perpendiculares al plano XY si $b = 0$, circunferencias si $d = 0$ y hélices si $b \neq 0 \neq d$.

Estas curvas además nos permiten definir un sistema de coordenadas natural que es muy útil a la hora de trabajar con variedades diferenciables; las coordenadas normales. Para definir las y estudiarlas primero deberemos presentar la aplicación exponencial.

Definición 3.43. Sea (M, ∇) una variedad diferenciable equipada con una conexión. Sean $p \in M$ un punto fijado en la variedad y $X \in T_p M$ un vector tangente en p y unitario. Sea c_X^p la única geodésica parametrizada por longitud de arco que pasa por p con vector tangente X en dicho punto. Entonces llamamos *aplicación exponencial* en el punto p a

$$\begin{aligned} \exp_p : U \subset T_p M &\longrightarrow M \\ (p, sX) &\longmapsto c_X^p(s). \end{aligned}$$

Para diferentes puntos p podemos definir de manera análoga

$$\begin{aligned} \exp : \tilde{U} \subset TM &\longrightarrow M \\ (p, sX) &\longmapsto \exp_p(sX) = c_X^p(s). \end{aligned}$$

Intuitivamente lo que hace esta aplicación es transformar las rectas que pasan por el origen en $T_p M$ en las geodésicas que pasan por p . Se puede probar que esta aplicación define una carta de la variedad.

Lema 3.44. La aplicación \exp_p restringida a un cierto entorno $U \subset T_p M$ del origen, es un difeomorfismo.

Demostración. Se puede encontrar en [Küho6, p. 285-286] □

Corolario 3.45. La aplicación inversa \exp_p^{-1} define una carta de la variedad.

Definición 3.46. Sean (X_1, \dots, X_n) una base ortonormal en $T_p M$ y \exp_p la aplicación definida en un entorno del origen $U \subset T_p M$. Llamamos *coordenadas normales* a las coordenadas definidas por dicha aplicación. Los elementos de la base de estas coordenadas se definen como

$$\partial_i|_{c_{X_i}^p(s)} = d(\exp_p)p(X_i).$$

Para cerrar el capítulo, demostraremos una generalización del resultado clásico como las fórmulas de Frenet-Serret para curvas, en este caso estarán contenidas en una variedad cualquiera.

Definición 3.47. Sea curva parametrizada regular $c : I \rightarrow M$ con M . Se dice curva de Frenet si está parametrizada por arco y, para cada $s \in I$, los siguientes vectores

$$\left\{ \nabla_{\dot{c}}^i \dot{c} := \overbrace{\nabla_{\dot{c}} \nabla_{\dot{c}} \cdots \nabla_{\dot{c}}}^{(i)} \dot{c} \right\}_{i=1}^{m-1}.$$

son linealmente independientes.

Proposición 3.48. Sea $c : I \rightarrow M$ una curva de Frenet que no es geodésica en una variedad diferenciable equipada con una métrica g no degenerada, y sea ∇ la única conexión inducida por g . Entonces existe una base ortonormal en cada punto de la curva tal que verifica las siguientes ecuaciones

$$\begin{cases} \nabla_{e_1} e_1(s) = \kappa_1(s) e_2(s), \\ \nabla_{e_1} e_j(s) = -\kappa_{j-1}(s) e_{j-1}(s) + \kappa_j e_{j+1} & 2 \leq j \leq m-1, \\ \nabla_{e_1} e_m(s) = -\kappa_{m-1} e_{m-1}(s) \end{cases}$$

donde $e_1 := \nabla_{\dot{c}} \dot{c}$ y $\kappa_i(s)$ son funciones diferenciables. A las funciones κ_i se les conoce curvatura i -ésimas. En el caso $i = m$ también se le llama torsión.

Demostración. Como c es una curva parametrizada por arco tenemos que $g(\dot{c}(s), \dot{c}(s)) = 1 \quad \forall s \in I$. Denotemos el vector tangente como $e_1 := \dot{c}$ Entonces por la condición de compatibilidad con la métrica la conexión tenemos que

$$\frac{d}{ds}(g(e_1, e_1)) = 0 \implies 2g(\nabla_{e_1} e_1, e_1) = 0.$$

Por tanto, $\nabla_{e_1} e_1 \perp e_1$. Ahora normalicemos este nuevo vector $e_2 := \frac{\nabla_{e_1} e_1}{\|\nabla_{e_1} e_1\|} = \frac{\nabla_{e_1} e_1}{\kappa_1}$ y hagamos un razonamiento similar al anterior. En primer lugar, al ser un vector unitario a lo largo de toda la trayectoria tenemos

$$\frac{d}{ds}(g(e_2, e_2)) = 0 \implies 2g(\nabla_{e_1} e_2, e_2) = 0.$$

Así que de nuevo tenemos $\nabla_{e_1} e_2 \perp e_2$. Por otro lado, tenemos de antes que $e_1 \perp e_2$ luego

$$\frac{d}{ds}(g(e_2, e_1)) = 0 \implies g(\nabla_{e_1} e_2, e_1) = -g(e_2, \nabla_{e_1} e_1) = -\kappa_1.$$

Por tanto, podemos escribir el nuevo vector como $\nabla_{e_1} e_2 = -\kappa_1 e_1 + W_3$, donde W_3 es ortogonal a e_1 y a e_2 , así que lo podemos reescribir como

$$\nabla_{e_1} e_2 = -\kappa_1 e_1 + \kappa_2 e_3 \quad \text{tal que } g(e_3, e_3) = 1.$$

Ahora podemos razonar de manera recursiva. Supongamos que para $2 \leq i \leq m-1$ (e_1, \dots, e_i) es una base ortonormal tal que $\nabla_{e_1} e_i = -\kappa_{i-1} e_{i-1} + \kappa_i e_{i+1}$. De la misma forma que antes tenemos que

$$\frac{d}{ds}(g(e_i, e_i)) = 0 \implies 2g(\nabla_{e_1} e_i, e_i) = 0 \implies \nabla_{e_1} e_i \perp e_i.$$

Y además para $j < i$ se verifica

$$\frac{d}{ds}(g(e_i, e_j)) = 0 \implies g(\nabla_{e_1} e_i, e_j) = -g(e_i, \nabla_{e_1} e_j) = \begin{cases} 0, & \text{si } j \neq i-1, \\ -\kappa_i, & \text{si } j = i-1. \end{cases}$$

De esta forma tenemos que $\nabla_{e_1} e_i = -\kappa_{i-1} e_{i-1} + W_{i+1}$ donde $W_{i+1} \in (\text{span}\{e_1, \dots, e_{i-2}, e_i\})^\perp$. Definamos entonces $e_{i+1} := \frac{W_{i+1}}{\kappa_{i+1}}$. Así, finalmente obtenemos

$$\nabla_{e_1} e_i = -\kappa_{i-1} e_{i-1} + \kappa_{i+1} e_{i+1} \quad 2 \leq i \leq m-1$$

Para terminar debemos estudiar el caso $i = m$. Siguiendo el mismo razonamiento podemos ver que $\nabla_{e_1} e_m \perp e_m$. Derivando

$$\frac{d}{ds}(g(e_m, e_j)) = 0 \implies g(\nabla_{e_1} e_m, e_j) = -g(e_m, \nabla_{e_1} e_j) = \begin{cases} 0, & \text{si } j \neq m-1, \\ -\kappa_{m-1}, & \text{si } j = m-1. \end{cases}$$

Podemos volver a escribir el vector como ya hicimos $\nabla_{e_1} e_m = -\kappa_{m-1} e_{m-1} + W_{m+1}$ donde $W_{m+1} \in (\text{span}\{e_1, \dots, e_{m-2}, e_{m-1}, e_m\})^\perp$. Es decir que $W_{m+1} = 0$ y por tanto $\nabla_{e_1} e_m = -\kappa_{m-1} e_{m-1}$.

De esta forma, hemos obtenido las ecuaciones de Frenet-Serret en el caso general de una variedad diferenciable

$$\begin{cases} \nabla_{e_1} e_1(s) = \kappa_1(s) e_2(s), \\ \nabla_{e_1} e_j(s) = -\kappa_{j-1}(s) e_{j-1}(s) + \kappa_j(s) e_{j+1}(s) & 2 \leq j \leq m-1 \\ \nabla_{e_1} e_m(s) = -\kappa_{m-1}(s) e_{m-1}(s) \end{cases}$$

□

Estas ecuaciones, al igual que en el caso de \mathbb{R}^m también las podemos escribir para curvas que no estén parametrizadas por longitud de arco.

Proposición 3.49. *Sea (M, g) una variedad métrica con g no degenerada. Sea $c : I \rightarrow M$ una curva diferenciable regular que no es geodésica tal que $\{\nabla_c^i \dot{c}\}_{i=1}^{m-1}$ son linealmente independientes. Sea ∇ la única conexión inducida por g . Entonces existe una base ortonormal en cada punto de la curva tal que verifica*

las siguientes ecuaciones

$$\begin{cases} \nabla_{e_1} e_1(s) = \frac{\kappa_1(s)}{v} e_2(s), \\ \nabla_{e_1} e_j(s) = -\frac{\kappa_{j-1}(s)}{v} e_{j-1}(s) + \frac{\kappa_j}{v} e_{j+1} & 2 \leq j \leq m-1, \\ \nabla_{e_1} e_m(s) = -\frac{\kappa_{m-1}}{v} e_{m-1}(s) \end{cases}$$

donde $\|\dot{c}\| = v \in C^\infty$.

Demostración. Usando la misma notación que en la proposición anterior definimos el vector normalizado $e_1 := \frac{\dot{c}}{v}$, y siguiendo el mismo razonamiento que antes

$$\frac{d}{ds}(g(e_1, e_1)) = 0 \implies 2g(\nabla_{\dot{c}} e_1, e_1) = 0.$$

Por otro lado, por las propiedades de la conexión tenemos que

$$\nabla_{\dot{c}} e_1 = \frac{1}{v} \nabla_{\dot{c}} \dot{c} - \frac{v'}{v^2} \dot{c} = \frac{1}{v} \nabla_{\dot{c}} \dot{c} - \frac{v'}{v} e_1$$

Sin embargo, hemos visto que $\nabla_{\dot{c}} e_1$ no tiene ninguna componente en la dirección de e_1 . Eso quiere decir que podemos escribir $\nabla_{\dot{c}} \dot{c} = \kappa_1 e_2 + v' e_1$, donde $e_1 \perp e_2$ y κ_1 es una función diferenciable, ya que así se verifica que

$$\nabla_{\dot{c}} e_1 = \frac{1}{v} \nabla_{\dot{c}} \dot{c} - \frac{v'}{v} e_1 = \frac{1}{v} (\kappa_1 e_2 + v' e_1) - \frac{v'}{v} e_1 = \frac{\kappa_1}{v} e_2$$

Ahora, de manera análoga a la proposición anterior tenemos que $\nabla_{e_1} e_2 \perp e_2$ y

$$\frac{d}{ds}(g(e_2, e_1)) = 0 \implies g(\nabla_{e_1} e_2, e_1) = -g(e_2, \nabla_{e_1} e_1) = -\frac{\kappa_1}{v}$$

Por tanto

$$\nabla_{e_1} e_2 = -\frac{\kappa_1}{v} e_1 + \frac{\kappa_2}{v} e_3,$$

donde e_3 es unitario tal que $e_3 \perp e_1$ y $e_3 \perp e_2$ y κ_2 es función diferenciable.

Razonando de manera recursiva como hicimos en la proposición anterior llegamos a que

$$\frac{d}{ds}(g(e_i, e_j)) = 0 \implies g(\nabla_{e_i} e_j, e_j) = -g(e_i, \nabla_{e_i} e_j) = \begin{cases} 0, & \text{si } j \neq i-1, \\ -\frac{\kappa_{i-1}}{v}, & \text{si } j = i-1, \end{cases}$$

para $2 \leq i \leq m$. Así, para $i \neq m$ podemos escribir los vectores como

$$\nabla_{e_i} e_i = -\frac{\kappa_{i-1}}{v} e_{i-1} + W_{i+1} = -\frac{\kappa_{i-1}}{v} e_{i-1} + \frac{\kappa_{i+1}}{v} e_{i+1}$$

donde $W_{i+1} \in (\text{span}\{e_1, \dots, e_{i-2}, e_i\})^\perp$. En el caso $i = m$, tenemos que $W_{m+1} = 0$ y por tanto $\nabla_{e_1} e_m = -\kappa_{m-1} e_{m-1}$. Finalmente concluimos que

$$\begin{cases} \nabla_{\dot{c}} e_1(s) = \frac{\kappa_1(s)}{v} e_2(s), \\ \nabla_{\dot{c}} e_j(s) = -\frac{\kappa_{j-1}(s)}{v} e_{j-1}(s) + \frac{\kappa_j}{v} e_{j+1}, & 2 \leq j \leq m-1, \\ \nabla_{\dot{c}} e_m(s) = -\frac{\kappa_{m-1}}{v} e_{m-1}(s). \end{cases}$$

□

4

Tensor de Curvatura

En este capítulo nos centraremos especialmente en la construcción del tensor de curvatura, que haremos de dos formas diferentes. Cabe resaltar el primer enfoque en el que vemos cómo surge de manera natural al hacer el límite de la diferencia entre un vector y su imagen por el transporte paralelo a lo largo de lazos infinitesimalmente pequeños; lo que hemos llamado “holonomía puntual”. Esta construcción diferente a la usual pone de relieve cómo la curvatura emerge a partir de las propiedades que tienen los flujos y los campos debido a su estructura de álgebra de Lie. Así, lo que entendemos *a priori* como una propiedad geométrica es en realidad una manifestación de una estructura algebraica más profunda. Una vez visto esto, comenzando con un ejemplo con superficies en \mathbb{R}^3 estudiaremos el porqué el tensor obtenido extiende la idea de curvatura. También daremos especial importancia al tensor de Ricci y la curvatura escalar. Haciendo énfasis en la interpretación del tensor, el cuál nos informa cómo se ve afectado un volumen según lo trasladamos por la variedad debido a la curvatura de ésta.

Entender bien estos conceptos será fundamental a la hora de trabajar en relatividad general, ya que tienen un papel central a la hora de relacionar la distribución de materia con la curvatura del espacio-tiempo.

4.1. Curvatura como “holonomía puntual”

En el capítulo anterior hemos construido toda la teoría de transporte paralelo y así poder trasladar vectores a lo largo de curvas sobre la variedad a lo largo de curvas. Ahora cabe preguntarnos qué pasa si en vez de al trasladarlo a otro punto, lo movemos a lo largo de la variedad para luego volver al inicio, siguiendo una curva cerrada o lazo.

Es razonable pensar que obtenemos debe ser el mismo, pues toda transformación que se aplica al vector debe deshacerse eventualmente al tratarse de un camino cerrado. Aunque este razonamiento es correcto en el caso de un plano, falla en otras variedades. El siguiente caso más es la esfera. Si tomamos

un triángulo esférico con un vértice en el polo y los otros dos sobre el Ecuador y trasladamos un vector a lo largo de esta curva cerrada, es fácil verificar que no obtenemos el original. Es decir, no siempre¹, obtenemos el vector original de vuelta. A este fenómeno se le conoce como holonomía y se estudia con el grupo de holonomía.

Definición 4.1. Sea (M, ∇) una variedad diferenciable equipada con una conexión. Sea $\gamma : [0, 1] \rightarrow M$ un lazo basado en $p \in M$, es decir una curva cerrada que empieza y termina en un punto p . Entonces llamamos grupo de holonomía en p , denotado $\text{Hol}_p(\nabla)$, al conjunto de todas las transformaciones posibles que surgen de recorrer todos los lazos basados en p .

Observación. El concepto de holonomía es global, pues se pueden tomar lazos tan grandes como se quieran. Podemos considerar su versión local tomando lazos infinitesimalmente pequeños.

Observación. Se puede demostrar que el grupo de holonomía es un grupo de Lie.

Calculemos ahora Para empezar, estudiaremos qué sucede al la diferencia entre un vector y el que se obtiene al transportarlo por un lazo infinitesimalmente pequeño.

Proposición 4.2. Sean (M, ∇) una variedad equipada con una conexión y $X, Y \in \mathfrak{X}(M)$ dos campos que conmutan, es decir que $[X, Y] = 0$. Sea $\varepsilon > 0$ y $\gamma_\varepsilon = \phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_\varepsilon^Y \circ \phi_\varepsilon^X(p)$ una familia de lazos basados en $p \in M$. Entonces se cumple que :

$$\lim_{\varepsilon \rightarrow 0} \frac{P_0^{4\varepsilon}(\gamma_\varepsilon)(v) - v}{\varepsilon^2} = \nabla_Y \nabla_X v - \nabla_X \nabla_Y v,$$

donde $v \in T_p M$.

Demostración. En primer lugar hay que notar que como los campos X, Y conmutan el lazo se cierra exactamente. Así, el transporte paralelo total alrededor del lazo γ_ε es:

$$P_0^{4\varepsilon}(\gamma_\varepsilon)(v) = P_0^{-\varepsilon}(\phi^Y) \circ P_0^{-\varepsilon}(\phi^X) \circ P_0^\varepsilon(\phi^Y) \circ P_0^\varepsilon(\phi^X)$$

Sea $v \in T_p M$, calculemos $T(\varepsilon)v$ paso a paso usando el desarrollo de Taylor que vimos en 3.20:

■ **Primer tramo:** $P_0^\varepsilon(\phi^X)(v)$:

$$P_0^\varepsilon(\phi^X)(v) = v + \varepsilon \nabla_X v + \frac{1}{2} \varepsilon^2 \nabla_X^2 v + o(\varepsilon^2).$$

¹En realidad la mayoría de las situaciones.

- **Segundo tramo:** $P_0^\varepsilon(\phi^Y) \circ P_0^\varepsilon(\phi^X)(v)$:

Sea $w_1 = P_0^\varepsilon(\phi^X)(v)$. Entonces:

$$P_0^\varepsilon(\phi^Y)(w_1) = w_1 + \varepsilon \nabla_Y w_1 + \frac{1}{2} \varepsilon^2 \nabla_Y^2 w_1 + o(\varepsilon^2).$$

Sustituyendo:

$$P_0^\varepsilon(\phi^Y)(w_1) = v + \varepsilon(\nabla_X v + \nabla_Y v) + \varepsilon^2 \left(\nabla_Y \nabla_X v + \frac{1}{2} \nabla_X^2 v + \frac{1}{2} \nabla_Y^2 v \right) + o(\varepsilon^2).$$

- **Tercer tramo:** $P_0^{-\varepsilon}(\phi^X) \circ P_0^\varepsilon(\phi^Y) \circ P_0^\varepsilon(\phi^X)(v)$:

Sea $w_2 = P_0^\varepsilon(\phi^Y) \circ P_0^\varepsilon(\phi^X)(v)$, de esta forma a partir del desarrollo de Taylor

$$P_0^{-\varepsilon}(\phi^X)(w_2) = w_2 - \varepsilon \nabla_X w_2 + \frac{1}{2} \varepsilon^2 \nabla_X^2 w_2 + o(\varepsilon^2),$$

tenemos que

$$P_0^{-\varepsilon}(\phi^X)(w_2) = v + \varepsilon \nabla_Y v + \varepsilon^2 \left(\nabla_Y \nabla_X v - \nabla_X \nabla_Y v + \frac{1}{2} \nabla_Y^2 v \right) + o(\varepsilon^2).$$

- **Cuarto tramo:** Sea $w_3 = P_0^{-\varepsilon}(\phi^X) \circ P_0^\varepsilon(\phi^Y) \circ P_0^\varepsilon(\phi^X)(v)$, entonces de la misma forma que en los pasos anteriores llegamos a

$$P_0^{4\varepsilon}(\gamma_\varepsilon)(v) = P_0^{-\varepsilon}(\phi^Y)(w_3) = v + \varepsilon^2 (\nabla_Y \nabla_X v - \nabla_X \nabla_Y v) + o(\varepsilon^2).$$

Despejando y tomando el límite llegamos a:

$$\lim_{\varepsilon \rightarrow 0} \frac{P_0^{4\varepsilon}(\gamma_\varepsilon)(v) - v}{\varepsilon^2} = \nabla_Y \nabla_X v - \nabla_X \nabla_Y v$$

□

Esto lo hemos hecho para campos que conmutan, veamos qué pasa si tomamos un par de campos cualquiera. Para ello primero debemos notar que si dos campos no conmutan, entonces el lazo no tiene por qué cerrar. Entonces para calcular el límite de antes necesitaremos saber cuánto difiere de ser un lazo. Para ello, tenemos el siguiente resultado.

Proposición 4.3. Sean X, Y campos vectoriales suaves sobre una variedad diferenciable. Sea ϕ_t^X el flujo de X y ϕ_t^Y el flujo de Y . Entonces,

$$\phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_{\varepsilon}^Y \circ \phi_{\varepsilon}^X(p) = \phi^{-\varepsilon^2[X, Y]}(p) + O(\varepsilon^3),$$

para todo punto p , donde $[X, Y]$ es el corchete de Lie de X y Y .

Demostración. Una forma de estudiar los flujos de un campo es estudiar su acción sobre funciones diferenciables. Es decir, evaluaremos cómo actúa esta composición sobre una función test $f \in C^\infty(M)$. De esta manera podemos comparar la composición de flujos con el flujo del corchete de Lie usando desarrollos de Taylor, que escribiremos de la siguiente forma:

$$f(\phi_{\varepsilon}^X(p)) = f(p) + \varepsilon X(f)(p) + \frac{\varepsilon^2}{2} X^2(f)(p) + O(\varepsilon^3).$$

Ahora, lo que vamos a hacer es ir componiendo cada uno de los tramos, como hemos hecho en la prueba anterior, para evaluar f y desarrollar en serie de Taylor. Así, componiendo con el flujo de Y

$$\begin{aligned} f(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p))) &= f(\phi_{\varepsilon}^X(p)) + \varepsilon Y(f)(\phi_{\varepsilon}^X(p)) + \frac{\varepsilon^2}{2} Y^2(f)(\phi_{\varepsilon}^X(p)) + O(\varepsilon^3) \\ &= f(p) + \varepsilon X(f)(p) + \frac{\varepsilon^2}{2} X^2(f)(p) \\ &\quad + \varepsilon (Y(f)(p) + \varepsilon X(Y(f))(p)) + \frac{\varepsilon^2}{2} Y^2(f)(p) + O(\varepsilon^3) \\ &= f(p) + \varepsilon(X + Y)(f)(p) + \frac{\varepsilon^2}{2} (X^2 + 2XY + Y^2)(f)(p) + O(\varepsilon^3). \end{aligned}$$

Ahora aplicamos el flujo inverso de X :

$$f(\phi_{-\varepsilon}^X(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p)))) = f(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p))) - \varepsilon X(f)(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p))) + \frac{\varepsilon^2}{2} X^2(f)(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p))) + O(\varepsilon^3).$$

De manera análoga al anterior obtenemos:

$$f(\phi_{-\varepsilon}^X(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p)))) = f(p) + \varepsilon Y(f)(p) + \frac{\varepsilon^2}{2} (2XY + Y^2 - 2YX)(f)(p) + O(\varepsilon^3).$$

Finalmente, aplicamos el flujo inverso de Y llegamos a la expresión que buscamos:

$$\begin{aligned} f(\phi_{-\varepsilon}^Y(\phi_{-\varepsilon}^X(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p)))))) &= f(\phi_{-\varepsilon}^X(\phi_{\varepsilon}^Y(\phi_{\varepsilon}^X(p)))) - \varepsilon Y(f)(p) + \frac{\varepsilon^2}{2} Y^2(f)(p) + O(\varepsilon^3) \\ &= f(p) + \varepsilon^2[X, Y](f)(p) + O(\varepsilon^3). \end{aligned}$$

Como esto se cumple para toda función diferenciable f , podemos concluir que

$$\phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_{\varepsilon}^Y \circ \phi_{\varepsilon}^X(p) = \phi^{-\varepsilon^2[X, Y]}(p) + O(\varepsilon^3).$$

□

Es decir, cuando los campos no conmutan la diferencia entre el punto inicial y final de la trayectoria viene dado por $\phi^{-\varepsilon^2[X,Y]}$.

Corolario 4.4. Sean $X, Y \in \mathfrak{X}(M)$ dos campos cualesquiera, entonces la curva $\gamma_\varepsilon = \phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_\varepsilon^Y \circ \phi_\varepsilon^X \circ \phi^{\varepsilon^2[X,Y]}(p)$ es una curva cerrada.

Este resultado se puede obtener de forma inmediata a partir de una fórmula conocida de álgebra no conmutativa.

Proposición 4.5 (Fórmula de Baker-Campbell-Hausdorff). Sean G un grupo de Lie y \mathfrak{g} su álgebra de Lie. Sean $X, Y \in \mathfrak{g}$. Entonces existe $Z \in \mathfrak{g}$ tal que

$$e^X e^Y = e^Z,$$

donde Z se expresa como una serie en X, Y y sus corchetes de Lie:

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \dots$$

Aquí, e^X denota la aplicación exponencial que relaciona \mathfrak{g} con G .

Demostración. En el libro [AB12] se puede ver este resultado además de un estudio sobre su evolución a lo largo de la Historia. \square

En A.10 se puede ver una demostración alternativa de la proposición 4.3 empleando el resultado anterior.

Observación. La estructura como álgebra de Lie de los campos es la causa del resultado anterior. No hay ninguna hipótesis algebraica.

Con esto en mente ya sí podemos calcular la holonomía en el caso general para campos que no conmutan.

Proposición 4.6. Sea (M, ∇) una variedad diferenciable con conexión y $X, Y \in \mathfrak{X}(M)$ dos campos vectoriales tales que $[X, Y] \neq 0$. Sea $\varepsilon > 0$ y definamos la curva cerrada

$$\gamma_\varepsilon = \phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_\varepsilon^Y \circ \phi_\varepsilon^X \circ \phi^{\varepsilon^2[X,Y]}(p),$$

basada en $p \in M$. Entonces, para todo $v \in T_p M$ se cumple que

$$\lim_{\varepsilon \rightarrow 0} \frac{P_0^{5\varepsilon}(\gamma_\varepsilon)(v) - v}{\varepsilon^2} = \nabla_Y \nabla_X v - \nabla_X \nabla_Y v - \nabla_{[X,Y]} v.$$

Demostración. Por el corolario 4.4 sabemos que γ_ε así definida efectivamente es una curva cerrada. Por tanto, podemos calcular la holonomía por este lazo. Así, el transporte paralelo total a lo largo de γ_ε es

$$P_0^{5\varepsilon}(\gamma_\varepsilon) = P_{4\varepsilon}^{5\varepsilon}(\phi^{\varepsilon^2[X,Y]}) \circ P_{3\varepsilon}^{4\varepsilon}(\phi_{-\varepsilon}^Y) \circ P_{2\varepsilon}^{3\varepsilon}(\phi_{-\varepsilon}^X) \circ P_\varepsilon^{2\varepsilon}(\phi_\varepsilon^Y) \circ P_0^\varepsilon(\phi_\varepsilon^X).$$

En la demostración de 4.2 vimos que el transporte paralelo a lo largo del lazo sin cerrar, es decir sin considerar la corrección que debemos añadir verifica

$$P_0^{4\varepsilon}(v) = v + \varepsilon^2(\nabla_Y \nabla_X v - \nabla_X \nabla_Y v) + o(\varepsilon^2).$$

Sabiendo esto, lo único que hay que hacer es componer con el transporte paralelo a lo largo de $\phi^{\varepsilon^2[X,Y]}$:

$$P_{4\varepsilon}^{5\varepsilon}(\phi^{\varepsilon^2[X,Y]})(v) = v - \varepsilon^2 \nabla_{[X,Y]} v + o(\varepsilon^2).$$

Así, para calcular el transporte paralelo completo basta con hacer la siguiente composición:

$$\begin{aligned} P_0^{5\varepsilon}(\gamma_\varepsilon)(v) &= P_{4\varepsilon}^{5\varepsilon}(\phi^{\varepsilon^2[X,Y]})(P_0^{4\varepsilon}(v)) \\ &= (\text{Id} - \varepsilon^2 \nabla_{[X,Y]} + o(\varepsilon^2))(v + \varepsilon^2(\nabla_Y \nabla_X v - \nabla_X \nabla_Y v) + o(\varepsilon^2)) \\ &= v + \varepsilon^2(\nabla_Y \nabla_X v - \nabla_X \nabla_Y v - \nabla_{[X,Y]} v) + o(\varepsilon^2). \end{aligned}$$

Finalmente, reorganizando la expresión y tomando el límite concluimos:

$$\lim_{\varepsilon \rightarrow 0} \frac{P_0^{5\varepsilon}(\gamma_\varepsilon)(v) - v}{\varepsilon^2} = \nabla_Y \nabla_X v - \nabla_X \nabla_Y v - \nabla_{[X,Y]} v.$$

□

Este resultado que acabamos de ver confirma que tiene sentido hablar de "holonomía puntual", pues al tomar lazos cada vez más pequeños el límite converge a un tensor. Es decir, obtenemos un objeto matemático que definido puntualmente. Ahora el objetivo es comprobar que efectivamente este tensor generaliza la idea que tenemos de curvatura para variedades de dimensión arbitraria.

Definición 4.7. Llamamos *tensor de curvatura*² R al tensor de tipo (1, 3) definido por:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z,$$

donde $X, Y, Z \in \mathfrak{X}(M)$. Si lo escribimos en componentes tenemos $R(e_k, e_l)e_j = \sum_{i=1}^n R_{klj}^i e_i$.

Observación. Por construcción se trata de un concepto intrínseco; depende únicamente de la conexión. Si desarrollamos las derivadas podemos expresar los coeficientes en términos de los símbolos de Christoffel:

$$R_{klj}^i = \partial_k \Gamma_{lj}^i - \partial_l \Gamma_{kj}^i + \sum_n \Gamma_{lj}^n \Gamma_{kn}^i - \Gamma_{kj}^n \Gamma_{ln}^i.$$

²También se le conoce como tensor de curvatura de Riemann

En el caso de que tengamos una métrica no degenerada podemos expresar R como un tensor de tipo $(0, 4)$ como se muestra en la siguiente definición.

Definición 4.8. Sea (M, g) una variedad diferenciable equipada con una métrica no degenerada. Entonces se define el tensor de curvatura de tipo $(0, 4)$ como:

$$R(X, Y, Z, W) = g(R(X, Y)Z, W).$$

En componentes lo podemos escribir como $R(e_k, e_l, e_j, e_i) = g(R(e_k, e_l)e_j, e_i) = R_{klji}$.

Veamos ahora mediante un ejemplo que este objeto efectivamente tiene una relación con la curvatura que conocemos para superficies.

Ejemplo 4.9. Sea una superficie $S \subset \mathbb{R}^3$ con métrica inducida g y conexión de Levi-Civita. Fijemos ahora un punto en S y consideremos una base (∂_1, ∂_2) ortonormal de $T_p S$. Evaluemos el tensor de curvatura:

$$R(\partial_1, \partial_2)\partial_1 = \sum_{i=1}^2 \left(\partial_1 \Gamma_{12}^i - \partial_2 \Gamma_{11}^i + \sum_{m=1}^2 (\Gamma_{12}^m \Gamma_{m1}^i - \Gamma_{11}^m \Gamma_{m2}^i) \right) \partial_i,$$

y

$$R(\partial_1, \partial_2, \partial_1, \partial_2) = \sum_{i=1}^2 \left(\partial_1 \Gamma_{12}^i - \partial_2 \Gamma_{11}^i + \sum_{m=1}^2 (\Gamma_{12}^m \Gamma_{m1}^i - \Gamma_{11}^m \Gamma_{m2}^i) \right) g_{i2}.$$

Al estar trabajando en una base ortonormal la matriz de la métrica es la identidad y por tanto:

$$R(\partial_1, \partial_2, \partial_1, \partial_2) = \partial_1 \Gamma_{12}^2 - \partial_2 \Gamma_{11}^2 + \Gamma_{12}^1 \Gamma_{11}^2 - \Gamma_{11}^1 \Gamma_{12}^2 + \Gamma_{12}^2 \Gamma_{21}^2 - \Gamma_{11}^2 \Gamma_{22}^2$$

Si nos fijamos con atención este es precisamente el miembro izquierdo de la ecuación de Gauss para el caso de superficies; ecuación (3,17) en [HC-PG19, p. 139]. Por tanto podemos afirmar que:

$$R(\partial_1, \partial_2, \partial_1, \partial_2) = \kappa,$$

donde κ es la curvatura gaussiana. Es decir, si tomamos una base ortonormal del espacio tangente, podemos obtener la curvatura a partir de este tensor. Es más, ni siquiera es necesario partir de un par de vectores ortonormales; basta con tomar vectores linealmente independientes. En ese caso, llegamos a

$$R(e_1, e_2, e_1, e_2) = \det(g)K.$$

El desarrollo se encuentra en [KoNo63, Vol. I, Ch. V, Prop. 1.3]. Es decir, no hace falta tomar una base ortonormal para obtener una relación con la curvatura.

Una vez hemos visto esto, comprobar que este tensor nos da una generalización al concepto de curvatura para variedades de mayor dimensión es sencillo. Supongamos una variedad M de dimensión $m > 2$ con una métrica no degenerada y fijemos $p \in M$. Consideremos una subvariedad de dimensión 2 que contiene a p , con espacio tangente $\pi \subset T_p M$. Entonces por lo que acabamos de ver, si tomamos una base (e_1, e_2) de π , podemos calcular su curvatura:

$$R(e_1, e_2, e_1, e_2) = \det(g)K.$$

Es decir, en el caso general este tensor mide la curvatura correspondiente a la superficie que tiene como espacio tangente a π , conocida como curvatura seccional. De esta forma el tensor codifica todas las curvaturas correspondientes a todos los posibles planos tangentes que se puedan considerar.

Definición 4.10. Sea (M, g, ∇) una variedad diferenciable equipada con una métrica no degenerada y la conexión de Levi-Civita. Sea $p \in M$ un punto fijado en la variedad y $\pi \subset T_p M$ un plano. Entonces llamamos *curvatura seccional* a

$$K(\pi) = \frac{R(e_1, e_2, e_1, e_2)}{\det(g)},$$

donde (e_1, e_2) es base de π .

En esta sección hemos obtenido el tensor de curvatura diferente de la usual, partiendo de un concepto geométrico como la holonomía y viendo que en realidad es una manifestación de la estructura algebraica de los campos tangentes a la variedad. En la siguiente sección lo haremos viendo cómo cambian las bases de los espacios tangentes y cotangentes según nos movemos por la variedad, llegando a las conocidas como ecuaciones de estructura de Cartan.

4.2. Curvatura a partir de las ecuaciones de estructura de Cartan

Definición 4.11. Sea (M, ∇) una variedad diferenciable de dimensión n equipada con una conexión, y sea (e_1, e_2, \dots, e_n) una base del espacio tangente. Entonces llamamos formas de conexión a la familia de 1-formas tales que:

$$\nabla e_j = \sum_{i=1}^n \omega_j^i \otimes e_i.$$

Observación. Nótese que al evaluar los vectores de la base en estas formas obtenemos los símbolos de Christoffel:

$$\sum_{i=1}^n \Gamma_{kj}^i e_i = \nabla_{e_k} e_j = \sum_{i=1}^n \omega_j^i(e_k) e_i \implies \Gamma_{kj}^i = \omega_j^i(e_k).$$

Proposición 4.12. Sean (M, ∇) una variedad diferenciable de dimensión n equipada con una conexión, y sea Θ el tensor de torsión de ∇ . Sea (e_1, e_2, \dots, e_n) una base local del espacio tangente y $(\theta^1, \dots, \theta^n)$ su base dual. Entonces se verifican las siguientes ecuaciones:

- $d\theta^i = -\sum_{j=1}^n \omega_j^i \wedge \theta^j + \Theta^i$
- $\sum_{j=1}^n (d\omega_j^i + \sum_{m=1}^n \omega_m^i \wedge \omega_j^m) \wedge \theta^j = d\Theta^i + \sum_{j=1}^n \omega_j^i \wedge \Theta^j.$

La primera se le conoce como *primera ecuación de estructura de Cartan*.

Demostración. La primera se halla aplicando el lema A.11 a las formas θ^i . La segunda tomando la diferencial exterior sobre la primera. Por ello, se encuentran en el apéndice A.12. \square

Corolario 4.13. Sea (M, ∇) una variedad diferenciable de dimensión n equipada con una conexión sin torsión. Entonces se verifican

- $d\theta^i = -\sum_{j=1}^n \omega_j^i \wedge \theta^j$
- $\sum_{j=1}^n (d\omega_j^i + \sum_{m=1}^n \omega_m^i \wedge \omega_j^m) \wedge \theta^j = 0.$

Definición 4.14. Llamamos formas de curvatura a las siguientes 2-formas:

$$\Omega_j^i := d\omega_j^i + \sum_{m=1}^n \omega_m^i \wedge \omega_j^m.$$

Esta se le conoce como *segunda ecuación de estructura de Cartan*. Por otro lado, al ser 2-formas podemos expresarlas como sigue

$$\Omega_j^i = \frac{1}{2} \sum_{k,l=1}^n R_{jkl}^i \theta^k \wedge \theta^l, \quad R_{jkl}^i \in C^\infty(M).$$

Observación. Por la antisimetría de Ω_j^i tenemos que $R_{jkl}^i = -R_{jlk}^i$. Además, en el caso en el que la conexión no tenga torsión, entonces tenemos $\sum_{j=1}^n \Omega_j^i \wedge \theta^j = 0$.

Ahora intentemos darle una interpretación geométrica a Ω_j^i , que como su nombre indica, contiene información sobre la curvatura de la variedad. Para ello, veamos cómo obtener los coeficientes R_{jkl}^i .

Proposición 4.15. *Los coeficientes R_{klj}^i de las formas de curvatura se pueden calcular de la siguiente forma*

$$R_{klj}^i = \theta^i(\nabla_{e_k}\nabla_{e_l}e_j - \nabla_{e_l}\nabla_{e_k}e_j - \nabla_{[e_k, e_l]}e_j),$$

donde e_j, e_k, e_l son vectores de la base de T_pM .

Demostración. La prueba se por cálculo directo evaluando $\Omega_j^i(e_k, e_l)$ y empleando el lema A.II, por lo que se muestra en el apéndice A.13. \square

De nuevo llegamos al tensor de curvatura por una vía completamente distinta a la algebraica mostrada en la sección anterior.

Observación. Podemos entender las formas de curvatura³ como proyecciones de dicho tensor fijado una dirección, ya que $\Omega_j^i(X, Y) = \theta^i(R(X, Y)e_j)$. Así, $R(X, Y)e_j = \sum_i \Omega_j^i(X, Y)e_i$. Esto lo que nos está diciendo es que podemos ver el tensor de curvatura como un tensor de tipo $(0, 4)$ como presentamos en la definición 4.8, o bien como una 2-forma con valores en $\mathcal{M}_{n \times n}(\mathbb{R})$.

