



---

---

# UNIVERSIDAD DE GRANADA

## FACULTAD DE CIENCIAS

### Grado de Física

#### Apuntes de

## RELATIVIDAD GENERAL

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G_N T_{\mu\nu}$$

Autor:

Rubén Carrión Castro

Profesores de la asignatura: Dr. Javier Olmedo

Dra. Mar Bastero Gil

Cursado en la Universidad de Granada

16 de septiembre de 2025

# Índice general

<b>Índice general</b>	<b>II</b>
<b>Prefacio</b>	<b>1</b>
<b>1. Álgebra de Tensores en espacios vectoriales</b>	<b>2</b>
1.1. Preliminares . . . . .	3
1.1.1. Espacios vectoriales . . . . .	3
1.1.2. Aplicaciones lineales . . . . .	5
1.1.3. Espacios duales . . . . .	11
1.1.3.1. Anulador de un subespacio . . . . .	21
1.1.3.2. Aplicación lineal traspuesta . . . . .	22
1.1.3.3. Una aplicación de la teoría del espacio dual: Interpretación de Lagrange . . . . .	23
1.1.4. Notación de Einstein . . . . .	26
1.1.5. Formas bilineales, productos escalares y formas cuadráticas . . . . .	27
1.1.5.1. Formas bilineales y producto escalar . . . . .	28
1.1.5.2. Formas cuadráticas . . . . .	32
1.1.5.3. Signatura de una forma cuadrática real . . . . .	36
1.2. Álgebra de Tensores . . . . .	40
1.2.1. Producto tensorial: caso de dos términos . . . . .	40
1.2.2. El espacio de tensores $(r,s)$ : Definición, propiedades y ejemplos . . . . .	49
1.2.3. Contracción de tensores . . . . .	58
1.2.4. Leyes de transformación . . . . .	61
1.2.5. Invariantes . . . . .	65
<b>2. Teoría de la Relatividad Especial</b>	<b>67</b>
2.1. Repaso histórico . . . . .	68
2.1.1. Relatividad Galileana . . . . .	68
2.1.2. Transformaciones de Lorentz . . . . .	70
2.2. Postulado de la Relatividad Especial . . . . .	70
2.2.1. Aspectos a tener en cuenta . . . . .	71
2.2.2. Adición de velocidades . . . . .	72
2.3. Dilatación temporal . . . . .	73

2.4.	Contracción de longitudes . . . . .	74
2.5.	Elemento de línea finito . . . . .	75
2.5.1.	Elemento de línea infinitesimal . . . . .	75
2.5.2.	Cono de luz . . . . .	76
2.5.3.	Sistema de referencia propio . . . . .	76
2.5.4.	Tiempo propio . . . . .	76
2.6.	Métrica de Minkowski y cuadrivectores . . . . .	77
2.6.1.	Grupo de Poincaré . . . . .	77
2.7.	Repasso de álgebra . . . . .	78
2.7.1.	Tensor simétrico y antisimétrico . . . . .	80
2.7.2.	Transformaciones de Lorentz. Versión covariante . . . . .	80
2.8.	Cinemática relativista . . . . .	82
2.8.1.	Vector cuadrivelocidad . . . . .	82
2.8.2.	Vector cuadrimomento . . . . .	83
2.8.3.	Vector cuadriaceleración . . . . .	83
2.8.4.	Derivación . . . . .	84
2.8.5.	Operador D'Alembertiano . . . . .	84
2.8.6.	Tensor de Levi-Civita . . . . .	84
2.9.	Grupo de Poincaré . . . . .	85
2.10.	Dinámica relativista . . . . .	87
2.10.1.	Principio variacional de acción estacionaria . . . . .	87
2.10.2.	Cantidades Conservadas . . . . .	88
2.10.3.	Partícula libre relativista . . . . .	88
2.10.4.	Sistema de N-partículas . . . . .	89
2.11.	Campos relativistas . . . . .	91
2.11.1.	Cantidades conservadas . . . . .	92
<b>3.</b>	<b>Geometría Diferenciable</b>	<b>93</b>
3.1.	Breve repaso histórico . . . . .	94
3.2.	Introducción . . . . .	94
3.2.1.	Prefacio . . . . .	95
3.3.	Variedad diferenciable . . . . .	97
3.3.1.	Definiciones . . . . .	98
3.3.2.	Transformaciones de coordenadas . . . . .	100
3.3.2.1.	Escalares . . . . .	100
3.3.2.2.	Vectores . . . . .	100
3.3.2.3.	Campo vectorial en la variedad . . . . .	101
3.3.2.4.	Uno-formas . . . . .	102
3.3.2.5.	Tensores . . . . .	103
3.4.	Aplicaciones diferenciables y difeomorfismos . . . . .	105
3.5.	Derivada de Lie . . . . .	107
3.5.1.	Derivada covariante . . . . .	109
3.6.	Conexión de Levi-Civita . . . . .	111
3.7.	Transporte paralelo . . . . .	112

3.8.	Geodésicas . . . . .	113
3.8.1.	Geodésicas como Principio Variacional . . . . .	114
3.8.2.	Derivación de los términos de Christoffel mediante las geodésicas	115
3.8.3.	Densidad tensorial . . . . .	116
3.9.	Tensores de Curvatura . . . . .	116
<b>4.</b>	<b>Relatividad General</b>	<b>118</b>
4.1.	Principio de Equivalencia . . . . .	119
4.1.1.	Principio de Equivalencia débil . . . . .	119
4.1.2.	Principio de Equivalencia de Einstein . . . . .	120
4.1.3.	Principio de Equivalencia fuerte . . . . .	121
4.2.	Ecuaciones de Einstein . . . . .	122
4.3.	Principio de Covariancia General . . . . .	125
4.4.	Campos materiales . . . . .	125
4.4.1.	Fluido perfecto . . . . .	126
4.4.2.	Campo escalar . . . . .	128
4.4.3.	Campo electromagnético . . . . .	128
4.5.	Condiciones de energía . . . . .	128
4.5.1.	Teoremas de Singularidad . . . . .	129
4.6.	Ecuaciones de Einstein y Principio de Mínima Acción . . . . .	131
4.7.	Constante cosmológica . . . . .	133
4.8.	Soluciones dinámicas de las ecuaciones de Einstein . . . . .	134
4.8.1.	Espacio-tiempo de Schwarzschild . . . . .	135
4.8.2.	Geometría del interior de una estrella estática y esféricamente simétrica . . . . .	138
4.8.3.	Estrellas de densidad uniforme . . . . .	141
4.8.4.	Colapso gravitatorio. Modelo de Oppenheimer-Snyder . . . . .	141
4.8.5.	Propiedades de la métrica de Schwarzschild . . . . .	143
4.8.6.	Métrica de Kerr. Agujero negro en rotación. 1963 . . . . .	144
4.9.	Testeos clásicos de la Relatividad General . . . . .	145
4.9.1.	Geodésicas en una geometría de Schwarzschild . . . . .	145
4.9.2.	Variación del perihelio de Mercurio . . . . .	147
4.9.3.	Corrimiento al rojo gravitacional . . . . .	149
4.9.4.	Deflexión de la luz . . . . .	149
4.9.5.	Dilatación temporal gravitatoria . . . . .	151
<b>5.</b>	<b>Ondas gravitacionales</b>	<b>153</b>
5.1.	Contexto histórico . . . . .	154
5.2.	Ondas gravitacionales . . . . .	154
5.3.	Ecuación de la desviación geodésica . . . . .	157
5.3.1.	Efecto de la una onda gravitacional que pasa cerca del observador	157
5.3.2.	Generación de las ondas gravitacionales . . . . .	159
5.4.	Energía de una onda gravitacional. Formalismo de Bondi . . . . .	161
<b>6.</b>	<b>Cosmología</b>	<b>163</b>

6.1. Conceptos básicos . . . . .	164
6.1.1. Observadores comóviles . . . . .	166
6.1.2. Contenido de materia . . . . .	167
6.1.3. Tensores de Killing . . . . .	171
6.1.4. Ley de Hubble . . . . .	172
<b>Appendices</b>	<b>174</b>
<b>A. Espacios topológicos</b>	<b>174</b>
<b>B. Variedades</b>	<b>183</b>
B.1. Definición de Variedad . . . . .	183
B.2. Ejemplos de Variedades . . . . .	187
B.2.1. Espacio Cartesiano . . . . .	187
B.2.2. Subvariedad Abierta . . . . .	187
B.2.3. Producto de Variedades . . . . .	188
B.3. Mapas diferenciables . . . . .	189
B.4. Subvariedades . . . . .	192
B.5. Curvas diferenciables . . . . .	195
<b>Bibliografía</b> . . . . .	<b>196</b>
<b>Bibliografía de Figuras</b> . . . . .	<b>196</b>

# Prefacio

*Todos somos muy ignorantes. Lo que ocurre es que no todos ignoramos las mismas cosas.*  
(A. Einstein)

Estos apuntes están hechos para acercar a estudiantes como yo a la Relatividad General, empezaremos con unas nociones de tensores, que hice en la Universidad de Córdoba con el doctor Jónatan Herrera. Seguiremos con la Teoría Especial de la Relatividad, luego un repaso de Geometría Diferencial y finalmente introduciremos la Relatividad General.

# Capítulo 1

## Álgebra de Tensores en espacios vectoriales

*“Toda la ciencia no es más que un refinamiento del pensamiento cotidiano”.*  
(A. Einstein)

Para poder abordar este capítulo, es recomendable hacer una lectura del **Apéndice A. Álgebra: Lógica, Conjuntos y Grupos.**

## 1.1. Preliminares

En esta sección hacemos un breve repaso a los conceptos básicos de álgebra lineal que necesitaremos para definir correctamente la noción de tensor.

Supondremos que ya existe un conocimiento previo sobre conjuntos, funciones y grupos.

### 1.1.1. Espacios vectoriales

Comenzamos revisionando la noción de espacio vectorial.

**Definición 1.1.1.** Consideremos  $V$  un conjunto y  $\mathbb{K}$  un cuerpo. Supongamos que tenemos definidas una operación interna  $+$ :  $V \times V \rightarrow V$  y una operación externa  $\cdot$ :  $V \times \mathbb{K} \rightarrow V$ . Diremos que  $(V, \mathbb{K}, +, \cdot)$  (o simplemente  $(V, \mathbb{K})$ ) tiene **estructura de espacio vectorial** si se satisfacen las siguientes propiedades:

(I)  $(V, +)$  es un grupo abeliano, es decir:

(I)-1) Se cumple que  $+$ :  $V \times V \rightarrow V$  es una operación cerrada, es decir,  $\forall v, w \in V$  se tiene que  $v + w \in V$ .