Proposición 4.16. *Sea (M, g, ∇) una variedad diferenciable equipada con una métrica y una conexión de Levi-Civita, y sean $X, Y, Z, V \in \mathfrak{X}(M)$ campos tangentes a dicha variedad. Entonces el tensor de curvatura verifica las siguientes propiedades:*

1. *Primera identidad de Bianchi:*

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0$$

2. *Segunda identidad de Bianchi:*

$$(\nabla_X R)(Y, Z)V + (\nabla_Y R)(Z, X)V + (\nabla_Z R)(X, Y)V = 0$$

3. *Simetría en el intercambio de pares:*

$$g(R(X, Y)Z, V) = g(R(Z, V)X, Y)$$

4. *Antisimetría en los dos primeros argumentos:*

$$R(X, Y)Z = -R(Y, X)Z$$

5. *Antisimetría en el intercambio del orden del segundo par:*

$$g(R(X, Y)Z, V) = -g(R(X, Y)V, Z)$$

Demostración. Ver [Küh06, p. 249-251]. \square

³Además, a partir de estas formas es muy fácil ver algunas de las simetrías del tensor de curvatura. Como la alternancia en los dos primeros índices.

4.3. Tensor de Ricci

Ahora nos enfocaremos en estudiar el tensor de Ricci. Este es una contracción del tensor de curvatura que nos indica cómo se deforma un pequeño volumen transportado a lo largo de curvas geodésicas. Este tensor juega un papel fundamental en relatividad general pues aparece en la ecuación de campo que determina cómo la energía curva el espacio-tiempo.

Definición 4.17. Sean M una variedad diferenciable y R su tensor de curvatura. Sean $X, Y, Z \in \mathfrak{X}(M)$ tres campos tangentes a M . Entonces definimos el *tensor de Ricci* como:

$$\text{Ric}(X, Y) = \text{tr}(Z \mapsto R(X, Z)Y).$$

Intentemos entender un poco mejor lo que estamos diciendo y cómo trabajamos con esto. Si recordamos, el tensor de curvatura R es un tensor de tipo $(1, 3)$, por lo que si fijamos los campos X e Y , entonces obtenemos un endomorfismo del espacio T_pM :

$$R : \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M), \quad Z \longmapsto R(X, Z)Y.$$

De esta forma, el tensor de Ricci simplemente consiste en tomar la traza de este endomorfismo. Así:

$$\text{Ric}(e_j, e_k) = \sum_i g(R(e_j, e_i)e_k, e_i) = \sum_{i=1}^n R_{jik}^i = \sum_{i=1}^n R_{jiki},$$

donde $(e_l)_{l=1}^n$ es una base ortonormal del espacio tangente. En coordenadas podemos escribir el tensor de Ricci como $\text{Ric} = \sum_{i,j=1}^n R_{jik}^i \theta^j \otimes \theta^k = \sum_{i,j=1}^n R_{ij} \theta^i \otimes \theta^j$.

Observación. Nótese que lo que estamos haciendo es definir un endomorfismo fijando dos entradas y tomando su traza. Es decir, su invariante más sencillo. Quizás esta fue la intuición que se tuvo en primer lugar para estudiar un tensor que *a priori* no tendría por qué tener una interpretación geométrica tan clara.

Una propiedad clave del tensor de Ricci y que facilita mucho su manejo es que es un tensor simétrico.

Proposición 4.18. El tensor de Ricci es simétrico, es decir, $\text{Ric}(X, Y) = \text{Ric}(Y, X) \quad \forall X, Y \in \mathfrak{X}(M)$. Equivalentemente en componentes, $R_{kl} = R_{lk}$.

Demostración. Aplicando la propiedad 3 de la proposición 4.16 tenemos a la definición del tensor de Ricci concluimos:

$$\text{Ric}(X, Y) = \sum_i g(R(X, e_i)Y, e_i) = \sum_i g(R(Y, e_i)X, e_i) = \text{Ric}(Y, X).$$

□

Ahora que ya entendemos la definición del tensor de Ricci podemos pasar a estudiar su interpretación geométrica. Para ello, debemos ver en primer lugar la estrecha relación que tiene con las curvaturas seccionales.

Proposición 4.19. *Sea (M, g) una variedad diferenciable con una métrica g . Entonces, para un campo $X \in \mathfrak{X}(M)$ tal que $g(X, X) = 1$ se cumple que:*

$$\text{Ric}(X, X) = \sum_{i=1}^{n-1} \kappa(\pi_i),$$

donde $\pi_i = \text{span}(e_i, X)$ con (X, e_1, \dots, e_{n-1}) base ortonormal de T_pM

Demostración. Aplicando la propiedad 4 de 4.16 tenemos que $R(X, X) = 0$. Aplicándolo a la definición del tensor de Ricci y por definición de la curvatura seccional tenemos que:

$$\text{Ric}(X, X) = g(R(X, X)X, X) + \sum_{i=1}^{n-1} g(R(X, e_i)X, e_i) = \sum_{i=1}^{n-1} g(R(X, e_i)X, e_i) = \kappa(\pi_i).$$

□

Observación. Podemos interpretar $\text{Ric}(X, X)$ como un “promedio” de las curvaturas seccionales en las direcciones que contienen a X .

Por otro lado, al igual que pudimos resumir la información del tensor de curvatura en el tensor de Ricci tomando la traza, podemos hacerlo una vez más para obtener un escalar que nos resuma la información de la curvatura en un punto. Este número recibe el nombre de curvatura escalar.

Definición 4.20. La *curvatura escalar* κ se define como la traza del tensor de Ricci con respecto a la métrica:

$$\kappa = \sum_{k,l=1}^n g^{kl} R_{kl}.$$

Intentemos entender más en profundidad qué información nos da la curvatura escalar. Fijemos un punto $p \in M$ y consideremos una base $(e_k)_{k=1}^n$ ortonormal de T_pM . Entonces por el resultado anterior podemos decir que $\text{Ric}(e_k, e_k) = \sum_{i \neq k} K(\text{span}(e_k, e_i))$. Por lo tanto, la curvatura escalar queda

$$\kappa = \sum_{k=1}^n \sum_{i \neq k} \kappa(\text{span}(e_k, e_i)) = 2 \sum_{1 \leq i < j \leq n} \kappa(\text{span}(e_i, e_j)).$$

Es decir, la curvatura escalar coincide con dos veces la suma de todas las curvaturas seccionales definidas por los pares de vectores de una base ortonormal de T_pM .

4.3.1. Interpretación geométrica del tensor de Ricci

El tensor de Ricci mide cómo la curvatura de una variedad afecta el volumen de pequeños elementos cuando estos se mueven a lo largo de trayectorias en la variedad. Para entender esto, necesitamos analizar cómo la métrica de la variedad se comporta cerca de un punto $p \in M$. Para ello, presentamos el desarrollo del Taylor de la métrica en coordenadas normales.

Proposición 4.21. *Sea (M, g) una variedad diferenciable equipada con una métrica. Sea $p \in M$ un punto fijado en la variedad. Entonces en coordenadas normales y en un entorno de p podemos expresar la métrica como:*

$$g_{ij}(x) = \delta_{ij} - \frac{1}{3} \sum_{k,l=1}^n R_{ikjl}(0) x^k x^l + O(|x|^3),$$

donde δ_{ij} es la delta de Kronecker.

Demostración. Consultar [Sak96, p. 41]. □

Proposición 4.22. *En un sistema de coordenadas normales centrado en $p \in M$, el elemento de volumen se expande como*

$$dV = \sqrt{\det(g_{ij}(x))} dx^1 \cdots dx^n = \left(1 - \frac{1}{6} \sum_{k,l=1}^n R_{kl}(0) x^k x^l + O(|x|^3)\right) dx^1 \cdots dx^n,$$

donde $R_{kl}(0)$ son los coeficientes del tensor de Ricci en p .

Demostración. Por la proposición 4.21 sabemos que en coordenadas normales tenemos

$$g_{ij}(x) = \delta_{ij} - \frac{1}{3} \sum_{k,l=1}^n R_{ikjl}(0) x^k x^l + O(|x|^3).$$

Por otro lado, definimos $g(x) = I + H(x)$, y $H_{ij}(x) = -\frac{1}{3} \sum_{k,l=1}^n R_{ikjl}(0) x^k x^l + O(|x|^3)$. Como H es simétrica y $\|H\|$ es de orden $O(|x|^2)$ si nos restringimos a un entorno de p podemos aplicar el lema A.14 y concluir que

$$\det(g_{ij}(x)) = 1 + \sum_{i=1}^n H_{ii}(x) + O(|x|^4) = 1 - \frac{1}{3} \sum_{i,k,l=1}^n R_{ikil}(0) x^k x^l + O(|x|^4).$$

Podemos reescribirlo en términos del tensor de Ricci ya que $\sum_i R_{ikil} = R_{kl}$, de modo que

$$\det(g_{ij}(x)) = 1 - \frac{1}{3} \sum_{k,l=1}^n R_{kl}(0) x^k x^l + O(|x|^4).$$

Finalmente con el desarrollo de la raíz cuadrada concluimos:

$$dV = \sqrt{\det(g_{ij}(x))} dx^1 \cdots dx^n = \left(1 - \frac{1}{6} \sum_{k,l=1}^n R_{kl}(0) x^k x^l + O(|x|^4)\right) dx^1 \cdots dx^n.$$

□

Con este resultado ya sí que podemos darle una interpretación al tensor de Ricci. Consideremos una dirección X tal que $\text{Ric}(X, X) > 0$. Entonces el término cuadrático será positivo y el volumen se contraerá a lo largo de la dirección X . En cambio si $\text{Ric}(X, X) < 0$, el término cuadrático es positivo y volumen se expande. En el caso de $\text{Ric}(X, X) = 0$ el volumen permanece constante.

Además, hemos visto en la proposición 4.19 que el tensor de Ricci nos informa sobre la las curvaturas seccionales ya que $\text{Ric}(X, X) = \sum_{i=1}^{n-1} \kappa(\pi_i)$. Por tanto, vemos que cada $\kappa(\pi_i)$ mide cómo la curvatura distorsiona el volumen en el plano π_i , y la suma $\text{Ric}(X, X)$ da un efecto promedio considerando todas las direcciones que incluyen X . Es decir, si consideramos un pequeño volumen y lo transportamos en algunos planos π_i , el volumen puede comprimirse si $\kappa(\pi_i) > 0$, ya que tiene geometría esférica que concentra las trayectorias. En otros, tendremos $\kappa(\pi_i) < 0$, y por tanto se expande ya que tiene geometría hiperbólica que separa las trayectorias. De esta forma el valor de $\text{Ric}(X, X)$ nos da una idea general de cómo se comporta un volumen transportado por la variedad dirección X .

5

Postulados de la Relatividad

Una vez establecidas las bases geométricas, pasamos al estudio de la Teoría de la Relatividad. Es conocido que esta se basa en dos principios, como lo son que la velocidad de la luz en el vacío y las leyes de la física son las mismas para todos los observadores inerciales. Empezaremos esta sección tratando de establecer unos postulados mínimos en términos matemáticos y a partir de nuestra experiencia en el mundo físico. Posteriormente, veremos que ambas formulaciones son equivalentes. También demostraremos que los modelos de la mecánica clásica y la relativista son los únicos compatibles con estos postulados.

De este capítulo cabe destacar especialmente la sección 5.2 en la que estudiaremos cómo se organizan las estructuras que dan lugar a los modelos newtoniano y relativista en una variedad con borde. Como consecuencia, veremos que en el caso de vivir en un universo newtoniano no podríamos confirmarlo experimentalmente. Por el contrario, confirmar que vivimos en un universo relativista sí es posible; como se comprobó con el experimento de Michelson-Morley.

Definición 5.1. Llamamos *espacio-tiempo* a una variedad diferenciable \mathcal{M} de dimensión 4. A un punto $p \in \mathcal{M}$ de la variedad lo llamamos evento.

Fijemos ahora un evento $p \in \mathcal{M}$. Entonces, por nuestra propia experiencia sabemos que un observador i cualquiera puede distinguir claramente entre espacio y tiempo en su entorno. Dicho de otro modo, cada observador descompone el espacio tangente $T_p\mathcal{M}$ como suma directa de dos espacios vectoriales:

$$T_p\mathcal{M} = \mathbb{T}_i \oplus \mathbb{S}_i,$$

donde \mathbb{S}_i es un subespacio vectorial de dimensión 3 que se corresponde con el espacio y \mathbb{T}_i un subespacio vectorial de dimensión 1, correspondiente al tiempo respectivamente. Además, cada observador es capaz de medir en dichos subespacios, por lo que estos deben ser euclídeos. Entonces podemos definir dos productos escalares: g_i sobre \mathbb{T}_i y \bar{g}_i sobre \mathbb{S}_i . De esta manera cada observador tendrá un tiempo y espacio propios, deshaciéndonos del absoluto newtoniano. Veremos más adelante que en algunos casos estos productos escalares surgen de la restricción de una métrica definida en todo $T_p\mathcal{M}$. Otro

hecho que percibimos de manera natural es que el tiempo está orientado pues distinguimos de pasado y futuro. En términos matemáticos, debemos ser capaces de diferenciar dos componentes conexas en \mathbb{T}_i . Llamaremos W_+ en $\mathbb{T}_i \setminus \{0\}$ a la componente conexa que apunta hacia el futuro. Además, por simplicidad a la hora de trabajar asumiremos que \mathbb{S}_i está orientado según la clase de equivalencia de 3-formas $[dX_i]$. Así, resumiremos todas estas condiciones en el primer postulado:

Postulado I. Sea $p \in M$ un evento e \mathcal{I} el conjunto de todos los posibles observadores en p . Entonces se tiene una familia de descomposiciones $\{(\mathbb{T}_i, g_i, W_+; \mathbb{S}_i, \bar{g}_i, [dX_i])\}_{i \in \mathcal{I}}$ del espacio tangente $T_p M$ tal que:

$$T_p M = \mathbb{T}_i \oplus \mathbb{S}_i.$$

La n -upla $(\mathbb{S}_i, \bar{g}_i, [dX_i])$ es el subespacio vectorial euclídeo orientado de dimensión 3 correspondiente al espacio y (\mathbb{T}_i, g_i, W_+) es el subespacio vectorial euclídeo orientado de dimensión 1 correspondiente al tiempo.

Observación. Nótese que de momento no estamos imponiendo ninguna restricción sobre cómo cada observador divide el espacio tangente, ni cómo lo mide ni su orientación. Las diferentes formas de hacerlo, y por tanto de modelizar la realidad física en la que vivimos, son lo que llamaremos estructuras sobre el espacio-tiempo y que estudiaremos en profundidad en la sección 5.1.

Evidentemente esta descomposición nos debería permitir definir de manera natural una aplicación que modeliza cómo mide el tiempo cada observador. Sea $e_i \in \mathbb{T}_i$ unitario y positivamente orientado, es decir que $g_i(e_i, e_i) = 1$, como \mathbb{T}_i es un subespacio vectorial de dimensión 1 entonces tenemos el isomorfismo natural $\mathbb{T}_i = \mathbb{R}e_i$.

Definición 5.2. Sea un observador $i \in \mathcal{I}$ cualquiera, y su correspondiente descomposición del espacio tangente $T_p M = \mathbb{T}_i \oplus \mathbb{S}_i$. Llamamos función tiempo del observador i o tiempo propio del observador i a la siguiente proyección:

$$t : T_p M = \mathbb{R}e_i \oplus \mathbb{S}_i \longrightarrow \mathbb{R}e_i = \mathbb{R}$$

Ahora que hemos entendido cómo un observador descompone el espacio tangente, veamos qué postulados necesitamos establecer sobre la familia de observadores. Sea $c_i : I \subset \mathbb{R} \longrightarrow M$ una curva diferenciable que define cómo se mueve el observador i por el espacio-tiempo. Al ser una curva diferenciable, existe un vector tangente para todo punto. Entonces el subespacio \mathbb{T}_i en cada punto es precisamente la recta definida por este vector. Nos referiremos por tanto a este subespacio como trayectoria¹. Por tanto, podemos considerar la colección $\{\mathbb{T}_i\}_{i \in \mathcal{I}}$ como la de todas las trayectorias u

¹A lo largo del texto nos referiremos como trayectoria tanto a la curva que hemos comentado, como al subespacio \mathbb{T}_i fijado un punto. Es evidente que la primera tiene un carácter global, mientras que la segunda es local. La necesidad de considerar el vector tangente en vez de la curva se debe a que nuestra construcción se basa en la descomposición del espacio tangente y por tanto, es local por definición.

observadores posibles en el evento p . Además, si seguimos profundizando en cómo hemos hecho la descomposición, podemos ver que dado un evento p y un observador i la correspondiente descomposición en realidad está definida unívocamente por el subespacio \mathbb{T}_i , pues los correspondientes \mathbb{S}_i son los subespacios ortogonales. Es decir, que $\{\mathbb{T}_i\}_{i \in \mathcal{I}}$ definen todos los observadores y sus correspondientes descomposiciones. Por lo anterior podemos intuir que la colección $\{\mathbb{S}_i\}_{i \in \mathcal{I}}$ también define todas las descomposiciones posibles. Veremos a partir de la proposición 5.5 que es equivalente formular los postulados en función de las trayectorias o de los espacios.

Ahora que ya conocemos cuáles son los posibles observadores, podría darse que en esta colección tengamos dos que su estado de movimiento sea el mismo. Como para nosotros solo es relevante su movimiento necesitamos imponer el siguiente postulado:

Postulado II. Si dados dos observadores i, j se verifica que $\mathbb{T}_i = \mathbb{T}_j$ entonces $i = j$.

Consideremos ahora un observador i cualquiera en un evento fijado. Si consideramos los objetos que desde su punto de vista se mueven a velocidades bajas son los que tienen un estado de movimiento similar, cuanto más similar más despacio los verá. En el caso de que sean iguales los verá parado. Esta idea de que existen trayectorias “próximas” se puede expresar matemáticamente en términos topológicos y del espacio proyectivo. Para ello, veamos antes el siguiente resultado:

Proposición 5.3. Sea V un espacio vectorial isomorfo a \mathbb{R}^n . Entonces V tiene una única topología inducida por la de \mathbb{R}^n que no depende del isomorfismo $\phi : V \rightarrow \mathbb{R}^n$ escogido.

Demostración. Sean $f : V \rightarrow \mathbb{R}^n$ y $g : V \rightarrow \mathbb{R}^n$ dos isomorfismos lineales. Sean τ_1, τ_2 sus topologías inducidas por f y g respectivamente. Entonces podemos construir el automorfismo

$$h = g \circ f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Sabemos que los automorfismos de \mathbb{R}^n se pueden identificar con $GL(n, \mathbb{R})$. Por tanto, h además es homeomorfismo. Sea U un abierto de \mathbb{R}^n , entonces $h(U)$ es abierto. Por otro lado, por definición de topología inducida $f^{-1}(U) \in \tau_1$. Además $h(U) = g(f^{-1}(U)) \in \tau_2$. Por tanto, también tenemos que $f^{-1}(U) \in \tau_2$. Con lo que concluimos que $\tau_1 = \tau_2$ \square

Definición 5.4. Sea V un espacio vectorial. Llamamos *grassmaniana* de dimensión k al conjunto $G_k(V) = \{S \subset V : S \text{ es subespacio vectorial con } \dim(S) = k\}$.

Observación. Para $k = 1$, $G_k(V)$ es el espacio proyectivo $\mathbb{P}(V)$.

Sabemos que $T_p M \cong \mathbb{R}^4$ por tanto, por el resultado anterior tenemos una única topología inducida en el espacio tangente. De esta forma, el conjunto de todas las rectas del espacio vectorial, es decir, $G_1(V)$ admite una única topología inducida; el cociente de la anterior. Por otro lado, es evidente que $\{\mathbb{T}_i\}_{i \in \mathcal{I}} \subseteq G_1(V)$, por lo que podemos expresar la idea anterior como el siguiente postulado.

Postulado III. La familia $\{\mathbb{T}_i\}_{i \in I}$ de trayectorias en p es un conjunto abierto U de $G_1(T_p M)$.

Veamos ahora que podemos formular un postulado equivalente en términos de $\{\mathbb{S}_i\}_{i \in I}$.

Proposición 5.5. La grassmaniana $G_k(\mathbb{R}^n)$ y el espacio proyectivo $\mathbb{P}(\mathbb{R}^{k(n-k)+1})$ son difeomorfos.

Demostración. Veamos en primer lugar cómo podemos describir los elementos de $G_k(\mathbb{R}^n)$. Es fácil ver que un elemento es solución del sistema lineal $Ax = 0$ con $A \in M_{(n-k) \times n}(\mathbb{R})$ y $\text{rg}(A) = n - k$. Sin embargo, esta descripción da lugar a redundancias pues si cambiamos A de base sigue representando el mismo subespacio pero la matriz es diferente. Sabemos por tanto que :

$$A' = PA \quad \text{con } P \in GL(n - k) \iff \{A'x = 0\} \equiv \{Ax = 0\}$$

Veamos cuál es la dimensión de este subespacio sabiendo esto. Sea $M \in M_{(n-k) \times n}(\mathbb{R})$ tal que $\text{rg}(M) = n - k$. Sabemos que existe $T \in GL(n - k)$ tal que:

$$TM = \left(A \mid B \right)$$

con $A \in M_{(n-k) \times (n-k)}(\mathbb{R})$ de rango máximo, es decir invertible. Así, $P = A^{-1}T \in GL(n - k)$ y por tanto:

$$PM = \left(Id \mid A^{-1}B \right)$$

De esta forma dos matrices A, A' representan el mismo elemento de la grassmaniana si al expresarlo de esta forma el segundo bloque coincide, es decir, la dimensión de este subespacio es $k(n - k)$ y podemos hacer la siguiente identificación:

$$G_k(\mathbb{R}^n) \longrightarrow G_k(\mathbb{R}^n) \longrightarrow \mathbb{P}(\mathbb{R}^{k(n-k)+1*})$$

$$\{Ax = 0\} \longmapsto \{x_i + \sum_{j=n-k+1}^n b_{ij}x_j = 0\}_{i=1, \dots, n-k} \longmapsto [1 : b_{ij}]_{\substack{i=1, \dots, n-k \\ j=n-k+1, \dots, n}}$$

A esta aplicación la denotaremos como Φ . Por otro lado, sabemos que dos espacios vectoriales de la misma dimensión son difeomorfos. Así, podemos concluir que $\mathbb{P}(V)$ y $\mathbb{P}(V^*)$ lo son mediante el difeomorfismo Δ del teorema de la dualidad canónica de la geometría proyectiva [ZB22, p. 280]. Así, podemos construir el siguiente difeomorfismo:

$$\Delta \circ \Phi : G_k(\mathbb{R}^n) \longrightarrow \mathbb{P}(\mathbb{R}^{k(n-k)+1})$$

□

Corolario 5.6. La dimensión de la grassmaniana $G_k(\mathbb{R}^n)$ es $k(n - k)$.

Corolario 5.7. Sean $G_{k_1}(\mathbb{R}^n)$, $G_{k_2}(\mathbb{R}^n)$ dos grassmanianas tales que $k_1 + k_2 = n$ entonces son difeomorfos.

Demostración. Basta ver que la dimensión del espacio proyectivo al que ambos son difeomorfos es la misma. Así:

$$k_1(n - k_1) = k_2(n - k_2) \iff n(k_1 - k_2) = k_1^2 - k_2^2 \iff n = k_1 + k_2 \quad \square$$

Para nuestro caso, podemos ver que $\{\mathbb{T}_i\}_{i \in \mathcal{I}} = G_1(T_p M)$ y $\{\mathbb{S}_i\}_{i \in \mathcal{I}} = G_3(T_p M)$, por lo que $\{\mathbb{T}_i\}_{i \in \mathcal{I}}$ y $\{\mathbb{S}_i\}_{i \in \mathcal{I}}$ son difeomorfos. Es decir, el postulado III es equivalente al siguiente:

Postulado III.2. La familia $\{\mathbb{S}_i\}_{i \in \mathcal{I}}$ de los posibles espacios todos los observadores en p es un conjunto abierto U de $G_1(T_p M)$.

Con esto ya sabemos cuáles son los observadores posibles, cómo caracterizarlos y su descomposición asociada. Sin embargo, al igual que en la mecánica newtoniana deberíamos tener una forma de cambiar de referencia, es decir, de un observador a otro. Para definir estos cambios debemos de emplear las mínimas imposiciones pero que sean coherentes con la estructura que hemos construido.

Definición 5.8. Un automorfismo lineal $\phi : T_p M \rightarrow T_p M$ se dice que es un automorfismo de la familia de observadores \mathcal{I} cuando transforma cualquier observador $i \in \mathcal{I}$ en otro $j \in \mathcal{I}$, denotado por $\phi(i)$. Es decir, tenemos $\phi(\mathbb{T}_i) = \mathbb{T}_j$, $\phi(\mathbb{S}_i) = \mathbb{S}_j$, y los isomorfismos lineales

$$\phi : (\mathbb{T}_i, g_i, W_i^+) \rightarrow (\mathbb{T}_j, g_j, W_j^+), \quad \phi : (\mathbb{S}_i, \bar{g}_i, [dX_i]) \rightarrow (\mathbb{S}_j, \bar{g}_j, [dX_j])$$

son isometrías orientadas.

Los automorfismos de la familia de observadores \mathcal{I} forman un subgrupo G del grupo lineal completo $GL(4, \mathbb{R})$ y decimos que G es el *grupo estructural*. Además, el *grupo de un observador* $i \in \mathcal{I}$ se define como el subgrupo H_i de todos los automorfismos $\phi \in G$ tales que $\phi(i) = i$.

Observación. Al transformar un observador por cualquiera de los automorfismos anteriores, al ser isometrías orientadas el nuevo observador emplea las mismas unidades y tiene la misma orientación que el primero. Evidentemente esto también sucede cuando transformamos un observador en sí mismo mediante un automorfismo del grupo de dicho observador.

Postulado IV.

1. El grupo estructural G actúa de forma transitiva sobre los observadores. En otras palabras, sean $i, j \in \mathcal{I}$ dos observadores, entonces existe un automorfismo $\phi \in G$ tal que $j = \phi(i)$
2. El grupo de cualquier observador verifica que $\phi \in H_i$ si y solo si ϕ es la identidad en \mathbb{T}_i y ϕ restringida en \mathbb{S}_i es una isometría orientada. Es decir, $H_i = \{Id\} \times SO(3)$

Nótese que el postulado IV.1 es en realidad uno de los postulados originales de la relatividad; la equivalencia de observadores o que las leyes de la física son las mismas para todos los observadores. En el caso de la relatividad especial el grupo estructural se le conoce como .

Con esta definición podemos imponer un postulado que hable del propio espacio-tiempo y no de sus posibles descomposiciones.

Postulado V. El espacio-tiempo es homogéneo. Es decir, dados $p, q \in \mathcal{M}$, y sus correspondientes espacios tangentes $T_p\mathcal{M}$ y $T_q\mathcal{M}$ existe un isomorfismo lineal $\phi : T_p\mathcal{M} \rightarrow T_q\mathcal{M}$ que respeta la estructura completa de los observadores. Dicho de otra forma, que verifica:

- Transforma un observador $i \in \mathcal{I}_p$ en otro $j \in \mathcal{I}_q$, es decir, $\phi(i) = j$.
- Preserva las descomposiciones asociadas a los observadores: $\phi(\mathbb{T}_i) = \mathbb{T}_j$, $\phi(\mathbb{S}_i) = \mathbb{S}_j$.
- Y además, respeta la métrica y las orientaciones tanto en la parte temporal como espacial:

$$\phi : (\mathbb{T}_i, g_i, W_i^+) \rightarrow (\mathbb{T}_j, g_j, W_j^+), \quad \phi : (\mathbb{S}_i, \bar{g}_i, [dX_i]) \rightarrow (\mathbb{S}_j, \bar{g}_j, [dX_j]).$$

Con el postulado IV.1 teníamos la invarianza de las leyes físicas para los observadores fijado un punto. Por otro lado, este último impone que las leyes tienen que ser las mismas para todo punto. Es decir, obtenemos la idea fundamental en relatividad de que las leyes de la física son las mismas en todo punto del espacio-tiempo y para todos los observadores.

5.1. Caracterización de las estructuras sobre el espacio-tiempo

Una vez hemos establecido los postulados de la relatividad, a continuación trataremos de construir las estructuras que nos permitirán definir las descomposiciones que satisfagan los postulados del I al IV y por tanto modelizar el espacio-tiempo sobre el espacio-tiempo. Antes de comenzar debemos imponer un postulado más sobre cómo son las propias estructuras que podemos definir sobre p .

Postulado VI. Una estructura sobre p tal que define descomposiciones $\{(\mathbb{T}_i, g_i, W_i^+; \mathbb{S}_i, \bar{g}_i, [dX_i])\}_{i \in \mathcal{I}}$ que satisface los postulados I a IV cambia de manera diferenciable respecto a p .

Empecemos definiendo la estructura que se corresponde con la mecánica relativista.

Definición 5.9. Una *estructura lorentziana* en un espacio vectorial E de dimensión 4 está conformada por una métrica g sobre E de signatura $(+, -, -, -)$, dX una 4-forma no nula y por tanto $[dX]$ una orientación en E y una constante positiva c . Además, deberemos escoger una orientación temporal W^+ como una de las componentes conexas de $\{e \in E : g(e, e) > 0\}$. La denotaremos como $(g, c, W^+, [dX])$.

Proposición 5.10. Una *estructura lorentziana* $(g, c, W^+, [dX])$ definida sobre el espacio-tiempo define una familia de descomposiciones compatible con todos los postulados salvo el **V** en cada evento.

Demostración. En primer lugar, debemos ser capaces de distinguir los subespacios correspondientes al tiempo y al espacio propio de cada observador. Para ello, basta ver que la propia métrica g permite hacer esta división. La familia de las trayectorias de los observadores \mathbb{T}_i , son los subespacios vectoriales de $T_p\mathcal{M}$ donde g es definida positiva. Por otro lado, dada una trayectoria \mathbb{T}_i el espacio correspondiente en esta descomposición es el espacio ortogonal:

$$\mathbb{S}_i = \{e \in T_p\mathcal{M} : g(e_i, e) = 0, \quad \forall e_i \in \mathbb{T}_i\}.$$

Así, esta definición de \mathbb{T}_i y \mathbb{S}_i es compatible con el postulado **I** ya que la signatura de g es $(+, -, -, -)$. Además, como veremos g define un cono cuyo interior son las trayectorias, por lo que también verifica el postulado **III**.

Una vez obtenidas las descomposiciones de $T_p\mathcal{M}$ a partir de la métrica, debemos de saber medir sobre dichos espacios. Para ello, volvemos a hacer uso de g . Fijada una trayectoria \mathbb{T}_i , como g tiene signatura $(+, -, -, -)$, define una estructura euclídea sobre este subespacio, donde $g_i := g|_{\mathbb{T}_i}$. Por otro lado, para medir sobre \mathbb{S}_i podemos considerar la métrica $\bar{g}_i := -c^2 g|_{\mathbb{S}_i}$.

Además de lo anterior, debemos ser capaces de definir las orientaciones. Para la orientación temporal basta considerar $W_i^+ = \mathbb{T}_i \cap W^+$. En cuanto a la espacial $[dX_i]$ se debe escoger de manera que sea coherente con la temporal y con la del espacio-tiempo, es decir $[dX] = [W_i^+ \wedge dX_i]$.

Por lo visto hasta ahora no podemos verificar el postulado **IV**. Sin embargo, el primero de ellos es fácil una vez nos damos cuenta que el grupo estructural es el grupo $O(1, 3)$ que veremos en 5.20. Finalmente, el postulado **VI** pues todos los elementos que definen la estructura son diferenciables y **II** es trivial. \square

Veamos ahora cómo se define la estructura correspondiente con la mecánica newtoniana.

Definición 5.11. Llamamos *estructura galileana* sobre un espacio vectorial E de dimensión 4 a una forma lineal no nula $\omega : E \rightarrow \mathbb{R}$, un producto escalar \bar{g} definido sobre el núcleo de la aplicación ω y una orientación $[dX]$ sobre E . Esta aplicación debe verificar que la dimensión del núcleo es 3. Denotamos esta estructura como $(\omega, \bar{g}, [dX])$.

Proposición 5.12. *Una estructura galileana $(\omega, \bar{g}, [dX])$ definida sobre el espacio-tiempo define una familia de descomposiciones compatible con todos los postulados salvo el V en cada evento.*

Demostración. Definimos las trayectorias \mathbb{T}_i como los subespacios donde ω no se anula. Por tanto, ω define un producto escalar en dichos subespacios $g_i := (\omega \otimes \omega)|_{\mathbb{T}_i}$ y una orientación temporal $W^+ := \{e \in \mathbb{T}_i : \omega(e) > 0\}$ ² sobre las trayectorias.

En cuanto al espacio, tenemos que para todo observador se puede definir el espacio como $\mathbb{S}_i := \ker(\omega)$. Sobre este subespacio se considera la única orientación $[d\bar{X}]$ tal que $[dX] = [\omega \wedge d\bar{X}]$. Con lo que es compatible con el postulado I. Además, vemos que $\mathbb{T}_i = E \setminus \ker(\omega)$. Por tanto, es compatible con el postulado III.

En cuanto a los postulados IV basta considerar $GL(4, \mathbb{R})$ como grupo estructural, ya que como veremos esta estructura se corresponde con la mecánica newtoniana. El postulado II es trivial y el VI se cumple ya que todos los elementos empleados para definir la estructura son diferenciables. \square

Observación. No hemos incluido el postulado V pues para ello sería necesario considerar una conexión ∇ sobre la variedad. Con esta definida basta ver que el transporte paralelo inducido respeta la estructura.

Veamos cuáles son las diferencias fundamentales entre ambas estructuras. Como adelantamos al principio del capítulo, las métricas en los subespacios del espacio y del tiempo, en general, iban a definirse a partir de una métrica en $T_p\mathcal{M}$. Estas aplicaciones nos permiten definir explícitamente cómo es la proyección que modeliza el tiempo propio de un observador³. Sean dos eventos $p_1, p_2 \in X$ y dos observadores $i, j \in \mathcal{I}$ en p_1 con trayectorias \mathbb{T}_i y \mathbb{T}_j diferentes generadas por los vectores e_i y e_j respectivamente. Entonces, para el caso en el que tengamos una métrica g en $T_p\mathcal{M}$, podremos calcular el tiempo que percibe cada uno de los observadores entre ambos eventos como $t_i = e_i \cdot (p_1 - p_2) = g(e_i, (p_1 - p_2))$ y $t_j = e_j \cdot (p_1 - p_2) = g(e_j, (p_1 - p_2))$. Es decir, el observador i mide el tiempo mediante la aplicación lineal $i_{e_i, g}$ y el observador j mediante $i_{e_j, g}$. Por tanto, en el caso de una estructura lorentziana, como los vectores son diferentes, los tiempos que mide cada observador entre ambos eventos también lo es. Es decir, se rompe la idea de simultaneidad. Por otro lado, la estructura galileana divide el espacio-tiempo en dos semiespacios por un hiperplano $\ker(\omega)$. De esta manera en la descomposición del espacio tangente, el espacio es el mismo para todos los observadores $T_p\mathcal{M} = \mathbb{T}_i \oplus \mathbb{S}_i = \mathbb{T}_i \oplus \ker(\omega)$. Es decir, con esta construcción recuperamos la simultaneidad de la mecánica newtoniana, pues independientemente de la trayectoria que siga el observador, el espacio para todos los observadores está conformado por los mismos eventos.

No sólo eso, sino que para el caso lorentziano, vemos que no todas las trayectorias son posibles. Sea $p \in \mathcal{M}$ un evento y sea $v \in T_p\mathcal{M}$ con coordenadas (v_0, v_1, v_2, v_3) en la base que diagonaliza g en dicho

²Es fácil ver que esto es equivalente a $W^+ := \{e \in \mathbb{T}_i : g(e, e) > 0\}$

³Una propiedad que en la sección anterior se indicó que nuestro modelo debía tener.

evento. Como $\tilde{g}_i := -c^2 g|_{\mathbb{S}_i}$ y $g_i := g|_{\mathbb{T}_i}$ tenemos que $g = dx_0^2 - \frac{1}{c^2} \sum_{i=1}^3 dx_i^2$. Así, como las trayectorias válidas verificar $g(v, v) \geq 0$, vemos que:

$$v_0^2 \geq \frac{1}{c^2} (v_1^2 + v_2^2 + v_3^2)$$

Es decir, las trayectorias válidas deben estar contenidas en un cono definido por la métrica⁴. Si consideramos $c_i : I \subset \mathbb{R} \rightarrow X$ una curva diferenciable que define cómo se mueve el observador i por el espacio-tiempo, entonces debe verificar que su vector tangente en cada punto esté contenido el cono definido por la métrica en dicho evento.

Definición 5.13. El cono que define el conjunto de trayectorias viables bajo una estructura lorentziana se le llama *cono de luz*.

En cambio, en la estructura galileana todas las trayectorias son posibles salvo las contenidas en el núcleo de ω . Además, esto nos da una intuición de que la estructura galileana es en realidad el límite de las lorentzianas. Pues si tomamos el límite⁵ con $c \rightarrow \infty$ el cono de luz se ensancha hasta que termina degenerando en el plano $\mathbb{S}_i = \{e \in T_p M : g(e_i, e) = 0, \quad \forall e_i \in \mathbb{T}_i\}$.

Una vez presentadas estas estructuras, el objetivo en lo que resta de capítulo es probar que estas dos son las únicas que dan lugar a una colección de descomposiciones que satisfacen los cuatro primeros postulados.

Definición 5.14. Sea e_0 la trayectoria de un observador fijado y $\mathbb{T}_0 \oplus \mathbb{S}_0$ su descomposición del espacio tangente. Si consideramos otro observador i tal que $\mathbb{T}_i \subset \mathbb{S}_0$ decimos que su velocidad aparente es infinita⁶. En cambio, si $e_0 + \vec{v}_i \in \mathbb{T}_i$ con $v_i \in \mathbb{S}_i$ único, decimos que v_i es la *velocidad aparente* de dicha trayectoria.

Lema 5.15. *Para todo observador no hay objetos con velocidad aparente infinita.*

Demostración. Razonemos por reducción al absurdo. Sea $\mathbb{R}e_0 \oplus \mathbb{S}_0$ la descomposición del espacio tangente de un observador fijado. Supongamos que existe otro observador i con velocidad aparente infinita, es decir, $e_i \in \mathbb{S}_0$. Por el postulado **IV** podemos tomar un automorfismo de la forma $\tau = Id_{\mathbb{R}e_0} \oplus \sigma_{\mathbb{S}_0}$ donde $\sigma_{\mathbb{S}_0}$ es una isometría en \mathbb{S}_0 . Es decir τ es un elemento del grupo del observador 0. Como el observador i tiene una velocidad infinita aparente, $e_i \in \mathbb{S}_0$. Así, si tomamos $\sigma_{\mathbb{S}_0}$ como aquella isometría que verifica $\sigma(e_i) = -e_i$, τ sigue perteneciendo al grupo del observador 0. Por otro lado, si aplicamos dicho automorfismo al observador i , $(\tau(e_i) = -e_i, \tau(\mathbb{S}_i))$ vemos que obtenemos otro observador diferente pero con la misma trayectoria que i . Esto contradice el postulado **II**. \square

⁴Este es el cono que adelantamos en la proposición 5.10

⁵El desarrollo riguroso de esta idea es el objetivo del libro [NS17]. Este se presenta en el capítulo 8.