(I)-2) Propiedad asociativa:

$$\forall u, v, w \in V, \quad u + (v + w) = (u + v) + w$$

(I)-3) Existe el elemento neutro:

$$\forall v \in V, \exists 0 \in V; \quad v + 0 = 0 + v = v$$

(I)-4) Existe elemento simétrico:

$$\forall v \in V, \exists -v \in V; \quad v + (-v) = (-v) + v = 0$$

(I)-5) Propiedad conmutativa:

$$\forall v, w \in V; \quad v + w = w + v$$

(II) Doble propiedad distributiva:

$$(II)-1) \quad \forall \lambda, \mu \in \mathbb{K}, \forall v \in V; \quad (\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v)$$

$$(II)-2) \quad \forall \lambda \in \mathbb{K}, \forall v, w \in V; \quad \lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w)$$

(III) Propiedad pseudo-asociativa:

$$\forall \lambda, \mu \in \mathbb{K}, \forall v \in V; \quad \lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$$

(IV) Se verifica que:

$$\forall v \in V; \quad 1 \cdot v = v \cdot 1 = v, \text{ donde } 1 \in \mathbb{K} \text{ es el elemento unitario de } \mathbb{K}$$

Los elementos del espacio vectorial suelen denominarse **vectores** mientras que a los del cuerpo  $\mathbb{K}$ , los llamaremos **escalares**. La operación externa recibe el nombre de **producto por escalares**.

El siguiente concepto que necesitaremos sobre un espacio vectorial es la noción de dimensión. Para poder definir dicha noción, requerimos dos conceptos adicionales: la noción de sistema generador y de lineal dependencia/independencia. El primero nos permite establecer cuando un conjunto de vectores genera todo el espacio vectorial. Esto quiere decir, que todo vector  $v \in V$  puede expresarse como **combinación lineal** de los vectores de mi sistema generador, los cuales pueden ser **linealmente dependientes o independientes** entre ellos mismos, recalmando el caso particular de que sean independientes. Veamos una definición más formal.

**Definición 1.1.2.** Sea  $V$  un  $\mathbb{K}$ -espacio vectorial y sean  $v_1, v_2, \dots, v_n \in V$ :

- Se dice que  $\{v_1, v_2, \dots, v_n\}$  es un conjunto generador si  $\langle v_1, v_2, \dots, v_n \rangle = V$ .
- Se dice que  $v_1, v_2, \dots, v_n$  son linealmente independientes si  $\lambda^1 \cdot v_1 + \lambda^2 \cdot v_2 + \dots + \lambda^n \cdot v_n = 0$  con  $\lambda^i \in \mathbb{K}$ , si y solo si  $\lambda^1 = \lambda^2 = \dots = \lambda^n = 0$ .
- Se dice que  $\{v_1, v_2, \dots, v_n\}$  es una base si es a la vez conjunto generador y los vectores son linealmente independientes.

Una vez obtenido el concepto de base, podemos ir ya al concepto de dimensión del espacio vectorial.

**Definición 1.1.3.** *Se llama dimensión de un espacio vectorial al cardinal de cualquiera de sus bases, es decir,  $\dim G = \#B_G$ .*

### 1.1.2. Aplicaciones lineales

Veamos ahora las aplicaciones lineales.

**Definición 1.1.4.** *Sean  $V$  y  $V'$  dos espacios vectoriales sobre el mismo cuerpo  $\mathbb{K}$ . Se dice que en una aplicación  $f : V \rightarrow V'$  es una aplicación lineal, o también llamado homomorfismo de espacios vectoriales, si se verifica:*

$$(I) \quad f(x + y) = f(x) + f(y), \forall x, y \in V$$

$$(II) \quad f(\lambda \cdot x) = \lambda \cdot f(x), \forall \lambda \in \mathbb{K}, \forall x \in V$$

*Diremos además que  $f$  es un isomorfismo lineal si es biyectiva, que  $f$  es un endomorfismo si  $V = V'$  y que es un automorfismo si es un endomorfismo biyectivo.*

Las aplicaciones lineales tienen asociados dos conjuntos cuyas características son de interés, a saber, el núcleo y la imagen.

**Definición 1.1.5.** *Sea  $f : V \rightarrow W$  definimos el núcleo o kernel de la aplicación  $f$  como*

$$\text{Ker } f = \{v \in V : f(v) = 0\}$$

*y la imagen como*

$$\text{Im } f = \{w \in W : \exists v \in V / f(v) = w\}.$$

Veamos algunas propiedades básicas de ambos conjuntos.

**Proposición 1.1.1.** *Sea  $f : V \rightarrow V'$  una aplicación lineal, se tienen las siguientes propiedades:*

(I)  $\text{Im } f$  es un subespacio de  $V'$  y que  $\text{Ker } f$  es un subespacio de  $V$ .

(II) Si  $W$  es un subespacio vectorial de  $V$ , entonces  $f(W) := \{f(w) : w \in W\}$  es un subespacio de  $V'$ .

(III) Si  $W'$  es un subespacio de  $V'$ , entonces  $f^{-1}(W') := \{v \in V : f(v) \in W'\}$  es también un subespacio de  $V$ .

**Demostración.** (I) Por definición, como los elementos de la  $\text{Im}f$  son pertenecientes a  $V'$ , entonces la  $\text{Im}f$  es subespacio de  $V'$ . De igual forma ocurre con el  $\text{Ker}f$ , pues sus elementos pertenecen a  $V$  y por tanto, este es subespacio de  $V$ .

(II) Como  $W$  es subespacio de  $V$ , tenemos que  $w \in V$  también, por tanto, los  $f(w)$  pertenecerán a  $V'$ , cosa que implica que  $f(W)$  es subespacio de  $V'$ , pues los  $f(w)$  de  $f(W)$  pertenecen a  $V'$ .

(III) Por analogía a (II) vemos que  $f^{-1}(W)$  es subespacio de  $V$ .

■

Ahora veamos algunas propiedades esenciales de las aplicaciones lineales.

**Proposición 1.1.2.** Sea  $f : V \rightarrow V'$  una aplicación lineal,

(I) entonces  $f$  es inyectiva si y solo si  $\text{Ker}f = \{0\}$ .

(II) si  $G$  es un conjunto generador de  $V$ ,  $\langle G \rangle = V$ , entonces  $f(G)$  es conjunto generador de  $\text{Im}f$ ,  $\langle f(G) \rangle = \text{Im}f$ .

(III) si  $S \subset V$  es un conjunto de vectores linealmente independientes, si  $f$  es inyectiva, entonces  $f(S)$  es linealmente independiente.

(IV)  $f$  es inyectiva  $\Leftrightarrow$  conserva la independencia lineal.

(V) si  $f$  es biyectiva y  $B$  es una base de  $V$ , entonces  $f(B)$  es base de  $V'$ .

(VI)  $f$  es sobreyectiva  $\Leftrightarrow \text{Im}f = V'$

**Demostración.** (I)  $\Rightarrow$  Suponiendo que  $f$  es inyectiva, sabemos que su Kernel es,

$$\text{Ker}f = \{v \in V : f(v) = 0\}$$

pero como la inyectividad nos implica que la imagen debe provenir de un único vector de entrada, entonces este vector será  $v = 0$ , y por tanto,  $\text{ker}f = \{0\}$ . ✓

$\Leftarrow$  Suponiendo que  $\ker f = \{0\}$ , esto nos quiere decir que únicamente el vector  $v = 0$  satisface  $f(v) = 0$ , luego como un vector tiene una única imagen, decimos que  $f$  es inyectiva. ■

(II) Veamos que el conjunto  $f(G)$  es sistema generador de la imagen, es decir,

$$\langle f(G) \rangle = \text{Im } f \Leftrightarrow \forall y \in \text{Im } f, \exists \lambda^1, \dots, \lambda^n \in \mathbb{K}, y_1, \dots, y_n \in f(G) \text{ tales que } y = \lambda^1 y_1 + \dots + \lambda^n y_n.$$

Sabemos que  $G$  es conjunto generador, luego sea  $y \in \text{Im } f$ . Entonces por definición se tiene que existe  $x \in V$  tal que  $f(x) = y$ . Como  $\langle G \rangle = V$ , existen  $\lambda^1, \dots, \lambda^n \in \mathbb{K}$ ,  $v_1, \dots, v_n \in G$  tales que

$$x = \sum_{i=1}^n \lambda^i v_i.$$

Tenemos entonces que:

$$y = f(x) = f(\lambda^1 v_1 + \dots + \lambda^n v_n) = \lambda^1 f(v_1) + \dots + \lambda^n f(v_n).$$

Por lo tanto,  $y$  es combinación lineal de elementos de  $f(G)$ , es decir,  $\langle f(G) \rangle = \text{Im } f$ . ■

(III) Sea  $S$  un conjunto linealmente independiente en  $V$ . Supongamos que  $f$  es inyectiva, vamos a probar que  $f(S)$  es linealmente independiente, es decir,

$$\lambda^1 y_1 + \dots + \lambda^n y_n = 0 \Rightarrow \lambda^1 = \lambda^2 = \dots = \lambda^n = 0 \quad \forall y_1, \dots, y_n \in f(S), \quad \forall \lambda^1, \dots, \lambda^n \in \mathbb{K}$$

Supongamos  $\lambda^1 y_1 + \dots + \lambda^n y_n = 0$ . Como  $y_j \in f(S)$ ,  $\exists x_j \in S / f(x_j) = y_j$ .

$$\left. \begin{array}{l} \lambda_1 f(x_1) + \dots + \lambda_n f(x_n) = 0 \\ f(\lambda_1 x_1 + \dots + \lambda_n x_n) = 0 \end{array} \right\} f(0) = 0 \Rightarrow \lambda_1 x_1 + \dots + \lambda_n x_n = 0 \Rightarrow f \text{ inyectiva} \Rightarrow \lambda_1 = \dots = \lambda_n = 0$$

(IV)  $\implies$  Trivial por (iii) ✓

$\Leftarrow$  Por reducción al absurdo:

Supongamos que existen  $v_1, v_2 \in V$  distintos, tales que  $f(v_1) = f(v_2) \Leftrightarrow f(v_1) - f(v_2) = 0 \Leftrightarrow f(v_1 - v_2) = 0$ . Luego,  $v = v_1 - v_2 \neq 0$  verifica que  $f(v) = 0$ ,  $\{v\}$  es un conjunto linealmente independiente,  $f(\{v\})$  tendría que ser un conjunto l.i. por hipótesis, pero  $f(\{v\}) = \{0\}$  que no es un conjunto l.i. cosa absurda. ■

(v) Sea una aplicación lineal biyectiva  $f : V \rightarrow V'$  y una base de  $V$ ,  $B = \{v_1, \dots, v_n\}$ . Entonces, si aplicamos

$$f(B) = \{f(v_1), \dots, f(v_n)\} = \{v'_1, \dots, v'_n\}$$

y entonces, estos  $v'_i \in V'$  van a formar una base de  $V'$ , pues al ser  $f$  biyectiva, los vectores serán linealmente independientes, pues los de  $B$  lo son; y además, como tienen la misma dimensión que  $V'$ , pasan de ser conjunto generador a base. ■

(vi)  $\Rightarrow$  Suponiendo que  $f$  es sobreyectiva, tendremos que para cada  $y \in V'$ , existe al menos un  $x \in V$ , tal que  $f(x) = y$ . Por consiguiente, cada elemento de  $V'$  es la imagen de un elemento de  $V$ , es decir,  $Im f = V'$ . ✓

$\Leftarrow$  Suponiendo que  $im f = V'$ , tenemos que todos los elementos de  $V'$  son imagen de los elementos de  $V$ , siendo esta la propia definición de sobreyectividad, luego  $f$  es sobreyectiva. ■

Una vez visto estas propiedades, de (v) podemos obtener un resultado interesante, que es la siguiente proposición.