⁶Esto se debe a que dicho objeto se encontraría en varios sitios a la vez

Proposición 5.16. *Sea M una variedad diferenciable de dimensión 4 equipada con una estructura lorentziana. Para todo observador, el supremo del módulo de las velocidades aparentes es c . Si M está equipada con una estructura galileana, entonces el supremo es infinito.*

Demostración. Comencemos por el caso lorentziano. Sean e_0 un observador cualquiera y $e_0 + \vec{v}$ una trayectoria viable, es decir, $g(e_0 + \vec{v}, e_0 + \vec{v}) > 0$. Entonces como e_0 y \vec{v} son ortogonales

$$0 < g(e_0, e_0) + 2g(e_0, \vec{v}) + g(\vec{v}, \vec{v}) = 1 + g(v, v) = 1 - \frac{1}{c^2}g|_{\mathbb{S}_i}(\vec{v}, \vec{v})$$

Por tanto, $g|_{\mathbb{S}_i}(\vec{v}, \vec{v}) \in (0, c)$; la velocidad máxima aparente que mide el observador i es c .

Veamos ahora el caso galileano. Sea e_0 un observador cualquiera entonces $e_0 + \vec{v}$ con $\vec{v} \in \ker(\omega)$ define una trayectoria ya que $\omega(e_0 + \vec{v})$ no se anula. Por tanto, no hay velocidad aparente máxima. \square

Corolario 5.17. *El conjunto de trayectorias con velocidades aparentes máximas definidas por una estructura lorentziana es la frontera del cono de luz $\partial(\bigcup_{i \in \mathcal{I}} \mathbb{T}_i)$. Además, son las mismas independientemente del observador.*

En el caso lorentziano, tenemos que las trayectorias con velocidades máximas aparentes se corresponden con la frontera del cono. Mientras que en el caso galileano las únicas trayectorias no permitidas son las de velocidad aparente infinita; las que fijado un observador 0 , la trayectoria está en \mathbb{S}_0 . Por lo que de nuevo volvemos a ver como la estructura galileana parece ser el límite de la lorentziana, ya que si c aumenta el cono se “abre” hasta que en el caso extremo, cuando $c \rightarrow \infty$ se degenera convirtiéndose en el hiperplano espacial. Permitiendo así velocidades aparentes tan altas como queramos.

Lema 5.18. *El conjunto de todas las trayectorias definen un espacio afín en $\mathbb{P}(T_p M)$. O bien definen el complementario de un plano o el interior de un elipsoide.*

Demostración. Ver lema 4.2.3 en [NS17, p. 68,69]. \square

Observación. Vemos que, este resultado es la generalización de lo que hemos visto para el caso lorentziano en el que tenemos el elipsoide $g|_{\mathbb{S}_i}(\vec{v}, \vec{v}) \in (0, c)$ de las velocidades aparentes posible y con el caso galileano en el que tenemos el complementario a $\ker(\omega)$.

Teorema 5.19. *Si una familia $\{(\mathbb{T}_i, g_i, W_i^+; \mathbb{S}_i, \bar{g}_i, [dX_i])\}_{i \in \mathcal{I}}$ de descomposiciones del espacio tangente $T_p M$ satisface los postulados I-IV, esta colección está definida o bien por una estructura lorentziana o por una galileana sobre $T_p M$.*

Demostración. Por el lema anterior, tenemos dos casos.

El primero, que el conjunto de trayectorias V' del espacio afín define un elipsoide. Podemos reescalar los elementos de la base para obtener una esfera. Así:

$$\begin{cases} x_0 = 1 \\ x_1^2 + x_2^2 + x_3^2 = c^2 \end{cases}$$

Es decir, que la esfera es la intersección del hiperplano anterior con el cono

$$x_0^2 - \frac{1}{c^2} \sum_{i=1}^3 x_i^2 = 0$$

Como hemos visto anteriormente, este cono está definido por una métrica de la forma

$$g = dx_0^2 - \frac{1}{c^2} \sum_{i=1}^3 dx_i^2$$

que a su vez define la estructura lorentziana. Este cono debe ser el mismo para todos los observadores salvo reescalamientos pues como vimos anteriormente está definido por $\bigcup \{\mathbb{T}_i\}_{i \in \mathcal{I}}$. Por tanto, lo que debemos verificar es que g y c están en las mismas unidades y que tienen la misma orientación. Es decir, que ninguno de los elementos de $(g, c, W^+, [dX])$ está modificado por una constante. Como hemos comentado en la observación anterior al **IV** el grupo H_0 del observador 0 debe preservar las unidades. Así, tomamos otro observador i cualquiera y lo movemos por el grupo. Por el postulado **IV** obtenemos un conjunto de observadores que tiene las mismas unidades que i . Evidentemente contiene a i y que llamaremos S_i .

Sabemos que los elementos de H_0 son de la forma $Id_{\mathbb{R}_{e_0}} \oplus \sigma_{\mathbb{S}_0}$ con $\sigma_{\mathbb{S}_0}$ una isometría orientada. Por tanto, $\sigma_{\mathbb{S}_0}$ sólo puede tratarse de una rotación alrededor de e_0 o de una simetría que mantenga la orientación espacial, es decir que cambie la orientación en dos de los tres ejes espaciales para el observador e_0 . Es evidente que una simetría que no mantuviese la orientación no se puede obtener mediante composición de rotaciones, sin embargo, las que mantienen la orientación sí. Por ello, si movemos i por el grupo H_0 obtenemos un elipsoide que contiene al propio i y cuyo centro es el propio e_0 . De forma análoga, rotando 0 por el grupo H_i obtenemos un elipsoide que contiene a 0 y cuyo centro es i .

Por tanto, podemos concluir que ambos elipsoides tienen intersección no vacía y por tanto 0 e i tienen las mismas unidades.

En el segundo caso, tenemos que V es todo el hiperplano y coincide con $e_0 + \mathbb{S}_0$ en el caso galileano y que como hemos visto antes es independiente del observador. Además, es fácil ver que existe una única forma lineal tal que $\omega(e_0) = 1$ y $\omega(\mathbb{S}_0) = 0$. Esta forma lineal induce la estructura galileana en el espacio tangente, sin embargo el resto de observadores podrían emplear formas proporcionales a

esta. Por tanto, al igual que antes debemos demostrar que todos los observadores emplean las mismas unidades, es decir, la misma forma lineal y producto escalar, y misma orientación sobre $T_p\mathcal{M}$

Sin embargo, ahora no podemos razonar de la misma manera. En el caso lorentziano tenemos una métrica sobre todo el espacio tangente y las métricas espaciales y temporales resultan de las restricciones de dicha métrica en los correspondientes subespacios. Por tanto, si podemos cambiar de un observador a otro necesariamente emplean las mismas unidades. Por el contrario, en el caso galileano solo imponemos que la dirección temporal es la misma para todos los observadores, y la métrica espacial en principio se define de manera independiente para cada observador. De esta forma que podamos transformar un observador en otro no nos garantiza que las unidades sean las mismas

Para ello, debemos razonar de forma ligeramente diferente. Consideremos un observador fijo $0 \in \mathcal{I}$ y otro observador cualquiera i . El producto escalar de este observador define un elipsoide $\{v \in \mathbb{S} : \bar{g}_i(v, v) = 1\}$ en \mathbb{S} .

Por otro lado, pensemos cómo funcionan los cambios de observadores para una estructura galileana. Tenemos que todos los observadores son la suma de un vector con la misma dirección que la temporal más otro espacial. También sabemos que el grupo de cada observador deja invariante dicho vector y en el espacio, que es común a todos, le aplica una rotación o simetría sin cambiar la orientación. Por tanto, por construcción alguno de los automorfismos de H_0 debe corresponderse con alguna de las rotaciones que deja invariante el elipsoide anterior. Sin embargo, estos automorfismos aplicados sobre cualquier observador diferente a 0 sí cambian que lo transforman en otro pues lo rotan alrededor de e_0 . Más concretamente, tenemos que la componente paralela de e_i a e_0 sí que queda fija y el resto rota hasta que volvemos a tener el elipsoide que teníamos en un principio. Es decir, hemos obtenido un observador j diferente de los dos anteriores pero que verifica que $\bar{g}_j = \bar{g}_i$ pues el elipsoide que definen es el mismo. De esta forma obtenemos la condición análoga al caso lorentziano, pudiendo afirmar que todos los automorfismos de H_j preservan el producto escalar de i .

Razonando igual que en el caso anterior, podemos rotar i alrededor de j y viceversa podemos concluir que existe un entorno de i en el que todos emplean el mismo producto escalar. Finalmente como el conjunto de las trayectorias es conexo concluimos que todos los observadores definen la misma estructura euclídea sobre \mathbb{S} . Por tanto, todos los observadores deben utilizar la misma forma lineal ω y por tanto este caso se corresponde a una estructura galileana. \square

5.2. Variedad de métricas del espacio-tiempo

Una vez hemos demostrado que sólo existen dos estructuras que verifiquen los postulados de la relatividad, veremos que el conjunto de todas estas estructuras define una variedad de métricas. Además,

comprobaremos que las estructuras se organizan de una manera muy concreta sobre ella. Para estudiar esta variedad necesitamos presentar algunas definiciones y resultados previos.

En el postulado **IV** presentamos los automorfismos que transformaban un observador en otro. Además en la demostración del teorema 5.19 comentamos que la métrica de las estructuras debía tener las mismas unidades para todo los observadores. Es decir, ser un invariante bajo las transformaciones de dicho grupo. Así, podemos reescribir todo esto en términos matriciales.

Definición 5.20. Sea $p \in M$ un evento y sea $(g, c, W^+, [dX])$ una estructura Lorentziana definida sobre p . Denotamos como $O(1, 3)$ al grupo de transformaciones que dejan invariante la métrica, es decir:

$$O(1, 3) := \{A \in GL(4, \mathbb{R}) : A^T \eta A = \eta\},$$

donde $\eta := \text{diag}(1, -1, -1, -1)$ es la matriz de la métrica lorentziana en una base que la diagonaliza.

Observación. El grupo $O(1, 3)$ representa matricialmente las mismas transformaciones que lo que habíamos llamado grupo estructural. A este grupo también se le conoce como *grupo de Lorentz*.

Ahora veamos que $O(1, 3)$ se trata de un grupo de Lie. Para ello, sólo nos falta ver que se trata de una variedad diferenciable. Lo haremos aplicando la siguiente proposición:

Proposición 5.21. Sea $F : V \rightarrow \mathbb{R}^{n-m}$, con $V \subseteq \mathbb{R}^n$ abierto, una función diferenciable tal que el rango del jacobiano JF satisface $\text{rg}(JF) = n - m$ en todo punto de

$$M = \{x \in V \mid F(x) = 0\},$$

entonces M es una variedad diferenciable de dimensión m .

Proposición 5.22. El grupo $O(1, 3)$ es una variedad diferenciable de dimensión 6.

Demostración. Para ver que $O(1, 3)$ se trata de una variedad diferenciable de dimensión 6, construiremos una aplicación cuyo núcleo sea dicho grupo y comprobaremos que se verifican las condiciones de la proposición 5.21. Sea la aplicación

$$\begin{aligned} f : GL(4, \mathbb{R}) &\longrightarrow M_{4 \times 4}(\mathbb{R}) \\ A &\longmapsto A^T \eta A - \eta \end{aligned}$$

Es fácil ver que $\ker(f) = O(1, 3)$ y que la imagen de f son las matrices simétricas $\text{Sym}(4, \mathbb{R})$:

$$f(A)^T = \left(A^T \eta A - \eta\right)^T = A^T \eta A - \eta = f(A).$$

Por tanto, tenemos que $\dim(\text{Im}(f)) = 10$. Ahora debemos calcular la diferencial de la aplicación anterior. Sea $\Lambda_0 \in \text{GL}(4, \mathbb{R})$, entonces

$$\begin{aligned} df_{\Lambda_0} &= \left. \frac{d}{dt} \right|_{t=0} f(\Lambda_0 + tA) = \left. \frac{d}{dt} \right|_{t=0} (\Lambda_0 + tA)^T \eta (\Lambda_0 + tA) - \eta \\ &= \left. \frac{d}{dt} \right|_{t=0} \Lambda_0^T \eta \Lambda_0 + t(A^T \eta \Lambda_0 + \Lambda_0 \eta A) + t^2(A^T \eta A) - \eta \\ &= A^T \eta \Lambda_0 + \Lambda_0 \eta A. \end{aligned}$$

Como era de esperar, vemos que $\text{Im}(df_{\Lambda_0}) = \text{Sym}(4, \mathbb{R})$ ya que el espacio tangente de un espacio vectorial es él mismo. Ahora debemos verificar el rango de esta aplicación para toda matriz de $\text{GL}(4, \mathbb{R})$. Empecemos viendo el rango de la diferencial en la identidad. Así, tenemos que

$$df_{\text{Id}}(A) = A^T \eta + \eta A.$$

Veamos que es sobreyectiva. Sea $S \in \text{Sym}(4, \mathbb{R})$ una matriz simétrica intentemos encontrar la matriz $A \in \text{GL}(4, \mathbb{R})$ tal que $df_{\text{Id}_0}(A) = S$:

$$df_{\text{Id}_0}(A) = S \iff (\eta A)^T + \eta A = S$$

Nótese que $\eta^2 = \text{Id}$. Por tanto si tomamos $A := \frac{1}{2}\eta S$ y sustituimos en la expresión anterior vemos que efectivamente este valor de A es el que hace que se cumpla:

$$(\eta A)^T + \eta A = \frac{1}{2}(\eta^2 S)^T + \frac{1}{2}\eta^2 S = S.$$

Así, $\text{rg}(df_{\text{Id}}) = \dim(\text{Sym}(4, \mathbb{R})) = 10$. Ahora debemos ver que para toda matriz de $\text{O}(1, 3)$ el rango del diferencial sigue siendo 10. Para ello, consideremos la aplicación $\hat{f} := f \circ L_\Lambda$, donde $L_\Lambda : \text{GL}(4, \mathbb{R}) \rightarrow \text{GL}(4, \mathbb{R})$, $L_\Lambda(A) := \Lambda A$ es la aplicación de multiplicación por la izquierda. Entonces se tiene que

$$\hat{f}(A) = f(\Lambda A) = (\Lambda A)^T \eta (\Lambda A) - \eta = A^T \eta A - \eta = f(A).$$

Por tanto, $\hat{f} = f$. De esta forma podemos escribir el diferencial como:

$$d\hat{f}_{\Lambda^{-1}} = d(f \circ L_\Lambda)_{\Lambda^{-1}},$$

y aplicando la regla de la cadena:

$$d\hat{f}_{\Lambda^{-1}} = df_{\text{Id}} \circ dL_{\Lambda^{-1}}.$$

Como L_Λ es una aplicación lineal, $dL_{\Lambda^{-1}}$ es un isomorfismo lineal y además df_{Id} también es isomorfismo, entonces $\hat{d}f_{\Lambda^{-1}}$ lo es para todo $\Lambda^{-1} \in \text{O}(1, 3)$. Si particularizamos para $\Lambda = \text{Id}$:

$$\hat{d}f_{\text{Id}} = df_\Lambda \circ dL_{\Lambda_{\text{Id}}}.$$

Como la composición debe ser un isomorfismo y $dL_{\Lambda_{\text{Id}}}$ lo es, entonces df_Λ también y por tanto:

$$\text{rg}(df_\Lambda) = \text{rg}(df_{\text{Id}}) = 10.$$

Así, aplicando el resultado 5.21 tenemos que $\text{O}(1, 3) = \ker f$ es una subvariedad diferenciable de $\text{GL}(4, \mathbb{R})$ de dimensión 6. □

Corolario 5.23. *El grupo $\text{O}(1, 3)$ es un grupo de Lie.*

Corolario 5.24. *El espacio tangente $T_{\text{Id}} \text{O}(1, 3)$ es isomorfo al conjunto de matrices*

$$\{A \in \text{GL}(4, \mathbb{R}) : A^T \eta + \eta A = 0\}.$$

Diremos que estas matrices son η -antisimétricas.

Demostración. Hemos visto en la proposición anterior que $\ker(f) = \text{O}(1, 3)$ y que $df_{\text{Id}}(A) = A^T \eta + \eta A$. Entonces, $T_{\text{Id}} \text{O}(1, 3) = \ker(df_{\text{Id}}) = \{A \in \text{GL}(4, \mathbb{R}) : A^T \eta + \eta A = 0\}$. □

Proposición 5.25. *El grupo $\text{O}(1, 3)$ tiene cuatro componentes conexas.*

Demostración. Sea $A \in \text{O}(1, 3)$, entonces verifica que $A^T \eta A = \eta$ y por tanto, $\det A = \pm 1$. Es decir, $\text{O}(1, 3)$ al menos tiene dos componentes conexas diferentes:

$$\text{SO}(1, 3) := \{A \mid \det A = +1\}, \quad \text{O}_-(1, 3) := \{A \mid \det A = -1\}.$$

A su vez, cada una de estas se divide en otras dos. Para ello, observemos la entrada A_{00} . De nuevo, por la ecuación anterior

$$\sum_{i,j=0}^3 A_{i0} \eta_{ij} A_{j0} = \eta_{00} \implies (A_{00})^2 = 1 + \sum_{i=1}^3 (A_{i0})^2 \geq 1,$$

y en particular $A_{00} \neq 0$. Así, podemos definir la siguiente función $\tau : \text{O}(1, 3) \rightarrow \{+1, -1\}$, $\tau(A) = \text{sign}(A_{00})$. Por tanto $\tau(A) = \pm 1$ está bien definida y es continua. Así, vemos que hay dos componentes conexas más en cada una de las vistas anteriormente.

$$\begin{aligned} \text{SO}^+(1, 3) &= \{A \in \text{SO}(1, 3) \mid \tau(A) = +1\}, \\ \text{SO}^-(1, 3) &= \{A \in \text{SO}(1, 3) \mid \tau(A) = -1\}, \\ \text{O}_+^-(1, 3) &= \{A \in \text{O}^-(1, 3) \mid \tau(A) = +1\}, \\ \text{O}_-^-(1, 3) &= \{A \in \text{O}^-(1, 3) \mid \tau(A) = -1\}. \end{aligned}$$

□

Cada una de estas componentes se corresponde con cambios en las orientaciones espaciales y temporales:

- $\text{SO}^+(3, 1)$ mantiene las orientaciones espacial y temporal.
- $\text{SO}^-(3, 1)$ mantiene la orientación espacial y cambia la temporal.
- $\text{O}_+^-(3, 1)$ cambia la orientación espacial pero no la temporal.
- $\text{O}_-^-(3, 1)$ cambia las orientaciones espacial y temporal.

Corolario 5.26. Como el álgebra de Lie $\mathfrak{o}(3, 1)$ es isomorfo al espacio tangente $T_{\text{Id}} \text{O}(1, 3)$:

$$\mathfrak{o}(3, 1) = \{A \in \text{GL}(4, \mathbb{R}) : A^T \eta + \eta A = 0\}.$$

Antes de pasar finalmente a la caracterización de la variedad de métricas necesitamos hallar la dimensión del conjunto de las métricas lorentzianas y establecer el isomorfismo entre éstas y el cociente $\text{GL}(4, \mathbb{R})/\text{O}(1, 3)$.

Proposición 5.27. Sea \mathcal{L} el conjunto de todas las métricas lorentzianas. Entonces, $\dim(\mathcal{L}) = 10$.

Demostración. Es fácil ver que existe una identificación entre las formas bilineales simétricas 4×4 con signatura $(+, -, -, -)$ y el conjunto de métricas lorentzianas:

$$\mathcal{L} \equiv \{A \in \text{Sym}(4, \mathbb{R}) : \det(A) \neq 0, \quad \text{sig} = (1, 3)\}.$$

Podemos ver que la primera condición $\det(A) \neq 0$ define un abierto en $\mathcal{M}_{4 \times 4}(\mathbb{R})$. La segunda también pues los autovalores cambian de forma continua con los elementos de la matriz. Por tanto, dicho conjunto es un abierto y tiene la misma dimensión que el espacio de matrices simétricas. Así:

$$\dim(\mathcal{L}) = \dim(\text{Sym}(4, \mathbb{R})) = 10.$$

□

Como hemos indicado es natural identificar una métrica lorentziana con una matriz simétrica con la signatura correspondiente, veamos cómo hacer esta identificación. Para construir todas las métricas lorentzianas podríamos partir de la forma diagonal correspondiente y conjugando esta matriz mediante cambios de referencia definidos por $GL(4, \mathbb{R})$ obtener todas las métricas posibles. *A priori* podría parecer que cada cambio se corresponde con una métrica diferente y por tanto el isomorfismo que buscamos sería entre estos conjuntos. Sin embargo, esto no es cierto pues como hemos comentado existen cambios de referencia que preserven la métrica. Estos cambios están caracterizados por el grupo $O(1, 3)$, pues precisamente lo hemos definido como los cambios que preservan la métrica lorentziana. Así, el candidato natural a ser isomorfo al conjunto de métricas lorentzianas es el cociente $GL(4)/O(1, 3)$.

Proposición 5.28. *El conjunto de métricas lorentzianas es difeomorfo a $GL(4, \mathbb{R})/O(1, 3)$.*

Demostración. En primer lugar vemos que las dimensiones coinciden. Como $\dim(GL(4, \mathbb{R})) = 16$ y $\dim(O(1, 3)) = 6$, tenemos que $\dim(GL(4, \mathbb{R})/O(1, 3)) = 10$.

Consideremos ahora la siguiente aplicación :

$$\begin{aligned}\phi : GL(4, \mathbb{R}) &\longrightarrow \mathcal{L} \\ A &\longmapsto A^T \eta A\end{aligned}$$

Podemos ver que se trata de la misma aplicación en la proposición 5.22 salvo una constante. Por tanto, al ϕ también es una aplicación sobreyectiva con imagen en las matrices simétricas. Por otro lado, podemos ver que no es inyectiva. Sean $A_1, A_2 \in GL(4, \mathbb{R})$ tales que $A_1 = BA_2$ con $B \in O(1, 3)$. Entonces

$$\phi(A_1) = A_1^T \eta A_1 = (BA_2)^T \eta BA_2 = A_2^T B^T \eta BA_2 = A_2^T \eta A_2 = \phi(A_2).$$

Como la penúltima igualdad sólo se da si la diferencia entre ambas matrices es una transformación del grupo de Lorentz, tenemos que

$$\begin{aligned}\bar{\phi} : GL(4, \mathbb{R})/O(1, 3) &\longrightarrow \mathcal{L} \\ [A] &\longmapsto A^T \eta A\end{aligned}$$

es un isomorfismo.

Veamos que la diferencial es también un isomorfismo en todo punto. Como ϕ es la misma aplicación que en la proposición 5.22 tenemos que $d\phi_{Id(A)} = df_{Id} = A^T \eta + \eta A$. También demostramos que es una aplicación sobreyectiva y en el corolario 5.24 vimos que

$$T_{Id} O(1, 3) = \ker(d\phi_{Id}) = \{A \in GL(4, \mathbb{R}) : A^T \eta + \eta A = 0\}.$$

Por tanto, si tomamos el cociente obtenemos el siguiente isomorfismo:

$$\overline{d\phi_{\text{Id}}} : T_{\text{Id}} \text{GL}(4, \mathbb{R}) / T_{\text{Id}} \text{O}(1, 3) \cong T_{\text{Id}} \left(\frac{\text{GL}(4, \mathbb{R})}{\text{O}(1, 3)} \right) \longrightarrow T_{\eta} \mathcal{L}.$$

Nótese que $T_{\text{Id}} \text{GL}(4, \mathbb{R}) / T_{\text{Id}} \text{O}(1, 3)$ y $T_{\text{Id}} \left(\frac{\text{GL}(4, \mathbb{R})}{\text{O}(1, 3)} \right)$ son isomorfos al ser espacios vectoriales de la misma dimensión.

Para ver que la diferencial evaluada en otro punto también es isomorfismo, seguiremos el mismo razonamiento que en la proposición 5.22.

Consideremos la aplicación $\hat{\phi} := \bar{\phi} \circ L_{\Lambda}$, donde L_{Λ} es la aplicación de multiplicación por la izquierda por un elemento $\Lambda \in \text{O}(1, 3)$

$$\begin{aligned} L_{\Lambda} : \text{GL}(4, \mathbb{R}) / \text{O}(1, 3) &\longrightarrow \text{GL}(4, \mathbb{R}) / \text{O}(1, 3) \\ [A] &\longmapsto [\Lambda A]. \end{aligned}$$

Entonces se tiene que

$$\hat{\phi}(A) = \bar{\phi}(\Lambda A) = ([\Lambda A])^T \eta([\Lambda A]) = [A]^T \eta[A] = \bar{\phi}([A]).$$

Por tanto, $\hat{\phi}$ está bien definida. De esta forma razonamos igual que en la proposición mencionada y concluimos que la aplicación $d\hat{\phi}_{\Lambda}$ es isomorfismo para todo $[\Lambda] \in \text{GL}(4, \mathbb{R}) / \text{O}(1, 3)$.

Finalmente, por el teorema de la función inversa podemos asegurar que para todo $[\Lambda] \in \text{GL}(4, \mathbb{R}) / \text{O}(1, 3)$ existe un entorno abierto $\mathcal{U} \subset \text{GL}(4, \mathbb{R}) / \text{O}(1, 3)$ tal que $\bar{\phi}|_{\mathcal{U}}$ es difeomorfismo. Con esto podemos concluir que :

$$\mathcal{L} \cong \text{GL}(4, \mathbb{R}) / \text{O}(1, 3)$$

□

Hasta ahora hemos trabajado con el cociente $\text{GL}(4, \mathbb{R}) / \text{O}(1, 3)$, el cual modela el conjunto de todas las estructuras Lorentzianas posibles sobre un espacio vectorial de dimensión 4. Hay que recalcar que este conjunto no tiene estructura de grupo, ya que $\text{O}(1, 3)$ no es un subgrupo normal de $\text{GL}(4, \mathbb{R})$. Además, es necesario recordar que como ya vimos anteriormente en la proposición 5.25, $\text{O}(1, 3)$ tiene cuatro componentes conexas, cada una correspondiente con un par de orientaciones espacial y temporal. Por lo que el cociente hereda esta misma estructura, de manera análoga a lo que sucede con $\text{O}(1, 3)$, cada una de estas componentes se corresponde a un par distinto de orientaciones espacial y temporal para la métrica Lorentziana. Una vez ya tenemos todo esto claro ya sí que podemos demostrar el teorema que caracteriza las estructuras que podemos definir en cada punto del espacio-tiempo restringiéndonos a una única componente conexa $\text{GL}(4, \mathbb{R}) / \text{SO}^+(3, 1)$, que es la que preserva simultáneamente ambas orientaciones.

Teorema 5.29. *El conjunto de todas las estructuras lorentzianas y galileanas con la misma orientación espacial y temporal sobre el espacio tangente a un evento forman un variedad conexa con borde de dimensión 11. Además, el borde se corresponde con las estructuras galileanas; los puntos interiores con estructuras lorentzianas.*

Demostración. Con el resultado anterior hemos visto que $\mathcal{L} \equiv \text{GL}(4, \mathbb{R})/\text{SO}^+(1, 3)$, por lo que nos queda ver cómo quedan organizadas todas las estructuras en la variedad. Esta organización se basará en los valores del parámetro c . Después veremos que se trata de una variedad conexa.

Con lo anterior, tenemos la siguiente identificación:

$$\begin{aligned}\mathcal{L} &\longleftrightarrow \text{GL}(4, \mathbb{R})/\text{SO}^+(1, 3) \\ [g] &\longleftrightarrow \Lambda \text{ tal que } g = \Lambda^T \eta \Lambda\end{aligned}$$

Nótese que aunque hayamos estado trabajando todo el rato con $\eta = \text{diag}(1, -1, -1, -1)$ η puede estar formada por cualquier matriz con esta signatura. Por tanto, podemos escoger $\eta_c := \text{diag}(1, -\frac{1}{c^2}, -\frac{1}{c^2}, -\frac{1}{c^2})$. Así, ya tenemos caracterizadas todas las estructuras lorentzianas (g, c) con $c > 0$ mediante la siguiente identificación

$$\begin{aligned}\mathcal{M}_p &:= \mathcal{L} \times \mathbb{R}^+ \longleftrightarrow \text{GL}(4, \mathbb{R})/\text{SO}^+(1, 3) \times \mathbb{R}^+ \\ g_c &:= (g, c) \longleftrightarrow (\Lambda, c) \text{ tal que } g_c = \Lambda^T \eta_c \Lambda,\end{aligned}$$

donde \mathcal{M}_p denota la *variedad de métricas* en el evento $p \in M$. De esta forma necesitamos 11 parámetros para determinar g_c : uno para c y diez para determinar Λ . Además también nos deja claro como podríamos definir una carta de esta variedad:

$$\begin{aligned}\psi : \mathcal{M}_p &\longrightarrow \mathbb{H}^{11} \\ g_c &\longrightarrow (a_1, \dots, a_{10}, \frac{1}{c})\end{aligned}$$

Así, ya tenemos que \mathcal{M}_p es una variedad con borde de dimensión 11.

Veamos ahora cuál es el borde. Primero necesitamos reescribir los puntos de esta variedad. Sea $\omega \in T_x^*X \setminus \{0\}$ tal que $\omega = (1, 0, 0, 0)$ en cierta base fijada. Entonces:

$$g_c \equiv (\Lambda, c) \equiv \Lambda^T \eta_c \Lambda = \Lambda^T (\omega \otimes \omega - \frac{1}{2} \text{diag}(0, 1, 1, 1)) \Lambda.$$

Así, el borde de la variedad es:

$$\partial \mathcal{M}_p = \psi^{-1}(a_1, \dots, a_{10}, 0) = \lim_{c \rightarrow \infty} \psi^{-1}(a_1, \dots, a_{10}, \frac{1}{c}) = \lim_{c \rightarrow \infty} g_c = \Lambda^T \omega \otimes \omega \Lambda.$$

Es fácil ver que si rotamos ω por la acción de $\text{GL}(4, \mathbb{R})/\text{SO}^+(3, 1)$ obtenemos todos los elementos de $T_x^*X \setminus \{0\}$ que dan lugar a la misma orientación temporal que ω . Así, como Λ es cualquier matriz

en $GL(4, \mathbb{R})/SO^+(3, 1)$, tenemos que $\partial\mathcal{M}_p$ se corresponde con todas las métricas temporales de las estructuras galileanas. Finalmente como $GL(4, \mathbb{R})/SO^+(3, 1)$ es conexo, la variedad también lo es.

□

Corolario 5.30. *Sea un evento $p \in M$ entonces las estructuras galileanas en T_pM son estructuras lorentzianas donde el parámetro c , la velocidad aparente máxima, es infinito.*

Corolario 5.31. *El borde de la variedad de las estructuras sobre T_pM es difeomorfo a las grassmanianas $G_3(\mathbb{R}^4)$ y $G_1(\mathbb{R}^4)$.*

Hay que recordar que estos resultados sólo nos hablan de las métricas en el espacio tangente a un evento. No nos dice cómo se distribuyen las métricas globales de todo el espacio-tiempo. Existe un resultado análogo al anterior que se desarrolla en el capítulo 8 de [NS17].

Volviendo de nuevo al teorema 5.29 podemos sacar varias conclusiones. La primera ya la adelantamos más arriba, y es que la estructura galileana en realidad se trata de un caso degenerado de la lorentziana en el que la velocidad máxima es infinita. Es por ser un caso degenerado que se encuentran en el borde de la variedad anteriormente definida. En segundo lugar, es que en el caso de vivir en un universo con una estructura galileana sería muy difícil descartar que en realidad se tratara de una lorentziana con una velocidad máxima permitida muy alta debido a que estas se encuentran en el borde de la variedad. En cambio, en un universo lorentziano sí que sería posible esta comprobación experimental, como efectivamente se ha hecho.

6

Espacio-tiempo de Minkowski

En este capítulo exploraremos el caso más sencillo dentro de la Teoría de la Relatividad, cuando no consideramos la fuerza gravitatoria, es decir, tomamos \mathcal{M} una variedad plana. Esto es lo que se conoce como espacio-tiempo de Minkowski o Relatividad Especial. Al principio mostramos los objetos definidos en el capítulo anterior particularizados a este caso, desde la métrica hasta el grupo de automorfismos. De este capítulo caben resaltar dos aspectos fundamentales. El primero es la modelización de la materia del trabajo para el caso más sencillo, como es el de las partículas puntuales, mediante el impulso. Además, veremos cómo de manera natural surge la idea de que la masa depende del estado de movimiento del observador. El segundo aspecto a destacar, es la obtención de una extensión de la Segunda Ley de Newton para el caso relativista y la equivalencia entre masa y energía junto a su interpretación. El capítulo acaba con la explicación de algunos fenómenos relativistas.

Definición 6.1. Un *espacio-tiempo de Minkowski* $(\mathcal{M} = \mathbb{A}_4, g, c, W^+, [dX])$ es un espacio afín de dimensión 4 \mathbb{A}_4 definido sobre un espacio vectorial E equipado con una estructura lorentziana.

Observación. En cierta base la métrica es de la forma $g = dx_0^2 - \frac{1}{c^2}(dx_1^2 + dx_2^2 + dx_3^2)$. A su vez, esta define una métrica contravariante que podemos escribir como $h_* := -c^2g = c^2dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2$. Mientras que la primera mide en unidades de tiempo, la segunda lo hace en unidades de longitud. Por ello, las llamaremos *métrica temporal* y *métrica espacial* respectivamente.

Observación. La orientación temporal $W^+ = \{e \in E : g(e, e) > 0\}$ se define tomando una componente conexa del cono de luz.

Como hemos visto en el capítulo anterior esta estructura nos da una familia de descomposiciones $\{(\mathbb{T}_i, g_i, W_+; \mathbb{S}_i, \bar{g}_i, [dX_i])\}_{i \in \mathcal{I}}$, que nos permite distinguir entre pasado y futuro y nos medir el espacio y el tiempo para cada observador. También define las trayectorias viables como aquellas contenidas en el cono de luz en cada punto. Así, clasificaremos los vectores en tres tipos.

Definición 6.2. Sea $e \in E$ un vector no nulo. Diremos que es un *vector temporal*, *espacial* o *luz* si

$g(e, e)$ es positivo, negativo o nulo respectivamente. También se les suele llamar vectores de tipo tiempo, espacio o luz.

De esta forma los vectores temporales son aquellos contenidos en el interior del cono. Por tanto, se tratan de trayectorias posibles y como vimos en la proposición 5.16 su velocidad aparente para cualquier otro observador es menor que c . Los vectores luz son otras trayectorias posibles correspondientes a las generatrices del cono, aquellas en las que la velocidad aparente es c . Estos dos tipos de vectores se pueden clasificar a su vez en orientados hacia el futuro o al pasado según pertenezcan o no a \mathcal{W}_+ . Por último, los vectores espaciales son los que están fuera del cono y definen trayectorias imposibles al superar la velocidad aparente máxima c . Así, podemos redefinir lo que es una trayectoria sobre X .

Definición 6.3. Una trayectoria o un observador en el espacio-tiempo se define como una curva conexa de tipo tiempo, es decir que para todo punto su vector tangente es de tipo tiempo.

Definición 6.4. Decimos que un observador es inercial cuando su trayectoria es una recta, es decir, trayectorias de velocidad constante.

Proposición 6.5. Si existe una trayectoria entre dos eventos, entonces existe una trayectoria inercial entre ambos.

Demostración. Sea $\gamma : [0, 1] \rightarrow \mathbb{R}^3$ una curva diferenciable tal que $\gamma(0) = p$ y $\gamma(1) = q$. Podemos reescribir el vector tangente como $\gamma'(\lambda) = v(\lambda) = (v_0(\lambda), \vec{v}(\lambda))$ en la base que diagonaliza la métrica. Al ser una trayectoria físicamente posible la curva verifica que

$$g(v(\lambda), v(\lambda)) = v_0^2(\lambda) - \|\vec{v}(\lambda)\|^2 > 0$$

Nótese que γ' está orientado hacia el futuro por lo que $v_0 > 0$ y por tanto al tomar raíces la desigualdad $v_0(\lambda) > \|\vec{v}(\lambda)\|$ se mantiene. Además,

$$q - p = \int_0^1 \gamma'(\lambda) d\lambda = \left(\int_0^1 v_0(\lambda) d\lambda, \int_0^1 \vec{v}(\lambda) d\lambda \right) =: (V_0, \vec{V})$$

Por otro lado tenemos que la única trayectoria inercial que va de p a q es

$$\gamma_i = p + \lambda(q - p) \quad \lambda \in [0, 1]$$

Por lo que basta demostrar que es una trayectoria con sentido físico, es decir que

$$g(q - p, q - p) = V_0^2 - \|\vec{V}\|^2 > 0$$

Finalmente, obtenemos la siguiente cadena de desigualdades

$$V_0 = \int_0^1 v_0(\lambda) d\lambda > \int_0^1 \|\vec{v}\|(\lambda) d\lambda > \left\| \int_0^1 \vec{v}(\lambda) d\lambda \right\| = \|\vec{V}\|$$

□

Veamos ahora una caracterización de los vectores temporales y espaciales.

Proposición 6.6. Sean $p, q \in X$ dos eventos del espacio-tiempo y e el vector que une los dos eventos. Entonces e es un vector temporal si y solo si existe un observador para el que p y q ocurren en el mismo punto de su espacio. Además, $g(e, e)$ es el tiempo medido entre p y q para dicho observador y es la cota inferior para todos los observadores.

Demostración. Consultar A.15. □

Proposición 6.7. Sean $p, q \in X$ dos eventos del espacio-tiempo y e el vector que une los dos eventos. Entonces e es un vector espacial si y solo si existe un observador para el que p y q son simultáneos. Además, $\sqrt{\langle e, e \rangle}$ es el espacio medido entre p y q por dicho observador y es la cota inferior para todos los observadores. Donde $\langle -, - \rangle$ denota el producto escalar habitual.

Demostración. Consultar A.16. □

Observación. A partir de esta proposición vemos cómo surge de manera natural la ruptura de la simultaneidad entre eventos, para este modelo.

Por último, debemos presentar el grupo estructural en este caso. Como ya vimos en la sección 5.2 para un evento fijado el grupo estructural es $O(1, 3)$. Sin embargo, para el espacio-tiempo de Minkowski, $O(1, 3)$ no se corresponde con todas las transformaciones que mantienen invariante la métrica pues los movimientos rígidos también la conservan.