**Proposición 1.1.3.** *Sea  $B = \{v_1, \dots, v_n\}$  una base de  $V$ , y sea  $f : V \rightarrow V'$  una aplicación lineal. Se tiene entonces que  $\{f(v_1), \dots, f(v_n)\}$  es un sistema generador de la imagen.*

**Demostración.** Supongamos que  $B$  es una base y que conocemos  $f(v_j), \forall v_j \in B$ . Sea  $v \in V$ , escrito en coordenadas de la base como  $v = \lambda^1 v_1 + \dots + \lambda^n v_n$ , con  $v_i \in B$ , y  $\lambda^i \in \mathbb{K}$ , entonces  $f(x) = f(\lambda^1 v_1 + \dots + \lambda^n v_n) = \lambda^1 f(v_1) + \dots + \lambda^n f(v_n)$ , luego hemos puesto  $f(x)$  en coordenadas de  $\{f(v_1), \dots, f(v_n)\}$ . ■

Ahora vamos a ver un resultado bastante importante, el cuál nos permitirá representar aplicaciones lineales en matrices, denominadas **matrices asociadas a la aplicación  $f$** . Además, este resultado es importante para Física, pues los físicos no solemos trabajar con aplicaciones, sino que trabajamos con sus matrices asociadas, pues se puede decir que "tienen" la misma información que las aplicaciones.

**Proposición 1.1.4.** *Sean  $(V, +, \cdot)$  y  $(V', +, \cdot)$   $\mathbb{K}$ -espacios vectoriales de dimensión finita con  $\dim V = n$  y  $\dim V' = m$ . Sea  $f : V \rightarrow V'$  una aplicación lineal, entonces dadas*

$$\begin{cases} B = \{v_1, \dots, v_n\} \text{ base de } V \\ B' = \{v'_1, \dots, v'_n\} \text{ base de } V' \end{cases}$$

*$f$  se representa en esas bases como una matriz en  $\mathcal{M}_{m \times n}(\mathbb{K})$ .*

**Demostración.** Como  $f$  es lineal, me basta con conocer  $f(B)$ , para ello, tenemos que conocer  $f(v_1), f(v_2), \dots, f(v_n)$ , teniendo:

$$\begin{aligned} f(v_1) &= a_1^1 v'_1 + a_1^2 v'_2 + \cdots + a_1^m v'_m, \quad a_1^i \in \mathbb{K} \\ f(v_2) &= a_2^1 v'_1 + a_2^2 v'_2 + \cdots + a_2^m v'_m, \quad a_2^i \in \mathbb{K} \\ &\vdots && \vdots && \vdots \\ f(v_n) &= a_n^1 v'_1 + a_n^2 v'_2 + \cdots + a_n^m v'_m, \quad a_n^i \in \mathbb{K} \end{aligned}$$

Sea  $v \in V : v = \lambda^1 v_1 + \cdots + \lambda^n v_n$ ,  $\lambda^i \in \mathbb{K}$ , si le aplicamos  $f$  tenemos,

$$\begin{aligned} f(v) &= \lambda^1 f(v_1) + \cdots + \lambda^n f(v_n) \\ &= \lambda^1(a_1^1 v'_1 + \cdots + a_1^m v'_m) + \lambda^2(a_2^1 v'_1 + \cdots + a_2^m v'_m) + \cdots + \lambda^n(a_n^1 v'_1 + \cdots + a_n^m v'_m) \\ &= (a_1^1 \lambda^1 + a_2^1 \lambda^2 + \cdots + a_n^1 \lambda^n)v'_1 + (a_1^2 \lambda^1 + \cdots + a_n^2 \lambda^n)v'_2 + \cdots + (a_1^m \lambda^1 + \cdots + a_n^m \lambda^n)v'_m \end{aligned}$$

Luego,  $f(v) = \mu^1 v'_1 + \mu^2 v'_2 + \cdots + \mu^m v'_m$ , siendo  $\mu^i = (a_1^i \lambda^1 + \cdots + a_n^i \lambda^n)$ , luego, para construir la matriz  $A$ , ponemos las coordenadas de  $v_1$  en la primera columna, las de  $v_2$  en la segunda y así sucesivamente, tal que:

$$\begin{pmatrix} \mu^1 \\ \mu^2 \\ \vdots \\ \mu^m \end{pmatrix} = \begin{pmatrix} a_1^1 & a_1^2 & \cdots & a_1^m \\ a_2^1 & a_2^2 & \cdots & a_2^m \\ \vdots & \vdots & \ddots & \vdots \\ a_n^1 & a_n^2 & \cdots & a_n^m \end{pmatrix} \begin{pmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{pmatrix} \Rightarrow \mu = A \cdot \lambda$$

■

Vamos a introducir ahora el concepto de **rango** de una aplicación lineal, que puede extenderse al rango de su matriz asociada.

**Definición 1.1.6.** Se llama rango de una aplicación lineal (matriz) a la dimensión de su imagen y se denota por  $rg()$ .

Como un mismo espacio vectorial puede estar generado por varias bases, es lógico pensar que debe haber una relación entre estas bases o al menos una forma de cambiar de una base a otra, lo que se conoce como **cambio de base**. Esto es posible y una forma sencilla de hacerlo es mediante las matrices asociadas.

**Proposición 1.1.5.** -Sean  $V$  y  $V'$  dos espacios vectoriales en  $\mathbb{K}$ , sea  $f : V \rightarrow V'$  lineal.

-Sea  $B_1 = \{v_1, \dots, v_n\}$  base de  $V$ ,  $B'_1 = \{v'_1, \dots, v'_m\}$  base de  $V'$ .

-Sea  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$  la matriz que representa a  $f$  en  $B_1, B'_1$ .

-Sea  $B_2 = \{u_1, \dots, u_n\}$  base de  $V$ ,  $B'_2 = \{u'_1, \dots, u'_m\}$  base de  $V'$ .

-Sea  $\tilde{A} \in \mathcal{M}_{m \times n}(\mathbb{K})$  la matriz que representa a  $f$  en  $B_2, B'_2$ .

-Sea  $P$  la matriz de cambio de base de  $B_1$  en  $B_2$ .

-Sea  $Q$  la matriz de cambio de base de  $B'_1$  en  $B'_2$ .

Entonces  $\tilde{A} = Q^{-1} \cdot A \cdot P$ .

**Demostración.** Sea  $f : V \rightarrow V'$  una aplicación lineal con  $n = \dim V$  y  $m = \dim V'$ . Si  $A$  y  $\tilde{A}$  son las matrices asociadas a  $f$  respecto de distintas bases, entonces

$$rg(A) = \dim(Im f) = rg(\tilde{A})$$

Luego  $A$  y  $\tilde{A}$  tienen igual rango, y por tanto, son matrices equivalentes. Concretemos más esta situación:

Sean  $B_1$  y  $B_2$  bases de  $V$  con cambio de base de  $B_1$  a  $B_2$  dado por  $X_1 = PX_2$  y sean  $B'_1$  y  $B'_2$  bases de  $V'$ , con cambio de  $B'_1$  a  $B'_2$  dado por  $Y_1 = QY_2$ .

Consideremos la matriz asociada a  $f$  respecto de  $B_1$  y  $B'_1$ ,  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ , tal que  $A = \mathcal{M}_{B_1, B'_1}(f)$  y la ecuación matricial

$$Y_1 = AX_1$$

De igual forma, sea  $\tilde{A} \in \mathcal{M}_{m \times n}(\mathbb{K})$  la matriz asociada a  $f$  respecto de  $B_2$  y  $B'_2$ , tal que  $\tilde{A} = \mathcal{M}_{B_2, B'_2}(\mathbb{K})$  y la ecuación matricial de  $f$  respecto de estas bases,

$$Y_2 = \tilde{A}X_2$$

Gráficamente,

$$\begin{array}{ccc} V & \rightarrow & V' \\ & A & \\ B_1 & \longrightarrow & B'_1 \\ P & \uparrow & \uparrow Q \\ & \tilde{A} & \\ B_2 & \longrightarrow & B'_2 \end{array}$$

Entonces,

$$Y_2 = \begin{cases} \tilde{A}X_2 \\ Q^{-1}Y_1 = Q^{-1}AX_1 = Q^{-1}APX_2 \end{cases}$$

y en consecuente,

$$\tilde{A} = Q^{-1}AP$$

O bien,

$$X_2 = \begin{cases} \tilde{A}^{-1}Y_2 \\ P^{-1}X_1 = P^{-1}A^{-1}Y_1 = P^{-1}A^{-1}QY_2 \end{cases}$$

y en consecuente,

$$\tilde{A}^{-1} = P^{-1}A^{-1}Q$$

■

Ahora vamos a enunciar el **Primer Teorema de Isomorfía**, del que obtendremos un Corolario muy importante a la hora de trabajar con aplicaciones lineales. Este teorema no se va a demostrar (si se quiere ver la prueba consultar [4, Chapter 6, Theorem 6.5, Page 77]).

**Teorema 1.1.7** (Primer teorema de isomorfismo de Noether). *Sea  $f : V \rightarrow V'$  una aplicación lineal, entonces:*

(I) *Existe una aplicación lineal sobreyectiva  $\pi : V \rightarrow V/\text{Ker } f$*

(II) *Existe un isomorfismo  $\bar{f} : V/\text{Ker } f \rightarrow \text{Im } f$*

(III) *Existe una aplicación lineal inyectiva  $i : \text{Im } f \rightarrow V'$ , tales que  $f = i \circ \bar{f} \circ \pi$ , tal que*

$$\begin{array}{ccc} & f & \\ V & \longrightarrow & V' \\ \pi & \downarrow & \uparrow i \\ & \bar{f} & \\ V/\text{Ker } f & \longrightarrow & \text{Im } f \end{array}$$

**Corolario 1.1.6.** *Si además  $V$  es finitamente generado,*

$$\dim V = \dim(\text{Ker } f) + \dim(\text{Im } f)$$

### 1.1.3. Espacios duales

Llegamos a un apartado un tanto especial, pues vamos a definir un nuevo espacio, formado por aplicaciones lineales (cosa que puede impresionar), pero veremos que vamos a poder trabajar de forma muy similar que con los espacios vectoriales. Este es el denominado **espacio dual**.