Proposición 6.8. Sea $(\mathbb{A}_4, g, c, W^+, [dX])$ un espacio-tiempo de Minkowski. Entonces su grupo estructural es $\mathbb{R}^4 \rtimes O(1, 3)$.

Demostración. Al trabajar en un espacio afín podemos considerar un sistema de coordenadas ortonormal global. De esta manera sustituyendo en la ecuación 3.9 podemos concluir que $\Gamma_{ij}^m = 0$ para todo i, j, m . Sea c una curva diferenciable sobre \mathbb{A}_4 y sea Y un campo paralelo sobre c , entonces verifica el sistema 3.10. Como los símbolos de Christoffel son todos nulos, entonces los campos paralelos a cualquier curva son los campos constantes. Además, por compatibilidad con la métrica, $\nabla g = 0$. Es decir, el producto escalar definido por g se conserva bajo transporte paralelo. Así, las traslaciones afines sobre preservan la métrica, y por tanto son isometrías del espacio de Minkowski.

Por tanto, el conjunto de todas¹ las transformaciones que preservan la estructura geométrica del espacio-tiempo de Minkowski está formado por las definidas por el grupo de Lorentz $O(1, 3)$, que dejan fijo el origen, y las traslaciones afines, que desplazan el origen. Este conjunto se denota como $\mathbb{R}^4 \rtimes O(1, 3)$. □

Definición 6.9. Llamamos *grupo de Poincaré* al conjunto completo de transformaciones entre observadores inerciales $\mathbb{R}^4 \rtimes O(1, 3)$ que respetan la estructura geométrica del espacio-tiempo de Minkowski.

¹Podríamos considerar cualquier curva diferenciable entre dos puntos, pero como el campo paralelo es el mismo y para toda curva entre dos puntos existe una transformación afín, éstas últimas nos caracterizan todas las transformaciones.

6.1. Cinemática

Hasta ahora nos hemos centrado en el estudio de la estructura del espacio-tiempo de Minkowski y sus propiedades. En las secciones siguientes trataremos de desarrollar la cinemática y la dinámica para este escenario, además de ver las principales diferencias con la formulación newtoniana.

Definición 6.10. Un *sistema de referencia inercial* en un evento p_0 , que llamamos origen temporal, en el espacio-tiempo de Minkowski es un observador inercial cuya velocidad es e_0 y una base ortonormal² (e_0, e_1, e_2) de $\mathbb{S}_0 = (\mathbb{R}e_0)^\perp$ respecto de la métrica temporal g .

Una vez hemos construido las referencias, lo primero que debemos hacer para estudiar la cinemática en este nuevo marco es ver cómo podemos generalizar la idea de tiempo propio que vimos en 5. Para ese caso vimos que dado un observador e_0 y dos eventos $p_1, p_2 \in X$ para calcular el tiempo entre ambos eventos debíamos proyectar. Es decir, $t = g(e_0, p_2 - p_1)$. Por tanto, la forma de tiempo propio viene dada por la contracción $i_{e_0}g$. Con esto en mente podemos pensar que la generalización a trayectorias cualesquiera entre dos eventos es precisamente la longitud de la trayectoria en el espacio-tiempo. Para ello, necesitamos en primer lugar ser capaces de expresar en coordenadas las trayectorias y en segundo lugar construir una 1-forma que nos permita medir la longitud de la trayectoria en el espacio-tiempo. Empezamos viendo cómo escribir las trayectorias en coordenadas.

Proposición 6.11. Sea $(\mathbb{A}_4, g, c, W^+, [dX])$ un espacio-tiempo de Minkowski. Sea $(p_0; e_0, e_1, e_2, e_3)$ un sistema de referencia inercial entonces toda trayectoria se puede expresar en esas coordenadas como

$$T := e_0 + x'(t)e_1 + y'(t)e_2 + z'(t)e_3 = e_0 + \vec{v}.$$

Demostración. Sabemos que los vectores tangentes a la trayectoria están contenidos en el cono de luz, donde g es definida positiva, por tanto no pueden ser ortogonales a e_0 para todo punto de la trayectoria³. Por tanto, podemos afirmar que localmente es la gráfica de una función y que localmente es homeomorfa \mathbb{R} . Es decir, la podemos parametrizar como $(t, x(t), y(t), z(t))$ con $0 \leq t \leq T$. Como además, suponemos que las trayectorias son inyectivas y conexas tenemos que la aplicación que la define es un difeomorfismo. \square

Observación. Esta condición es fundamental, pues es la que nos garantiza que podamos escribir el campo de velocidades sobre la curva como la suma de la velocidad del observador y la velocidad aparente

$$T := e_0 + x'(t)e_1 + y'(t)e_2 + z'(t)e_3 = e_0 + \vec{v}.$$

²A lo largo del texto usaremos indistintamente (e_0, e_1, e_2, e_3) y $(\partial_t, \partial_1, \partial_2, \partial_3)$ para referirnos a los vectores de la base de un sistema inercial.

³Esto es precisamente lo que nos dice la proposición 5.16 sobre las velocidades aparentes

Definición 6.12. Dado un sistema inercial y una trayectoria $T = e_0 + \vec{v}$ llamamos *factor de Lorentz* al escalar

$$\gamma := \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \quad \text{con } v = \|\vec{v}\|.$$

Observación. Se tiene que $g(T, T) = 1 - \left(\frac{v}{c}\right)^2 = \frac{1}{\gamma^2}$.

Con esto ya podemos pasar a construir la 1-forma anteriormente mencionada.

Proposición 6.13. Sean $(\mathbb{A}_4, g, c, W^+, [dX])$ un espacio-tiempo de Minkowski y $(p_0; e_0, e_1, e_2, e_3)$ un sistema de referencia inercial. Entonces para toda trayectoria con campo de velocidades T existe una única 1-forma tal que $\bar{\omega}(T) = \sqrt{g(T, T)} = \frac{1}{\gamma}$.

Demostración. Consideremos la base dual $(e^i)_{i=0}^3$ a la que define la referencia. Entonces la 1-forma $\bar{\omega} := \frac{1}{\gamma} e^0$ verifica que:

$$\bar{\omega}(T) = \frac{1}{\gamma} e^0(e_0 + \vec{v}) = \frac{1}{\gamma}.$$

Evidentemente por construcción esta forma es única. □

Observación. La forma $\bar{\omega}$ también es la única que verifica $\bar{\omega}(U) = 1$, con $U := \gamma T = \gamma(e_0 + \vec{v})$ el único campo unitario que tiene la misma dirección que T a lo largo de toda la trayectoria.

Definición 6.14. Sea e_0 un observador inercial fijo. Sea otro observador con una trayectoria diferenciable σ cualquiera entre dos eventos $a, b \in \mathbb{A}_4$. Es decir, que $\sigma(0) = a$ y $\sigma(T) = b$ para un cierto $T \in \mathbb{R} \setminus \{0\}$. Sea U el campo tangente a σ normalizado orientado hacia el futuro. Entonces podemos calcular el *tiempo propio* de un observador que siga esa trayectoria como

$$\tau = \int_a^b \bar{\omega}$$

donde $\bar{\omega}$ es la única 1-forma tal que $\bar{\omega}(U) = 1$.

Ejemplo 6.15. (Cálculo del tiempo propio de una trayectoria.) Así, para calcular el tiempo propio basta sustituir lo obtenido en la proposición 6.13 en la definición anterior:

$$\tau = \int_0^T \bar{\omega}(T) = \int_0^T \sqrt{g(T, T)} dt = \int_0^T \sqrt{1 - \left(\frac{v}{c}\right)^2} dt.$$

La interpretación y consecuencias de esta fórmula junto a otros aspectos de la relatividad especial la dejamos para la sección 6.3.

6.2. Masa y energía: Dos caras de la misma moneda

De entre las consecuencias más relevantes y también más conocidas de la relatividad es la equivalencia entre masa y energía⁴. En esta sección nos dedicaremos a estudiarla en profundidad. Para ello, en primer lugar daremos una descripción de la masa para una partícula puntual que sea coherente con toda la teoría construida hasta ahora. Esto lo haremos generalizando la idea del momento. Con ello, hablaremos de algunas leyes de conservación que de esta formulación se deriva y llegaremos a una versión relativista de la Segunda Ley de Newton. Finalmente presentaremos la equivalencia entre masa y energía.

En la proposición 6.II vimos que para un sistema de referencia inercial podemos expresar una trayectoria en coordenadas como $T = \partial_t + \vec{v}$. Es decir, mientras que la velocidad aparente depende del sistema de referencia que escojamos, el vector T es absoluto. Esto nos indica que el momento $\vec{p} = m\vec{v}$ así definido, es también un concepto dependiente del observador; por lo que debemos de buscar un concepto análogo. En vez de la velocidad aparente podemos considerar el vector T para definir su homólogo absoluto como $I := mT$. Sin embargo, este razonamiento no es del todo correcto ya que pasamos por alto un detalle importante.

Consideremos una trayectoria T y en un sistema de referencia inercial tal que $T = \partial_t$. Ahora tomamos I como acabamos de proponer $I_1 := m_1 T = \partial_t$. Ahora si escogemos otro sistema de referencia donde T tiene una velocidad aparente no nula. Entonces tenemos que $I_2 := m_2 \tilde{\partial}_t + m_2 \vec{v}$. De esta forma, como el impulso debe ser una magnitud absoluta, debe verificarse la igualdad $I_1 = I_2$ y por tanto su módulo también debe coincidir. Así obtenemos que:

$$g(I_1, I_1) = g(I_2, I_2) \implies m_1^2 = m_2^2 \left(1 - \frac{v^2}{c^2}\right).$$

Como m_1 y m_2 deben ser positivos, entonces podemos tomar la raíz y escribir en términos del factor de Lorentz:

$$m_1 = \frac{m_2}{\gamma}.$$

Este razonamiento muestra que si queremos que el impulso sea un concepto invariante del observador entonces el coeficiente que multiplica al vector tangente T no puede ser la masa tal y como la entendíamos hasta ahora, pues pasa a depender del estado de movimiento relativo entre partícula y

⁴Aunque fenómenos como la contracción espacial como la dilatación temporal son consecuencia de esta teoría, en realidad ya habían eran efectos conocidos antes de la publicación de la teoría de Einstein, y que fueron interpretados hasta entonces como errores. Por tanto el papel que jugó Einstein fue afirmar que no se trataban de errores, sino que aquellas predicciones eran consecuencias inevitables de la constancia de la velocidad de la luz y de la equivalencia de todos los sistemas inerciales. En cambio, la equivalencia entre masa y energía no apareció hasta que el propio Einstein lo enunciara.

observador. Es decir, la idea de que la masa es una magnitud inmutable se rompe. Ahora bien, esto no impide que al igual que antes podamos encontrar una magnitud que sí sea absoluta.

Es evidente que si todos los observadores con el mismo estado de movimiento deben medir lo mismo. En ese caso, la masa en reposo, m_1 en el ejemplo, es un absoluto. De esta forma, partir de la velocidad aparente y mediante la fórmula anterior podemos relacionar la *masa en reposo* que denotaremos como m con la que llamaremos *masa aparente* m_{ap} ; en el ejemplo m_2 . Así, podemos definir el impulso como sigue.

Definición 6.16. Sean un espacio-tiempo de Minkowski y un sistema de referencia inercial. Sea una trayectoria con T vector velocidad en dicho espacio-tiempo. Entonces llamamos *impulso*

$$I := m_{ap}T = m_{ap}\partial_t + m_{ap}\vec{v},$$

donde m_{ap} es la masa medida por el observador en cada punto de la trayectoria.

Este nuevo concepto además de ser independiente del observador, codifica la información que nos permite caracterizar una partícula. Empleando la proposición 6.13 tenemos una forma de calcular la masa en reposo

$$\bar{\omega}(I) = \bar{\omega}(m_{ap}T) = \frac{m_{ap}}{\gamma} = m.$$

Por otro lado, podemos calcular la velocidad como $T = m_{ap}^{-1}I$. Es decir, a partir de la trayectoria y el impulso tenemos perfectamente definida una partícula ⁵.

Observación. Nótese que así calculada m siempre será positiva, por lo que está bien definida de esta forma. También por la proposición 6.13 el módulo de I coincide con la masa en reposo

$$g(I, I) = m_{ap}^2 \cdot g(T, T) = m_{ap}^2 \left(1 - \frac{v^2}{c^2}\right) = m^2.$$

Sin embargo, veremos que al no ser $\sqrt{g(I, I)}$ aditiva no es una descripción generalizable al caso de más de una partícula. Por lo que aunque en este caso sí coincida con el módulo de I nunca nos referiremos a la masa en estos términos.

Habiendo definido ya el impulso como una magnitud absoluta ya podemos enunciar la relación entre la masa en reposo y la aparente en forma de teorema.

Teorema 6.17. Sea I el impulso de una partícula y T su vector tangente en un sistema de referencia inercial. Entonces la masa en reposo $m = \bar{\omega}(I)$ y la aparente $m_{ap} = g(\partial_t, I)$ verifican la siguiente relación

$$m_{ap} = \gamma m.$$

⁵Esto sería equivalente a conocer la trayectoria y la masa en reposo, pues conociendo la trayectoria podemos calcular γ y con ello m_{ap} .

En consecuencia, la relatividad nos obliga a reconocer que la masa en reposo y la masa medida por un observador depende de su estado de movimiento; no es una magnitud absoluta. Sin embargo, este fenómeno no lo apreciamos en la vida diaria, en la que la masa es independiente del estado de movimiento del objeto. Esto se debe a que γ es muy próximo a 1 para las velocidades aparentes a las que estamos acostumbrados. Para ver algunos cálculos consultar el ejemplo [A.17](#)

Hasta ahora siempre que hemos hablado de trayectorias en realidad nos referíamos a observadores, no objetos con masa pues nos faltaba el concepto de impulso. Una vez visto esto ya podemos definir una partícula en el espacio-tiempo basándonos en la definición [6.3](#) y empezar a estudiar objetos más allá de los propios observadores.

Definición 6.18. Una *partícula* en un espacio-tiempo M es una curva $\sigma : \mathbb{R} \rightarrow M$ junto con un campo vectorial tangente I tal que $\omega(I) > 0$ en todo instante de la trayectoria. Llamamos trayectoria a σ , masa de la partícula a $\omega(I)$ e impulso de la partícula al campo I .

Observación. Nótese que esta definición se presenta para un espacio-tiempo cualquiera y σ debe ser coherente con la estructura de dicha variedad. Además, no suponemos que la masa sea constante, sólo positiva por la observación anterior.

Además de poder caracterizar las partículas independientemente del observador, la formulación de la mecánica en términos del impulso nos permite condensar la primera ley de Newton y la conservación de la masa en una sola expresión.

Teorema 6.19. *Sea una partícula en un espacio-tiempo de Minkowski $(\mathbb{A}_4, g, c, W^+, [dX])$ con impulso I . Entonces si el impulso se conserva, es decir $\nabla_I I = 0$, la partícula sigue una trayectoria inercial y su masa en reposo no varía.*

Demostración. Para comprobar que la trayectoria que sigue la partícula es inercial basta darse cuenta de que I es tangente a la trayectoria por definición. Entonces $\nabla_I I = 0$ implica que el vector tangente no cambia a lo largo de la trayectoria; el movimiento es rectilíneo.

Veamos ahora que la masa en reposo no cambia. Por simplicidad en los cálculos estudiaremos su cuadrado. Así, por compatibilidad de la métrica con la conexión:

$$\nabla_I m^2 = \nabla_I g(I, I) = 2g(\nabla_I I, I) = 0$$

□

No sólo eso, sino que para una única partícula la conservación del impulso implica la conservación del momento. Es más, se puede aplicar también para un sistema de partículas.

Teorema 6.20. *Sea un conjunto de partículas en un espacio-tiempo de Minkowski con impulsos I_j . Entonces si estas partículas chocan dando lugar a un nuevo conjunto de impulsos I'_k y el impulso total se conserva*

$$\sum_j I_j = \sum_k I'_k$$

entonces tanto la masa aparente total del sistema como el momento se conserva.

Demostración. Por un lado tenemos que $m_{ap,j} = g(\partial_t, I_j)$, $m'_{ap,k} = g(\partial_t, I'_k)$. Entonces aplicando la conservación del impulso total y la linealidad de la métrica tenemos que

$$\sum_j m_{ap,j} = \sum_j g(\partial_t, I_j) = g(\partial_t, \sum_j I_j) = g(\partial_t, \sum_k I'_k) = \sum_k g(\partial_t, I'_k) = \sum_k m'_{ap,k}$$

Por otro lado, para demostrar la conservación del momento el razonamiento es análogo aplicando $\vec{p} = g(\partial_1 + \partial_2 + \partial_3, I)$. \square

Observación. Pese a que se verifiquen todas estas conservaciones en un sistema de partículas, no tiene porqué darse que $\sum_j m_j = \sum_j \sqrt{g(I_j, I_j)} \neq \sum_k \sqrt{g(I'_k, I'_k)} = \sum_k m'_k$ ya que no se tratan de funciones aditivas. Es decir, la masa no coincide con el módulo del impulso total. Es por ello que aunque la masa coincida con el módulo del impulso en el caso de una partícula, no la hemos definido así. El estudio de estos casos lo haremos en el capítulo 7.

Veamos ahora qué pasa en una situación donde el impulso no sea constante lo largo de la trayectoria. Es decir, tenemos que $K = \nabla_T I$, con, K un vector no nulo. Es fácil ver que

$$K = \nabla_T(m_{ap} T) = \frac{dm_{ap}}{d\tau} T + m_{ap} \nabla_T T = \frac{dm_{ap}}{d\tau} T + m_{ap} \vec{a}.$$

Es decir, K en realidad es una fuerza y por tanto llegamos a una ecuación análoga a la Segunda Ley de Newton: el cambio de impulso siempre es debido a una fuerza externa.

Teorema 6.21 (Segunda Ley de Newton Relativista). *Sea una partícula que sigue una trayectoria T y de impulso $I = m T$ en un espacio-tiempo de Minkowski. Si se ejerce una fuerza externa entonces se verifica que*

$$F = \nabla_T I.$$

En forma desarrollada

$$F = \nabla_T(m_{ap} T) = \frac{dm_{ap}}{d\tau} T + m_{ap} \vec{a}.$$

Salvo que tengamos una situación de aceleraciones muy grandes la masa aparente podemos aproximarla por una constante. En ese caso obtenemos la Segunda ley de Newton.

Corolario 6.22 (Segunda Ley de Newton). Si $\frac{dm_{ap}}{dt} = 0$, entonces tenemos que

$$F = \nabla_T(m_{ap}T) = m\vec{a}.$$

Una vez entendido qué es el impulso y algunas propiedades que se derivan de este, podemos pasar a la equivalencia entre masa y energía que anunciamos al principio de la sección.

Teorema 6.23. Sea una partícula en un espacio-tiempo de Minkowski $(\mathbb{A}_4, g, c, W^+, [dX])$ con impulso I . Si estamos en un caso en el que la masa en reposo se conserva, entonces se verifica la siguiente relación

$$E = \gamma mc^2,$$

o equivalentemente

$$E = \left(m + \frac{mv^2}{2c^2} + \frac{3mv^4}{8c^4} + \frac{5mv^6}{16c^6} + \dots\right)c^2.$$

Demostración. Como estamos en el caso en el que la masa intrínseca se conserva tenemos que $\frac{d}{dt}g(I, I) = 0$. Tomando un sistema inercial podemos expandir esta expresión:

$$0 = 2g\left(\frac{d}{dt}I, I\right) = 2g(m_{ap}\partial_t + m_{ap}\vec{v}, \frac{d}{dt}m_{ap}\partial_t + \frac{d}{dt}\vec{p}),$$

donde (\cdot) denota el producto escalar habitual. Simplificando tenemos que

$$0 = m_{ap}\frac{d}{dt}m_{ap} - \frac{m_{ap}}{c^2}(\vec{v} \cdot \frac{d}{dt}\vec{p}) \implies \frac{d}{dt}m_{ap}c^2 = \vec{v} \cdot \frac{d}{dt}\vec{p}.$$

Si ahora recordamos que una forma alternativa de escribir la segunda ley de Newton es $\vec{F} = \frac{d}{dt}\vec{p}$, y que el trabajo se define como $W = \vec{F} \cdot \vec{v}$. De esta forma tenemos que

$$\frac{d}{dt}m_{ap}c^2 = W.$$

A su vez el trabajo es precisamente el cambio de energía, por tanto:

$$\frac{d}{dt}m_{ap}c^2 = \frac{d}{dt}E \implies E = m_{ap}c^2 = \gamma mc^2.$$

Finalmente, desarrollando en serie de Taylor concluimos:

$$E = \left(m + \frac{mv^2}{2c^2} + \frac{3mv^4}{8c^4} + \frac{5mv^6}{16c^6} + \dots\right)c^2.$$

□

Observación. Nótese que esta expresión se puede interpretar como la suma de energía en reposo y cinética $E = mc^2 + K(v)$. Además, vemos que $K(v)$ coincide hasta tercer orden con la expresión a la que estamos habituados. Además, empleando que $E = \gamma mc^2$ y sustituyendo en lo anterior tenemos que :

$$K(v) = (\gamma(v) - 1)mc^2.$$

Corolario 6.24. Cuando $v \ll c$ podemos hacer la aproximación que da lugar a la famosa fórmula $E = mc^2$.

Antes de continuar intentemos dar una interpretación a lo obtenido. Como hemos dicho, en el caso más sencillo volvemos al caso newtoniano. En cambio, si existiera una fuerza con un componente temporal no nula para ese observador, entonces lo que vemos es que la componente temporal de la fuerza se invierte en aumentar la masa aparente. Además, si despejamos T vemos que de manera instantánea se rompe la simultaneidad. Esto junto con el teorema 6.23 lo podemos interpretar como que lo que percibimos como un cambio en la masa aparente, es en realidad proporcional a la energía empleada para ralentizar los relojes del objeto sobre el que se ejerce una fuerza.

6.3. Fenómenos relativistas

En esta sección estudiaremos algunas de las consecuencias de la relatividad más famosas, la contracción del espacio y el tiempo al aumentar la velocidad y la regla de la suma de velocidades

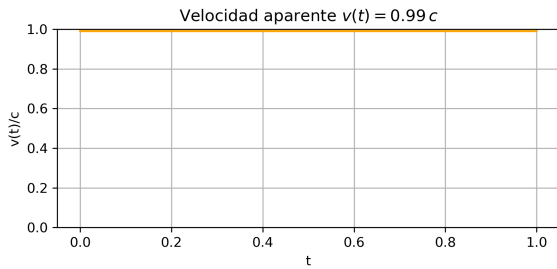
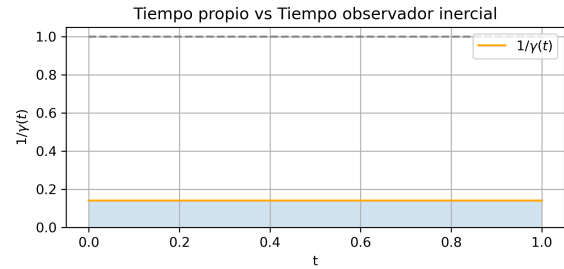
6.3.1. Dilatación temporal

Empecemos esta sección estudiando el efecto de la *dilatación temporal*, un fenómeno en el que el ritmo de los relojes se desacelera progresivamente al incrementar la velocidad respecto a un observador inercial. Para ello, recordemos que en el ejemplo 6.15 vimos que el tiempo propio se calcula mediante la siguiente integral

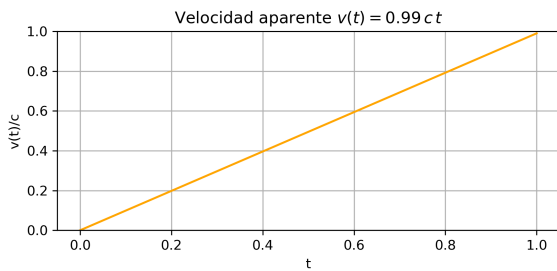
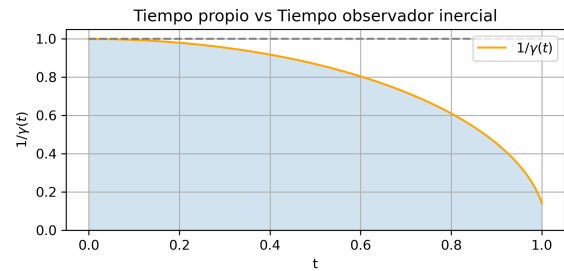
$$\tau = \int_0^T \sqrt{1 - \left(\frac{v}{c}\right)^2} dt.$$

A partir de esta, podemos entender de manera geométrica el fenómeno de la dilatación temporal. Consideremos una trayectoria y su correspondiente función de velocidad aparente. Entonces, podemos dibujar el factor de Lorentz en función del tiempo para dicha trayectoria. Este factor, por definición, oscila entre 0 y 1. Así, el tiempo propio se corresponde con el área bajo la curva de la función $\gamma(t)$ para $0 \leq t \leq T$. Este área es máxima cuando $\gamma(t) = 1$, o equivalentemente, $v(t) = 0$. A medida que el factor de Lorentz disminuye, el área bajo la curva también lo hace, y por tanto, para un mismo valor de T , el tiempo propio del observador no inercial es menor. En el límite, cuando nos aproximamos

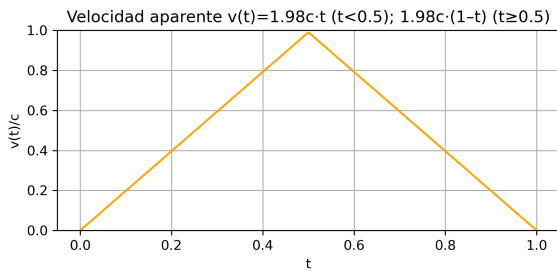
a la velocidad de la luz, el área tiende a 0 y el tiempo medido por dicho observador se detiene. A continuación, se muestran gráficas de tres casos diferentes. En la columna de la izquierda se muestra la función de velocidad aparente y en la de la derecha $\gamma(t)$ para dicha trayectoria y para el caso inercial.

(a) Velocidad aparente $v(t) = 0,99 c$.

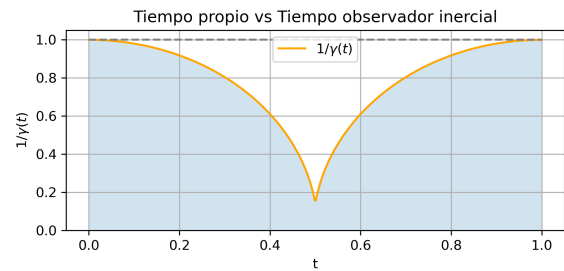
(b) Tiempo propio vs tiempo inercial.

(c) Velocidad aparente $v(t) = 0,99 c t$.

(d) Tiempo propio vs tiempo inercial.



(e) Velocidad aparente triangular.



(f) Tiempo propio vs tiempo inercial.

Figura 6.1. Comparación de distintas velocidades aparentes con el tiempo propio asociado en cada caso.

Para cuantificar cuánto se ralentizan los relojes supongamos que tomamos un objeto que sigue una trayectoria inercial. Entonces, el tiempo propio queda

$$\tau = \gamma^{-1} \int_0^T dt = \gamma^{-1} T,$$

y si recordamos algunos de los valores de γ que vimos en el ejemplo A.17 llegamos de nuevo a que para apreciar este fenómeno necesitaremos al menos un 25 % de la velocidad de la luz.

Ejemplo 6.25. Intentemos ahora sacar explícitamente cuál es la velocidad necesaria para ralentizar los relojes proporcionalmente a un factor $\lambda \in (0, 1)$ en una trayectoria inercial. Para ello, simplemente debemos considerar $T - \tau = \lambda T$. Al estar en un movimiento uniforme, hemos visto que $\tau = \gamma^{-1} T$ y sustituyendo en lo anterior tenemos que $\gamma^{-1} = 1 - \lambda$. Finalmente, despejando llegamos a

$$\frac{v}{c} = \sqrt{\lambda(2 - \lambda)}.$$

A continuación mostramos la gráfica de dicha función.

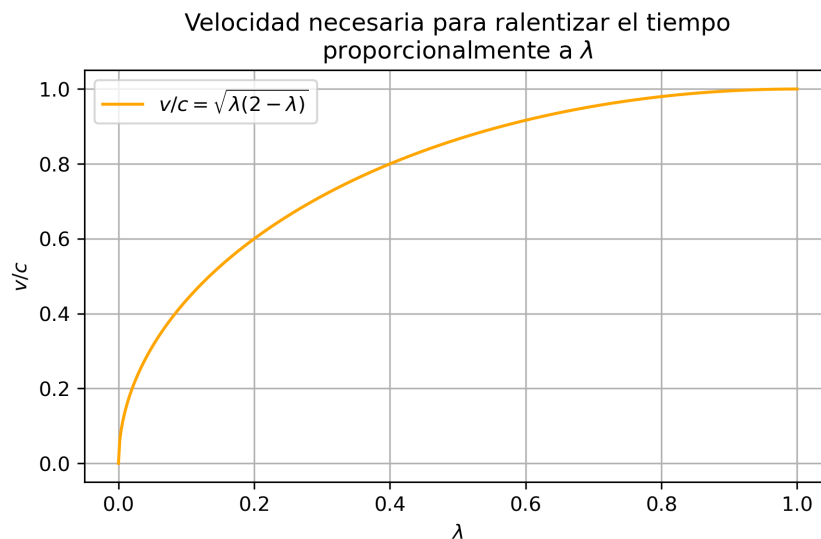


Figura 6.2. Velocidad necesaria para ralentizar el tiempo proporcionalmente a λ , donde $v/c = \sqrt{\lambda(2 - \lambda)}$.

Se puede observar que necesitamos alcanzar grandes velocidades para poder empezar a ver estos efectos, como ya habíamos indicado. Como novedad, apreciamos claramente que al principio, pequeñas variaciones de velocidad producen grandes cambios en la medida del tiempo. Por ejemplo, basta con alcanzar $v = 0,4c$ para ralentizar el tiempo un 80 %. Sin embargo, a partir de ahí, los incrementos se vuelven más costosos. Por ejemplo, para reducir el tiempo desde un 80 % al 100 % restante, es necesario incrementar la velocidad desde $0,4c$ hasta prácticamente la velocidad de la luz. Es decir, cuanto mayor sea el tiempo propio respecto a un observador inercial, más difícil es ralentizarlo aún más.

Por último, antes de terminar con el fenómeno de la dilatación temporal veamos una forma alternativa de expresar cuánto se ralentiza el tiempo en una trayectoria en función de la energía.

Proposición 6.26. *Sea una partícula con un factor de Lorentz $\gamma(t)$ definido a lo largo de toda su trayectoria en un espacio-tiempo de Minkowski $(\mathbb{A}_4, g, c, W^+, [dX])$ para un sistema de referencia inercial fijado. Entonces se puede calcular la dilatación temporal entre esta trayectoria y la inercial como*

$$T - \tau = \int_0^T \frac{K(t)}{E(t)} dt.$$

Demostración. Teniendo en cuenta el teorema 6.23 y la observación 6.2 tenemos que la energía total de la partícula es $E(t) = \gamma(t)mc^2$ y la cinética $K(t) = E - mc^2$ respectivamente. Así, para concluir basta sustituir en la diferencia de tiempos la fórmula vista en el ejemplo 6.15:

$$T - \tau = \int_0^T (1 - \gamma(t)^{-1}) dt = \int_0^T \left(1 - \frac{mc^2}{E}\right) dt = \int_0^T \frac{K(t)}{E(t)} dt.$$

□

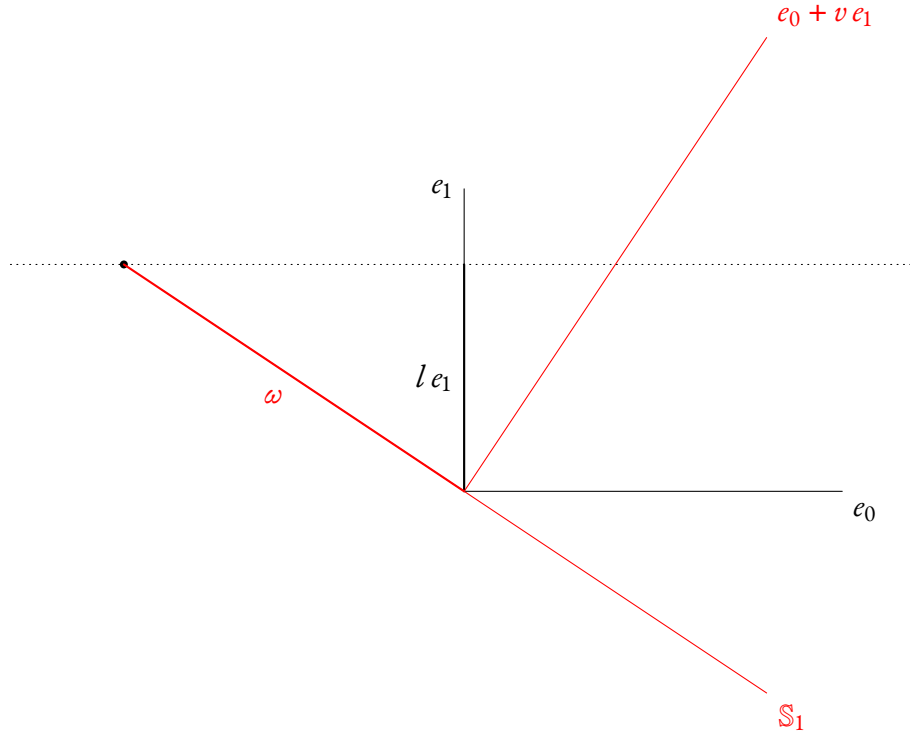
6.3.2. Contracción Lorentz

Acabamos de ver cómo el tiempo se dilata cuando una partícula se mueve a velocidades altas. Resulta que en este caso también se da una contracción del espacio en la dirección del movimiento. Esto es lo que se conoce como *contracción Lorentz*.

Para estudiarlo consideremos un sistema de referencia inercial $(p_0; e_0, e_1, e_2, e_3)$ y una barra de longitud l con un extremo en p_0 y en dirección del vector e_1 . De esta manera es evidente que este observador al medir la barra dirá que esta mide l unidades.

Consideremos ahora otro observador inercial que se mueve con una velocidad aparente v respecto del original en dirección $e_0 + ve_1$ según el sistema de coordenadas del observador original. Ahora nuestro objetivo será medir la longitud de la barra desde el punto de vista del nuevo observador. Para ello, primero debemos obtener las coordenadas del vector que representa la barra para el nuevo observador. Es evidente que este vector no puede tener coordenadas en e_2 o e_3 no nulas, pues el observador no se mueve en ninguna de esas direcciones y la barra no rota. Es decir, el vector que buscamos debe ser de la forma $\lambda e_0 + \mu e_1$. Por otro lado, este debe ser un vector espacial para el nuevo observador. A continuación se muestra un esquema en el que se pueden ver los dos observadores junto con sus respectivos espacios contenidos en el plano generado por e_0, e_1 , pues los otros dos vectores espaciales no son relevantes en este problema⁶. Además, sobre el espacio del primer observador se puede ver la barra que estamos estudiando y una recta punteada que pasa por su extremo. La intersección de esa recta con el espacio del primer observador es el vector que se corresponde a la barra para el segundo observador.

⁶Los espacios se muestran perpendiculares respecto a las trayectorias de los observadores, sin embargo esto no es cierto porque son ortogonales respecto a la métrica de Minkowski. Se ha dibujado así por claridad y para ver que la segunda coordenada de ω está fijada y que éste es ortogonal al segundo observador.



Así, es evidente que ω es de la forma $\lambda e_0 + l e_1$. Además, este vector debe ser espacial, por tanto:

$$0 = g(\lambda e_0 + l e_1, e_0 + v e_1) = \lambda - \frac{vl}{c^2} \implies \lambda = \frac{vl}{c^2}.$$

Finalmente, para saber cuánto mide la barra para el segundo observador simplemente hay que calcular el módulo de ω en la métrica espacial $h_* = -c^2 g$. Así:

$$h_*(\omega, \omega) = -c^2 g(\omega, \omega) = -c^2 g(l(\frac{v}{c^2} e_0 + e_1), l(\frac{v}{c^2} e_0 + e_1)) = -l^2 c^2 [(\frac{v}{c^2})^2 - \frac{1}{c^2}] = l^2 [1 - (\frac{v}{c})^2].$$

Por tanto, la longitud que mide el segundo observador es $\sqrt{h_*(\omega, \omega)} = l\gamma^{-1}$. Es decir, el factor de contracción del espacio es exactamente el mismo que el de la dilatación temporal.

6.3.3. Suma de velocidades

Otro fenómeno que va en contra de nuestra intuición física es la suma de velocidades. Por ejemplo, supongamos que nos encontramos parados en un andén y vemos pasar un tren con una persona que se desplaza en el sentido del movimiento dentro del tren. Por nuestra experiencia diaria diríamos que la velocidad de esa persona es la suma de la velocidad del tren respecto a nosotros más la de la persona respecto del tren. En esta sección veremos que esta regla en verdad es incorrecta, aunque en velocidades bajas como las que experimentamos en nuestro día a día es una buena aproximación de la regla correcta.

Por simplicidad consideremos el espacio-tiempo de Minkowski definido sobre un espacio afín de dimensión 2. De esta forma la métrica tendrá signatura $(+, -)$. Consideremos una base (e_0, e_1) ortonormal. Ahora consideremos otro vector e'_0 unitario. En este sistema sus coordenadas (t, x) verifican

$$1 = g(e'_0, e'_0) = g(te_0 + xe_1, te_0 + xe_1) = t^2 - x^2.$$

Es decir, e'_0 es un vector cuyo extremo está en una hipérbola. Sabiendo esto, podemos reescribir las coordenadas como $t = \cosh(\alpha)$, $x = \sinh(\alpha)$, donde α es la longitud del arco entre e_0 y e'_0 . Así, podemos calcular el módulo de las velocidades como:

$$v = \tanh(\alpha).$$

Con esta idea en mente podemos utilizar las propiedades de \tanh para hallar la fórmula de la suma de velocidades.

Proposición 6.27. Sean tres observadores y definamos $v_{i|j}$ la velocidad aparente del observador i medida por j . Sea α la longitud de arco entre el primer observador y el segundo y β entre el segundo y el tercero. Entonces

$$v_{3|1} = \frac{v_{3|2} + v_{2|1}}{1 + \frac{v_{3|2}v_{2|1}}{c^2}}.$$

Demostración. Por lo comentado anteriormente tenemos que $v_{3|1} = c \tanh(\alpha + \beta)$. Esto lo podemos reescribir con la siguiente fórmula que no vamos a probar:

$$\tanh(\alpha + \beta) = \frac{\tanh \alpha + \tanh \beta}{1 + \tanh \alpha \tanh \beta}.$$

Por tanto, sustituyendo concluimos

$$v_{3|1} = \frac{c \tanh \alpha + c \tanh \beta}{1 + (\tanh \alpha)(\tanh \beta)} = \frac{v_{2|1} + v_{3|2}}{1 + \frac{v_{2|1} v_{3|2}}{c^2}}.$$

□

Ahora veamos que esta regla cumple las condiciones que debería cumplir. En primer lugar, a diferencia de la suma a la que estamos habitual el resultado obtenido debe pertenecer a $(0, c)$. Además, la forma en que cambian los sistemas de referencia debe ser coherente. Por ejemplo, si consideramos un observador en reposo respecto a otro, la composición nos debe devolver la velocidad original, es decir existe un elemento neutro. También cada velocidad debe tener un inverso, correspondiente con el cambio de sentido del movimiento. Por último, la composición de tres velocidades en cualquier orden ha de dar el mismo resultado; debe ser asociativa. Por tanto, la regla de suma de velocidades ha de definir un grupo abeliano.