**Definición 1.1.8.** Al conjunto de todas las formas lineales de un espacio vectorial  $V$ , se le denomina espacio dual de  $V$  y se le designa por  $V^*$ , es decir,

$$V^* = \{f : V \rightarrow \mathbb{K} : \text{lineales}\}$$

Al igual que con los espacios vectoriales, podemos definir la **base dual**, de estos espacios duales y, en consecuencia, podemos determinar la dimensión de este espacio.

**Definición 1.1.9.** Dadas  $B = \{e_1, e_2, \dots, e_n\}$  base de  $V$  y  $B^* = \{f^1, f^2, \dots, f^n\}$  base de  $V^*$ , decimos que  $B^*$  es la base dual de  $B$  si para cada  $i = 1, 2, \dots, n$  se verifica que

$$f^i(e_j) = \delta_i^j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

**Proposición 1.1.7.** Si  $V$  tiene dimensión  $n$ , entonces  $V^*$  es un espacio vectorial de dimensión  $n$ .

**Demostración.** Sea  $B = \{e_1, e_2, \dots, e_n\}$  base de  $V$ , tendremos que  $\dim V = n$ . Construimos la base dual como  $B^* = \{f^1, f^2, \dots, f^n\}$ , tal que para cada  $j$ , tenemos

$$\begin{aligned} f^j : \quad V &\longrightarrow \mathbb{K} \\ v = \sum_{j=1}^n \lambda^j e_j &\mapsto f^j(v) = \lambda_j \end{aligned}$$

Tenemos que verificar que las formas sean linealmente independientes, para ello debe cumplirse que

$$\lambda_1 f^1 + \lambda_2 f^2 + \cdots + \lambda_n f^n = 0 \iff \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$$

Para ello, lo aplicamos a un  $e_i \in B$ , tal que

$$(\lambda_1 f^1 + \lambda_2 f^2 + \cdots + \lambda_n f^n)(e_i) = 0$$

Al ser aplicación lineal,

$$\lambda_1 f^1(e_i) + \lambda_2 f^2(e_i) + \cdots + \lambda_n f^n(e_i) = 0$$

pero sabemos que  $f^j(e_i) = \delta_j^i = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$ , luego la ecuación anterior se reduce a

$$\lambda_i f^i(e_i) = \lambda_i = 0$$

luego, como hemos cogido un  $i$  arbitrario, tendremos que  $\lambda_i = 0$  para todo  $i = 1, 2, \dots, n$ . Por tanto, son linealmente independientes. ✓

Ahora vamos comprobar que los elementos de  $B^*$  son un conjunto generador que genera el espacio  $V^*$ . Para ello, tomamos un elemento  $g \in V^*$  y se deberá poder escribir como combinación lineal de los elementos de la base de  $V^*$ . Sabemos que  $g$  debe satisfacer que  $g(e_j) = \alpha_j$ , pues debe estar determinado por la imagen de la base. Podemos tomar un  $v \in V$ , que sea  $v = \sum_{j=1}^n \lambda^j e_j = \lambda^j e_j$ , así,

$$g(v) = g(\lambda^j e_j) = \lambda^j g(e_j) = \lambda^j \alpha_j$$

siendo  $\alpha_j \in \mathbb{K}$  y usando que  $\lambda^j = f^j(v)$ , tenemos que

$$g(v) = \alpha_j f^j(v)$$

y por tanto,

$$g = \alpha_1 f^1 + \alpha_2 f^2 + \cdots + \alpha_n f^n$$

Luego, hemos concluido que  $B^*$  es también conjunto generador, y por tanto es base de  $V^*$ , que posee  $n$  elementos, así  $\dim V = \dim V^* = n$ . ■

Este será el caso que nos interese, cuando  $V$  y, por tanto,  $V^*$  son espacios vectoriales de dimensión finita. Así, para la matriz asociada a una forma lineal  $f : V \rightarrow \mathbb{K}$  usaremos siempre la base  $\{1\}$  de este espacio y escribiremos  $\mathcal{M}_B(f) \equiv \mathcal{M}_{B,\{1\}}(f)$ .

Veremos ahora varias propiedades de las bases duales.

**Proposición 1.1.8** (1<sup>a</sup> Propiedad de las bases duales). *Si  $B^*$  es la base dual de  $B$ , entonces para cada forma lineal  $f$ , los elementos de su matriz asociada en la base  $B$  coinciden con sus coordenadas en la base  $B^*$ .*

**Demostración.** Llamemos  $A = (a_1 \ a_2 \ \dots \ a_n)$  a la matriz asociada a  $f$  en la base  $B$ . Entonces  $a_i = f(e_i)$ , siendo los  $e_i$  los elementos de la base  $B$ . Por otra parte, si

$f = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix}_{B^*}$  entonces,

$$f = b_1 f^1 + b_2 f^2 + \dots + b_n f^n$$

luego,

$$a_i = f(e_i) = (b_1 f^1 + b_2 f^2 + \dots + b_n f^n)(e_i) = b_1 f^1(e_i) + b_2 f^2(e_i) + \dots + b_i f^i(e_i) + \dots + b_n f^n(e_i)$$

Usando la definición de la base dual, que es

$$f^j(e_i) = \delta_j^i = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

nos queda entonces que  $a_i = b_i$ , luego, los elementos de la matriz asociada de  $f$  en  $B$  se corresponden con sus coordenadas en  $B^*$ . ■

Hemos visto que  $V^*$  es un espacio vectorial con la misma dimensión que el espacio  $V$ , por tanto, es lógico pensar que debería existir una relación entre sus bases, cosa que nos da las siguientes proposiciones.

**Proposición 1.1.9.** *Para cada base  $B$  de un espacio vectorial  $V$ , existe una base de  $V^*$  que es dual de la base  $B$ .*

**Demostración.** Dada la base  $B = \{e_1, e_2, \dots, e_n\}$  de  $V$ , y sabiendo que una forma lineal está completamente determinada conociendo las imágenes de los vectores de  $B$ , por tanto, para cada  $i = 1, 2, \dots, n$  existe una única  $f^i$  verificando

$$f^i(e_i) = 1 \quad \text{y} \quad f^j(e_i) = 0, \quad \forall i \neq j$$

Luego, un conjunto de formas lineales  $\{f^1, f^2, \dots, f^n\}$  verificando las condiciones para ser base dual (lo anterior), siempre existe.

Debemos comprobar que realmente sea una base de  $V^*$ . Como conocemos que la dimensión del espacio es  $n$ , solo necesitamos demostrar que los  $f^i$  son linealmente independientes:

$$\lambda_1 f^1 + \lambda_2 f^2 + \dots + \lambda_n f^n = 0 \iff \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

Evaluamos un  $e_i$  arbitrario, tal que

$$0 = (\lambda_1 f^1 + \lambda_2 f^2 + \dots + \lambda_n f^n)(e_i) = \lambda_1 f^1(e_i) + \lambda_2 f^2(e_i) + \dots + \lambda_i f^i(e_i) + \dots + \lambda_n f^n(e_i) = \lambda_i$$

Luego, son linealmente independientes y por tanto, son base dual. ■

**Nota 1.1.10.** *El siguiente Teorema lo vamos a necesitar para demostrar la Proposición 1.1.11, ver su demostración en [3, Chapter I, Section 4, Theorem 1, Page 38]*

**Teorema 1.1.10.** *Para un matriz cuadrada  $A \in \mathcal{M}_n(\mathbb{K})$ , las siguientes afirmaciones son equivalentes:*

1. *A es invertible.*
2. *A es regular a derecha (esto es,  $BA = 0 \Rightarrow B = 0$ ).*
3. *A es regular a izquierda (esto es,  $AB = 0 \Rightarrow B = 0$ ).*
4.  $rg(A) = n.$
5. *La forma de Hermite por filas de A es la identidad.*
6. *La forma de Hermite por columnas de A es la identidad.*
7. *A es un producto de matrices elementales.*

**Proposición 1.1.11.** *Sea  $V$  un espacio vectorial de dimensión  $n$  y sea  $B$  una base de  $V$  respecto de la cual todos los vectores y formas lineales vienen dados. Si  $B = \{e_1, e_2, \dots, e_n\}$  es una base cuyos vectores, escritos por columnas, forman la matriz  $A$ , entonces la base dual de  $B$ ,  $B^* = \{f^1, f^2, \dots, f^n\}$  viene dada por las filas de  $A^{-1}$  y viceversa.*

**Demostración.** En primer lugar, como  $B$  es base, entonces la matriz  $A$  es regular, es decir, tiene inversa<sup>1</sup>. Si  $\begin{pmatrix} b_{i_1} & b_{i_2} & \dots & b_{i_n} \end{pmatrix}$  es la matriz asociada a  $f^i$  y  $e_j = \begin{pmatrix} a_{j_1} & a_{j_2} & \dots & a_{j_n} \end{pmatrix}_B$ , entonces debe cumplirse que

$$f^i(e_j) = \begin{pmatrix} b_{i_1} & b_{i_2} & \dots & b_{i_n} \end{pmatrix} \cdot \begin{pmatrix} a_{j_1} \\ a_{j_2} \\ \vdots \\ a_{j_n} \end{pmatrix} = \delta_{ij}$$

Es decir, las matrices asociadas a  $f^i$ , con  $i = 1, 2, \dots, n$ , son las filas de  $A^{-1}$ , o lo que es lo mismo, sus coordenadas en la base  $B^*$ . Además, puesto que  $A^{-1}$  es regular, sus filas son linealmente independientes. ■

<sup>1</sup>Cosa que se cumple, pues al ser  $A$  la matriz asociada a una  $f$  en la base  $B$ , que tiene dimensión  $n$ , por tanto  $rg(A) = n$ . Así, usando el siguiente Teorema vemos que  $A$  es regular por el Teorema 1.1.10

**Nota 1.1.12.** Esta proposición nos sirve para calcular el problema inverso, es decir, dada una base de  $V^*$ , calcular la base de  $V$  de la cual es dual.