Proposición 6.28. *El conjunto $G = (-c, c)$ con la operación*

$$u \oplus v = \frac{u + v}{1 + \frac{uv}{c^2}}$$

es un grupo abeliano.

Demostración. Es evidente que 0 es el elemento neutro y que cada v tiene inverso $-v$, y también que $u \oplus v = v \oplus u$. Por tanto, queda demostrar la asociatividad. Sean $u, v, w \in (-c, c)$, entonces

$$(u \oplus v) \oplus w = \frac{\frac{u+v}{1 + \frac{uv}{c^2}} + w}{1 + \frac{1}{c^2} \left(\frac{u+v}{1 + \frac{uv}{c^2}} \right) w} = \frac{\frac{u+v+w(1 + \frac{uv}{c^2})}{1 + \frac{uv}{c^2}}}{\frac{1 + \frac{uv}{c^2} + \frac{uv}{c^2} + \frac{vw}{c^2}}{1 + \frac{uv}{c^2}}} = \frac{u + v + w + \frac{uvw}{c^2}}{1 + \frac{uv+uw+vw}{c^2}}.$$

Por otro lado,

$$u \oplus (v \oplus w) = \frac{u + \frac{v+w}{1 + \frac{vw}{c^2}}}{1 + \frac{u}{c^2} \frac{v+w}{1 + \frac{vw}{c^2}}} = \frac{u(1 + \frac{vw}{c^2}) + (v+w)}{1 + \frac{vw}{c^2} + \frac{u(v+w)}{c^2}} = \frac{u + v + w + \frac{uvw}{c^2}}{1 + \frac{uv+uw+vw}{c^2}}.$$

Como ambas expresiones coinciden podemos afirmar que se cumple la propiedad asociativa. \square

7

Tensor de Materia

Como hemos visto, la teoría de la relatividad surge de la necesidad de explicar la mecánica clásica con los mínimos postulados posibles. Como resultado de esta forma de pensar, obtenemos una generalización de la teoría conocida que cambia por completo la visión que tenemos de ciertas magnitudes fundamentales en la mecánica como la masa. Mientras que en la visión newtoniana esta es invariante e independiente del observador, en el nuevo marco teórico es una manifestación de la energía y además tiene una relación muy estrecha con el momento. Por tanto, como el objetivo principal de la nueva teoría es generalizar estas ideas, debemos desarrollar una representación adecuada para la nueva visión que tenemos de la masa. Esta no solo debe coincidir con su representación newtoniana cuando particularicemos a sistemas no relativistas, sino que también debe mostrar su relación con la energía y el momento. Hasta ahora ya hemos presentado una representación de la masa que cumple todas estas condiciones mediante el impulso. Sin embargo, vimos que esta solo era adecuada para partículas individuales. Si recordamos, el problema que presenta esta formulación es que, aunque el impulso sea independiente del observador y su módulo sea una buena representación de la masa para una única partícula, no es una función aditiva. Por tanto, no podemos simplemente sumar los módulos de los impulsos individuales para obtener una representación adecuada del impulso total en sistemas compuestos. Esto nos obliga a buscar una formulación más general y que sí respete esta propiedad.

En este capítulo abordaremos la transición comenzando desde la 3-forma de masa hasta definir la 3-forma de impulso, capaz de estudiar sistemas de múltiples partículas y con distribuciones continuas de materia. Veremos su relación con el tensor de materia, la descripción más habitual en el contexto de la relatividad. Además, exploraremos cómo las simetrías del espacio-tiempo, representadas mediante formas de Killing, permiten identificar y calcular las cantidades conservadas.

7.1. 3-forma de masa

Empecemos por la representación más fácil que podemos hacer para una distribución continua de materia, la 3-forma de masa.

Definición 7.1. Sea M un espacio-tiempo de Minkowski. Sean $p \in \mathbb{A}_4$ y $D_1, D_2, D_3 \in T_p M$, entonces definimos la *3-forma de materia* como aquella tal que

$$\omega_{M,p}(D_1, D_2, D_3) = \left[\begin{array}{l} \text{Suma con signo de las masas de las partículas} \\ \text{contenidas en el paralelogramo infinitesimal} \\ \text{orientado } (D_1, D_2, D_3) \end{array} \right].$$

El signo de las masas se corresponde con la orientación calculada como el signo de $dX_p(T, D_1, D_2, D_3)$, donde T es la velocidad de la partícula. En coordenadas podemos expresarla como

$$\omega_M = \rho dx \wedge dy \wedge dz + w_1 dt \wedge dz \wedge dy + w_2 dt \wedge dx \wedge dz + w_3 dt \wedge dy \wedge dx,$$

donde $\rho = \omega_M(\partial_1, \partial_2, \partial_3)$ es la densidad de masa, y $w_i = \omega_M(\partial_t, \partial_j, \partial_k)$ es la densidad de flujo de masa en la dirección x_i . Es decir, la suma con signo de la masa de todas las partículas que atraviesan una unidad de área perpendicular al eje x_i por unidad de tiempo.

Observación. Nótese que $\omega_{3,p}$ está definido sobre el espacio tangente a p , por tanto esta definición es válida para un espacio-tiempo cualquiera.

Proposición 7.2. Sea un espacio-tiempo $(M, g, c, W^+, [dX])$ y supongamos que la distribución de materia se mueve según $T \in TM$, entonces la 3-forma de materia se calcula como $\omega_M = \rho i_T(dX)$. Fijado un sistema inercial, si $T = \partial_t + v_1 \partial_1 + v_2 \partial_2 + v_3 \partial_3$ entonces

$$\omega_M = \rho(dx \wedge dy \wedge dz - v_1 dt \wedge dy \wedge dz + v_2 dt \wedge dx \wedge dz - v_3 dt \wedge dx \wedge dy)$$

Demostración. Empecemos considerando la situación en la que todas las partículas tienen una velocidad constante T . Entonces la densidad de flujo a través de cualquier plano coordenado es nulo, pues ninguna partícula cruza alguno de ellos. Es decir, $\omega_i = 0$ para $i = 1, 2, 3$. Si consideramos un sistema de referencia inercial tal que $T = \partial_t$ tenemos que

$$\begin{cases} \omega_M(\partial_1, \partial_2, \partial_3) = \rho, \\ \omega_M(\partial_t, \partial_i, \partial_j) = 0, \quad i, j = 1, 2, 3. \end{cases}$$

De esta forma podemos escribir la 3-forma como

$$\omega_M = \rho(t, x, y, z) dx \wedge dy \wedge dz = \rho(i_{\partial_t} dX) = \rho(i_T dX).$$

Ahora consideremos un campo T cualquiera. Fijada una región pequeña del espacio sabemos que la materia se moverá siguiendo las curvas integrales. Tomemos un sistema de referencia con el mismo estado de movimiento que el definido por la curva integral en dicha región. Así, por lo anterior tenemos que $\omega_M = \rho (i_T dX)$. Como $T = \partial_t + v_1 \partial_1 + v_2 \partial_2 + v_3 \partial_3$, sustituyendo obtenemos

$$\begin{aligned}\omega_M = \rho (i_T dX) &= \rho \left(i_{\partial_t + v_1 \partial_1 + v_2 \partial_2 + v_3 \partial_3} (dt \wedge dx \wedge dy \wedge dz) \right) \\ &= \rho \left(dx \wedge dy \wedge dz - v_1 dt \wedge dy \wedge dz + v_2 dt \wedge dx \wedge dz - v_3 dt \wedge dx \wedge dy \right).\end{aligned}$$

□

Ejemplo 7.3. Sea un cilindro de densidad $\rho(x, y, z)$ constante en el tiempo, centrado en el eje Z y que rota alrededor de dicho eje con velocidad angular ω . Entonces este caso, el campo de velocidad

$$T = \partial_t + \omega (-y \partial_x + x \partial_y).$$

Entonces, sustituyendo en la fórmula anterior tenemos

$$\omega_M = \rho dx \wedge dy \wedge dz + \omega \rho (y dt \wedge dx \wedge dz - x dt \wedge dy \wedge dz).$$

7.2. 3-forma de impulso

Ahora definiremos una 3-forma análoga a la anterior, pero en vez de con la masa, lo haremos con el impulso.

Definición 7.4. Sea \mathbb{A}_4 un espacio-tiempo de Minkowski. Para cada evento $p \in \mathbb{A}_4$ y tres vectores $D_1, D_2, D_3 \in T_p \mathbb{A}_4$, definimos la *3-forma de impulso*

$$\Pi_{3,p}(D_1, D_2, D_3) = \left[\begin{array}{c} \text{Suma con signo de los impulsos de las partículas} \\ \text{que atraviesan el paralelogramo infinitesimal} \\ \text{orientado } (D_1, D_2, D_3) \end{array} \right],$$

donde el signo de cada impulso I se determina como el signo de $dX_p(I, D_1, D_2, D_3)$.

Observación. De nuevo, nótese que $\Pi_{3,p}$ está definido sobre el espacio tangente a p , por tanto esta definición es válida para un espacio-tiempo cualquiera.

En este caso la intuición del criterio del signo es ligeramente más complicada pues el impulso es una magnitud vectorial. Consideremos un hipervolumen K , si tomamos un vector N que apunte hacia afuera podemos orientar el borde mediante la contracción $i_N dX$. De esta forma orientamos el

paralelepípedo (D_1, D_2, D_3) considerando un K tal que dicho volumen sea una de sus caras. Cuando el paralelepípedo (D_1, D_2, D_3) esté positivamente orientado, tomaremos el impulso con signo positivo si este apunta hacia afuera de (D_1, D_2, D_3) . De esta forma podemos interpretar la forma de impulso de manera equivalente como

$$\Pi_3(D_1, D_2, D_3) = \left[\begin{array}{c} \text{Densidad de impulso total de } K \\ \text{saliente por la cara } (D_1, D_2, D_3) \end{array} \right] - \left[\begin{array}{c} \text{Densidad de impulso total de } K \\ \text{entrante por la cara } (D_1, D_2, D_3) \end{array} \right].$$

Esta interpretación se puede afinar aún más evaluando en unos campos concretos. Supongamos un sistema de referencia $(p; \partial_t, \partial_1, \partial_2, \partial_3)$ ortonormal. Entonces si tomamos $\Pi_3(\partial_t, \partial_2, \partial_3)$, estamos midiendo el flujo de impulso que cruza un volumen espacial perpendicular al eje temporal. Es decir, es la densidad de impulso en un instante dado. Así, la coordenada temporal es ρ y en las espaciales tendremos la densidad de momento. Por tanto, denotando como $w_i := \rho v_i$ a la densidad de momento en la dirección x_i podemos escribirlo como

$$\Pi_3(\partial_t, \partial_2, \partial_3) = \rho \partial_t + w_1 \partial_1 + w_2 \partial_2 + w_3 \partial_3.$$

Análogamente, $\Pi_3(\partial_t, \partial_i, \partial_j)$ mide el flujo de impulso que atraviesa el plano generado por ∂_i, ∂_j durante un desplazamiento infinitesimal en la dirección temporal. Como dicho flujo debe ser proporcional a la velocidad en la dirección x_k y a la densidad de impulso tendremos que

$$\Pi_3(\partial_t, \partial_i, \partial_j) = v_k(\rho \partial_t + \rho v_1 \partial_1 + \rho v_2 \partial_2 + \rho v_3 \partial_3), \quad \text{con } \{i, j, k\} = \{1, 2, 3\}.$$

Evidentemente, intercambiar el orden en las evaluaciones sólo afecta a la orientación del paralelepípedo y por tanto al signo de Π_3 .

Observación. Nótese que $\omega_M(D_1, D_2, D_3) = dt(\Pi_3(D_1, D_2, D_3))$.

Definición 7.5. Un *sistema de partículas libres* es un modelo continuo de la materia en el que las partículas no colisionan entre sí y se mueven según un campo de velocidades continuo.

Proposición 7.6. *Sea un sistema libre de partículas en un espacio-tiempo cualquiera. Sean $p \in M$ un evento y T el campo de velocidad del sistema. Sea (D_1, D_2, D_3) una base de vectores espaciales en $T_p M$ ortonormal y orientada y $(\theta_0, \theta_1, \theta_2, \theta_3)$ la base dual correspondiente a (T, D_1, D_2, D_3) . Entonces, la 3-forma de impulso Π_3 en dicha región viene dada por*

$$\Pi_3 = (\theta_1 \wedge \theta_2 \wedge \theta_3) \otimes (\rho T) = \rho (i_T dX) \otimes T,$$

donde ρ representa la densidad de masa.

Demostración. Como todas las partículas se mueven en la dirección de T entonces el impulso por unidad del volumen (D_1, D_2, D_3) es $\Pi_3(D_1, D_2, D_3) = \rho T$. Como T es tangente al movimiento de las partículas, ninguna atraviesa una superficie generada por (T, D_i, D_j) . Por tanto, $\Pi_3(T, D_i, D_j) = 0$. Como $(\theta_0, \theta_1, \theta_2, \theta_3)$ es la base dual de (T, D_1, D_2, D_3) , entonces podemos considerar la forma volumen $dX = \theta_0 \wedge \theta_1 \wedge \theta_2 \wedge \theta_3$. Vemos que al contraer con T se obtiene $i_T dX = \theta_1 \wedge \theta_2 \wedge \theta_3$. Finalmente, si consideramos la 3-forma $\Pi_3 = (\theta_1 \wedge \theta_2 \wedge \theta_3) \otimes (\rho T) = \rho (i_T dX) \otimes T$, vemos que verifica todas las condiciones anteriores:

$$\begin{cases} \Pi_3(D_1, D_2, D_3) = \rho T, \\ \Pi_3(T, D_i, D_j) = 0, \quad i, j = 1, 2, 3. \end{cases}$$

Por lo que concluimos que $\Pi_3 = \rho (i_T dX) \otimes T$. □

Ejemplo 7.7. Sea un cilindro de densidad $\rho(x, y, z)$ constante en el tiempo, que rota alrededor del eje Z con velocidad angular constante ω . Entonces el campo de velocidad espacial que describe el movimiento del fluido es $T = \partial_t + \omega(-y \partial_x + x \partial_y)$.

Para este caso vimos en el ejemplo 7.3 que

$$\rho (i_T dX) = \rho dx \wedge dy \wedge dz + \omega \rho (y dt \wedge dx \wedge dz - x dt \wedge dy \wedge dz).$$

Como $\Pi_3 = \rho (i_T dX) \otimes T$, sustituyendo llegamos a

$$\begin{aligned} \Pi_3 = & \rho dx \wedge dy \wedge dz \otimes (\partial_t + \omega(-y \partial_x + x \partial_y)) \\ & + \omega \rho (y dt \wedge dx \wedge dz - x dt \wedge dy \wedge dz) \otimes (\partial_t + \omega(-y \partial_x + x \partial_y)). \end{aligned}$$

A partir del ejemplo del cilindro en rotación, claramente la diferencia entre ambas representaciones. La primera solo tiene en cuenta la densidad y el flujo de masa pero no su dirección. En cambio, la 3-forma de impulso contiene además la dirección del movimiento. Así, mientras que ω_M responde a la pregunta de cuánta masa atraviesa una región infinitesimal, Π_3 nos dice, además, con qué velocidad y en qué dirección lo hace. Esto resulta crucial en relatividad, pues como hemos visto la masa depende del observador, por lo que ω_M también. Mientras que el impulso es independiente del observador y contienen toda la información necesaria para caracterizar el sistema.

7.3. Tensor de materia

Una vez hemos entendido qué es la 3-forma de impulso y porqué esta representa adecuadamente la materia, es momento de presentar el tensor de materia. Este tensor aunque es menos intuitivo que la forma de impulso, es el objeto más ampliamente usado para codificar la información acerca de la materia en relatividad.

Teorema 7.8. Sea (M, dX) una variedad orientada entonces a cada $\Pi_3 \in \Lambda^3 M \otimes TM$ le corresponde un tensor 2-contravariante $T^2 \in TM \otimes TM$, tal que

$$\Pi_3 = C_1^1(dX \otimes T^2).$$

Demostración. Tomemos la forma de volumen dX y un campo $D \in \mathfrak{X}(M)$. Como dX es no degenerada la aplicación $D \mapsto i_D dX$ define un isomorfismo entre TM y $\Lambda^3 M$. Haciendo el producto tensorial con TM , obtenemos un isomorfismo entre $TM \otimes TM$ y $\Lambda^3 M \otimes TM$. Es decir, a cada forma de impulso le corresponde un tensor 2-contravariante.

Ahora construyamos explícitamente la correspondencia. Si consideramos el siguiente tensor

$$dX \otimes T^2 \in \Lambda^4 M \otimes TM \otimes TM$$

y contraemos el primer índice covariante con el primer contravariante obtenemos un elemento en $\Lambda^3 M \otimes TM$. Es decir,

$$\Pi_3 = C_1^1(dX \otimes T^2) \in \Lambda^3 M \otimes TM.$$

□

Definición 7.9. Se denomina *tensor de materia contravariante* al tensor 2-contravariante $T^2 \in TM \otimes TM$ para el cual la forma impulso Π_3 verifica:

$$\Pi_3 = C_1^1(dX \otimes T^2).$$

Aunque la forma de impulso Π_3 y el tensor de materia T^2 contengan la misma información, el uso de T^2 es más extendido que el de la forma de impulso. Quizás el motivo detrás de esto es que Π_3 depende de la orientación del espacio-tiempo, mientras que T^2 no. Sin embargo, aunque Π_3 no sea la forma más empleada, la hemos presentado en primer lugar pues tiene una interpretación física más directa al ser contravariante. Veamos ahora qué información nos aporta el nuevo tensor definido.

Proposición 7.10. En un espacio-tiempo los coeficientes del tensor T^2 representan las siguientes magnitudes físicas:

$$\left\{ \begin{array}{l} T^2(dt, dt) = \text{densidad de masa} = \rho(t, x_1, x_2, x_3) \\ T^2(dt, dx_j) = \text{densidad de momento en la dirección } x_j = w_j(t, x_1, x_2, x_3) \\ T^2(dx_i, dt) = \text{densidad de flujo de masa en la dirección } x_i \\ T^2(dx_i, dx_j) = \text{densidad de flujo de momento en la dirección } x_j \text{ a través de una superficie perpendicular a } x_i \end{array} \right.$$

Además, el tensor T^2 es simétrico y verifica la $i_{dt} T^2 = \Pi_3(\partial_1, \partial_2, \partial_3)$.

Demostración. Para evaluar el tensor T^2 , debemos hacerlo por definición. Así:

$$\Pi_3 = C_1^1(dX \otimes T^2) = \sum_{\mu, \nu \in \{t, x_1, x_2, x_3\}} T^{\mu\nu} [i_{\partial_\mu}(dt \wedge dx_1 \wedge dx_2 \wedge dx_3)] \otimes \partial_\nu.$$

Empecemos evaluando en los vectores espaciales $\Pi_3(\partial_1, \partial_2, \partial_3)$:

$$\Pi_3(\partial_1, \partial_2, \partial_3) = \sum_{\mu, \nu} T^{\mu\nu} [i_{\partial_\mu}(dt \wedge dx_1 \wedge dx_2 \wedge dx_3)](\partial_1, \partial_2, \partial_3) \partial_\nu.$$

Calculemos las contracciones. Para $\mu = t$ es fácil ver que

$$i_{\partial_t}(dt \wedge dx_1 \wedge dx_2 \wedge dx_3)(\partial_1, \partial_2, \partial_3) = (dx_1 \wedge dx_2 \wedge dx_3)(\partial_1, \partial_2, \partial_3) = 1.$$

Veamos para el resto de índices:

$$\begin{aligned} i_{\partial_i}(dt \wedge dx_1 \wedge dx_2 \wedge dx_3)(\partial_1, \partial_2, \partial_3) &= -(dt \wedge i_{\partial_i}(dx_1 \wedge dx_2 \wedge dx_3))(\partial_1, \partial_2, \partial_3) \\ &= -dt(\partial_1) \cdot i_{\partial_i}(dx_1 \wedge dx_2 \wedge dx_3)(\partial_2, \partial_3) \\ &= 0. \end{aligned}$$

Es decir, los únicos términos no nulos son los que tienen el índice $\mu = t$. Así

$$\Pi_3(\partial_1, \partial_2, \partial_3) = T^{tt} \partial_t + T^{t1} \partial_1 + T^{t2} \partial_2 + T^{t3} \partial_3 = i_{dt} T^2.$$

Si ahora recordamos la interpretación de la 3-forma de impulso, vimos que

$$\Pi_3(\partial_1, \partial_2, \partial_3) = \rho \partial_t + w_1 \partial_1 + w_2 \partial_2 + w_3 \partial_3,$$

donde w_i es la densidad de momento en la dirección x_i . Ahora, si comparamos coeficientes, llegamos a

$$\boxed{T^2(dt, dt) = T^{tt} = \rho, \quad T^2(dt, dx_j) = T^{tj} = w_j \quad (j = 1, 2, 3).}$$

Ahora hagamos lo mismo evaluando en $(\partial_t, \partial_i, \partial_j)$ con $\partial_i \neq \partial_j$ y ∂_k ortogonal a los anteriores. Es fácil ver que los términos con $\mu = t$ son nulos. Así, podemos reescribir la expresión como

$$\Pi_3(\partial_t, \partial_i, \partial_j) = - \sum_{\substack{\mu \neq t \\ \nu}} T^{\mu\nu} [i_{\partial_\mu}(dx_1 \wedge dx_2 \wedge dx_3)](\partial_t, \partial_i, \partial_j) \partial_\nu = \sum_{\substack{\mu \neq t \\ \nu}} (-1)^k T^{\mu\nu} \partial_\mu^k \partial_\nu = \sum_{\nu} (-1)^k T^{k\nu} \partial_\nu,$$

De nuevo, llegados a este punto comparamos con la expresión de Π_3 que vimos en su interpretación:

$$\Pi_3(\partial_t, \partial_i, \partial_j) = v_k (\rho \partial_t + \rho v_1 \partial_1 + \rho v_2 \partial_2 + \rho v_3 \partial_3), \quad \text{con } \{i, j, k\} = \{1, 2, 3\}.$$

Así, concluimos que

$$T^2(dx_k, dt) = T^{kt} = (-1)^k \rho v_k, \quad T^2(dx_k, dx_i) = T^{ki} = (-1)^k \rho v_k v_i, \quad (i = 1, 2, 3).$$

Es decir, $T^2(dx_k, dt)$ es el flujo de la densidad de masa y $T^2(dx_k, dx_i)$ se puede interpretar de dos maneras según agrupes. Si lo agrupas como $(-1)^k (\rho v_k) v_i$ entonces es la densidad del k -ésimo momento en la dirección i , y si lo agrupas como $(-1)^k (\rho v_i) v_k$ es la densidad del i -ésimo momento en la dirección k . Esto nos indica además que el tensor T^2 es simétrico.

Para verlo en los cálculos basta evaluar en la segunda entrada de T^2 en vez de la primera y se ve que simplemente los índices intercambian sus papeles. Es decir, T^2 es simétrico. \square

Veamos un ejemplo sencillo del tensor de materia volviendo al caso de un sistema de partículas libre.

Proposición 7.11. *Sea un sistema de partículas libres de densidad ρ y campo de velocidades U en un espacio-tiempo $(X, g, c, W^+, [dX])$. Entonces su tensor de materia es $T^2 = \rho U \otimes U$.*

Demostración. Reescribiendo la expresión de la forma de impulso de un sistema de partículas libres que vimos en la proposición 7.6 tenemos que $\Pi_3 = \rho (i_U dX) \otimes U = C_1^1(dX \otimes (\rho U \otimes U))$. Identificando términos con la definición 7.9 concluimos que $T^2 = \rho U \otimes U$. \square

Definición 7.12. *Sea un espacio-tiempo $(M, g, c, W^+, [dX])$ el cual contiene materia según indica el tensor T^2 . Sean dos campos $D_1, D_2 \in \mathfrak{X}(X)$. Entonces definimos el *tensor de materia covariante* como*

$$T_2(D_1, D_2) := T^2(g(D_1, -), g(D_2, -)).$$

Proposición 7.13. *El tensor T_2 de materia de un sistema de partículas libres se puede expresar como*

$$T_2 = \rho dt \otimes dt.$$

Demostración. Expresando los campos en un sistema de coordenadas inercial usando la notación de Einstein tenemos que

$$g(D_1, -) = g(D_1^t \partial_t + D_1^i \partial_i, -) = D_1^t dt,$$

$$g(D_2, -) = g(D_2^t \partial_t + D_2^j \partial_j, -) = D_2^t dt.$$

Entonces por definición y sustituyendo lo que hemos visto en la proposición 7.10

$$T_2(D_1, D_2) = T^2(D_1^t dt, D_2^t dt) = D_1^t D_2^t T^2(dt, dt) = (\rho dt \otimes dt)(D_1, D_2).$$

\square

Observación. Nótese que cuando el isomorfismo entre el fibrado tangente y cotangente existe, entonces ambas formas del tensor de materia tiene la misma información. Sin embargo, esto no sucede en la mecánica clásica, la versión contravariante mucho más informativa que la covariante.

Ahora hagamos un análisis dimensional de todos los objetos vistos hasta ahora. En el caso de los tensores sólo sus evaluaciones las tienen, no ellos mismos. Sin embargo podemos estudiar cuál es el factor por el que su resultado se ve afectado cuando hacemos un reescalamiento en los elementos a evaluar. Denotaremos como T , M y L a las magnitudes tiempo, masa y longitud respectivamente.

Empecemos con la dimensión de la 1-forma ω que mide el tiempo de forma lineal, por tanto $[\omega] = T$. Además, si $[U] = T^{-1}$ pues definimos la forma temporal como aquella tal que $\omega(U) = 1$. Como Π_3 mide el impulso por unidad de volumen y por lo anterior, el impulso tiene unidades $\frac{M}{T}$, entonces $[\Pi_3] = \frac{M}{T}$. Por otro lado, es fácil ver que la dimensión de dX es TL^3 . Veamos ahora la dimensión de T^2 . Por definición tenemos que $\Pi_3 = C_1^1(dX \otimes T^2)$, por lo que podemos escribir que

$$[T^2] = \frac{[\Pi_3]}{[dX]} = ML^{-3}T^{-2}.$$

Por último, debemos estudiar el tensor covariante T_2 . Para ello, debemos darnos cuenta de que la dimensión de g es T^2 , y por tanto también la de la forma $g(D, -)$. Así, como por definición $T_2 = T^2(g(D, -), g(D, -))$, entonces tenemos que $[T_2] = [T^2]T^4 = ML^{-3}T^2$.

Estas expresiones se pueden simplificar aún más si hacemos una elección inteligente de las unidades. Como la idea es estudiar los campos gravitatorios, podemos darnos cuenta de que la aceleración provocada por la fuerza gravitatoria no depende de su masa. Así, si escogemos unidades de forma que $G = 1$ tenemos que dicha aceleración es $\frac{m}{r^2}$, y por tanto podemos considerar que la masa en este contexto es de magnitud $\frac{L^3}{T^2}$. Sustituyendo esto en las expresiones anteriores vemos que $[T_2] = 1$. Es decir, en el contexto de campos gravitatorios este tensor es el único invariante por cambios de escala.

Observación. Si la fuerza de la Ley de Gravitación Universal fuera proporcional a r^{-n} tendríamos que $[T_2] = L^{n-2}$; no sería un tensor natural.

Antes de seguir estudiando otros aspectos de la formulación de la materia, vamos una distribución algo más realista que el sistema de partículas libres.

Definición 7.14. Un *fluido perfecto* es un modelo continuo de materia en el que se describe el movimiento de las partículas mediante un campo de velocidades continuo. En este modelo las partículas del fluido sí interactúan entre sí, de forma que para todo evento existe un observador para el cual la velocidad media del fluido en un evento dado es $U = \partial_t$. Además, no incluye efectos de viscosidad ni de conducción de calor.

Proposición 7.15. *Sea un fluido perfecto en un espacio-tiempo. En el sistema inercial donde la velocidad media del fluido es ∂_t la forma de impulso es:*

$$\Pi_3 = \rho dx \wedge dy \wedge dz \otimes \partial_t + p (dt \wedge dz \wedge dy \otimes \partial_1 + dt \wedge dx \wedge dz \otimes \partial_2 + dt \wedge dy \wedge dx \otimes \partial_3),$$

donde ρ es la densidad de masa y p es lo que llamamos presión escalar del fluido.

Demostración. Como por hipótesis la velocidad media aparente es nula, al evaluar en la forma de impulso tenemos:

$$\Pi_3(\partial_t, \partial_2, \partial_3) = \rho \partial_t.$$

Es decir que el primer término de Π_3 es $\rho dx \wedge dy \wedge dz \otimes \partial_t$.

Ahora consideremos ahora el flujo de momento a través de una cara ortogonal a ∂_1 . Como en el sistema de referencia escogido la velocidad media del sistema es ∂_t , el flujo de densidad de impulso que sale es igual al que entra. Entonces como tienen sentidos contrarios y suponiendo que el paralelepípedo está orientado positivamente tenemos que

$$\Pi_3(\partial_t, \partial_1, \partial_2) = 2 \left[\begin{array}{c} \text{Densidad de impulso total entrante} \\ \text{por la cara } (\partial_t, \partial_1, \partial_2) \end{array} \right].$$

Esta magnitud es la presión que ejerce todo el fluido sobre las caras del paralelepípedo que estamos considerando. Así:

$$\Pi_3(\partial_t, \partial_2, \partial_3) = -p\partial_1,$$

con $p > 0$. Análogamente, tenemos que la presión sobre los paralelepípedos generados por $(\partial_t, \partial_1, \partial_3)$ y $(\partial_t, \partial_2, \partial_3)$ es $-p\partial_2$ y $-p\partial_1$ respectivamente. Sin embargo, ahora debemos comprobar la orientación de cada una de las caras. Así, evaluando dX sobre las cuaternas $(\partial_3, \partial_t, \partial_1, \partial_2)$, $(\partial_2, \partial_t, \partial_1, \partial_3)$ y $(\partial_1, \partial_t, \partial_2, \partial_3)$ comprobamos que sólo la segunda está orientada positivamente, por lo que los signos correctos son

$$\Pi_3(\partial_t, \partial_1, \partial_2) = p\partial_3$$

$$\Pi_3(\partial_t, \partial_1, \partial_3) = -p\partial_2,$$

$$\Pi_3(\partial_t, \partial_2, \partial_3) = p\partial_1.$$

Por tanto, podemos escribir estos términos de Π_3 como:

$$p (dt \wedge dy \wedge dz \otimes \partial_1 + dt \wedge dz \wedge dx \otimes \partial_2 + dt \wedge dx \wedge dy \otimes \partial_3).$$

Finalmente sumando todos los términos concluimos.

$$\Pi_3 = \rho dx \wedge dy \wedge dz \otimes \partial_t + p (dt \wedge dy \wedge dz \otimes \partial_1 + dt \wedge dz \wedge dx \otimes \partial_2 + dt \wedge dx \wedge dy \otimes \partial_3).$$

□

Corolario 7.16. *El tensor de materia de un fluido perfecto es $T^2 = \rho U \otimes U + pb$, donde $b := -\partial_1 \otimes \partial_1 - \partial_2 \otimes \partial_2 - \partial_3 \otimes \partial_3$ es la métrica espacial en su forma contravariante.*

Demostración. Para ello, simplemente hay que intentar escribir la expresión anterior en términos de dX y de la contracción C_1^1 . Así :

$$\begin{aligned}\Pi_3 &= \rho dx \wedge dy \wedge dz \otimes \partial_t + p (dt \wedge dy \wedge dz \otimes \partial_1 + dt \wedge dz \wedge dx \otimes \partial_2 + dt \wedge dx \wedge dy \otimes \partial_3) \\ &= C_1^1 (dX \otimes \rho \partial_t \otimes \partial_t + p (-\partial_1 \otimes \partial_1 - \partial_2 \otimes \partial_2 - \partial_3 \otimes \partial_3)) \\ &= C_1^1 (dX \otimes (\rho U \otimes U + pb)),\end{aligned}$$

□

Es decir, podemos ver cómo el tensor de materia para un fluido perfecto difiere del correspondiente al sistema de partículas libres. Así, podemos decir generalizar esto y decir que la diferencia entre el tensor de materia de un modelo y el de un sistema libre se corresponde a las fuerzas internas.

Definición 7.17. El *tensor de tensión* de un modelo continuo de materia con campo de velocidades U y densidad ρ se define como $\mathcal{T}^2 := T^2 - \rho U \otimes U$.

7.4. Conservación del impulso

Una vez ya sabemos tratar con distribuciones continuas de materia, podemos estudiar ciertas leyes de conservación que encontramos en los sistemas dinámicos. Comenzaremos viendo cómo se formula matemáticamente la conservación del impulso en el contexto continuo, generalizando lo visto para el caso de un conjunto discreto

7.4.1. Caso galileano

Empecemos en un espacio-tiempo galileano $(\mathbb{A}_4, \omega, \bar{g}, [dX])$ con una distribución de materia que está descrita por una forma de impulso Π_3 . Consideremos una región compacta y conexa K del espacio-tiempo. Al ser un espacio tiempo galileano existe la forma diferencial ω que determina la dirección temporal absoluta. De esta manera, podemos pensar K como un tubo de radio variable cuya dirección longitudinal es la indicada por ω . Así, para cada valor de ω tenemos una sección tridimensional de K . Por construcción, cada una de las secciones es una región espacial. Así, podemos aplicar el teorema de Stokes:

$$\int_K d\Pi_3 = \int_{K_{t_1}} \Pi_3 - \int_{K_{t_0}} \Pi_3 + \int_{\text{lateral}} \Pi_3.$$

El signo negativo en el segundo término se debe a que el impulso de las partículas apunta hacia el interior de K . En la parte lateral del tubo, si el impulso apunta hacia fuera, corresponde a una partícula

saliente y tienen signo positivo. Mientras que si apunta hacia dentro, representa partícula entrante. Por tanto, podemos escribir esa integral como:

$$\int_{\text{lateral}} \Pi_3 = \begin{bmatrix} \text{Impulso} \\ \text{saliente} \end{bmatrix} - \begin{bmatrix} \text{Impulso} \\ \text{entrante} \end{bmatrix}.$$

Por tanto, la expresión original puede interpretarse como

$$\int_K d\Pi_3 = \begin{bmatrix} \text{Variación total del impulso} \\ \text{de todas las partículas contenidas} \\ \text{en las secciones } K_t, t_0 \leq t \leq t_1 \end{bmatrix}.$$

Es decir, la condición $d\Pi_3 = 0$ expresa una ley infinitesimal de conservación del impulso. Podemos ahora generalizar esta observación y tratar de obtener la ley de conservación para el universo completo. Supongamos que existen dos tubos disjuntos K y K' , de modo que uno contiene al menos una trayectoria completa de una partícula y el otro contiene toda la materia restante. Por construcción Π_3 se anula fuera de la unión $K \cup K'$ y sobre la superficie lateral de K . Si ahora suponemos la ley de conservación $d\Pi_3 = 0$, por lo visto antes:

$$\int_{K_t} \Pi_3 \text{ no depende de } t.$$

Ahora, si se reduce progresivamente el tamaño de K' mientras se aumenta el de K , manteniendo en todo momento contenida toda la materia en la unión de ambos tubos, se sigue conservando el impulso total. En el caso extremo en que $K' = \emptyset$, obtenemos una ley de conservación global:

$$\int_{K_t} \Pi_3 \text{ no depende de } t.$$

Esta formulación unifica de manera continua la conservación del impulso.

Teorema 7.18. *Dado un espacio-tiempo galileano y una distribución continua de materia definida sobre él. Si $d\Pi_3 = 0$ el impulso se conserva.*

7.4.2. Caso lorentziano

Ahora veremos que la conservación del impulso también se puede escribir como $d\Pi_3 = 0$ en el caso de una estructura lorentziana. La idea detrás de la demostración es exactamente la misma. La diferencia con lo anterior es que no tenemos un tiempo absoluto, por lo que hay que hacer un trabajo previo en la definición de las secciones de K . Empecemos enunciando un teorema conocido de geometría diferencial que será fundamental en esta construcción.

Teorema 7.19 (Teorema del flujo tubular). *Sea M una variedad diferenciable y $X \in \mathfrak{X}(M)$ un campo vectorial diferenciable tal que $X_p \neq 0$ para un punto $p \in M$. Entonces, existe una carta (U, ϕ) alrededor de p , con coordenadas locales (x^1, \dots, x^m) , tal que en U se cumple:*

$$X|_U = \frac{\partial}{\partial x^1}.$$

Con esto en mente, ya podemos demostrar la condición de conservación del impulso.

Teorema 7.20 (Conservación del impulso). *Sea $K \subset M$ una región compacta y conexa en un espacio-tiempo de Minkowski equipado con una estructura lorentziana $(\mathbb{A}_4, g, c, W^+, [dX])$ con materia distribuida según la forma de impulso Π_3 . Entonces $d\Pi_3 = 0$ modeliza la conservación del impulso en K .*

Demostración. Al estar en \mathbb{A}_4 podemos definir un campo de vectores T de tipo tiempo y orientados hacia el futuro¹ en un abierto $U \supset K$. Al ser K compacto, existe un recubrimiento finito $\{(U_i)\}_{i=1}^N$ de K por abiertos en U . Como T está formado por vectores de tipo tiempo orientados hacia el futuro podemos aplicar el teorema del flujo tubular en cada abierto (U_i) . Así, tenemos que para cada una de ellas existen coordenadas locales $(x_i^1, x_i^2, x_i^3, x_i^4)$ tales que

$$T|_{U_i} = \frac{\partial}{\partial x_i^1}.$$

A partir de estos sistemas de coordenadas podemos definir una función $t_i := x_i^1 : U_i \rightarrow \mathbb{R}$, en cada abierto tal que $T(t_i) = 1$. Es decir, dada una curva integral de T , la función t_i es estrictamente creciente, y podemos definir las secciones

$$\Sigma_\tau^i := \{p \in U_i : t_i(p) = \tau\}.$$

Así, sobre cada abierto U_i definimos $K_\tau^i = K \cap \Sigma_\tau^i$. Ahora que tenemos las secciones razonamos como antes aplicando el teorema de Stokes a Π_3 sobre la porción $K \cap U_i$:

$$\int_{K \cap U_i} d\Pi_3 = \int_{\partial(K \cap U_i)} \Pi_3 = \int_{K_{t_1}^i} \Pi_3 - \int_{K_{t_0}^i} \Pi_3 + \int_{\text{lateral} \cap U_i} \Pi_3,$$

y llegamos a que

$$\int_{K \cap U_i} d\Pi_3 = \left[\begin{array}{c} \text{Variación total del impulso} \\ \text{de todas las partículas contenidas} \\ \text{en las secciones } K_{t_0}^i, t_0 \leq t \leq t_1 \end{array} \right].$$

¹Aquí la hipótesis de estar en el espacio-tiempo de Minkowski no es fundamental, esto sigue siendo posible en el caso más general pues siempre consideramos la variedad M orientable

Si $d\Pi_3 = 0$, tenemos que el impulso se conserva en cada abierto.