**Ejemplo 1.1.13** (Ejemplo de las proposiciones 1.3 y 1.4).

(a) **Considerando en  $\mathbb{R}^3$  la base  $B = \{(1, -1, 1), (-1, 2, -1), (-1, 1, 0)\}$ , calcular la base dual de  $B$ .** (b) **Hacerlo también a la inversa.**

(a) Tenemos  $B = \{(1, -1, 1), (-1, 2, -1), (-1, 1, 0)\} \in \mathbb{R}^3$ .  $\mathcal{B}^* = \{f^1, f^2, \dots, f^n\}$ ?

Como debemos obtener 3 formas lineales (componentes de  $B^*$ ), bastará con obtener las matrices asociadas a la base canónica:

- Comenzamos obteniendo la matriz asociada de  $f^1$ :

Por definición de base dual, sabemos que  $f^i(e_j) = \delta_{ij}$

$$\Rightarrow \begin{cases} f^1[(1, -1, 1)] = 1 \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 1 \\ f^2[(-1, 2, -1)] = 0 \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = 0 \\ f^3[(-1, 1, 0)] = 0 \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 0 \end{cases}$$

Entonces, tenemos el siguiente sistema de ecuaciones:

$$\left. \begin{array}{l} a_{11} - a_{12} + a_{13} = 1 \\ -a_{11} + 2a_{12} - a_{13} = 0 \\ -a_{11} + a_{12} + 0 = 0 \end{array} \right\}$$

Resolviendo el sistema tenemos,  $a_{11} = a_{12} = a_{13} = 1$ .

- Ahora calculamos la de  $f^2$ :

$$\Rightarrow \begin{cases} f^1 [(1, -1, 1)] = 0 \Rightarrow \begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 0 \\ f^2 [(-1, 2, -1)] = 1 \Rightarrow \begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = 1 \\ f^3 [(-1, 1, 0)] = 0 \Rightarrow \begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 0 \end{cases}$$

Entonces, tenemos el siguiente sistema de ecuaciones:

$$\left. \begin{array}{l} a_{21} - a_{22} + a_{23} = 0 \\ -a_{21} + 2a_{22} - a_{23} = 1 \\ -a_{21} + a_{22} + 0 = 0 \end{array} \right\}$$

Resolviendo el sistema tenemos  $a_{21} = a_{22} = 1$  y  $a_{23} = 0$ .

- Ahora calculamos la de  $f^3$ :

$$\Rightarrow \begin{cases} f^1 [(1, -1, 1)] = 0 \Rightarrow \begin{pmatrix} a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 0 \\ f^2 [(-1, 2, -1)] = 0 \Rightarrow \begin{pmatrix} a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = 0 \\ f^3 [(-1, 1, 0)] = 1 \Rightarrow \begin{pmatrix} a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 1 \end{cases}$$

Entonces, tenemos el siguiente sistema de ecuaciones:

$$\left. \begin{array}{l} a_{31} - a_{32} + a_{33} = 0 \\ -a_{31} + 2a_{32} - a_{33} = 0 \\ -a_{31} + a_{32} + 0 = 1 \end{array} \right\}$$

Resolviendo el sistema tenemos  $a_{31} = -1$ ,  $a_{32} = 0$  y  $a_{33} = 1$ .

Por tanto la matriz asociada en la base canónica será,

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

donde la primera fila será la matriz asociada a  $f^1 \sim \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$ , la segunda fila será la matriz asociada a  $f^2 \sim \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$  y la tercera fila será la matriz asociada a  $f^3 \sim \begin{pmatrix} -1 & 0 & 1 \end{pmatrix}$ .

Entonces tendremos:

$$f^1, f^2, f^3 : \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$\left. \begin{array}{l} f^1(x, y, z) = x + y + z \\ f^2(x, y, z) = x + y \\ f^3(x, y, z) = -x + z \end{array} \right\} (\Delta)$$

(b) Partiendo de  $(\Delta)$ , vamos a llegar a la base  $B$ .

Sabemos que  $\{f^1, f^2, f^3\}$  forman una base dual de  $\mathbb{R}^3$ , pues lo acabamos de obtener, pero como queremos hacerlo como un ejercicio genérico, debemos comprobar que son linealmente independientes, ya que  $\dim(\mathbb{R}^3) = 3$ . La forma más sencilla de ver que son l.i. es introducir los vectores en una matriz y calcular su rango.

Usando la primera propiedad de las bases duales, la matriz asociada a  $f^i$  en la base canónica nos proporciona también las coordenadas de  $f^i$  en la base dual de la base canónica. Así, las matrices asociadas a  $f^1, f^2, f^3$  son:

$$f^1 \sim \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \quad f^2 \sim \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}, \quad f^3 \sim \begin{pmatrix} -1 & 0 & 1 \end{pmatrix}$$

Así, ver que son l.i. se reduce al cálculo del determinante,

$$|A| = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = 1 + 0 + 0 + 1 - 1 - 0 = 1 \neq 0$$

por tanto, son linealmente independiente.

Para encontrar la base de  $\mathbb{R}^3$  de la cual es dual  $\{f^1, f^2, f^3\}$ , solo tenemos que calcular la inversa de la matriz cuyo determinante acabamos de calcular:

$$A^{-1} = \frac{\text{Adj}(A)^t}{|A|^1}, \quad \text{Adj}(A) = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \quad \text{Adj}(A)^t = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix} = A^{-1}$$

Por tanto, los vectores de la base de  $\mathbb{R}^3$  de la cual es dual  $\{f^1, f^2, f^3\}$ , serán las columnas de la matriz  $A^{-1}$ , tal que

$$B = \{(1, -1, 1), (-1, 2, -1), -1, 1, 0\}$$

Sigamos viendo más propiedades de las bases duales.

**Proposición 1.1.14** (2<sup>a</sup> Propiedad de las bases duales). *Si  $B^* = \{f^1, \dots, f^n\}$  es la base dual de  $B$ , entonces dado un vector  $x \in V$ , si  $x = (x_1, x_2, \dots, x_n)_B$ , se verifica que  $x^i = f^i(x); \forall i = 1, \dots, n$ .*

**Demostración.** Como  $x \in V$  y  $B = \{e_1, e_2, \dots, e_n\}$ , base de  $V$ , entonces podremos escribir  $x = x^1e_1 + x^2e_2 + \dots x^ne_n$ , luego

$$f^i(x) = f^i(x^1e_1 + x^2e_2 + \dots x^ne_n) = x^1f^i(e_1) + x^2f^i(e_2) + \dots + x^if^i(e_i) + \dots + x^nf^i(e_n) = x^i$$

Por tanto,  $x^i = f^i(x)$ . ■

**Teorema 1.1.11.**  *$V$  es isomorfo con  $(V^*)^*$ , siendo el espacio bidual ( $V \cong (V^*)^*$ ).*

**Demostración.** Sea

$$\begin{aligned} \phi : \quad V &\rightarrow (V^*)^* \\ v &\mapsto \phi(v) \end{aligned}$$

donde  $\phi(v) \in (V^*)^*$ . Se define como

$$\begin{aligned}\phi(v) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto f(v)\end{aligned}$$

Veamos que  $\phi$  es lineal,  $\forall u, v \in V$ ;

$$\begin{aligned}\phi(u) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto f(u);\end{aligned}\quad \begin{aligned}\phi(v) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto f(v)\end{aligned}$$

$$\begin{aligned}\phi(u + v) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto f(u + v);\end{aligned}\quad \begin{aligned}\phi(v) + \phi(u) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto f(v) + f(u)\end{aligned}$$

Como  $f \in V^*$  es lineal, entonces se cumple que  $f(u + v) = f(u) + f(v)$  y por tanto,  $\phi(u + v) = \phi(u) + \phi(v)$ . También,

$$\forall u \in V \quad \forall \lambda \in \mathbb{K}; \quad \begin{aligned}\phi(\lambda \cdot u) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto f(\lambda \cdot u);\end{aligned}\quad \begin{aligned}\lambda \cdot \phi(u) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto \lambda \cdot f(u)\end{aligned}$$

como  $f \in V^*$  es lineal, se cumple que  $f(\lambda \cdot u) = \lambda \cdot f(u)$  y por tanto,  $\phi(\lambda \cdot u) = \lambda \cdot \phi(u)$ . Luego,  $\phi$  es lineal.

Veamos que  $\text{Ker } \phi = \{0\}$ :

si  $u \in \text{ker } \phi$ , entonces  $\phi(u) = 0$ , tal que

$$\begin{aligned}\phi(u) : & V^* \rightarrow \mathbb{K} \\ & f \mapsto 0\end{aligned}$$

Sea  $\{e_1, e_2, \dots, e_n\}$  base de  $V$ , sabemos que

$$\begin{aligned}\phi_{e_j} : & V \rightarrow \mathbb{K} \\ & v \mapsto \phi_{e_j}(v) = x_j\end{aligned}$$

tal que  $v = x^1 e_1 + x^2 e_2 + \dots + x^n e_n$ , se verifica que  $\phi_{e_j} \in V^*$ . Por tanto,  $\phi(u)(\phi_{e_j}) = 0$ , pues  $u \in \text{Ker } \phi$  y  $\phi(u)$  es la función nula, tal que

$$u = x^1 e_1 + \dots + x^n e_n = 0e_1 + \dots + 0e_n = 0$$

Por tanto,  $\text{Ker } \phi = \{0\}$ , luego  $\phi : V \rightarrow (V^*)^*$  es inyectiva y lineal.

Por la Proposición 1.1.7, tenemos que  $\dim V = \dim V^* = n$  y por tanto,  $\dim(V^*)^* = n =$

$\dim V^*$ . Luego,  $\phi : V \rightarrow V^*$  verifica

$$\dim \text{Im } \phi = \dim V - \dim \text{Ker } \phi = n - 0 = n \Rightarrow \dim \text{Im } \phi = \dim (V^*)^* = n \Rightarrow \text{Im } \phi = (V^*)^*$$

luego, la aplicación es sobreyectiva y por tanto, es biyectiva. Luego, al ser biyectiva y aplicación lineal, son isomorfos. ■

### 1.1.3.1. Anulador de un subespacio

El **anulador de un subespacio** nos permite, como su propio nombre indica, anular subespacios.

**Definición 1.1.12.** Consideremos un espacio vectorial  $V$  y sea  $S$  un subconjunto de  $V$ , entonces se define el anulador de  $S$  como el conjunto

$$an(S) = \{f \in V^* \mid f(v) = 0, \forall v \in S\}$$

**Proposición 1.1.15.** Sea  $V$  un  $\mathbb{K}$ -espacio vectorial y sea  $S$  un subconjunto de  $V$ , entonces se verifica:

(I)  $an(S)$  es un subespacio vectorial de  $V^*$ .