Ahora debemos ver que lo mismo sucede en todo K . Para ello, debemos estudiar qué pasa en las intersecciones de los abiertos. Razonemos una vez más de la misma manera sobre la intersección $K \cap (U_i \cap U_j)$. Como estamos en la intersección, tenemos dos formas de dar secciones sobre este abierto dependiendo de qué función utilicemos para describir el tiempo:

$$\Sigma_\tau^{ij} := \{p \in U_i \cap U_j : t_i(p) = \tau\}, \quad \Sigma_{\tau'}^{ij} := \{p \in U_i \cap U_j : t_j(p) = \tau'\}.$$

Denotemos las secciones como $K_\tau^{ij} := K \cap \Sigma_\tau^{ij}$, $K_{\tau'}^{ij} := K \cap \Sigma_{\tau'}^{ij}$. Así, podemos aplicar Stokes sobre dos conjuntos de secciones diferentes

$$\int_{K \cap (U_i \cap U_j)} d\Pi_3 = \int_{K_{t_1}^{ij}} \Pi_3 - \int_{K_{t_0}^{ij}} \Pi_3 + \int_{\text{lateral} \cap (U_i \cap U_j)} \Pi_3 = \left[\begin{array}{l} \text{Variación total del impulso} \\ \text{de todas las partículas contenidas} \\ \text{en las secciones } K_t^{ij}, t_0 \leq t \leq t_1 \end{array} \right].$$

Podemos razonar de manera análoga con las secciones $K_{t'}^{ij}$. De esta forma si $d\Pi_3 = 0$, independientemente de las secciones escogidas la variación total del impulso es nula.

Finalmente, veamos que el flujo también es nulo en la unión de dos abiertos con intersección no vacía. Para ello, en primer lugar vemos que la frontera se puede descomponer en dos partes: la que está contenida en otro abierto U_j , la frontera compartida, y la que no, la frontera externa. Ahora razonemos por contradicción. Supongamos que el flujo de impulso no se conserva en la unión. Entonces al considerar sus fronteras externas, existirá un flujo neto saliente² en ambas. Esto implicaría que, en la intersección $U_i \cap U_j$, la dirección del flujo sería entrante en ambas partes que forman dicha frontera común, ya que el flujo en cada abierto sabemos que es nulo. Esto da lugar a un flujo neto entrante en $K \cap (U_i \cap U_j)$. Pero esto contradice la conservación local del impulso en la intersección³. Por tanto, el flujo de impulso a través de las fronteras externas debe compensarse mutuamente.

Finalmente, podemos aplicar este argumento de forma iterativa, considerando un abierto la unión anterior y un tercer abierto que lo interseque. Así, concluimos que el flujo total de impulso a través de la frontera de K es necesariamente cero, garantizando la conservación global del impulso en K . \square

Este razonamiento sólo se puede hacer sobre variedades que sean planas. Calculemos $d\Pi_3$ empleando la notación de Einstein

$$d\Pi_3 = d(\omega^i \otimes e_i) = d\omega^i \otimes e_i + \omega^i \otimes de_i = 0.$$

Vemos que en el caso de ser plana el término de_i es nulo y entonces

$$d\Pi_3 = 0 \iff d\omega^i = 0 \quad \forall i.$$

²Podemos considerar entrante también, el razonamiento es análogo

³Nótese que si flujos netos en la cara externa fueran de distinto signo pero diferentes en magnitud, también generan un flujo neto en la intersección no nula.

Por tanto, el impulso sí se conserva. En el caso de que haya curvatura $de_i \neq 0$ y por tanto no tenemos que $d\omega^i = 0 \quad \forall i$; no podemos decir nada del impulso. Sin embargo, aunque haya curvatura este razonamiento es válido en el espacio tangente, por ello decimos que es una *ley de conservación infinitesimal*.

Por otro lado, veamos que la condición de conservación del impulso también puede formularse en términos del tensor de materia. No demostraremos este resultado, pues no nos da una mayor comprensión de lo que estamos estudiando.

Proposición 7.21. *Sea $K \subset M$ una región compacta y conexa en un espacio-tiempo equipado con una estructura lorentziana $(M, g, c, W^+, [dX])$ con materia distribuida según el tensor de materia T^2 . Entonces $\text{div } T^2 = 0$ modeliza la conservación del impulso en K .*

Demostración. Se obtiene a partir de la definición del tensor de materia $\Pi_3 = C_1^1(dX \otimes T^2)$, que tomando la diferencial exterior queda $d\Pi_3 = dX \otimes \text{div } T^2$. Los cálculos se pueden encontrar en [NS17, Prop. F.2.2 pg. 190]. \square

Corolario 7.22. *Para el caso de un sistema libre de partículas de velocidad U la conservación del impulso se expresa mediante $\text{div}(\rho U \otimes U) = 0$. Para un fluido perfecto de velocidad U y presión p se tiene: $\text{div}(\rho U \otimes U + p h) = 0$.*

7.5. Formas de Killing y leyes de conservación

A partir de la conservación del impulso podemos estudiar otras magnitudes conservadas empleando el resultado anterior. Para hacerlo, buscaremos 1-formas para evaluar la forma de impulso y obtener 3-formas a la que estamos habituados que nos den la información que buscamos. Con esta idea, debemos ver cómo escoger ω de una forma inteligente para que esto nos informe sobre las conservaciones del sistema. Estas formas son las que se conocen como formas de Killing y para poder estudiarlas, debemos ver algunos resultados previos.

Definición 7.23. *Sea M una variedad diferenciable de dimensión $n + 1$ entonces llamamos *formas de impulso* a las n -formas con valores vectoriales $\Pi_n \in \Lambda^n M \otimes TM$ tales que $C_1^1(\Pi_n) = 0$, o análogamente, $\Pi_n \wedge I = 0$, donde I es la 1-forma identidad con valores vectoriales.*

Veamos ahora que las formas de impulso con las que hemos trabajado anteriormente efectivamente cumplen estas condiciones equivalentes y, por tanto, la definición que hemos escogido es apropiada.

Proposición 7.24. *Sea M variedad diferenciable de dimensión $n + 1$ y sea Π_n una forma de impulso. Entonces verifica que $C_1^1(\Pi_n) = 0$ y $\Pi_n \wedge I = 0$, donde el producto exterior se considera respecto a $TM \otimes T^*M$.*

Demostración. La forma de impulso en un sistema de coordenadas locales (x^0, x^1, \dots, x^n) se escribe:

$$\Pi_n = \sum_{\nu=0}^n \sum_{0 \leq i_1 < \dots < i_n \leq n} \Pi_{i_1 \dots i_n}^{\nu} dx^{i_1} \wedge \dots \wedge dx^{i_n} \otimes \partial_{\nu}.$$

La contracción del primer vector del tensor con el primer índice de la forma se escribe:

$$C_1^1(\Pi_n) = \sum_{\nu=1}^n \sum_{i_1 < \dots < i_n} \sum_{k=1}^n (-1)^{k-1} \Pi_{i_1 \dots i_n}^{\nu} \partial_{i_k}^{\nu} dx^{i_1} \wedge \dots \wedge \widehat{dx^{i_k}} \wedge \dots \wedge dx^{i_n},$$

donde $\widehat{dx^{i_k}}$ denota que dicha forma no aparece en el producto. Evaluando $C_1^1(\Pi_n)$ sobre una n -upla arbitraria de vectores $(\partial_{i_1}, \dots, \partial_{i_{n-1}})$, es fácil ver que

$$C_1^1(\Pi_n)(\partial_{j_1}, \dots, \partial_{j_{n-1}}) = 0 \quad \text{para todo } 1 \leq j_1 < \dots < j_{n-1} \leq n.$$

Veamos ahora que las formas de impulso también verifican que $\Pi_n \wedge I = 0$, con $I = \sum_k dx^k \otimes \partial_k$. Escribiendo en coordenadas el producto anterior tenemos

$$(\Pi_n \wedge I)(D_1, \dots, D_{n+1}) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_{n+1}} \text{sgn}(\sigma) \Pi_n(D_{\sigma(1)}, \dots, D_{\sigma(n)}) D_{\sigma(n+1)}^k,$$

con D_1, \dots, D_{n+1} campos vectoriales (véase, [N24, pg. 395]).

Dado que Π_n es alternada respecto a sus índices de forma, es decir, $\Pi_{i_1 \dots i_n}^k$ cambia de signo ante cualquier transposición de los índices i_j . Por tanto, esta suma recoge todas las permutaciones posibles de los vectores y, como Π_n es alternada, dicho sumatorio se cancela, con lo que se concluye. \square

Presentamos ahora el siguiente lema fundamental a la hora de trabajar con las formas de Killing.

Lema 7.25. *Si Π_n es una forma de impulso y H es un tensor 2-covariante de orden alternado, entonces se cumple*

$$\Pi_n \wedge H = 0,$$

donde H es una 1-forma con valores en T^*M y el producto exterior se toma respecto a $TM \otimes T^*M$.

Demostración. [NS17, Lema F.1.2 pg.187] \square

Con todo esto ya podemos empezar con las formas de Killing.

Definición 7.26. Sea (M, ∇) una variedad diferenciable equipada con una conexión. Decimos que $\omega \in T^*M$ es una *forma Killing* cuando el tensor 2-covariante $\nabla\omega$ es alternado.

Proposición 7.27. *Nótese que si ∇ no tiene torsión, entonces ω es de Killing si y sólo si $2\nabla\omega = d\omega$.*

Demostración. Teniendo en cuenta que ω es de Killing, y por tanto alternado, y que ∇ no tiene torsión:

$$\begin{aligned} 2(\nabla\omega)(X, Y) &= (\nabla_X\omega)(Y) - (\nabla_Y\omega)(X) \\ &= X(\omega(Y)) - \omega(\nabla_X Y) - Y(\omega(X)) + \omega(\nabla_Y X) \\ &= X(\omega(Y)) - Y(\omega(X)) - \omega(\nabla_X Y - \nabla_Y X) \\ &= X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]) \end{aligned}$$

Finalmente, por el lema A.11 concluimos $\nabla\omega = d\omega$. \square

Veamos ahora la propiedad fundamental de las formas de Killing.

Proposición 7.28. *Sea Π_n una forma impulso en una variedad diferenciable M equipada con una conexión sin torsión tal que $d\Pi_n = 0$. Si ω es una 1-forma de Killing, entonces se cumple que $d(\omega(\Pi_n)) = 0$.*

Demostración. Sea $p \in M$ y (D_1, \dots, D_n) una base de $T_p M$. En una base fijada podemos escribir la forma de impulso en coordenadas como $\Pi_n = \sum_i \eta^i \otimes D_i$, donde $\eta^i \in \Lambda^n M$. Evaluando $\omega(\Pi_n)$ en una n -upla de campos (V_1, \dots, V_n) tenemos

$$\omega(\Pi_n(V_1, \dots, V_n)) = \omega\left(\sum_i \eta^i(V_i) D_i\right) = \sum_i \eta^i(V_i) \omega(D_i).$$

Es decir, podemos expresar $\omega(\Pi_n)$ como

$$\omega(\Pi_n) = \sum_i \eta^i \otimes \omega(D_i).$$

Por otro lado, podemos identificar mediante un isomorfismo ω con $1 \otimes \omega$. De esta forma si hacemos $\Pi_n \wedge (1 \otimes \omega)$ con el producto exterior respecto a $TM \otimes T^*M$ tenemos que

$$\Pi_n \wedge (1 \otimes \omega) = \sum_i (\eta^i \otimes D_i) \wedge (1 \otimes \omega) = \sum_i \eta^i \otimes \omega(D_i),$$

es decir, $\omega(\Pi_n) = \Pi_n \wedge (1 \otimes \omega)$. Tomando la diferencial exterior y recordando que ω es una forma de Killing tenemos que

$$\begin{aligned} d(\omega(\Pi_n)) &= d(\Pi_n \wedge (1 \otimes \omega)) = (d\Pi_n) \wedge (1 \otimes \omega) + (-1)^n \Pi_n \wedge d(1 \otimes \omega) \\ &= 0 + (-1)^n \Pi_n \wedge (2\nabla\omega). \end{aligned}$$

De nuevo, por ser ω de Killing tenemos que es un tensor 2-covariante alternado, y por el lema 7.25 concluimos que $d(\omega(\Pi_n)) = 0$. \square

Es decir, al aplicar las formas de Killing sobre formas con valores vectoriales cerradas, obtenemos otra forma cerrada. Como ya vimos, el lenguaje apropiado para expresar conservación de magnitudes distribuidas de forma continua en el espacio-tiempo es precisamente que son cerradas. Por tanto, cada forma Killing nos indica una conservación de una magnitud.

Corolario 7.29. *Sea M un espacio-tiempo con una estructura galileana o lorentziana. Sea una distribución de materia determinada por Π_3 tal que $d\Pi_3 = 0$ y K un compacto y conexo en M . Entonces $\omega(\Pi_3)$ es una cantidad conservada.*

Demostración. Basta ver que podemos identificar $\omega(\Pi_3)$ con forma diferenciable con valores vectoriales en los que todas las entradas salvo una son nulas. Evidentemente su diferencial seguiría siendo nula y así podemos aplicar los teoremas 7.18 y 7.20 según corresponda para concluir que $\omega(\Pi_n)$ es una cantidad conservada. \square

Ahora intentemos caracterizar todas las formas Killing de un espacio afín.

Teorema 7.30. *Las formas de Killing en el espacio afín \mathbb{A}^n forman un espacio vectorial de dimensión $\binom{n+1}{2}$. Fijada una base, las formas Killing están generadas por las siguientes*

$$dx^i, \quad x^i dx^j - x^j dx^i.$$

Demostración. Sea $\omega = \sum_i f_i dx^i$ una 1-forma cualquiera en coordenada. Entonces se tiene

$$\nabla\omega = \sum_i df_i \otimes dx^i = \sum_{i,j} (\partial_j f_i) dx^j \otimes dx^i.$$

La condición de que $\nabla\omega$ sea alternada es $\partial_i f_j = -\partial_j f_i$. Derivando esta condición obtenemos que las segundas derivadas se anulan:

$$\partial_k \partial_j f_i = -\partial_k \partial_i f_j = \partial_j \partial_i f_k = -\partial_j \partial_k f_i.$$

Por tanto, las funciones deben ser de la forma $f_i = \sum_j a_{ij} x^j + b_i$. Además, si sustituimos en la condición $\partial_i f_j = -\partial_j f_i$, tenemos que $a_{ij} = -a_{ji}$ y por tanto:

$$\omega = \sum_i b_i dx^i + \sum_{i < j} a_{ij} (x^i dx^j - x^j dx^i).$$

\square

Observación. Aunque el resultado anterior se obtenga sobre espacios afines, esto no perdemos generalidad, pues al integrar en variedades lo hacemos sobre los espacios tangentes, que son espacios vectoriales.

Teorema 7.31. *En un espacio-tiempo M equipada con una estructura galileana o lorentziana, con una distribución de materia dada por una forma de impulso Π_3 . Si hay conservación del impulso entonces para todo observador inercial también se conserva la masa, el momento lineal y el momento angular. Además, el centro de masas sigue un movimiento rectilíneo uniforme; un movimiento inercial.*

Demostración. Empecemos fijando un observador inercial con coordenadas (t, x_1, x_2, x_3) . Tomemos ahora un compacto conexo K y consideremos las secciones K_t siguiendo la idea de los teoremas 7.18 y 7.20 según corresponda. Por otro lado, sabemos por el teorema 7.30 que dx_i y $x_i dx_j - x_j dx_i$ son formas Killing. Además, por el corolario 7.29 conocemos que dada una forma Killing hay una cantidad conservada. Para ver a cuál se corresponde cada una debemos recordar que por la interpretación de la forma de impulso vista en la sección 7.2 esta verifica que

$$\Pi_3(\partial_1, \partial_2, \partial_3) = \rho \partial_t + w_1 \partial_1 + w_2 \partial_2 + w_3 \partial_3.$$

Así podemos tomar las formas Killing e integrar $\omega(\Pi_3)$ en un K_t cualquiera.

- Si tomamos dt entonces la siguiente integral es constante

$$\int_{K_t} dt(\Pi_3) = \int_{K_t} \rho dx_1 dx_2 dx_3.$$

Es decir, la masa se conserva.

- Si tomamos dx_i con $i = 1, 2, 3$ la siguiente integral es constante

$$\int_{K_t} dx_i(\Pi_3) = \int_{K_t} w_i dx_1 dx_2 dx_3.$$

Es decir, el momento lineal se conserva.

- Si tomamos $x_i dx_j - x_j dx_i$ con $i, j = 1, 2, 3$ la siguiente integral es constante

$$\int_{K_t} (x_i dx_j - x_j dx_i)(\Pi_3) = \int_{K_t} (x_i \omega_j - x_j \omega_i) dx_1 dx_2 dx_3.$$

Es decir, el momento angular se conserva.

- Si tomamos $x_i dt - t dx_i$ entonces la siguiente integral es constante

$$\int_{K_t} (x_i dt - t dx_i)(\Pi_3) = \int_{K_t} (x_i \rho - t \omega_i) dx_1 dx_2 dx_3 = C'.$$

Ahora sí dividimos entre la masa total m medida por el observador y reorganizamos términos vemos que

$$\frac{1}{m} \int_{K_t} (x_i \rho dx_1 dx_2 dx_3) = C + \frac{t}{m} \int_{K_t} \omega_i dx_1 dx_2 dx_3.$$

Vemos que el miembro izquierdo es el centro de masas y por lo anterior, la otra integral es un vector constante. Por tanto, esta es la ecuación del movimiento del centro de masas; que sigue un movimiento rectilíneo uniforme.

□

Observación. El hecho de que el centro de masas siga una trayectoria inercial junto con la Segunda Ley de Newton relativista 6.21, indica que la situación en la que el impulso se conserva es cuando hay ausencia de fuerzas.

Para terminar la sección vamos a ver como para el caso de un fluido perfecto la conservación del impulso codifica la ecuación del movimiento y la conservación de la masa. Para ello, necesitamos primero el siguiente lema.

Lema 7.32. *Sea (M, ∇) una variedad diferenciable equipada con un conexión. Sea $\rho \in C^\infty(X)$ es una función diferenciable y U un campo de vectores tangentes, entonces*

$$\operatorname{div}_\nabla(\rho U \otimes U) = (U\rho + \rho \operatorname{div}_\nabla U) U + \rho \nabla_U U.$$

Demostración. Véase [NS17, Prop. F.2.3 pg. 190]

□

Proposición 7.33. *Consideremos un fluido perfecto de densidad ρ , presión p y velocidad U . Entonces la conservación del impulso $\operatorname{div}(\rho U \otimes U + p\mathbf{h}) = 0$ es equivalente a*

$$\begin{cases} U\rho + \rho \operatorname{div} U = 0, & \text{Conservación de la masa} \\ \rho \nabla_U U = -\operatorname{grad} p. & \text{Ecuación del movimiento} \end{cases}$$

Demostración. Por el lema anterior tenemos calculado $\operatorname{div}(\rho U \otimes U) = (U\rho) U + \rho(\operatorname{div}_\nabla U)U + \rho \nabla_U U$. Por otro lado es fácil ver que $\operatorname{div}(p \sum_i \partial_i \otimes \partial_i) = \sum_i (\partial_i p) \partial_i = \operatorname{grad} p$. Así:

$$\operatorname{div}(\rho U \otimes U) = ((U\rho) + \rho \operatorname{div}_\nabla U)U + \rho \nabla_U U + \operatorname{grad} p$$

Como U tiene la primera coordenada constante e igual a 1, entonces como $\nabla_U U$ debe tener dicha coordenada nula; es un vector espacial. Por otro lado, como p es espacial su gradiente también lo es. Por tanto, $\operatorname{div}(\rho U \otimes U) = 0$ si y sólo si se cumple el siguiente sistema:

$$\begin{cases} U\rho + \rho \operatorname{div} U = 0, \\ \rho \nabla_U U = -\operatorname{grad} p, \end{cases}$$

donde la primera ecuación es la de la conservación de la masa y la segunda la ecuación de movimiento.

□

7.6. Movimiento con fuerzas externas

En la sección anterior una vez vimos todas las magnitudes conservadas nos dedicamos a estudiar el movimiento de la materia en dos casos sencillos. El primero, el movimiento del centro de masas, salía de manera inmediata junto con las cantidades conservadas. El segundo, fue el caso de un fluido perfecto. Ambos los estudiamos siempre en el caso en el que el impulso se conserva. Como vimos con la Segunda Ley de Newton Relativista 6.21, el caso en el que se conserva el impulso es cuando hay ausencia de fuerzas; cosa que parece que se verifica para distribuciones continuas de materia pues el centro de masas cuando se conserva el impulso sigue una trayectoria inercial. En esta sección veremos en primer lugar que efectivamente son las fuerzas externas las causantes de que el impulso no se conserve y además veremos las ecuaciones de la dinámica.

Consideremos un observador inercial y un paralelepípedo generado por $\partial_1, \partial_2, \partial_3$ en un espacio-tiempo en el que existe un campo espacial D . Al ser este paralelepípedo infinitamente pequeño por continuidad del campo podemos asumir que D es constante en él. Ahora, si dejamos moverse al paralelepípedo en un instante de tiempo infinitesimalmente pequeño según indique el campo D , veremos que se ha estirado en la dirección temporal. De esta manera, y abusando de notación, podemos decir que se habrá estirado $dt \cdot \|D\|$ unidades. De esta forma el volumen del paralelepípedo quedaría

$$dx_1 dx_2 dx_3 dt \|D\|.$$

Ahora recordemos que lo que queremos es representar la materia. Por tanto, en realidad lo que estamos diciendo es que en un volumen suficientemente pequeño durante un tiempo suficientemente pequeño podemos suponer que la trayectoria que siguen las partículas es la misma. Sin embargo, la descripción anterior no termina de ser útil pues no sólo nos interesa cuánto se mueve la materia si no hacia adonde. Para ello, simplemente consideramos el vector D completo; sin el módulo. De esta manera tenemos una buena representación de lo que sucede en cada dirección del espacio-tiempo. De esta forma estaríamos considerando

$$dt dx_1 dx_2 dx_3 D.$$

Si profundizamos aún más nos damos cuenta que D en realidad representa el impulso por unidad de volumen y tiempo. Esto se ve fácil suponiendo que tenemos una función de densidad ρ que en cada paralelepípedo suponemos constante, y viendo que

$$\rho dx_1 dx_2 dx_3 dt D = m_{ap} dt D,$$

es decir, es la masa trasladada por unidad de tiempo en la dirección D . Concretamente la variación del impulso, pues es cuánto se ha movido esa masa y en qué dirección desde que empezamos a medir el tiempo.

Ahora vamos a buscar una forma alternativa de escribir este producto. Si recordamos, Π_3 es una 3-forma alternada con valores vectoriales. Entonces si tomamos su diferencial $d\Pi_3$ es una 4-forma con valores vectoriales. Además, como vimos, esta diferencial indica precisamente la variación de momento. Por tanto, si evaluamos en el paralelepípedo podemos afirmar que

$$d\Pi_3(\partial_1, \partial_2, \partial_3) = dt \, dx_1 \, dx_2 \, dx_3 \, D,$$

donde estamos empleando dx_i para indicar la longitud infinitesimal de las aristas, no como forma diferencial. Por tanto, esto lo podemos escribir en forma tensorial como:

$$d\Pi_3(\partial_1, \partial_2, \partial_3) = dt \wedge dx_1 \wedge dx_2 \wedge dx_3 \otimes D = dX \otimes D,$$

donde dx_i denota la forma diferencial. Por último, debemos notar que la magnitud

$$[D] = \frac{[\text{Impulso}]}{V \, T} = \frac{M \frac{L}{T}}{V \, T} = \frac{1}{V} \frac{ML}{T^2} = \frac{1}{V} [\text{Fuerza}],$$

que dijimos que era el impulso por unidad de volumen y tiempo, haciendo un sencillo análisis dimensional vemos que es fuerza por unidad de volumen; esto es, densidad de fuerza que denotamos \vec{F} .

Así, podemos escribir que $d\Pi_3(\partial_1, \partial_2, \partial_3) = dt \wedge dx_1 \wedge dx_2 \wedge dx_3 \otimes D = dX \otimes \vec{F}$, y como comentamos en la demostración de 7.21 tenemos que $d\Pi_3 = dX \otimes \text{div } T^2$. Por tanto, podemos afirmar que la ecuación del movimiento en función del tensor de materia es

$$\text{div } T^2 = \vec{F}.$$

Sabiendo esto desarrollemos la ecuación para el caso de un fluido tal que la presión interna no es igual en todas las direcciones espaciales. Así, en un sistema de coordenadas (t, x_1, x_2, x_3) , tenemos que

$$T_2 = \rho \, \partial_t \otimes \partial_t + \sum_{i=1}^3 w_i (\partial_i \otimes \partial_t + \partial_t \otimes \partial_i) + \sum_{i,j=1}^3 h_{ij} \partial_i \otimes \partial_j.$$

Entonces por definición de divergencia llegamos a

$$\text{div } T_2 = (\partial_t \rho) \partial_t + \sum_{i=1}^3 (\partial_i w_i) \partial_t + \sum_{i=1}^3 (\partial_t w_i) \partial_i + \sum_{i,j=1}^3 (\partial_i h_{ij}) \partial_j.$$

Por otro lado, como la fuerza es un vector espacial, tenemos que $F = F_1 \partial_1 + F_2 \partial_2 + F_3 \partial_3$. Finalmente, igualando las coordenadas espaciales y temporales por separado, llegamos a

$$\begin{cases} 0 = \partial_t \rho + \text{div}(w), & \text{(Ecuación de continuidad)} \\ F_i = \partial_t w_i + \text{div}(h_i), & \text{(Ecuaciones de Euler)} \end{cases}$$

con $\mathbf{h}_i = h_{i1}\partial_1 + h_{i2}\partial_2 + h_{i3}\partial_3$. La ecuación de continuidad codifica la conservación de la masa, y las ecuaciones de Euler describen el balance de momento.

Por último, antes de acabar el capítulo y empezar con Relatividad General, es importante hacer un comentario sobre la última observación que hemos hecho. Hemos dicho que ya suponemos que todo lo que está en el espacio de un observador inercial tiene la misma componente temporal que lo hace simultáneo. Además, estamos diciendo que las fuerzas sólo pueden ser espaciales. Por tanto, no se debería dar la ruptura de la simultaneidad que tanto trabajamos más arriba, dando lugar a una gran contradicción en la teoría. Lo sorprendente es que en realidad no hay ninguna contradicción; veámoslo sólo para el caso de las partículas.

8

Relatividad General

Como ya vimos en el capítulo 6, el espacio-tiempo de Minkowski $(\mathbb{A}_4, g, c, W^+, [dX])$ es el caso más sencillo en el que podemos estudiar la relatividad. Vimos que se trata de una formulación de la mecánica en ausencia de gravedad coherente con los postulados del capítulo 5. Una situación más general es cuando consideramos esta estructura con una variedad diferenciable cualquiera $(M, g, c, W^+, [dX])$. Este es el marco en el que encontramos la teoría de la Relatividad General, y la diferencia fundamental con anterior es la curvatura del espacio-tiempo; cuya manifestación es la gravedad. Además, como el espacio tangente tiene estructura de espacio vectorial, podemos ver que el espacio-tiempo de Minkowski es en realidad la estructura infinitesimal de la Relatividad General. Como el espacio tangente hereda la estructura Lorentziana de la variedad, obtenemos el espacio-tiempo de Minkowski sobre el espacio tangente a cada evento ¹. Por tanto, en Relatividad General, se tiene una variedad cuya estructura infinitesimal es de un espacio-tiempo de Minkowski y la relación entre todos estos espacios viene determinada por la conexión. Esta conexión no puede ser otra que la de Levi-Civita, pues debe preservar la métrica y además no debe tener una dirección privilegiada; no puede tener torsión; por tanto es única. No sólo eso, sino que la métrica bajo ciertas condiciones define unívocamente el tensor T_2 y por tanto la distribución de materia. A su vez vimos en el capítulo 6 que el movimiento inercial se expresa como $\nabla_I I = 0$; geodésicas del espacio tiempo. Por tanto, como la métrica define la curvatura podremos resumirla relatividad en la famosa frase de que son la materia y la energía la que curva el espacio-tiempo y es el espacio-tiempo el que indica a la materia como moverse. Lo más importante de este capítulo será obtener la ecuación de campo de Einstein que nos mostrará la relación entre la curvatura del espacio-tiempo y la masa o energía contenido en él. Pero antes de tratar estos tenemos que ver cómo las ideas de la mecánica en el espacio-tiempo de Minkowski son las mismas que en el caso más general, y que el único cambio a tener en cuenta es que la conexión no es trivial.

¹Esta observación justifica por qué la simplificación aún sigue siendo válida cuando trabajamos localmente incluso en presencia de un campo gravitatorio.

8.1. Mecánica

Definición 8.1. Llamamos *variedad lorentziana* a una variedad diferenciable M de dimensión 4 equipada con una estructura lorentziana que denotaremos como $(M, g, c, W^+, [dX])$.

Corolario 8.2. Sea M una variedad Lorentziana y $p \in M$ un evento. Entonces $(T_p M, g, c, W^+, [dX])$ es un espacio-tiempo de Minkowski.

Definición 8.3. Sean dos variedades Lorentzianas $(M, g, c, W^+, [dX])$ y $(\bar{M}, \bar{g}, \bar{c}, \bar{W}^+, [d\bar{X}])$. Decimos que un difeomorfismo $\bar{\phi} : \bar{M} \rightarrow M$ es un isomorfismo² de variedades de Lorentz si

$$\bar{g} = \lambda^2(\phi^* g),$$

donde λ es una función constante no nula.

Con esta formulación, el concepto de orientación se preserva sin modificaciones, por lo que podemos trabajar exactamente del mismo modo que antes, sin que sea necesario introducir cambios en el tratamiento de los vectores de tipo tiempo, espacio o luz.

Por otro lado, si recordamos todos los conceptos, desde el tiempo propio hasta la conservación del impulso, se formulan en términos infinitesimales. Como en el caso general la estructura infinitesimal es la del espacio-tiempo de Minkowski entonces lo podemos definir de la misma manera que en el capítulo 6. Sin embargo, ahora no estamos en el caso de un espacio tiempo plano, por lo que la conexión no es la trivial y debemos emplear el transporte paralelo inducido por la conexión de Levi-Civita.

Así, podemos definir las trayectorias u observadores como curvas $\gamma \subset M$ cuyos vectores tangentes sean temporales. De igual manera existe un único campo unitario orientado hacia el futuro tangente a la trayectoria, lo que nos permite definir el tiempo propio entre dos eventos p y q como

$$\tau := \int_p^q i_U g.$$

Además, de igual manera podemos definir el impulso I y con ello la masa en reposo de una partícula como $m := \sqrt{g(I, I)}$. Donde encontramos la mayor diferencia es a la hora de definir los sistemas de referencia inerciales. Infinitesimalmente estamos en un espacio-tiempo de Minkowski y se define igual, pero a lo largo de las trayectorias estos deben moverse acorde a la curvatura de la variedad. Para ello, empleamos las coordenadas normales y la aplicación exponencial que vimos en la definición 3.43. Sean $p \in X$ un evento, $U_p \in T_p X$ un vector temporal orientado hacia el futuro y (U_p, D_1, D_2, D_3) una base ortonormal de $T_p X$. Entonces dicho observador puede definir un sistema inercial como:

$$\begin{aligned} \mathbb{R}^4 &\longrightarrow W \subset M \\ (t, x_1, x_2, x_3) &\longmapsto \exp_p(tU_p + \sum_i x_i D_i) \end{aligned}$$

²Añadir el término λ^2 es para tener en cuenta el reescalamiento de las unidades.

Así, también podemos definir parte del espacio para el observador U_p como $(\mathbb{S} := \langle U_p \rangle^\perp, -g)$. En el caso de conocer la trayectoria $\sigma(a, b) \rightarrow M$ con velocidad U_p en un punto p , podemos definir el sistema de referencia transportando la base $(U, D_1, D_2, D_3) \in T_p M$ como

$$\begin{aligned} \mathbb{R}^4 &\longrightarrow W \subset M \\ (t, x_1, x_2, x_3) &\longmapsto \exp_{\sigma(t)}(\sum_i x_i D_i). \end{aligned}$$

8.2. Ecuación de Einstein

En esta sección demostraremos el resultado más importante del trabajo, la ecuación de Einstein. Para ello trataremos de encontrar un tensor de materia 2-covariante que dependa de la métrica de una variedad lorentziana $T_2(g)$. Para hacerlo, supondremos tres hipótesis fundamentales:

1. El tensor $T_2(g)$ debe ser un **tensor natural**. Supongamos que tenemos dos variedades lorentzianas M y \bar{M} difeomorfas mediante el difeomorfismo $\tau : \bar{M} \rightarrow M$. Entonces el tensor que obtengamos para la métrica τ^*g debe ser igual al transformar $T_2(g)$ por τ^* . Es decir,

$$T_2(\tau^*g) = \tau^*(T_2(g)) \quad \text{para todo difeomorfismo } \tau.$$

2. El tensor T_2 debe ser **invariante ante reescalamientos de las unidades** usando unidades naturales para la masa. Ya vimos en la sección 7.3 que el tensor T_2 siempre cumplía dicha propiedad. Así:

$$T_2(\lambda^2 g) = T_2(g) \quad \text{para cualquier } \lambda > 0.$$

3. El tensor T_2 debe tener **divergencia nula**. Es decir, estamos en un caso sin fuerzas externas. Así

$$\operatorname{div} T_2 = 0$$

Una hipótesis extra es que $T_2(g)$ en cada punto depende sólo de las derivadas de g hasta un cierto orden k^3 . Una vez tenemos esto claro necesitaremos saber los siguientes dos lemas que utilizaremos en la demostración de la ecuación de campo.

Lema 8.4. *Sea f una función diferenciable definida sobre un producto $V_2 \times \cdots \times V_k$ de espacios vectoriales reales de dimensión finita. Si existe un número real b tal que $f(\lambda^2 v_2, \dots, \lambda^k v_k) = \lambda^b f(v_2, \dots, v_k)$ para todo número real positivo λ , entonces $f = 0$ o bien f es una suma finita de funciones polinomiales*

$$f(v_2, \dots, v_k) = p_2(v_2) \cdots p_k(v_k),$$

donde cada p_i es un polinomio homogéneo de grado d_i y $2d_2 + \cdots + kd_k = b$.

³En realidad se puede suponer que T_2 está definido localmente, pero por razones de extensión del trabajo no lo trataremos.

Demostración. Véase [NS17, Teorema 7.2.1, pg.98]. \square

Lema 8.5. *Sea (X, ∇) una variedad equipada con una conexión ∇ . Entonces se verifica la siguiente igualdad*

$$\operatorname{div}(\operatorname{Ric}) = \frac{1}{2}d\kappa,$$

donde κ es la curvatura escalar y d denota la diferencial exterior.

Demostración. Empecemos considerando la segunda identidad de Bianchi tomando $X = e_i$, $Y = e_j$ y $Z = e_i$ donde $(e_k)_{k=1}^{n+1}$ son campos coordenados en un sistema de coordenadas. Así, sumando en los índices i, j :

$$\sum_{i,j} \nabla_{e_i} R(e_j, Z, e_j, e_i) + \nabla_{e_j} R(Z, e_i, e_j, e_i) + \nabla_Z R(e_i, e_j, e_j, e_i) = 0.$$

Ahora estudiemos cada término por separado.

- **Primer término.** Si aplicamos las propiedades 3, 5 y 4 de la proposición 4.16 y la definición del tensor de Ricci tenemos que

$$\sum_{i,j} \nabla_{e_i} R(e_j, Z, e_j, e_i) = \sum_{i,j} \nabla_{e_i} R(e_j, e_i, e_j, Z) = - \sum_{i,j} \nabla_{e_i} R(e_j, e_i, Z, e_j) = \sum_{i,j} \nabla_{e_i} R(e_i, e_j, Z, e_j) = \sum_i \nabla_{e_i} \operatorname{Ric}(e_i, Z)$$

- **Segundo término.** Por la propiedad 4 definición, se ve inmediatamente que es $\sum_j \nabla_{e_j} \operatorname{Ric}(Z, e_j)$.
- **Tercer término.** Aplicando la propiedad 5 de la proposición 4.16 y por definición de tensor de Ricci y de curvatura escalar tenemos que

$$\sum_{i,j} \nabla_Z R(e_i, e_j, e_j, e_i) = - \sum_{i,j} \nabla_Z R(e_i, e_j, e_i, e_j) = - \sum_i \nabla_Z \operatorname{Ric}(e_i, e_i) = -\nabla_Z \kappa = -d\kappa(Z).$$

Sustituyendo en la segunda identidad de Bianchi tenemos que $2 \sum_i \nabla_{e_i} \operatorname{Ric}(e_i, Z) - d\kappa(Z) = 0$. y por la definición de divergencia tenemos que

$$\operatorname{div}(\operatorname{Ric}) = \frac{1}{2}d\kappa.$$

\square

Teorema 8.6 (Ecuación de campo). *Sea una variedad una variedad lorentziana, entonces el tensor de materia T_2 está unívocamente determinado por la métrica g mediante la siguiente ecuación*

$$\operatorname{Ric} - \frac{\kappa}{2}g = 8\pi T_2,$$

donde κ denota la curvatura escalar.

Demostración. La demostración la haremos sobre una variedad M de dimensión $n + 1$ y por tanto con una métrica de signatura $(1, -1, \dots, -1)$. Por la hipótesis extra que hicimos sobre T_2 , el tensor sólo depende de g hasta una derivada de orden k en cualquier punto. Entonces

$$T_2\left(\sum_{ij} g_{ij} dx_i \otimes dx_j\right) = \sum_{ab} T_{ab}(g_{ij}, \partial_k g_{ij}, \partial_\alpha g_{ij}) dx_a \otimes dx_b,$$

donde $T_{ab} \in C^\infty(M)$ y $2 \leq |\alpha| \leq k$ y $\partial_\alpha = \frac{\partial^{|\alpha|}}{\partial_1 x^{\alpha_1} \dots \partial_n x^{\alpha_n}}$.