(II)  $an(S) = an(L(S))$

**Demostración.** (i) Dadas dos formas de  $an(S)$ , digamos  $f, g \in an(S)$ , queremos probar que  $f + g \in an(S)$ , y sea  $\lambda \in \mathbb{K}$ , probar que  $\lambda f \in an(S)$ . Para lo primero, hacemos  $\forall v \in V$ ,

$$(f + g)(v) = f(v) + g(v) = 0 + 0 = 0 \Rightarrow f + g \in an(S)$$

Para lo segundo hacemos,

$$(\lambda f)(v) = \lambda f(v) = \lambda \cdot 0 = 0 \Rightarrow \lambda f \in an(S)$$

Por tanto, es subespacio vectorial de  $V^*$ . ■

(ii)  $\subseteq$  Sabemos que  $S \subseteq L(S)$ , entonces si  $f \in an(L(S))$  anula a todo vector de  $L(S)$ , en particular, anula a todo vector de  $S$  y así se tiene la inclusión  $an(L(S)) \subseteq an(S)$ .

$\supseteq$  Considerando  $g \in an(S)$  y cualquier vector de  $L(S)$  de la forma,

$$v = a^1 s_1 + a^2 s_2 + \dots + a^r s_r$$

con  $s_i \in S$  para cada  $i = 1, 2, \dots, r$ . Entonces,

$$f(v) = f(a^1 s_1 + a^2 s_2 + \dots + a^r s_r) = a^1 \cancel{f(s_1)}^0 + a^2 \cancel{f(s_2)}^0 + \dots + a^r \cancel{f(s_r)}^0 = 0$$

y por tanto,  $f \in an(S)$ , luego  $an(S) \subseteq an(L(S))$ . Luego,  $an(S) = an(L(S))$ . ■

**Nota 1.1.16.** Cuando consideramos un subespacio  $U \leq V$ , su anulador  $an(U)$  puede calcularse fácilmente usando la segunda de las propiedades anteriores.

**Nota 1.1.17.** En general, si llamamos  $n = \dim V$  y  $U$  es un subespacio de  $V$  con una base  $\{u_1, u_2, \dots, u_r\}$ , entonces  $an(U) = an(\{u_1, \dots, u_r\})$  (por la Proposición 1.1.15), lo que permite escribir las  $r$ -ecuaciones cartesianas de  $an(U)$ . Además, como  $\{u_1, \dots, u_r\}$  son l.i., entonces  $\dim(an(U)) = n - r$ , es decir,

$$\dim(an(U)) = \dim V - \dim U$$

### 1.1.3.2. Aplicación lineal traspuesta

Dados dos espacios vectoriales  $V$  y  $V'$  sobre el cuerpo  $\mathbb{K}$ , es posible considerar los respectivos espacios duales, y si  $f : V \rightarrow V'$  es una aplicación lineal con matriz asociada  $A$  respecto de ciertas bases  $B$  y  $B'$ , entonces se puede definir una aplicación lineal entre los duales mediante  $f$ . Para ello, tomemos  $\varphi \in (V')^*$ , es decir,  $\varphi : V' \rightarrow \mathbb{K}$ , entonces podemos considerar el diagrama:

$$\begin{array}{ccc} & f & \\ V & \rightarrow & V' \\ & \downarrow & \varphi \\ & & \mathbb{K} \end{array}$$

y la composición  $\varphi \circ f : V \rightarrow \mathbb{K}$  es un elemento de  $V^*$ . De esta manera, tenemos definida una aplicación a la que llamamos  $f^t$ , aplicación traspuesta de  $f$ , tal que  $f^t : (V')^* \rightarrow V^*$  dada por  $f^t(\varphi) = \varphi \circ f$ .

**Proposición 1.1.18.** Sea  $f : V \rightarrow V'$  una aplicación lineal y sean  $B$  y  $B^*$  bases de  $V$  y  $V'$  respectivamente, entonces se verifica:

(I)  $f^t$  es una aplicación lineal.

(II) Si la matriz asociada a  $f$  respecto de las bases  $B$  y  $B'$  es  $A$ , entonces la matriz asociada a  $f^t$  en las bases  $B^*$  y  $(B')^*$  es  $A^t$ .

**Demostración.** (i) Si  $\varphi_1, \varphi_2 \in (V')^*$  y  $a_1, a_2 \in \mathbb{K}$ , entonces

$$f^t(a_1\varphi_1 + a_2\varphi_2) = (a_1\varphi_1 + a_2\varphi_2) \circ f$$

Dado  $v \in V$ , tenemos

$$\begin{aligned} ((a_1\varphi_1 + a_2\varphi_2) \circ f)(v) &= (a_1\varphi_1 + a_2\varphi_2)(f(v)) = a_1\varphi_1(f(v)) + a_2\varphi_2(f(v)) = \\ &= a_1(\varphi_1 \circ f)(v) + a_2(\varphi_2 \circ f)(v) = (a_1(\varphi_1 \circ f) + a_2(\varphi_2 \circ f))(v) = (a_1f^t(\varphi_1) + a_2f^t(\varphi_2))(v) \end{aligned}$$

con lo que

$$f^t(a_1\varphi_1 + a_2\varphi_2) = a_1f^t(\varphi_1) + a_2f^t(\varphi_2)$$

luego, es aplicación lineal.

(ii) Calculamos las imágenes de las formas lineales de  $(B')^*$  por  $f^t$ : si  $\varphi_i$  es el  $i$ -ésimo elemento de  $(B')^*$ , su matriz asociada en la base  $B'$  es  $\begin{pmatrix} 0 & 0 & \dots & 1^{(i)} & \dots & 0 \end{pmatrix}$ , donde el 1 está en el  $i$ -ésimo lugar.

Entonces,  $f^t(\varphi_i) = \varphi_i \circ f$  tiene como matriz asociada en la base  $B$  el producto siguiente,

$$\begin{pmatrix} 0 & 0 & \dots & 1^{(i)} & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$$

que es la  $i$ -ésima fila de  $A$ . Nos da las coordenadas de  $f^t(\varphi_i)$  en la base  $B^*$ ; estas coordenadas constituyen a la  $i$ -ésima columna de la matriz asociada a  $f^t$ . ■

### 1.1.3.3. Una aplicación de la teoría del espacio dual: Interpretación de Lagrange

La interpretación de Lagrange es una de las tantas aplicaciones que tiene la teoría del espacio dual.

Sea  $a \in \mathbb{R}$ , llamamos **evaluar un polinomio**  $q(x)$  en  $a$ , al proceso de sustituir la indeterminada  $x$  por el valor real  $a$  en el polinomio  $p(x)$ ; al número real obtenido de esta forma lo denotamos por  $p(a)$ .

**Proposición 1.1.19.** *La aplicación evaluar en  $a$ , que denotamos por  $E_a$ , es una forma lineal en  $\mathbb{P}(\mathbb{R})$ .*

**Demostración.** Consideramos dos polinomios,

$$p(x) = a_0 + a_1x + \cdots + a_nx^n$$

$$q(x) = b_0 + b_1x + \cdots + b_nx^n$$

y  $\lambda_1, \lambda_2 \in \mathbb{R}$ . Entonces,

$$\begin{aligned} E_a(\lambda_1 p(x) + \lambda_2 q(x)) &= E_a((\lambda_1 a_0 + \lambda_2 b_0) + (\lambda_1 a_1 + \lambda_2 b_1)x + \cdots + (\lambda_1 a_n + \lambda_2 b_n)x^n) = \\ &= \lambda_1 a_0 + \lambda_2 b_0 + (\lambda_1 a_1 + \lambda_2 b_1)a + \cdots + (\lambda_1 a_n + \lambda_2 b_n)a^n \end{aligned}$$

y también,

$$\begin{aligned} \lambda_1 E_a(p(x)) + \lambda_2 E_a(q(x)) &= \lambda_1 E_a(a_0 + a_1x + \cdots + a_nx^n) + \lambda_2 E_a(b_0 + b_1x + \cdots + b_nx^n) = \\ &= \lambda_1(a_0 + a_1a + \cdots + a_na^n) + \lambda_2(b_0 + b_1a + \cdots + b_na^n) = \\ &= \lambda_1 a_0 + \lambda_2 b_0 + (\lambda_1 a_1 + \lambda_2 b_1)a + \cdots + (\lambda_1 a_n + \lambda_2 b_n)a^n \end{aligned}$$

Luego,  $E_a(\lambda_1 p(x) + \lambda_2 q(x)) = \lambda_1 E_a(p(x)) + \lambda_2 E_a(q(x))$ , y por tanto, es una aplicación lineal. ■

**Nota 1.1.20.** La matriz asociada a  $E_a$  en la base  $\{1, x, x^2, \dots, x^n\}$  de  $\mathbb{P}(\mathbb{R})$ , como  $E_a(1) = 1$ ,  $E_a(x) = a$ ,  $E_a(x^2) = a^2$ , ...,  $E_a(x^n) = a^n$ , y por tanto, la matriz asociada de  $E_a$  es  $A = \begin{pmatrix} 1 & x & x^2 & \dots & x^n \end{pmatrix}_{\mathbb{P}(\mathbb{R})}$ .

**Proposición 1.1.21.** Si  $a_0, a_1, \dots, a_n \in \mathbb{R}$  con  $a_i \neq a_j$ , entonces  $\{E_{a_0}, E_{a_1}, \dots, E_{a_n}\}$  es una base del espacio dual de  $\mathbb{P}(\mathbb{R})$ .

**Demostración.** Llamemos  $B^*$  a la base dual de la base  $B = \{1, x, x^2, \dots, x^n\}$  de  $\mathbb{P}(\mathbb{R})$ . Entonces las coordenadas de  $E_{a_0}, E_{a_1}, \dots, E_{a_n}$  en la base  $B^*$  son

$$(1, a_0, a_0^2, \dots, a_0^n), (1, a_1, a_1^2, \dots, a_1^n), \dots, (1, a_n, a_n^2, \dots, a_n^n)$$

respectivamente. Veamos que son linealmente independientes; para ello, vemos que el determinante e la matriz es no nulo:

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ a_0 & a_1 & \dots & a_n \\ a_0^2 & a_1^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_0^n & a_1^n & \dots & a_n^n \end{vmatrix} = \prod_{i>j} (a_i - a_j)$$

siendo este el Determinante de Vandermonde.

Como  $a_i \neq a_j$  para  $i \neq j$ , este determinante nunca se anula, luego son linealmente independientes, y por tanto, son una base. ■

**Proposición 1.1.22.** *Los polinomios*

$$p_j(x) = \prod_{i \neq j} \frac{x - a_j}{a_j - a_i}; \quad j = 0, 1, 2, \dots, n$$

forman la base de  $\mathbb{P}(\mathbb{R})$  de la cual es dual la formada por  $E_{a_0}, E_{a_1}, \dots, E_{a_n}$ .