Fijemos un punto $p \in M$ y consideremos un sistema de referencia en coordenadas normales. Entonces, por el lema A.9, sabemos que las primeras derivadas de g son nulas. Denotemos las derivadas como $g_{ij,\alpha} = (\partial_\alpha g_{ij})(x)$ y desarrollemos en serie de Taylor la métrica:

$$g = \sum_{ij} (\mu_{ij} + \sum_{2 \leq |\alpha| \leq k} g_{ij,\alpha} x^\alpha + \dots) dx_i \otimes dx_j,$$

donde $\mu_{00} = 1, \mu_{11} = \dots, \mu_{nn} = -1$ y en otro caso $\mu_{ij} = 0$.

Ahora consideremos el difeomorfismo local τ tal que su *pull-back* verifica $\tau^*(x_i) = \lambda x_i$. Aplicando esta transformación al tensor tenemos que

$$\tau^*(T_2(g)_x) = \lambda^2 \left(\sum_{ab} T_{ab}(g_{ij}, 0, \partial_\alpha g_{ij}) dx_a \otimes dx_b \right).$$

Por otro lado, si aplicamos el difeomorfismo τ a la métrica:

$$\tau^*g = \lambda^2 \sum_{ij} (\mu_{ij} + \sum_{2 \leq |\alpha| \leq k} \lambda^{|\alpha|} g_{ij,\alpha} x^\alpha + \dots) dx_i \otimes dx_j.$$

Ahora, como el tensor que buscamos es natural, por la hipótesis 1 tenemos que

$$\tau^*(T_2(g)_x) = T_2(\tau^*g_x).$$

Además, por la hipótesis 2 de invarianza bajo cambio de unidades tenemos que

$$\tau^*(T_2(g)_x) = T_2(\lambda^{-2}(\tau^*g_x)).$$

Escrito en componentes:

$$\lambda^2 T_{ab}(\mu_{ij}, 0, g_{ij,\alpha}) = T_{ab}(\mu_{ij}, 0, \lambda^{|\alpha|} g_{ij,\alpha}).$$

Podemos renombrar estas funciones como $f_{ab}(g_{ij}) := T_{ab}(\mu_{ij}, 0, g_{ij,\alpha})$. Entonces la expresión anterior la podemos escribir como $f_{ab}(\lambda^{|\alpha|} g_{ij,\alpha}) = \lambda^2 f_{ab}(g_{ij,\alpha})$.

Por el lema 8.4 las funciones $f_{ab}(g_{ij,\alpha})$ son una suma finita de funciones polinomiales $p_2(v_2) \dots p_k(v_k)$ donde cada p_i es un polinomio homogéneo de grado d_i tal que

$$2d_2 + \dots + kd_k = 2.$$

Como evidentemente $d_i \geq 1$ y entero, no queda otra que $d_2 = 1$ y el resto nulos. Es decir, $f_{ab}(g_{ij,\alpha})$ son funciones lineales de las segundas derivadas. Por otro lado, si aplicamos la proposición 4.21 e igualamos componente a componente al desarrollo de Taylor que hicimos anteriormente⁴

$$\mu_{ij} + \sum_{2 \leq |\alpha| \leq k} g_{ij,\alpha} x^\alpha + \dots = g_{ij}(x) = \mu_{ij} - \frac{1}{3} \sum_{k,l=1}^n R_{ikjl}(0) x^k x^l + O(|x|^3) \implies g_{ij,kl} = \frac{-1}{3} R_{ikjl}.$$

Es decir, las funciones $f_{ab}(g_{ij,\alpha})$ son funciones lineales de los coeficientes del tensor de curvatura.

Hasta aquí hemos demostrado que, efectivamente, el tensor de materia depende de un tensor de curvatura, como veníamos buscando. Ahora lo que queremos es determinar cuál es dicho tensor. Para ello, debemos considerar una aplicación lineal⁵ $\phi : R_{2,2} \rightarrow S_2$ que va del conjunto $R_{2,2}$ de tensores con las simetrías del tensor de curvatura, al conjunto S_2 de tensores simétricos 2-covariantes en \mathcal{M} . Además, esta aplicación debe cumplir una propiedad fundamental, siguiendo la misma idea que vimos en la sección 5.2 para estudiar la variedad de métricas. En ese caso si transformábamos una métrica por un elemento del grupo de Lorentz $O(1, 3)$ entonces teníamos una métrica equivalente; representan la misma métrica. Y por eso tomábamos el cociente por ese grupo para estudiarlas. En este caso, si tenemos una métrica en unas coordenadas y la transformamos por un elemento de $O(1, n)$ entonces el tensor de curvatura correspondiente debe ser el mismo. Es decir, la aplicación ϕ es una aplicación lineal $O(1, n)$ -invariante. En vez de la anterior, podemos considerar una aplicación $R_{2,2} \otimes S_2 \rightarrow \mathbb{R}$ con las mismas propiedades.

Así, por un resultado clásico de teoría de invariantes [GNS19, Teorema 2.4], tenemos que sólo hay dos aplicaciones lineales $O(1, n)$ -invariantes: $R_{ijij}S_{kk}$ y $R_{jiki}S_{jk}$, que se corresponden a la curvatura escalar $T_2 = \kappa g$ y al tensor de Ricci $T_2 = \text{Ric}$ respectivamente. Por tanto, todo tensor de materia natural e invariante bajo cambios de unidades es una combinación lineal de los anteriores:

$$T_2(g) = a\kappa g + b\text{Ric}, \quad a, b \in \mathbb{R}.$$

Finalmente, para que cumpla todas las hipótesis requeridas, T_2 debe tener divergencia nula. Así:

$$0 = \text{div}(a\kappa g + b\text{Ric}) = aC_1^1(d\kappa \otimes g + \kappa \otimes \nabla g) + b \text{div}(\text{Ric})$$

Como ∇ es de Levi-Civita entonces $\nabla g = 0$. Aplicando el lema 8.5 tenemos que

$$0 = a d\kappa + \frac{b}{2} d\kappa \implies a = -\frac{b}{2}.$$

⁴Nótese que en la proposición está escrito usando δ_i^j ya que está en una base ortonormal.

⁵Es lineal porque acabamos de ver es que los coeficientes T_2 dependen linealmente de los del tensor de curvatura.

Por tanto, sustituyendo en la ecuación de T_2 llegamos a que T_2 es proporcional a $\text{Ric} - \frac{1}{2}\kappa g$.

$$T_2(g) = b(\text{Ric} - \frac{\kappa}{2}g).$$

Para concluir, debemos hallar el valor de la constante. Esto se hace en [NS17, pg.100] tomando el límite de la estructura lorentziana obteniendo una galileana⁶. Una vez hecho eso sustituye en la ecuación el valor del tensor T^2 que para el caso de un sistema de partículas libres es conocido y comparando con la ecuación de Gauss concluye que $b = 8\pi$. Así, llegamos a la ecuación de campo:

$$T_2 = 8\pi(\text{Ric} - \frac{\kappa}{2}g).$$

□

Definición 8.7. Llamamos *tensor de Einstein covariante* G_2 de una métrica no degenerada al siguiente tensor simétrico:

$$G_2 := \text{Ric} - \frac{\kappa}{2}g,$$

donde κ es la curvatura escalar.

Esta forma de la ecuación no coincide con la habitual debido a que como ya comentamos estamos trabajando en unidades naturales para olvidarnos de las constantes. Veamos ahora cómo obtener la ecuación con unidades en el sistema internacional a partir de esta.

En primer lugar debemos analizar las unidades de G_2 , estas deben coincidir con la de la curvatura escalar. Por definición, κ es el producto de las curvaturas principales, es decir, es producto de curvaturas de dos curvas. Por definición de curvatura como el inverso de la circunferencia osculatriz concluimos que $[\kappa] = m^{-2}$. Por otro lado es fácil comprobar que el tensor de Ricci tiene las mismas unidades, pues es la suma de ciertas curvaturas seccionales como vimos en la proposición 4.19.

Por otro lado, ya vimos las magnitudes del tensor de materia en la proposición 7.10. Haciendo un rápido análisis dimensional se comprueba que para las distintas entradas tenemos

$$[T^2(dt, dt)] = \frac{\text{kg}}{\text{m}^3}, \quad [T^2(dt, dx_i)] = \frac{\text{kg}}{\text{m}^2\text{s}}, \quad [T^2(dx_i, dx_j)] = \frac{\text{kg}}{\text{m}\text{s}^2}.$$

Entonces a priori no podemos decir que tenga las mismas unidades en todas las entradas, algo que *a priori* no debería ser. Si observamos con atención, este resultado es consecuencia de la interpretación de la forma de impulso, para la que usamos una base ortonormal. Es decir, los vectores espaciales estaban

⁶No entramos en más detalle en estos dos puntos pues no hemos estudiado ni la ecuación de Gauss ni la teoría detrás de tomar el límite de una estructura lorentziana. Esto se explica en el capítulo 8 del libro. Aunque emplean muchas herramientas que no hemos visto sí podemos comentar que la idea es parecida a la vista en la sección 5.2. Ahí vimos que fijado un evento la variedad de métricas en ese punto tenía un borde definido por estructuras galileanas y al tomar el límite $c \rightarrow \infty$ la métrica lorentziana degeneraba en una galileana.

multiplicados por c . Así, si dividimos por potencias de c según el número de vectores espaciales que empleamos se puede comprobar que tienen las mismas unidades. Sin embargo, de manera equivalente lo que podemos hacer es reescalar los vectores temporales multiplicando por c . Así

$$[T^2(dt, dt)] \frac{\text{m}^2}{\text{s}^2} = \frac{\text{kg}}{\text{ms}^2}, \quad [T^2(dt, dx_i)] \frac{\text{m}}{\text{s}} = \frac{\text{kg}}{\text{ms}^2}, \quad [T^2(dx_i, dx_j)] = \frac{\text{kg}}{\text{ms}^2}.$$

Es decir, sí tiene las mismas unidades en todas las entradas. Concretamente unidades de densidad de energía. De nuevo, vemos la equivalencia entre masa y energía. Por ello, también se le conoce como tensor de energía-momento.

De esta manera, podemos hallar las unidades de la constante

$$\text{m}^{-2} = [\alpha] \frac{\text{kg}}{\text{ms}^2} \implies [\alpha] = \frac{\text{s}^2}{\text{m kg}}.$$

Finalmente, es fácil comprobar que estas son precisamente las unidades de $\frac{G}{c^4}$. Por tanto, podemos escribir la ecuación de campo en unidades del sistema internacional

Teorema 8.8. *La ecuación de campo en Relatividad General con unidades del sistema internacional es*

$$G_2 = \frac{8\pi G}{c^4} T_2$$

Al igualar curvatura, en el miembro izquierdo, con la distribución de materia, en el miembro derecho, esto nos indica que es la energía la que curva el espacio-tiempo. A su vez, en el caso de estar trabajando con una conexión de Levi-Civita, como es el caso, si conocemos curvatura, conexión o métrica, las otras están determinadas. Como además hemos visto que las trayectorias inerciales son aquellas que $\nabla_I I = 0$, entonces la distribución de materia o la curvatura son las que determinan cómo se mueven los objetos en ausencia de fuerzas externas⁷. Por tanto, esa fuerza sólo puede ser la gravedad. Dicho en otras palabras la gravedad no es una fuerza, es la manifestación de la curvatura del espacio-tiempo y las órbitas son geodésicas del espacio-tiempo. Los objetos no se sienten atraídos por una fuerza si no que siguen las trayectorias naturales de la variedad en las que están contenidos. Esta interpretación es idéntica a la ecuación de campo en unidades naturales, pero cuando la escribimos en unidades del sistema internacional podemos ver mejor cuánto cuesta curvar el espacio-tiempo. En la nueva expresión vemos que el tensor T_2 está multiplicado por el factor $\frac{8\pi G}{c^4}$. Es decir, para provocar que el espacio-tiempo se curve de manera significativa la cantidad de energía debe ser muy grande pues G es del orden de 10^{-11} y c^4 de 10^8 . Así que ya no sólo sabemos que es muy costoso curvar el espacio-tiempo sino que la energía también lo curva.

Una vez demostrada y entendida veamos que la ecuación de campo también tiene una forma contravariante equivalente. Para ello, debemos ver un poco más de teoría.

⁷Hay que recordar que una hipótesis sobre T_2 es $\text{div} T_2 = 0$.

Empecemos revisitando las simetrías del tensor de curvatura R . Este lo podemos ver como una aplicación multilineal $R : \mathfrak{X}(M)^4 \rightarrow \mathbb{R}$ tal que es antisimétrica en el primer y segundo par de entradas y simétrica si intercambiamos ambos pares entre sí. Por lo que podemos expresar esta simetrías como $R(D_1, D_2, D_3, D_4) = R_{2,2}(D_1 \wedge D_2, D_3 \wedge D_4)$. Además, si fijamos las dos últimas entradas entonces tenemos una aplicación $\mathfrak{X}^2(X) \rightarrow \Lambda^2 M$, que por lo anterior es alternada. Es decir, lo podemos ver como una 2-forma con valores en $\Lambda^2 M$ y por lo tanto tiene sentido calcular su diferencial exterior que nos servirá más adelante.

Proposición 8.9 (Segunda identidad de Bianchi en forma diferencial). *Sea R el tensor de curvatura entendido como una 2-forma con valores en $\Lambda^2 M$. Entonces podemos escribir la segunda identidad de Bianchi como*

$$dR = 0.$$

Demostración. Aplicando el lema A.19 a R y la segunda identidad de Bianchi, se concluye

$$(dR)(X, Y, Z)(V) = (\nabla_X R)(Y, Z)V + (\nabla_Y R)(Z, X)V + (\nabla_Z R)(X, Y)V = 0.$$

□

Una vez visto esto, construyamos una aplicación que tiene una relación muy estrecha con R . En primer lugar es fácil ver que si contraemos la métrica con un campo $i_D g$ obtenemos una 1-forma, por lo que podemos definir una p -forma alternada con valores en $\Lambda^p M$ tomando el siguiente producto:

$$\Lambda^p g(D_1, \dots, D_p) = i_{D_1} g \wedge \dots \wedge i_{D_p} g.$$

Es decir, estamos viendo g como la 1-forma con valores en T^*X ; el isomorfismo entre el fibrado tangente y cotangente. Así, podemos escribir la aplicación como $\Lambda^p g = \frac{1}{p!} g \wedge \dots \wedge g$, donde dividimos entre $p!$ para normalizar.

Proposición 8.10. $R \wedge (\Lambda g^{n-2})$ es una n -forma cerrada y simétrica.

Demostración. Podemos calcular dI para la 1-forma $I = \sum_i dx_i \otimes \partial_i$ con valores vectoriales, tomando la diferencial en cada coordenada. Por el lema A.11 vemos que $dI = 0$ y como $g = 0$ lo podemos ver como su correspondiente 1-forma con valores en T^*X , también tenemos que $dg = 0$. Por tanto $d(\Lambda^p g) = 0$ para todo p .

Así, aplicando la segunda identidad de Bianchi diferencial 8.9 tenemos

$$d(R \wedge (\Lambda g^{n-2})) = (dR) \wedge (\Lambda g^{n-2}) + R \wedge d(\Lambda g^{n-2}) = 0 + 0 = 0,$$

y es simétrica porque tanto g como el tensor de curvatura R lo son.

□

Teorema 8.11 (Ecuación de campo contravariante). *Sea (X, g, dX) una variedad orientada de dimensión $(1+n)$, equipada con una métrica no degenerada. Entonces*

$$\pm R \wedge (\Lambda g^{n-2}) = C_1^1(dX \otimes G^2).$$

Demostración. Véase [NS17, Teorema F.3.3, pg. 192]. □

Si recordamos la definición 7.9 este resultado nos define la forma de impulso de la relatividad general; la n -forma con valores en $\Lambda^n \mathcal{M}$ que veníamos trabajando $R \wedge (\Lambda g^{n-2})$. Además, por definición de tensor de materia también obtenemos el tensor de materia contravariante G^2 .

Definición 8.12. *Sea (X, g, dX) una variedad orientada de dimensión $(1+n)$, equipada con una métrica no degenerada. Si $n \geq 2$ definimos la *forma de Cartan* C_n con la n -forma $R \wedge (\Lambda g^{n-2})$ con el signo correspondiente tal que*

$$C_n = C_1^1(dX \otimes G^2).$$

De esta forma el resultado anterior es una forma equivalente de enunciar la ecuación de campo pues nos indica cómo el impulso se relaciona con la curvatura y la métrica de la variedad. Además, vemos que el tensor de Einstein contravariante así definido cumple con la hipótesis de divergencia nula.

Corolario 8.13. *El tensor contravariante G^2 tiene divergencia nula.*

Demostración. Repitiendo el razonamiento de [NS17, Prop. F.2.2 pg. 190] tenemos que $dC_3 = dX \otimes \text{div } G^2$. Por la proposición 8.10 sabemos que $dC_3 = 0$, con lo que se concluye $\text{div } G^2 = 0$. □

Observación. Cabe resaltar que ambas formulaciones de la ecuación de campo son equivalentes, aunque la primera tiene una interpretación física más clara.

Definición 8.14. Decimos que una variedad lorentziana no tiene materia si $G^2 = 0$.

A

Apéndice

A.1. Conexión transporte paralelo y geodésicas

Lema A.1. La aplicación ∇ definida en la fórmula de Koszul 3.8, define una conexión.

Demostración. La compatibilidad con la métrica y torsión nula se cumple por construcción. Es inmediato comprobar que se verifica la **aditividad en el primer y segundo argumento**. Comprobamos la linealidad en el primer argumento respecto al **producto por una función** $f \in C^\infty(M)$. Sustituyendo fX en la fórmula de Koszul (3.8), se tiene

$$\begin{aligned} g(\nabla_{fX} Y, Z) &= \frac{1}{2} \left(fX(g(Y, Z)) + fY(g(X, Z)) + Y(f)g(X, Z) \right. \\ &\quad \left. - fZ(g(X, Y)) - Z(f)g(X, Y) - fg(Z, [Y, X]) \right. \\ &\quad \left. - Y(f)g(Z, X) - fg(X, [Y, Z]) - fg(Y, [X, Z]) + Z(f)g(Y, X) \right) \\ &= \frac{1}{2} f \left(X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) - g(Z, [Y, X]) - g(X, [Y, Z]) - g(Y, [X, Z]) \right) \\ &= fg(\nabla_X Y, Z). \end{aligned}$$

Finalmente, veamos que se satisface la **regla de Leibniz**:

$$\begin{aligned} g(\nabla_X(fY), Z) &= \frac{1}{2} \left(X(f)g(Y, Z) + fX(g(Y, Z)) + fY(g(X, Z)) \right. \\ &\quad \left. - Z(f)g(X, Y) - fZ(g(X, Y)) - fg(Z, [Y, X]) \right. \\ &\quad \left. + X(f)g(Z, Y) - fg(Y, [X, Z]) - fg(X, [Y, Z]) + Z(f)g(X, Y) \right) \\ &= \frac{1}{2} \left(2X(f)g(Y, Z) + f \left(X(g(Y, Z)) + Y(g(X, Z)) \right. \right. \\ &\quad \left. \left. - Z(g(X, Y)) - g(Z, [Y, X]) - g(Y, [X, Z]) - g(X, [Y, Z]) \right) \right) \\ &= X(f)g(Y, Z) + fg(\nabla_X Y, Z) = g(X(f)Y + f\nabla_X Y, Z). \end{aligned}$$

□

Proposición A.2. *El transporte paralelo $P_{t_0}^{t_2}(c): T_{c(t_0)}M \rightarrow T_{c(t_2)}M$ satisface*

$$P_{t_0}^{t_2}(c) = P_{t_1}^{t_2}(c_2) \circ P_{t_0}^{t_1}(c_1),$$

$c: [t_0, t_2] \rightarrow M$ curva concatenación diferenciable de las curvas $c_1: [t_0, t_1] \rightarrow M$ y $c_2: [t_1, t_2] \rightarrow M$. Además, el transporte paralelo $P_{t_0}^t(c): T_{c(t_0)}M \rightarrow T_{c(t)}M$ es una isometría lineal.

Demostración. Sea V el único campo paralelo a lo largo de c tal que $V(t_0) = v \in T_pM$, donde a partir de ahora denotamos $V(t)$ a $V(c(t))$ en el caso de campo con soporte en una curva $c(t)$. Sea $V_i, i = 1, 2$ la restricción de V a $c_i, i = 1, 2$. Por la unicidad del campo V , tenemos $P_{t_0}^{t_1}(c_1)(v) = V_1(t_1) = V(t_1)$ y $P_{t_1}^{t_2}(c_2)(V(t_1)) = V_2(t_2) = V(t_2)$, con $V_2(t_1) = V(t_1)$. Dado que v es arbitrario, se concluye que $P_{t_0}^{t_2}(c) = P_{t_1}^{t_2}(c_2) \circ P_{t_0}^{t_1}(c_1)$.

Veamos que es una aplicación biyectiva. Sea $v \in T_{c(t_0)}M$ y sea Y el único campo paralelo a lo largo de c tal que $Y(t_0) = v$.

pues por unicidad de los campos paralelos el campo el que traslada los vectores de $c(t_0)$ a $c(t_1)$ y viceversa es el mismo. Así, geoméricamente la aplicación inversa consiste en recorrer la curva en sentido contrario. A la hora de operar, basta con evaluar el campo en el instante inicial. Así,

$$(P_{t_0}^{t_1}(c))^{-1} = P_{t_1}^{t_0}(c) = Y(t_0).$$

Veamos ahora que es una aplicación lineal. Sean $y, z \in T_{c(t_0)}M$, entonces

$$P_{t_0}^{t_1}(c)(y) = Y(t_1), \quad P_{t_0}^{t_1}(c)(z) = Z(t_1),$$

donde $Y, Z \in \mathfrak{X}M$ son campos paralelos a lo largo de c tales que $Y(t_0) = y, Z(t_0) = z$. Sean $a, b \in \mathbb{R}$, hemos visto entonces que $aY + bZ$ es otro campo paralelo. Como $(aY + bZ)(t_0) = aY(t_0) + bZ(t_0) = ay + bz$ tenemos que

$$P_{t_0}^{t_1}(c)(ay + bz) = (aY + bZ)(t_1) = aY(t_1) + bZ(t_1) = aP_{t_0}^{t_1}(c)(y) + bP_{t_0}^{t_1}(c)(z).$$

Es decir, es lineal. Veamos ahora que preserva el producto escalar. Sean $y, z \in T_{c(t_0)}M$. Sean $Y, Z \in \mathfrak{X}(c)$ los dos campos paralelos a lo largo de c con $Y(t_0) = y$ y $Z(t_0) = z$. Entonces, $P_{t_0}^{t_1}(c)(y) = Y(t_1)$ y $P_{t_0}^{t_1}(c)(z) = Z(t_1)$.

Como que Y y Z son paralelos, hemos probado que $g(Y(t), Z(t))$ es constante a lo largo de c . Por tanto,

$$g(P_{t_0}^{t_1}(c)(y), P_{t_0}^{t_1}(c)(z)) = g(Y(t_1), Z(t_1)) = g(Y(t_0), Z(t_0)) = g(y, z)$$

lo que muestra que $P_{t_0}^{t_1}(c)$ preserva el producto escalar g y, por ser un isomorfismo lineal, es una isometría lineal. □

Lema A.3. *Sea (M, g, ∇) una variedad diferenciable con una métrica no degenerada y una conexión compatible con la métrica. Sean los campos $X, Y, Z \in \mathfrak{X}(M)$ y sea la forma diferenciable $\alpha_Y = g(Y, Z)$, entonces $\nabla_X \alpha_Y = \alpha_{\nabla_X Y}$.*

Demostración. Por definición de la conexión sobre formas,

$$(\nabla_X \alpha_Y)(Z) = X(\alpha_Y(Z)) - \alpha_Y(\nabla_X Z) = X(g(Y, Z)) - g(Y, \nabla_X Z)$$

Por otro lado, teniendo en cuenta la compatibilidad con la métrica, tenemos

$$(\nabla_X \alpha_Y)(Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) - g(Y, \nabla_X Z) = g(\nabla_X Y, Z) = \alpha_{\nabla_X Y}(Z).$$

□

Proposición A.4. *Sea (M, ∇) una variedad diferenciable equipada con una conexión. Entonces la derivada tensorial dada en la definición 3.29 verifica las propiedades 1-4 dadas en la definición 3.1.*

Demostración. Sea $T \in T_q^p(M)$, la condición de linealidad en el primer argumento es inmediata teniendo en cuenta que la conexión es lineal pues:

$$\begin{aligned} & (\nabla_{fX+gY} T)(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) \\ &= fX(T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)) + gY(T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)) \\ &\quad - \sum_{i=1}^r T(\omega_1, \dots, \nabla_{X+Y} \omega_i, \dots, \omega_r, Y_1, \dots, Y_s) - \sum_{j=1}^s T(\omega_1, \dots, \omega_r, Y_1, \dots, \nabla_{fX+gY} Y_j, \dots, Y_s) \\ &= fX(T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)) + gY(T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)) \\ &\quad - \sum_{i=1}^r T(\omega_1, \dots, f\nabla_X \omega_i, \dots, \omega_r, Y_1, \dots, Y_s) - \sum_{i=1}^r T(\omega_1, \dots, g\nabla_Y \omega_i, \dots, \omega_r, Y_1, \dots, Y_s) \\ &\quad - \sum_{j=1}^s T(\omega_1, \dots, \omega_r, Y_1, \dots, g\nabla_Y Y_j, \dots, Y_s) - \sum_{j=1}^s T(\omega_1, \dots, \omega_r, Y_1, \dots, f\nabla_X Y_j, \dots, Y_s) \\ &= (f\nabla_X T)(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) + g(\nabla_Y T)(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) \end{aligned}$$

con lo que se concluye $\nabla_{fX+gY} T = f\nabla_X T + g\nabla_Y T$, para todo $f, g \in C^\infty(M)$, $X, Y \in \mathfrak{X}(M)$.

Comprobamos ahora la linealidad en el segundo argumento. Se tiene:

$$\begin{aligned} & (\nabla_X (fT_1 + gT_2))(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) \\ &= X((fT_1 + gT_2)(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)) - \sum_{i=1}^r (fT_1 + gT_2)(\omega_1, \dots, \nabla_X \omega_i, \dots, \omega_r, Y_1, \dots, Y_s) \\ &\quad - \sum_{j=1}^s (fT_1 + gT_2)(\omega_1, \dots, \omega_r, Y_1, \dots, \nabla_X Y_j, \dots, Y_s) \\ &= f\nabla_X T_1(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) + g\nabla_X T_2(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s), \end{aligned}$$

con lo que se concluye $(\nabla_X(T_1 + T_2)) = \nabla_X T_1 + \nabla_X T_2$, para todo $T_1, T_2 \in T_q^p(M)$ y $f, g \in C^\infty(M)$.

Por último, veamos que se satisface la regla de Leibniz. Teniendo en cuenta la definición y la regla del producto, tenemos que

$$\begin{aligned}
& (\nabla_X(f T))(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) \\
&= X\left(f T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)\right) - \sum_{i=1}^r (f T)(\omega_1, \dots, \nabla_X \omega_i, \dots, \omega_r, Y_1, \dots, Y_s) \\
&\quad - \sum_{j=1}^s (f T)(\omega_1, \dots, \omega_r, Y_1, \dots, \nabla_X Y_j, \dots, Y_s) \\
&= X(f) T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) + f X\left(T(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s)\right) \\
&\quad - \sum_{i=1}^r (f T)(\omega_1, \dots, \nabla_X \omega_i, \dots, \omega_r, Y_1, \dots, Y_s) \\
&\quad - \sum_{j=1}^s (f T)(\omega_1, \dots, \omega_r, Y_1, \dots, \nabla_X Y_j, \dots, Y_s).
\end{aligned}$$

Sacando factor común f y por definición de derivada tensorial se concluye $\nabla_X(f T) = X(f) T + f \nabla_X T$, para todo $f \in C^\infty(M)$ y $T \in T_q^p(M)$. \square

Lema A.5. *La derivada tensorial definida por la regla del producto es equivalente a la definida por componentes.*

Demostración. Sea $T \in T_q^p(M)$ cuya expresión en coordenadas es:

$$T = \omega_1 \otimes \dots \otimes \omega_q \otimes X_1 \otimes \dots \otimes X_p.$$

Por la definición de derivada covariante, tenemos que

$$\begin{aligned}
(\nabla_X T)(Y_1, \dots, Y_q, \alpha_1, \dots, \alpha_p) &= X\left(T(Y_1, \dots, Y_q, \alpha_1, \dots, \alpha_p)\right) \\
&\quad - \sum_{i=1}^q T(Y_1, \dots, \nabla_X Y_i, \dots, Y_q, \alpha_1, \dots, \alpha_p) \\
&\quad - \sum_{j=1}^p T(Y_1, \dots, Y_q, \alpha_1, \dots, \nabla_X \alpha_j, \dots, \alpha_p).
\end{aligned}$$

Teniendo en cuenta la regla del producto, el primer sumando del miembro de la derecha de la igualdad

anterior se escribe como sigue:

$$X(\omega_1(Y_1) \cdots \omega_q(Y_q) \alpha_1(X_1) \cdots \alpha_p(X_p)) = \sum_{i=1}^q \left[X(\omega_i(Y_i)) \prod_{m \neq i} \omega_m(Y_m) \prod_{j=1}^p \alpha_j(X_j) \right] + \sum_{j=1}^p \left[X(\alpha_j(X_j)) \prod_{i=1}^q \omega_i(Y_i) \prod_{n \neq j} \alpha_n(X_n) \right],$$

y, por la regla de derivación para formas, tenemos que:

$$X(T(Y_1, \dots, Y_q, \alpha_1, \dots, \alpha_p)) = \sum_{i=1}^q \left[((\nabla_X \omega_i)(Y_i) + \omega_i(\nabla_X Y_i)) \prod_{m \neq i} \omega_m(Y_m) \prod_{j=1}^p \alpha_j(X_j) \right] + \sum_{j=1}^p \left[((\nabla_X \alpha_j)(X_j) + \alpha_j(\nabla_X X_j)) \prod_{i=1}^q \omega_i(Y_i) \prod_{n \neq j} \alpha_n(X_n) \right].$$

Por otro lado, tenemos

$$-\sum_{i=1}^q T(Y_1, \dots, \nabla_X Y_i, \dots, Y_q, \alpha_1, \dots, \alpha_p) = -\sum_{i=1}^q \omega_i(\nabla_X Y_i) \prod_{m \neq i} \omega_m(Y_m) \prod_{j=1}^p \alpha_j(X_j),$$

$$-\sum_{j=1}^p T(Y_1, \dots, Y_q, \alpha_1, \dots, \nabla_X \alpha_j, \dots, \alpha_p) = -\sum_{j=1}^p (\nabla_X \alpha_j)(X_j) \prod_{i=1}^q \omega_i(Y_i) \prod_{n \neq j} \alpha_n(X_n).$$

Finalmente, sumando las tres expresiones anteriores, y teniendo en cuenta que los términos con $\omega_i(\nabla_X Y_i)$ y $(\nabla_X \alpha_j)(X_j)$ se cancelan, se obtiene:

$$(\nabla_X T)(Y_1, \dots, Y_q, \alpha_1, \dots, \alpha_p) = \sum_{i=1}^q (\nabla_D \omega_i)(Y_i) \prod_{m \neq i} \omega_m(Y_m) \prod_{j=1}^p \alpha_j(X_j) + \sum_{j=1}^p \alpha_j(\nabla_X X_j) \prod_{i=1}^q \omega_i(Y_i) \prod_{n \neq j} \alpha_n(X_n).$$

Ahora, es fácil ver que el lado derecho de esta nueva expresión no es más que el siguiente tensor evaluado sobre los mismos campos y formas que el primero:

$$\nabla_X T = \sum_{i=1}^q \omega_1 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_q \otimes X_1 \otimes \cdots \otimes X_p + \sum_{j=1}^p \omega_1 \otimes \cdots \otimes \omega_q \otimes X_1 \otimes \cdots \otimes \nabla_X X_j \otimes \cdots \otimes X_p.$$

□

Lema A.6. *La única extensión de ∇ para campos tensoriales verifica la regla de Leibniz.*

Demostración. Comprobemos que se verifica la regla de Leibniz, para tensores de cualquier orden, sabiendo que se verifica para tensores del tipo $(0, 0)$, $(1, 0)$ y $(0, 1)$. Es fácil ver que se verifica para tensores $(1, 1)$. Sea el campo $X \in \mathfrak{X}(M)$ y el tensor $\omega \otimes Y \in T_1^1(M)$, entonces:

$$\begin{aligned} C(\nabla_X(\omega \otimes Y)) &= \nabla_X(C(\omega \otimes Y)) = \nabla_X(\omega(Y)) \\ &= \nabla_X(\omega)(Y) + \omega(\nabla_X Y) = C(\nabla_X \omega \otimes Y + \omega \otimes \nabla_X Y). \end{aligned}$$

Por tanto, ya sabemos que se verifica esta regla para tensores de tipo (q, p) con $q + p \leq 2$.

Veamos ahora qué sucede con tensores que son del tipo $(0, p)$. Sea $T = \omega_1 \otimes \cdots \otimes \omega_p \in T_0^p(M)$ y sean $X, Y_1, \dots, Y_p \in \mathfrak{X}(M)$. Se verifica que:

$$\begin{aligned} X((\omega_1 \otimes \cdots \otimes \omega_p)(Y_1, \dots, Y_p)) &= X\left(\prod_{i=1}^p \omega_i(Y_i)\right) = \sum_{i=1}^p \left(X(\omega_i(Y_i)) \prod_{\substack{j=1 \\ j \neq i}}^p \omega_j(Y_j) \right) \\ &= \sum_{i=1}^p \left((\nabla_X \omega_i)(Y_i) \prod_{\substack{j=1 \\ j \neq i}}^p \omega_j(Y_j) \right) + \sum_{i=1}^p \left(\omega_i(\nabla_X Y_i) \prod_{\substack{j=1 \\ j \neq i}}^p \omega_j(Y_j) \right) \\ &= \sum_{i=1}^p ((\omega_1 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_p)(Y_1, \dots, Y_p)) \\ &\quad + \sum_{i=1}^p ((\omega_1 \otimes \cdots \otimes \omega_p)(Y_1, \dots, \nabla_X Y_i, \dots, Y_p)). \end{aligned}$$

Si despejamos el primer sumatorio del miembro derecho, llegamos a la versión por componentes de la regla de Leibniz; por lo tanto, se cumple. Análogamente, se obtiene el resultado para tensores $(q, 0)$.

Razonamos, ahora, por inducción sobre el grado del tensor. Consideremos que la regla de Leibniz se verifica hasta tensores de orden (q, p) tales que $q + p = k$ y $q \neq 0 \neq p$.¹

Sean $X, Y \in \mathfrak{X}(M)$ y $T \in T_p^q(M)$ con $q + p = k$ y $q \neq 0 \neq p$. Consideremos el tensor $T \otimes Y$ y la aplicación C que contrae ω_1 con el campo Y .² Como ∇ conmuta con la contracción, tenemos $C(\nabla_X(T \otimes Y)) = \nabla_X(C(T \otimes Y))$.

¹Nótese que la condición $q \neq 0 \neq p$ se debe a que ya lo hemos demostrado para todos los tensores con $q = 0$ o $p = 0$.

²A lo largo del desarrollo C se aplicará sobre tensores de distinto orden, pero siempre denotará la evaluación indicada. No se distinguen las distintas contracciones para facilitar la comprensión de la prueba y porque, por hipótesis, la conmutación se da para toda contracción.

Por otro lado, teniendo en cuenta la expresión de T , tenemos:

$$\begin{aligned} C(T \otimes Y) &= C(Z_1 \otimes \cdots \otimes Z_q \otimes \omega_1 \otimes \cdots \otimes \omega_p \otimes Y), \\ &= f \cdot Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p = T' \in T_{p-1}^q(M), \end{aligned}$$

con $f \in C^\infty(M)$.

Ahora, por hipótesis de inducción, se verifica la regla de Leibniz:

$$\begin{aligned} \nabla_X T' &= X(f) \cdot Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p \\ &\quad + \sum_{i=1}^q f \cdot Z_1 \otimes \cdots \otimes \nabla_X Z_i \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p \\ &\quad + \sum_{i=2}^p f \cdot Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_p. \end{aligned}$$

También tenemos que:

$$\begin{aligned} \nabla_X T' &= \nabla_X(C(\omega_1 \otimes Y)) \otimes Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p \\ &\quad + \sum_{i=1}^q Z_1 \otimes \cdots \otimes \nabla_X Z_i \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p \\ &\quad + \sum_{i=2}^p Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_p. \end{aligned}$$

Por conmutación con contracciones y la hipótesis de inducción, podemos aplicar la regla de Leibniz:

$$\begin{aligned} \nabla_X T' &= C(\nabla_X(\omega_1) \otimes Y + \omega_1 \otimes \nabla_X Y) \otimes Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p \\ &\quad + \sum_{i=1}^q C(\omega_1 \otimes Y) \cdot Z_1 \otimes \cdots \otimes \nabla_X Z_i \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \omega_p \\ &\quad + \sum_{i=2}^p C(\omega_1 \otimes Y) \cdot Z_1 \otimes \cdots \otimes Z_q \otimes \omega_2 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_p \\ &= C\left(\sum_{i=1}^q Y \otimes Z_1 \otimes \cdots \otimes \nabla_X Z_i \otimes \cdots \otimes Z_q \otimes \omega_1 \otimes \cdots \otimes \omega_p \right. \\ &\quad \left. + \nabla_X Y \otimes Z_1 \otimes \cdots \otimes Z_q \otimes \omega_1 \otimes \cdots \otimes \omega_p \right. \\ &\quad \left. + \sum_{i=2}^p Y \otimes Z_1 \otimes \cdots \otimes Z_q \otimes \omega_1 \otimes \cdots \otimes \nabla_X \omega_i \otimes \cdots \otimes \omega_p\right). \end{aligned}$$

□

Lema A.7. Sea $\varphi: (M, g) \rightarrow (N, h)$ una isometría. Entonces se tiene

$$X(g(Y, Z)) = d\varphi(X)\left(h(d\varphi Y, d\varphi Z)\right), \quad \forall X, Y, Z \in \mathfrak{X}(M).$$

Demostración. Dado que φ es una isometría, se cumple que $g(Y, Z) = h(d\varphi Y, d\varphi Z)$. Sea la función $f: N \rightarrow \mathbb{R}$ definida como sigue

$$f(q) = h_q\left((d\varphi Y)(q), (d\varphi Z)(q)\right), \quad \forall q \in N$$

De esta forma, como φ es una isometría, tenemos que

$$(f \circ \varphi)(p) = f(\varphi(p)) = h_{\varphi(p)}\left((d\varphi Y)(\varphi(p)), (d\varphi Z)(\varphi(p))\right) = g_p(Y(p), Z(p)), \quad \forall p \in M.$$

Finalmente, si aplicamos la regla de la cadena a la composición $f \circ \varphi$, obtenemos

$$X(f \circ \varphi)(p) = (df)_{\varphi(p)}(d\varphi_p(X(p))) = d\varphi(X)(f).$$

Por lo tanto, dado que $f \circ \varphi = g(Y, Z)$, se concluye que $X(g(Y, Z)) = d\varphi(X)(h(d\varphi Y, d\varphi Z))$. \square

Lema A.8. Sea $p \in M$ y \exp_p la aplicación exponencial en p . Entonces en coordenadas normales, los símbolos de Christoffel en p son nulos.