**Demostración.** Si evaluamos el polinomio  $p_j(x) = \prod_{i \neq j} \frac{x - a_j}{a_i - a_j}$  en  $a_i$  con  $i \neq j$ , como en el producto aparece el término  $(a_i - a_i) = 0$ , obtenemos  $p_j(a_i) = 0$ . Al evaluar en  $a_j$ , el numerador y el denominador de  $p_j(a_j)$  son idénticos, y se tiene  $\frac{a_j - a_i}{a_j - a_i} = 1$ , luego  $p_j(a_j) = 1$ . Así,  $E_{a_i}(p_j) = \delta_{ij}$ , que es la condición de base dual. ■

Los polinomios descritos en la Proposición 1.1.22, reciben el nombre de *polinomios de interpolación de Lagrange*.

**Teorema 1.1.13.** *Dados  $n + 1$ -números reales distintos  $a_0, a_1, \dots, a_n$ , para cualesquiera  $b_0, b_1, \dots, b_n \in \mathbb{R}$ , existe un único polinomio  $p(x)$  de grado menor o igual que  $n$ , de forma que  $p(a_i) = b_i$ , para cada  $i = 0, 1, 2, \dots, n$ . Dicho polinomio viene dado por*

$$p(x) = \sum_{j=0}^N b_j p_j(x)$$

siendo  $p_0(x), p_1(x), \dots, p_n(x)$  los polinomios de interpolación de Lagrange.

**Demostración.** Si  $p(x) = \sum_{i=0}^N b_i p_i(x)$  con  $p_j(x) = \prod_{i \neq j} \frac{x - a_i}{a_j - a_i}$ . Vamos a ver que  $p(x)$  es único:

Tomamos un polinomio de grado  $n$ ,  $q(x)$ , que satisface lo mismo que  $p(x)$ , tal que  $q(a_i) = b_i$ . Definimos otro polinomio de grado  $n$ , tal que

$$r(x) = p(x) - q(x), \quad \text{con} \quad \begin{cases} r(a_i) = 0 & i = 0, 1, \dots, n \\ \text{Como } r(a_i) = 0, & (x - a_i) | r(x) \end{cases}$$

luego,

$$r(x) = (x - a_i)s(x) = \Lambda(x)(x - a_0)(x - a_1) \dots (x - a_n)$$

pero vemos que el grado de  $r(x)$  es mayor o igual que  $n + 1$ , que es una contradicción, salvo que  $\Lambda(x) = 0$  y por tanto  $q(x) = p(x)$ , luego, el polinomio es único. Ahora, vamos a demostrar que este polinomio satisface la igualdad:

$$\begin{aligned} p(a_0) &= \sum_{i=0}^N b_i p_i(a_0) = \sum_{i=0}^N b_i \delta_0^i = b_0 \\ p(a_j) &= \sum_{i=0}^N b_i p_i(a_j) = \sum_{i=0}^N b_i \delta_j^i = b_j \end{aligned}$$

Luego, queda demostrado. ■

#### 1.1.4. Notación de Einstein

La notación de Einstein va a servir para facilitarnos la escritura, pues cada vez que tengamos un vector o una forma escrita como combinación lineal, vamos a poder redefinirlos como

$$w = \sum_{i=1}^n \lambda^i v_i \equiv \lambda^i v_i$$

esto para un vector. Para una forma, tendremos

$$p = \sum_{i=1}^n \mu_i f^i \equiv \mu_i f^i$$

Además, para simplificar aún más la notación y dejarnos de tantas letras, vamos a identificar los escalares de  $w$  como

$$\lambda^i \equiv w^i$$

Así, los vectores como combinación lineal de otros vectores, los escribiremos como

$$w = w^i v_i$$

Y para las formas, haremos la identificación

$$\mu_i \equiv p_i$$

Así, las formas como combinación lineal de otras formas se escribirán como

$$p = p_i f^i$$

**Ejemplo 1.1.23.** Un ejemplo de ello, será a la hora de identificar un vector en los términos de su base, pues suponiendo un  $V$  espacio vectorial sobre el cuerpo  $\mathbb{K}$  y cuya base sea  $B = \{v_1, v_2, \dots, v_n\}$ , tomando un  $u \in V$ , lo denominaremos como,

$$u = u^i v_i$$

**Ejemplo 1.1.24.** Otro ejemplo será a la hora de identificar una forma en términos de la base dual, pues suponiendo un  $V^*$  espacio dual de  $V$ , cuya base dual es  $B^* = \{f^1, f^2, \dots, f^n\}$ , tomando un  $q \in V^*$ , lo denominaremos como,

$$q = q_i f^i$$

**Nota 1.1.25.** En un artículo físico, se identifica directamente el escalar con el vector, es decir,

$$w^i \equiv w$$

pues se presupone que existe una base donde  $w$  está bien definido. Así, los físicos usaremos de forma indistinguible los vectores y sus componentes respecto de una base fijada.

### 1.1.5. Formas bilineales, productos escalares y formas cuadráticas

A continuación, vamos a estudiar las **formas bilineales**, el **producto escalar** y las **formas cuadráticas**. Cosa que será de especial interés a la hora de definir métricas y aplicaciones lineales más generales.

### 1.1.5.1. Formas bilineales y producto escalar

Comenzamos definiendo las aplicaciones bilineales, pues el producto escalar es una aplicación bilineal.

**Definición 1.1.14.** *Sea  $\mathbb{K}$  un cuerpo y  $V$  un  $\mathbb{K}$ -espacio vectorial, una forma bilineal en  $V$  es una aplicación*

$$f : V \times V \rightarrow \mathbb{K}$$

verificando:

$$(I) \quad f(u_1 + u_2, v) = f(u_1, v) + f(u_2, v)$$

$$(II) \quad f(u, v_1 + v_2) = f(u, v_1) + f(u, v_2)$$

$$(III) \quad f(\lambda u, v) = \lambda f(u, v)$$

$$(IV) \quad f(u, \lambda v) = \lambda f(u, v)$$

para cualesquiera  $\lambda \in \mathbb{K}$  y  $u, v, u_1, u_2, v_1, v_2 \in V$ .

**Proposición 1.1.26.** *Una aplicación  $f : V \times V \rightarrow \mathbb{K}$  es una forma bilineal si y solo si,*

$$(I) \quad f(\lambda u_1 + \mu u_2, v) = \lambda f(u_1, v) + \mu f(u_2, v)$$

$$(II) \quad f(u, \lambda v_1 + \mu v_2) = \lambda f(u, v_1) + \mu f(u, v_2)$$

para cualesquiera  $\lambda, \mu \in \mathbb{K}$  y  $u, v, u_1, u_2, v_1, v_2 \in V$ .

**Demostración.** Supongamos que  $f$  es bilineal, entonces

$$f(\lambda u_1 + \mu u_2, v) = f(\lambda u_1, v) + f(\mu u_2, v) = \lambda f(u_1, v) + \mu f(u_2, v) \checkmark$$

$$f(u, \lambda v_1 + \mu v_2) = f(u, \lambda v_1) + f(u, \mu v_2) = \lambda f(u, v_1) + \mu f(u, v_2) \checkmark$$

Recíprocamente, suponiendo que se verifican estas condiciones, entonces para  $\lambda = 1, \mu = 1$  se obtienen (i) y (ii), y con  $\mu = 0$ , obtenemos (iii) y (iv). ■

Veamos algunas propiedades de las aplicaciones bilineales.

**Proposición 1.1.27** (Propiedades). *Sea  $f : V \times V \rightarrow \mathbb{K}$  una forma bilineal, entonces se verifica:*

$$(I) \ f(u, 0) = f(0, u) = 0; \forall u, v \in V.$$

$$(II) \ f(-u, v) = f(u, -v) = -f(u, v); \forall u, v \in V$$

$$(III) \ f\left(\sum_i a^i u_i, \sum_j b^j v_j\right) = f(a^i u_i, b^j v_j) = a^i b^j f(u_i, v_j); \forall a^i, b^j \in \mathbb{K}, \forall u_i, v_j \in V.$$

**Demostración.**

$$(I) \ \text{Sea } 0 = v - v, \forall v \in V,$$

$$f(u, 0) = f(u, v - v) \stackrel{(ii), (iv)}{=} f(u, v) - f(u, v) = 0$$

$$f(0, u) = f(v - v, u) \stackrel{(i), (iii)}{=} f(v, u) - f(v, u) = 0$$

$$(II) \ \text{Sea } -v = (-1) \cdot v, (-1) \in \mathbb{K}, \forall u, v \in V,$$

$$f(-u, v) \stackrel{(iii)}{=} -f(u, v)$$

$$f(u, -v) \stackrel{(iv)}{=} -f(u, v)$$

$$(III) \ \text{Sean } \forall v_j, u_i \in V \text{ y } \forall a_i, b_j \in \mathbb{K},$$

$$f(a^i u_i, b^j v_j) \stackrel{(i), (ii)}{=} f(a^i u_i, b^j v_j) \stackrel{(iii), (iv)}{=} a^i b^j f(u_i, v_j)$$

donde hemos usado las condiciones de la definición de forma bilineal. ■

Veamos ahora una definición del **producto escalar**.

**Definición 1.1.15.** *Dado un espacio vectorial real  $V$ , definimos el producto escalar como una aplicación*

$$\langle , \rangle : V \times V \rightarrow \mathbb{R}$$

verificando,

$$(I) \ \langle u, v \rangle = \langle v, u \rangle, \forall u, v \in V$$

$$(II) \ \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle, \forall u, v, w \in V$$

$$(III) \ \langle \lambda u, v \rangle = \lambda \langle u, v \rangle, \forall u, v \in V, \forall \lambda \in \mathbb{R}$$

$$(IV) \ No \ degenerada, \ es \ decir, \ \nexists v \neq 0, v \in V, \ tal \ que \ \langle v, w \rangle = 0, \forall w \neq 0 \in V$$

**Proposición 1.1.28.** *El producto escalar es una forma bilineal.*

**Demostración.** Para ver que el producto escalar sea una forma bilineal, debe satisfacer:

$$\langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle$$

$$\langle u, v_1 + v_2 \rangle = \langle u, v_1 \rangle + \langle u, v_2 \rangle$$

$$\langle \lambda \cdot u, v \rangle = \lambda \cdot \langle u, v \rangle$$

$$\langle u, \lambda \cdot v \rangle = \lambda \cdot \langle u, v \rangle$$

pero es trivial ver que las dos primeras condiciones se satisfacen por (ii), y las dos últimas, por (iii), que vienen de la definición de producto escalar, luego es una forma bilineal. ■

Definimos ahora dos tipos particulares de formas bilineales, las **simétricas** y las **antisimétricas**.