Demostración. Sea $V \in T_p M$ un vector unitario y tangente en p . Consideremos la geodésica radial $c_V^p(t)$. En coordenadas normales tenemos que:

$$x^i(c_V^p(t)) = tv^i,$$

pues $\exp_p(tV) = c_V^p(t)$. De esta forma, derivando respecto al parámetro de longitud de arco t :

$$\frac{dx^i}{dt} = v^i, \quad \frac{d^2x^i}{dt^2} = 0.$$

Por otro lado, la ecuación de una geodésica es (Proposición 3.40):

$$\frac{d^2x^k}{dt^2} + \sum_{i,j=1}^n \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} = 0.$$

Sustituyendo para el caso de la geodésica $c_V^p(t)$, tenemos que: $\sum_{i,j=1}^n \Gamma_{ij}^k(x(t))v^i v^j = 0$. Evaluando en $t = 0$, y teniendo en cuenta que estamos considerando coordenadas normales se tiene: $x^i(0) = 0$ y por tanto:

$$\sum_{i,j=1}^n \Gamma_{ij}^k(0)v^i v^j = 0.$$

Esta ecuación debe cumplirse para cualquier vector unitario v . Esto implica que todos los coeficientes $\Gamma_{ij}^k(0)$ son nulos. Veámoslo explícitamente. Tomemos $v^i = \delta_a^i$, entonces: $\Gamma_{aa}^k(0) = 0$.

Por otro lado, sean $\Gamma_{ab}^k(0)$ con $a \neq b$ y consideremos el caso $v^a = v^b = \frac{1}{\sqrt{2}}$, $v^c = 0$ para $c \neq a, b$. Si ahora sustituimos en la ecuación y multiplicamos por 2:

$$\sum_{i,j=1}^n \Gamma_{ij}^k(0) v^i v^j = \Gamma_{aa}^k(0) + \Gamma_{ab}^k(0) + \Gamma_{ba}^k(0) + \Gamma_{bb}^k(0) = 0.$$

Como $\Gamma_{ab}^k(0) = \Gamma_{ba}^k(0)$ por la simetría de los símbolos de Christoffel, llegamos a:

$$\Gamma_{aa}^k(0) + 2\Gamma_{ab}^k(0) + \Gamma_{bb}^k(0) = 0,$$

con lo que se concluye, pues $\Gamma_{aa}^k(0) = 0$ y $\Gamma_{bb}^k(0) = 0$. \square

Lema A.9. *En un sistema de coordenadas normales, las primeras derivadas de g en el punto son nulas.*

Demostración. Por el lema A.8 y la proposición 3.II tenemos el siguiente sistema

$$\begin{cases} \partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij} = 0 \\ \partial_j g_{\ell i} + \partial_\ell g_{ji} - \partial_i g_{\ell j} = 0 \\ \partial_\ell g_{ij} + \partial_i g_{\ell j} - \partial_j g_{i\ell} = 0 \end{cases}$$

Sumando las tres ecuaciones anteriores y la simetría de la métrica y de sus derivadas, tenemos

$$\begin{aligned} 0 &= (\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}) + (\partial_j g_{\ell i} + \partial_\ell g_{ji} - \partial_i g_{\ell j}) + (\partial_\ell g_{ij} + \partial_i g_{\ell j} - \partial_j g_{i\ell}) \\ &= \partial_i g_{j\ell} + \partial_j g_{i\ell} + \partial_\ell g_{ij}. \end{aligned}$$

Restando la primera ecuación del sistema a la anterior, tenemos que $\partial_\ell g_{ij} = 0$. Finalmente, al sustituir en el sistema original concluimos: $\partial_i g_{j\ell} = \partial_j g_{i\ell} = \partial_\ell g_{ij} = 0$. \square

A.2. Apéndice: Tensor de Curvatura

Proposición A.10. *Sean X, Y campos vectoriales suaves sobre una variedad diferenciable. Sea ϕ_t^X el flujo de X y ϕ_t^Y el flujo de Y . Entonces,*

$$\phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_{\varepsilon}^Y \circ \phi_{\varepsilon}^X = \phi^{-\varepsilon^2[X,Y]} + O(\varepsilon^3),$$

donde $[X, Y]$ es el corchete de Lie de X y Y .

Demostración. Recordemos que para un grupo de Lie G y su álgebra de Lie \mathfrak{g} , la aplicación exponencial $\exp : \mathfrak{g} \rightarrow G$ asocia a cada elemento $X \in \mathfrak{g}$ el flujo ϕ_t^X evaluado en $t = 1$; es decir, $e^X = \phi_1^X$. Por tanto, podemos expresar los flujos que aparecen en la proposición como

$$\phi_\varepsilon^X = e^{\varepsilon X}, \quad \phi_\varepsilon^Y = e^{\varepsilon Y}$$

. De esta forma podemos reescribir la composición de flujos anterior como

$$\phi_{-\varepsilon}^Y \circ \phi_{-\varepsilon}^X \circ \phi_\varepsilon^Y \circ \phi_\varepsilon^X = e^{-\varepsilon Y} e^{-\varepsilon X} e^{\varepsilon Y} e^{\varepsilon X} = \left(e^{-\varepsilon Y} e^{-\varepsilon X} \right) \left(e^{\varepsilon Y} e^{\varepsilon X} \right),$$

, pues el producto en el grupo es asociativo. Ahora podemos aplicar la fórmula de Baker–Campbell–Hausdorff. Así, sabemos que para $X, Y \in \mathfrak{g}$ existen elementos $Z_1, Z_2 \in \mathfrak{g}$ tales que

$$\begin{aligned} e^{\varepsilon Y} e^{\varepsilon X} &= e^{Z_1}, \quad \text{con } Z_1 = \varepsilon(X + Y) + \frac{\varepsilon^2}{2}[X, Y] + O(\varepsilon^3), \\ e^{-\varepsilon Y} e^{-\varepsilon X} &= e^{Z_2}, \quad \text{con } Z_2 = -\varepsilon(X + Y) + \frac{\varepsilon^2}{2}[X, Y] + O(\varepsilon^3). \end{aligned}$$

Por tanto, $e^{-\varepsilon Y} e^{-\varepsilon X} e^{\varepsilon Y} e^{\varepsilon X} = e^{Z_1+Z_2}$. Teniendo en cuenta

$$Z_2 + Z_1 = \left(-\varepsilon(X + Y) + \frac{\varepsilon^2}{2}[X, Y] \right) + \left(\varepsilon(X + Y) + \frac{\varepsilon^2}{2}[X, Y] \right) + O(\varepsilon^3) = \varepsilon^2[X, Y] + O(\varepsilon^3),$$

y por otro lado

$$[Z_1, Z_2] = [-\varepsilon(X + Y) + O(\varepsilon^2), \varepsilon(X + Y) + O(\varepsilon^2)] = O(\varepsilon^3),$$

se concluye que $e^{-\varepsilon Y} e^{-\varepsilon X} e^{\varepsilon Y} e^{\varepsilon X} = e^{\varepsilon^2[X, Y]} + O(\varepsilon^3)$. \square

Lema A.II. La diferencial exterior de una 1-forma $\omega \in \Lambda^1(M)$ satisface:

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]), \quad \forall X, Y \in \mathfrak{X}(M).$$

Demostración. En coordenadas, la 1-forma ω se escribecomo $\omega = \omega_i dx^i$, con $\omega_i \in C^\infty(M)$. Calculamos su diferencial exterior:

$$d\omega = \sum_i d\omega_i \wedge dx^i = \sum_{i,j} \frac{\partial \omega_i}{\partial x^j} dx^j \wedge dx^i.$$

Evaluamos $d\omega$ en los campos vectoriales $X = X^i \frac{\partial}{\partial x^i}$ y $Y = Y^j \frac{\partial}{\partial x^j}$, y sumando y restando el término

$$\sum_{i,j} \left(\omega_i X^j \frac{\partial Y^i}{\partial x^j} - \omega_i Y^j \frac{\partial X^i}{\partial x^j} \right)$$

obtenemos

$$\begin{aligned} d\omega(X, Y) &= \sum_{i,j} \frac{\partial \omega_i}{\partial x^j} X^j Y^i - \sum_{i,j} \frac{\partial \omega_i}{\partial x^j} Y^j X^i \\ &= \sum_{i,j} X^j \left(\frac{\partial \omega_i}{\partial x^j} Y^i + \omega_i \frac{\partial Y^i}{\partial x^j} \right) - \sum_{i,j} Y^j \left(\frac{\partial \omega_i}{\partial x^j} X^i + \omega_i \frac{\partial X^i}{\partial x^j} \right) - \sum_{i,j} \omega_i \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right). \end{aligned}$$

Ahora, analicemos cada sumando del miembro de la derecha de la igualdad anterior. El primero sumando se rescribe como sigue:

$$\sum_{i,j} X^j \left(\frac{\partial \omega_i}{\partial x^j} Y^i + \omega_i \frac{\partial Y^i}{\partial x^j} \right) = X^j \frac{\partial}{\partial x^j} (\omega_i Y^i) = X(\omega(Y)).$$

y, análogamente, el segundo se escribe como

$$\sum_{i,j} Y^j \left(\frac{\partial \omega_i}{\partial x^j} X^i + \omega_i \frac{\partial X^i}{\partial x^j} \right) = Y(\omega(X)).$$

Recordando la expresión del corchete de Lie en coordenadas:

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

el tercer sumando se escribe:

$$\sum_{i,j} \omega_i \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right) = \sum_{i,j} \omega_i \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right) = \sum_i \omega_i ([X, Y]^i) = \omega([X, Y]).$$

Finalmente, sustituyéndolo todo en la expresión original, se concluye. \square

Proposición A.12. Sean (M, ∇) una variedad diferenciable de dimensión n equipada con una conexión, y sea Θ el tensor de torsión de ∇ . Sea (e_i) una base del espacio tangente y (θ^i) su base dual. Entonces se verifican las siguientes ecuaciones:

$$\begin{aligned} d\theta^i &= - \sum_{j=1}^n \omega_j^i \wedge \theta^j + \Theta^i \\ \sum_{j=1}^n \left(d\omega_j^i + \sum_{m=1}^n \omega_m^i \wedge \omega_j^m \right) \wedge \theta^j &= d\Theta^i + \sum_{j=1}^n \omega_j^i \wedge \Theta^j. \end{aligned}$$

Demostración. Empecemos probando la primera ecuación de estructura calculando la diferencial exterior $d\theta^i$ mediante el lema anterior:

$$d\theta^i(e_k, e_l) = e_k(\theta^i(e_l)) - e_l(\theta^i(e_k)) - \theta^i([e_k, e_l]).$$

Es fácil ver que $e_k(\theta^i(e_l)) = e_k(\delta_l^i) = 0$ y $e_l(\theta^i(e_k)) = e_l(\delta_k^i) = 0$, por tanto, sustituyendo en la expresión anterior se tiene que

$$d\theta^i(e_k, e_l) = -\theta^i([e_k, e_l]).$$

Ahora, por definición de Θ , podemos expresar el corchete de Lie en términos de la conexión y de la torsión:

$$[e_k, e_l] = \nabla_{e_k} e_l - \nabla_{e_l} e_k - \Theta(e_k, e_l).$$

Por definición de las formas de conexión, tenemos que $\nabla_{e_k} e_l - \nabla_{e_l} e_k = \sum_{m=1}^n (\omega_l^m(e_k) - \omega_k^m(e_l)) e_m$. De esta forma:

$$[e_k, e_l] = \sum_{m=1}^n (\omega_l^m(e_k) - \omega_k^m(e_l)) e_m - \Theta(e_k, e_l).$$

Aplicando θ^i :

$$\theta^i([e_k, e_l]) = \sum_{m=1}^n (\omega_l^m(e_k) - \omega_k^m(e_l)) \delta_m^i - \theta^i(\Theta(e_k, e_l)) = \omega_l^i(e_k) - \omega_k^i(e_l) - \Theta^i(e_k, e_l),$$

donde $\Theta^i(e_k, e_l) := \theta^i(\Theta(e_k, e_l))$. Así, sustituyendo en $d\theta^i(e_k, e_l) = -\theta^i([e_k, e_l])$ llegamos a:

$$d\theta^i(e_k, e_l) = -\theta^i([e_k, e_l]) = -(\omega_l^i(e_k) - \omega_k^i(e_l) - \Theta^i(e_k, e_l)) = -\omega_l^i(e_k) + \omega_k^i(e_l) + \Theta^i(e_k, e_l).$$

Ahora, intentemos reescribir el miembro derecho de la expresión anterior:

$$\begin{aligned} d\theta^i(e_k, e_l) &= -\omega_l^i(e_k) + \omega_k^i(e_l) + \Theta^i(e_k, e_l) = -\sum_{j=1}^n (\omega_j^i(e_k) \delta_l^j - \omega_j^i(e_l) \delta_k^j) + \Theta^i(e_k, e_l) \\ &= -\sum_{j=1}^n (\omega_j^i(e_k) \theta^j(e_l) - \omega_j^i(e_l) \theta^j(e_k)) + \Theta^i(e_k, e_l) \\ &= \left(-\sum_{j=1}^n \omega_j^i \wedge \theta^j \right) (e_k, e_l) + \Theta^i(e_k, e_l). \end{aligned}$$

Finalmente, como Θ^i es una 2-forma y lo anterior se verifica para todo par de campos podemos expresarlo como:

$$d\theta^i = -\sum_{j=1}^n \omega_j^i \wedge \theta^j + \Theta^i.$$

Demostremos ahora la segunda ecuación partiendo de la primera. Para ello, comenzamos aplicando la derivada exterior d a ambos lados:

$$d(d\theta^i) + d\left(\sum_{j=1}^n \omega_j^i \wedge \theta^j\right) = d\Theta^i.$$

Por un lado, es fácil ver que $d(d\theta^i) = 0$. Por otro lado, calculemos $d\left(\sum_{j=1}^n \omega_j^i \wedge \theta^j\right)$ usando la regla del producto:

$$d\left(\sum_{j=1}^n \omega_j^i \wedge \theta^j\right) = \sum_{j=1}^n d(\omega_j^i \wedge \theta^j) = \sum_{j=1}^n (d\omega_j^i \wedge \theta^j - \omega_j^i \wedge d\theta^j).$$

Sustituimos $d\theta^j$ aplicando la primera ecuación de estructura :

$$\omega_j^i \wedge d\theta^j = \omega_j^i \wedge \left(-\sum_{k=1}^n \omega_k^j \wedge \theta^k + \Theta^j\right) = -\sum_{k=1}^n \omega_j^i \wedge \omega_k^j \wedge \theta^k + \omega_j^i \wedge \Theta^j.$$

Sustituimos:

$$d\left(\sum_{j=1}^n \omega_j^i \wedge \theta^j\right) = \sum_{j=1}^n d\omega_j^i \wedge \theta^j + \sum_{j=1}^n \sum_{k=1}^n \omega_j^i \wedge \omega_k^j \wedge \theta^k - \sum_{j=1}^n \omega_j^i \wedge \Theta^j.$$

Sustituyendo en la expresión original tenemos que:

$$\sum_{j=1}^n d\omega_j^i \wedge \theta^j + \sum_{j=1}^n \sum_{k=1}^n \omega_j^i \wedge \omega_k^j \wedge \theta^k - \sum_{j=1}^n \omega_j^i \wedge \Theta^j = d\Theta^i.$$

Ahora, intentemos escribirla de la forma más sencilla posible. Así, intentemos agrupar los dos primeros sumatorios del miembro izquierdo en uno solo. Renombrando los índices tenemos:

$$\sum_{m=1}^n \sum_{j=1}^n \omega_m^i \wedge \omega_j^m \wedge \theta^j.$$

De esta forma vemos que:

$$\sum_{j=1}^n d\omega_j^i \wedge \theta^j + \sum_{m=1}^n \sum_{j=1}^n \omega_m^i \wedge \omega_j^m \wedge \theta^j = \sum_{j=1}^n \left(d\omega_j^i + \sum_{m=1}^n \omega_m^i \wedge \omega_j^m\right) \wedge \theta^j.$$

Ahora, escribiendo todos los términos con el tensor de torsión en el lado derecho, concluimos

$$\sum_{j=1}^n \left(d\omega_j^i + \sum_{m=1}^n \omega_m^i \wedge \omega_j^m\right) \wedge \theta^j = d\Theta^i + \sum_{j=1}^n \omega_j^i \wedge \Theta^j.$$

□

Proposición A.13. Los coeficientes R_{klj}^i se escriben como sigue

$$R_{klj}^i = \theta^i (\nabla_{e_k} \nabla_{e_l} e_j - \nabla_{e_l} \nabla_{e_k} e_j - \nabla_{[e_k, e_l]} e_j),$$

donde (e_i) es una base de $T_p M$.

Demostración. Aplicando el lema A.II calculamos $d\omega_j^i(e_k, e_l)$:

$$d\omega_j^i(e_k, e_l) = e_k(\omega_j^i(e_l)) - e_l(\omega_j^i(e_k)) - \omega_j^i([e_k, e_l]).$$

Por otro lado, los productos exteriores quedan:

$$\sum_{m=1}^n (\omega_m^i \wedge \omega_j^m)(e_k, e_l) = \sum_{m=1}^n (\omega_m^i(e_k)\omega_j^m(e_l) - \omega_m^i(e_l)\omega_j^m(e_k)).$$

Sustituyendo ambas expresiones en

$$R_{klj}^i = \Omega_j^i(e_k, e_l) = d\omega_j^i(e_k, e_l) + \sum_{m=1}^n (\omega_m^i \wedge \omega_j^m)(e_k, e_l).$$

y reagrupando términos, tenemos:

$$R_{klj}^i = e_k(\omega_j^i(e_l)) + \sum_{m=1}^n \omega_m^i(e_k)\omega_j^m(e_l) - \left(e_l(\omega_j^i(e_k)) + \sum_{m=1}^n \omega_m^i(e_l)\omega_j^m(e_k) \right) - \omega_j^i([e_k, e_l]). \quad (\text{A.I})$$

Expresamos los primeros dos términos del miembro derecho en (A.I), de una forma más sencilla como sigue

$$\begin{aligned} e_k(\omega_j^i(e_l)) + \sum_{m=1}^n \omega_m^i(e_k)\omega_j^m(e_l) &= \theta^i \left(\sum_{s=1}^n \left(e_k(\omega_j^s(e_l))e_s + \sum_{m=1}^n \omega_j^m(e_l)\omega_m^s(e_k)e_s \right) \right) \\ &= \theta^i \left(\sum_{s=1}^n e_k(\omega_j^s(e_l))e_s + \sum_{s=1}^n \sum_{m=1}^n \omega_j^m(e_l)\omega_m^s(e_k)e_s \right) \\ &= \theta^i \left(\sum_{s=1}^n e_k(\omega_j^s(e_l))e_s + \sum_{m=1}^n \omega_j^m(e_l) \sum_{s=1}^n \omega_m^s(e_k)e_s \right) \\ &= \theta^i \left(\sum_{s=1}^n e_k(\omega_j^s(e_l))e_s + \sum_{m=1}^n \omega_j^m(e_l) \nabla_{e_k} e_m \right) \\ &= \theta^i \left(\nabla_{e_k} \left(\sum_{m=1}^n \omega_j^m(e_l)e_m \right) \right) \\ &= \theta^i (\nabla_{e_k} \nabla_{e_l} e_j). \end{aligned}$$

De manera análoga, tenemos: $\theta^i (\nabla_{e_l} \nabla_{e_k} e_j) = e_l(\omega_j^i(e_k)) + \sum_{m=1}^n \omega_j^m(e_k)\omega_m^i(e_l)$. En cuanto al último término en (A.I), vemos que

$$\omega_j^i([e_k, e_l]) = \theta^i \left(\sum_{s=1}^n \omega_j^s([e_k, e_l])e_s \right) = \theta^i (\nabla_{[e_k, e_l]} e_j).$$

Finalmente, sustituyendo las expresiones anteriores en (A.1), tenemos:

$$\begin{aligned} R_{klj}^i &= e_k(\omega_j^i(e_l)) + \sum_{m=1}^n \omega_m^i(e_k)\omega_j^m(e_l) - \left(e_l(\omega_j^i(e_k)) + \sum_{m=1}^n \omega_m^i(e_l)\omega_j^m(e_k) \right) - \omega_j^i([e_k, e_l]) \\ &= \theta^i (\nabla_{e_k} \nabla_{e_l} e_j - \nabla_{e_l} \nabla_{e_k} e_j - \nabla_{[e_k, e_l]} e_j). \end{aligned}$$

□

Lema A.14. Sea H una matriz simétrica $n \times n$, entonces $\det(I + H) = 1 + \text{tr}(H) + O(\|H\|^2)$.

Demostración. Es fácil ver que los autovalores de $I + H$ son $1 + \lambda_i$ con λ_i autovalor de H , luego

$$\det(I + H) = \prod_{i=1}^n (1 + \lambda_i) = 1 + \sum_{i=1}^n \lambda_i + \sum_{1 \leq i < j \leq n} \lambda_i \lambda_j + \dots$$

Por tanto, como H es simétrica, su norma coincide con el máximo de los autovalores y concluimos que

$$\det(I + H) = 1 + \text{tr}(H) + O(\|H\|^2).$$

□

A.3. Espacio-tiempo de Minkowski

Proposición A.15. Sean $p, q \in X$ dos eventos del espacio-tiempo y e el vector que une los dos eventos. Entonces e es un vector temporal si y solo si existe un observador para el que p y q ocurren en el mismo punto de su espacio. Además, $g(e, e)$ es el tiempo medido entre p y q para dicho observador y es la cota inferior para todos los observadores.

Demostración. Sea e un vector temporal, es decir $g(e, e) > 0$, y por tanto define una trayectoria viable. Existe un hiperplano $(\mathbb{R}e)^\perp = \mathbb{S}$ que contiene a p y por tanto podemos expresar p en una base cuyo primer vector sea e como $p = (0, x, y, z)$. De esta forma, como $q = e + p$ tenemos que $q = (1, x, y, z)$. Concluimos que estos dos eventos ocurren en el mismo lugar para el observador e .

Supongamos ahora que p y q ocurren en el mismo lugar para un observador \tilde{e} , veamos que $g(e, e) > 0$. Dicho observador define una base en la que $p = (0, x, y, z)$ y $q = (t, x, y, z)$. Es fácil ver que $e = (t, 0, 0, 0) = t \cdot \tilde{e}$. Por tanto, $g(e, e) = t^2 g(\tilde{e}, \tilde{e}) = t^2$ ya que \tilde{e} es unitario. Además, es inmediato verificar que $\sqrt{e \cdot e}$ es el tiempo entre los dos eventos para ese observador.

Veamos ahora que este tiempo es el menor que puede medir cualquier otro observador $e' = e + \vec{v}$. Entonces el tiempo que mide es $t' = e' \cdot (q - p) = \tilde{e} \cdot (q - p) + \vec{v} \cdot (q - p)$. Se verifica que $\vec{v} \cdot (q - p) \geq 0$ pues $\vec{v} \in \mathbb{S}$ al ser una velocidad aparente, por lo que se concluye. □

Proposición A.16. Sean $p, q \in X$ dos eventos del espacio-tiempo y e el vector que une los dos eventos. Entonces e es un vector espacial si y solo si existe un observador para el que p y q son simultáneos. Además, $\sqrt{\langle e, e \rangle}$ es el espacio medido entre p y q por dicho observador y es la cota inferior para todos los observadores. Donde $\langle -, - \rangle$ denota el producto escalar habitual.

Demostración. Supongamos que e es un vector espacial, es decir que $g(e, e) < 0$. Para ver que existe un observador que percibe los eventos p y q como simultáneos debemos encontrar un vector unitario \tilde{e} tal que $g(e, \tilde{e}) = 0$.

Sea e_0 un observador arbitrario. Buscamos \tilde{e} de la forma $\tilde{e} = e_0 + \lambda e$ con $\lambda \in \mathbb{R}$. Así:

$$g(e, \tilde{e}) = g(e, e_0 + \lambda e) = g(e, e_0) + \lambda g(e, e).$$

Para que $g(e, \tilde{e}) = 0$, tomamos $\lambda = -\frac{g(e, e_0)}{g(e, e)}$. Nótese que $g(e, e) < 0$, por lo que el valor de λ está bien definido. Para dicho valor del parámetro tenemos que $g(e, \tilde{e}) = 0$, por tanto para el observador que define la trayectoria $\mathbb{R}\tilde{e}$ los eventos p y q son simultáneos.

Demostremos ahora que si hay un observador \tilde{e} que percibe los eventos p y q como simultáneos entonces e es un vector espacial. Como \tilde{e} que percibe los eventos p y q como simultáneos, tenemos que $g(\tilde{e}, e) = 0$. Por otro lado, sabemos que la restricción de la métrica g al subespacio ortogonal a e es definida negativa. Por tanto, $g(e, e) < 0$. Es decir, e es un vector espacial.

Es inmediato ver que $\sqrt{e \cdot e}$ es la distancia medida por \tilde{e} ya que en la definición de estructura lorentziana define un espacio euclídeo sobre el subespacio \mathbb{S} por lo que $\|e\| = \sqrt{\langle e, e \rangle}$

Veamos ahora que el resto de observadores al medir la distancia entre los puntos espaciales en los que ellos perciben los eventos p y q es siempre mayor que el valor medido por \tilde{e} . Sea otro observador e' . Entonces, la distancia medida por e' es

$$e_{\perp} = e - g(e, e')e'.$$

Así,

$$g(e_{\perp}, e_{\perp}) = g(e, e) - g(e, e')^2 \iff g(e, e) = g(e_{\perp}, e_{\perp}) + g(e, e')^2.$$

Por tanto, tenemos que $g(e, e) > g(e_{\perp}, e_{\perp})$. Nótese que la desigualdad estricta viene de que $g(e, e') \neq 0$ ya que que si no e' no sería una trayectoria válida al tener velocidad aparente infinita para el observador \tilde{e} , como vimos en el lema 5.15. Volviendo de nuevo a la definición de estructura lorentziana, la métrica euclídea que se define en cada espacio es $-c^2 g|_{\mathbb{S}}$. Por lo que concluimos que $\langle e, e \rangle < \langle e_{\perp}, e_{\perp} \rangle$. \square

Ejemplo A.17. Sabemos que

$$m_{ap} = \gamma m, \quad \text{donde} \quad 0 < \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \leq 1,$$

por tanto podemos entender el factor de Lorentz como la tasa en la que aumenta la masa aparente en tanto por uno. Así, vamos a calcular el valor de γ para distintas velocidades, expresadas como porcentaje de la velocidad de la luz c :

- Si $v = 0,01c$, entonces

$$\gamma(0,01c) = \frac{1}{\sqrt{1-0,01^2}} \approx 1,00005,$$

es decir, la masa relativista es apenas un 0,005 % mayor que la masa en reposo.

- Si $v = 0,1c$, entonces

$$\gamma(0,1c) = \frac{1}{\sqrt{1-0,1^2}} \approx 1,005,$$

la diferencia es del 0,5 %.

- Si $v = 0,25c$, entonces

$$\gamma(0,25c) = \frac{1}{\sqrt{1-0,25^2}} = \frac{1}{\sqrt{0,9375}} \approx 1,0328,$$

es decir, la masa relativista es aproximadamente un 3,28 % mayor que la masa en reposo.

- Si $v = 0,5c$, entonces

$$\gamma(0,5c) = \frac{1}{\sqrt{1-0,5^2}} = \frac{1}{\sqrt{0,75}} \approx 1,1547,$$

tenemos un aumento del 15,5 %.

- Si $v = 0,9c$, entonces

$$\gamma(0,9c) = \frac{1}{\sqrt{1-0,9^2}} = \frac{1}{\sqrt{0,19}} \approx 2,294,$$

la masa se ha más que duplicado.

- Si $v = 0,99c$, entonces

$$\gamma(0,99c) \approx 7,09.$$

Por tanto, el efecto relativista empieza a ser apreciable a partir de velocidades cercanas al $0,25c$, es decir una velocidad de 74948 km/s. En ese caso la diferencia ya supera el 3 %, y se vuelve muy notable según nos acercamos a la velocidad de la luz. Es decir, este efecto sólo es apreciable cuando nos movemos a velocidades extremadamente altas.

A.4. Relatividad General

Lema A.18. Sea $\omega \in \Lambda^2 M$. Entonces, para todo $X, Y, Z \in \mathfrak{X}(M)$, se cumple:

$$d\omega(X, Y, Z) = X(\omega(Y, Z)) + Y(\omega(Z, X)) + Z(\omega(X, Y)) - \omega([X, Y], Z) - \omega([Y, Z], X) - \omega([Z, X], Y).$$

Demostración. En coordenadas podemos escribir ω como:

$$\omega = \sum_{i < j} \omega_{ij} dx^i \wedge dx^j.$$

Entonces tomando la diferencial exterior queda,

$$d\omega = \sum_{i < j} \sum_k \partial_k \omega_{ij} dx^k \wedge dx^i \wedge dx^j.$$

Escribamos los campos como $X = \sum_r X^r \partial_r$, $Y = \sum_s Y^s \partial_s$ y $Z = \sum_t Z^t \partial_t$. Ahora, evaluemos $d\omega(X, Y, Z)$:

$$d\omega(X, Y, Z) = \sum_{i < j} [d\omega_{ij}(X) \wedge dx^i \wedge dx^j(Y, Z) + d\omega_{ij}(Y) \wedge dx^i \wedge dx^j(Z, X) + d\omega_{ij}(Z) \wedge dx^i \wedge dx^j(X, Y)].$$

Intentemos expresar $d\omega_{ij}(X)$ en función de los campos.

$$\begin{aligned} X(\omega(Y, Z)) &= X\left(\sum_{i < j} \omega_{ij}(Y^i Z^j - Y^j Z^i)\right) \\ &= \sum_{i < j} X(\omega_{ij})(Y^i Z^j - Y^j Z^i) + \sum_{i < j} \omega_{ij} X(Y^i Z^j - Y^j Z^i) \\ &= \sum_{i < j} X(\omega_{ij})(Y^i Z^j - Y^j Z^i) + \sum_{i < j} \omega_{ij} [X(Y^i)Z^j + Y^i X(Z^j) - X(Y^j)Z^i - Y^j X(Z^i)]. \end{aligned}$$

Despejando tenemos que

$$\sum_{i < j} X(\omega_{ij})(Y^i Z^j - Y^j Z^i) = X(\omega(Y, Z)) - \sum_{i < j} \omega_{ij} [X(Y^i)Z^j + Y^i X(Z^j) - X(Y^j)Z^i - Y^j X(Z^i)]$$

Análogamente tenemos

$$\sum_{i < j} Y(\omega_{ij})(Z^i X^j - Z^j X^i) = Y(\omega(Z, X)) - \sum_{i < j} \omega_{ij} [Y(Z^i)X^j + Z^i Y(X^j) - Y(Z^j)X^i - Z^j Y(X^i)]$$

y

$$\sum_{i < j} Z(\omega_{ij})(X^i Y^j - X^j Y^i) = Z(\omega(X, Y)) - \sum_{i < j} \omega_{ij} [Z(X^i)Y^j + X^i Z(Y^j) - Z(X^j)Y^i - X^j Z(Y^i)]$$

$$d\omega(X, Y, Z) = X(\omega(Y, Z)) + Y(\omega(Z, X)) + Z(\omega(X, Y)) - \omega([X, Y], Z) - \omega([Y, Z], X) - \omega([Z, X], Y).$$

Finalmente notando que $d\omega_{ij}(X) = X(\omega_{ij})$ podemos sustituir en la expresión de $d\omega(X, Y, Z)$ y recordando que $[X, Y]^i = X(Y^i) - Y(X^i)$ llegamos a

□

Lema A.19. Sea (X, ∇) una variedad diferenciable equipada con una conexión sin torsión. Sea $\omega \in \Lambda^2 M$, entonces para campos $X, Y, Z \in \mathfrak{X}(X)$, se cumple que

$$d\omega(X, Y, Z) = (\nabla_X \omega)(Y, Z) + (\nabla_Y \omega)(Z, X) + (\nabla_Z \omega)(X, Y).$$

Demostración. Por el lema A.18 sabemos que

$$\begin{aligned} d\omega(X, Y, Z) &= X(\omega(Y, Z)) + Y(\omega(Z, X)) + Z(\omega(X, Y)) \\ &\quad - \omega([X, Y], Z) - \omega([Y, Z], X) - \omega([Z, X], Y). \end{aligned}$$

Por otro lado, la definimos como

$$(\nabla_X \omega)(Y, Z) = X(\omega(Y, Z)) - \omega(\nabla_X Y, Z) - \omega(Y, \nabla_X Z).$$

Consideramos la siguiente suma

$$\begin{aligned} &(\nabla_X \omega)(Y, Z) + (\nabla_Y \omega)(Z, X) + (\nabla_Z \omega)(X, Y) \\ &= X(\omega(Y, Z)) + Y(\omega(Z, X)) + Z(\omega(X, Y)) \\ &\quad - [\omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z) + \omega(\nabla_Y Z, X) + \omega(Z, \nabla_Y X) + \omega(\nabla_Z X, Y) + \omega(X, \nabla_Z Y)]. \end{aligned}$$

Como ω es una forma alternada podemos agrupar los términos del corchete como

$$[-\omega(\nabla_X Y, Z) + \omega(\nabla_Y X, Z)] + [-\omega(\nabla_Y Z, X) + \omega(\nabla_Z Y, X)] + [-\omega(\nabla_Z X, Y) + \omega(\nabla_X Z, Y)].$$

Teniendo en cuenta que la conexión tiene torsión nula, podemos expresar dichos pares en términos del corchete de Lie y así:

$$\begin{aligned} &(\nabla_X \omega)(Y, Z) + (\nabla_Y \omega)(Z, X) + (\nabla_Z \omega)(X, Y) \\ &= X(\omega(Y, Z)) + Y(\omega(Z, X)) + Z(\omega(X, Y)) \\ &\quad - \omega([X, Y], Z) - \omega([Y, Z], X) - \omega([Z, X], Y) = d\omega(X, Y, Z). \end{aligned}$$

□

Bibliografía

- [AB12] R. ACHILLES, A. BONFIGLIOLI *The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin*. Springer-Verlag, 2012.
- [BG80] R.L. BISHOP, S. I. GOLDBERG *Tensor Analysis on Manifolds*. Dover Publications, New York, 1980.
- [Car04] S.M. CARROLL *Spacetime and Geometry: An Introduction to General Relativity*. Addison Wesley, 2004.
- [GNS19] A. GORDILLO, J. NAVARRO AND P. J. SANCHO *A remark on the invariant theory of real Lie groups*. Colloq. Math. 159 (2019), no.2, 295–300. <https://arxiv.org/pdf/1711.02444>
- [HE94] S.W. HAWKING, G.F.R. ELLIS *The large scale structure of space-time*. Cambridge Monographs on Mathematical Physics, 1994.
- [HC-PG19] M.A. HERNÁNDEZ CIFRE, J.A. PASTOR GONZÁLEZ *Un curso de Geometría Diferencial: Teoría, problemas, soluciones y prácticas con ordenador*. Consejo Superior de Investigaciones Científicas, 2019.
- [KN51] M.S. KNEBELMAN *Spaces of Relative Parallelism*. Annals of Mathematics, 53(3), 387–399, mayo 1951. <https://doi.org/10.2307/1969562>
- [KoNo63] S. KOBAYASHI, K. NOMIZU *Foundations of Differential Geometry. Vol. I. II*. Interscience Publishers, New York-London, 1963.
- [Küh06] W. KÜHNEL *Differential Geometry: Curves, Surfaces, Manifolds*. American Mathematical Society, 2006.
- [NS04] J.A. NAVARRO, J.B. SANCHO *Gravitación Newtoniana y Relatividad*. <https://matematicas.unex.es/~navarro/relatividad/apuntrel.pdf>
- [NS17] J.A. NAVARRO, J.B. SANCHO *Newtonian Gravitation: Its Geometry and Relation with the Relativistic Gravitation*. No publicado, 2017.
- [N24] J.A. NAVARRO *Notes for a Licenciatura: Based on Lectures of J. Sancho*.

-
- [SW77] R.K. SACHS, H. WU *General Relativity for Mathematicians*. Springer-Verlag, New York, 1977.
- [Sak96] T. SAKAI *Riemannian Geometry*. American Mathematical Society, 1996.
- [ZB22] A. ZAMORA, H. BARGE *Geometría Proyectiva y Afín*. Sanz y Torres, 2022.
- [1] Math StackExchange. *Taylor series with covariant derivative*, 2025.
<https://math.stackexchange.com/questions/5032217/taylor-series-with-covariant-derivative>.

Índice alfabético

- 3-forma de impulso, 91
- 3-forma de materia, 90
- Aplicación exponencial, 31
- Conexión, 6
- Conexión de Levi-Civita, 10
- Cono de luz, 59
- Contracción Lorentz, 84
- Coordenadas normales, 31
- Curvatura escalar, 48
- Curvatura seccional, 44
- Derivada covariante, 9
- Dilatación temporal, 81
- Espacio-tiempo, 51
- espacio-tiempo de Minkowski, 71
- Estructura galileana, 57
- Estructura lorentziana, 57
- Factor de Lorentz, 75
- Fluido perfecto, 97
- Forma de Cartan, 122
- Forma de Killing, 104
- Fórmula de Koszul, 9
- Geodésica, 27
- Grassmaniana, 53
- Grupo de Lorentz, 63
- grupo de Lorentz, 56
- Grupo de Poincaré, 73
- Grupo de un observador, 55
- Grupo estructural, 55
- Impulso, 77
- Masa aparente, 77
- Masa en reposo, 77
- Métrica espacial, 71
- Métrica temporal, 71
- Presión escalar, 98
- Primera ecuación de estructura de Cartan, 45
- Pseudotransporte paralelo, 16
- Segunda ecuación de estructura de Cartan, 45
- Sistema de partículas libres, 92
- Sistema de referencia inercial, 74
- Símbolos de Christoffel, 8
- Tensor de curvatura, 42
- Tensor de Einstein covariante, 119
- Tensor de materia covariante, 96
- Tensor de Ricci, 47
- Tensor de tensión, 99
- Tiempo propio, 58, 75
- Torsión, 6
- Transporte paralelo, 14, 18
- Variedad de métricas, 69
- Variedad lorentziana, 114
- Vector espacial, 71
- Vector luz, 71
- Vector temporal, 71
- Velocidad aparente, 59



EPES

Escuela Politécnica
de Enseñanza Superior