**Definición 1.1.16.** *Una forma bilineal  $f : V \times V \rightarrow \mathbb{K}$  es simétrica si verifica:*

$$f(x, y) = f(y, x); \quad \forall x, y \in V$$

**Definición 1.1.17.** *Una forma bilineal  $f : V \times V \rightarrow \mathbb{K}$  es antisimétrica si verifica:*

$$f(x, y) = -f(y, x); \quad \forall x, y \in V$$

**Proposición 1.1.29.** *El producto escalar es una forma bilineal simétrica.*

**Demostración.** Por la definición de producto escalar, la condición (i) nos dice que es una forma simétrica. ■

**Lema 1.1.30.** *Sea  $V$  un  $\mathbb{K}$ -espacio vectorial y sea  $f : V \times V \rightarrow \mathbb{K}$  una forma bilineal. Si  $\mathbb{K} = \mathbb{R}$  ó  $\mathbb{K} = \mathbb{C}$ , entonces se verifica:*

$$f \text{ es antisimétrica} \Leftrightarrow f(x, x) = 0, \text{ para cada } x \in V$$

**Demostración.** Si  $f$  es antisimétrica, entonces  $f(x, x) = -f(x, x)$ , luego, sumando  $f(x, x)$  a ambos lados, tenemos  $2f(x, x) = 0$ , con  $2 \in \mathbb{C}$ , luego  $f(x, x) = 0$ .

Si  $f(x, x) = 0, \forall x \in V$ , tal que

$$0 = \underline{f(u+v, u+v)}^0 = \cancel{f(u, u)}^0 + f(u, v) + f(v, u) + \cancel{f(v, v)}^0$$

luego,  $f(u, v) + f(v, u) = 0$ , y entonces  $f(u, v) = -f(v, u)$ , luego es antisimétrica. ■

Ahora veremos que cualquier forma bilineal puede descomponerse en una forma **simétrica** y otra **antisimétrica**, cosa que también pasará con los tensores.

**Proposición 1.1.31.** *Toda forma bilineal puede descomponerse como suma de una forma bilineal simétrica y una antisimétrica.*

**Demostración.** Sea una forma bilineal  $f : V \times V \rightarrow \mathbb{K}$ , consideramos las aplicaciones  $f_S, f_T : V \times V \rightarrow \mathbb{K}$  definidas por

$$f_S(x, y) = \frac{1}{2} [f(x, y) + f(y, x)]$$

$$f_T(x, y) = \frac{1}{2} [f(x, y) - f(y, x)]$$

Vamos a ver que son formas bilineales:

Empezamos por  $f_S$ ,

$$f_S(x_1 + x_2, y) = \frac{1}{2} [f(x_1 + x_2, y) + f(y, x_1 + x_2)] = \frac{1}{2} [f(x_1, y) + f(x_2, y) + f(y, x_1) + f(y, x_2)]$$

$$= \frac{1}{2} [f(x_1, y) + f(y, x_1)] + \frac{1}{2} [f(x_2, y) + f(y, x_2)] = f_S(x_1, y) + f_S(x_2, y) \checkmark$$

$$f_S(x, y_1 + y_2) = \frac{1}{2} [f(x, y_1 + y_2) + f(y_1 + y_2, x)] = \frac{1}{2} [f(x, y_1) + f(x, y_2) + f(y_1, x) + f(y_2, x)]$$

$$= \frac{1}{2} [f(x, y_1) + f(y_1, x)] + \frac{1}{2} [f(x, y_2) + f(y_2, x)] = f_S(x, y_1) + f_S(x, y_2) \checkmark$$

$$\begin{aligned} f_S(\lambda \cdot x, y) &= \frac{1}{2} [f(\lambda \cdot x, y) + f(y, \lambda \cdot x)] = \frac{1}{2} [\lambda \cdot f(x, y) + \lambda f(y, x)] = \\ &= \lambda \frac{1}{2} [f(x, y) + f(y, x)] = \lambda f_S(x, y) \checkmark \end{aligned}$$

$$\begin{aligned} f_S(x, \lambda \cdot y) &= \frac{1}{2} [f(x, \lambda \cdot y) + f(\lambda \cdot y, x)] = \frac{1}{2} [\lambda \cdot f(x, y) + \lambda f(y, x)] \\ &= \lambda \frac{1}{2} [f(x, y) + f(y, x)] = \lambda f_S(x, y) \checkmark \end{aligned}$$

Veamos que  $f_S$  es simétrica:

$$f_S(x, y) = \frac{1}{2} [f(x, y) + f(y, x)] = \frac{1}{2} [f(y, x) + f(x, y)] = f(y, x) \checkmark$$

Seguimos con  $f_T$ ,

$$f_T(x_1 + x_2, y) = \frac{1}{2} [f(x_1 + x_2, y) - f(y, x_1 + x_2)] = \frac{1}{2} [f(x_1, y) + f(x_2, y) - f(y, x_1) - f(y, x_2)] =$$

$$\begin{aligned}
 &= \frac{1}{2} [f(x_1, y) - f(y, x_1)] + \frac{1}{2} [f(x_2, y) - f(y, x_2)] = f_T(x_1, y) + f_T(x_2, y) \checkmark \\
 f_T(x, y_1+y_2) &= \frac{1}{2} [f(x, y_1 + y_2) - f(y_1 + y_2, x)] = \frac{1}{2} [f(x, y_1) + f(x, y_2) - f(y_1, x) - f(y_2, x)] = \\
 &= \frac{1}{2} [f(x, y_1) - f(y_1, x)] + \frac{1}{2} [f(x, y_2) - f(y_2, x)] = f_T(x, y_1) + f_T(x, y_2) \checkmark \\
 f_T(\lambda \cdot x, y) &= \frac{1}{2} [f(\lambda \cdot x, y) - f(y, \lambda \cdot x)] = \frac{1}{2} [\lambda \cdot f(x, y) - \lambda f(y, x)] \\
 &= \lambda \frac{1}{2} [f(x, y) - f(y, x)] = \lambda f_T(x, y) \checkmark \\
 f_T(x, \lambda \cdot y) &= \frac{1}{2} [f(x, \lambda \cdot y) - f(\lambda \cdot y, x)] = \frac{1}{2} [\lambda \cdot f(x, y) - \lambda f(y, x)] \\
 &= \lambda \frac{1}{2} [f(x, y) - f(y, x)] = \lambda f_T(x, y) \checkmark
 \end{aligned}$$

Veamos que  $f_T$  es antisimétrica:

$$f_T(x, y) = \frac{1}{2} [f(x, y) - f(y, x)] = \frac{1}{2} [-f(y, x) + f(x, y)] = -\frac{1}{2} [f(y, x) - f(x, y)] = -f(y, x) \checkmark$$

Para cada par de vectores  $x, y \in V$  tenemos:

$$f_S(x, y) + f_T(x, y) = \frac{1}{2} [f(x, y) + f(y, x)] + \frac{1}{2} [f(x, y) - f(y, x)] = f(x, y)$$

■

### 1.1.5.2. Formas cuadráticas

En este apartado vamos a ver qué son las **formas cuadráticas**.

**Definición 1.1.18.** Sea  $V$  un  $\mathbb{K}$ -espacio vectorial y sea  $f : V \times V \rightarrow \mathbb{K}$  una forma bilineal en  $V$ . Se llama **forma cuadrática** asociada a  $f$ , a la aplicación,

$$\begin{aligned}
 \Phi : \quad V &\rightarrow \quad \mathbb{K} \\
 \Phi(x) &\mapsto f(x, x)
 \end{aligned}$$

$$\forall x \in V.$$

Veamos algunas propiedades.

**Proposición 1.1.32** (Propiedades). Sea  $\Phi : V \rightarrow \mathbb{K}$  la forma cuadrática asociada a la función bilineal  $f$ , para cualesquiera  $\lambda \in \mathbb{K}$  y  $x, y \in V$  se verifica:

$$(I) \quad \Phi(0) = 0$$

$$(II) \quad \Phi(\lambda x) = \lambda^2 \Phi(x)$$

$$(III) \quad \Phi(x, y) = \Phi(x) + \Phi(y) + f(x, y) + f(y, x)$$

**Demostración.** (I)  $\Phi(0) = f(0, 0) = 0$  ■

$$(II) \quad \Phi(\lambda x) = f(\lambda x, \lambda x) = \lambda(x, \lambda x) = \lambda^2 f(x, x) = \lambda^2 \Phi(x) \quad \blacksquare$$

$$(III) \quad \Phi(x + y) = f(x + y, x + y) = f(x, x) + f(y, y) + f(x, y) + f(y, x) = \Phi(x) + \Phi(y) + f(x, y) + f(y, x) \quad \blacksquare$$

**Nota 1.1.33.** *Distintas formas bilineales pueden dar lugar a la misma forma cuadrática.*

Veamos que una forma cuadrática puede descomponerse como **suma** de una forma bilineal simétrica y otra antisimétrica.

**Proposición 1.1.34.** *Si  $f$  es una forma bilineal simétrica y  $\Phi$  es la forma cuadrática asociada a  $f$ , entonces para cada forma bilineal antisimétrica  $g$  se tiene,*

$$\Phi(x) = f(x, x) = f(x, x) + g(x, x) = (f + g)(x, x)$$

**Demostración.** Si  $g$  es antisimétrica, entonces  $g(x, x) = -g(x, x) \Rightarrow g(x, x) = 0$ , luego,

$$\Phi(x) = f(x, x) = f(x, x) + 0 = f(x, x) + g(x, x)$$

al ser  $f$  y  $g$  aplicaciones bilineales, entonces

$$(f + g)(x, x) = f(x, x) + g(x, x)$$

luego,

$$\Phi(x) = f(x, x) = (f + g)(x, x) \quad \blacksquare$$

Por tanto,  $\Phi$  también será la forma cuadrática asociada a  $f + g$ . Luego, podemos decir que la forma cuadrática asociada a una forma bilineal solo depende de la parte simétrica de ésta.

Vemos entonces que de entre todas las formas bilineales que dan lugar a la misma forma cuadrática (existen infinitas), solo existe una forma que es simétrica.

**Proposición 1.1.35.** *Dada una forma cuadrática  $\Phi$  en  $V$ , existe una única forma bilineal simétrica  $f_P$ , cuya forma cuadrática asociada es  $\Phi$ . Llamada forma polar de  $\Phi$ .*

**Demostración.**  $\Rightarrow$  Para cada forma bilineal  $f$  cuya forma cuadrática asociada sea  $\Phi$ , se tiene por la propiedad (iii),  $f(x, y) + f(y, x) = \Phi(x + y) - \Phi(x) - \Phi(y)$ .

Si imponemos que  $f$  sea simétrica, entonces  $2f(x, y) = \Phi(x + y) - \Phi(x) - \Phi(y)$ , con lo cual,  $f$  está unívocamente determinada por:

$$f(x, y) = \frac{1}{2} [\Phi(x + y) - \Phi(x) - \Phi(y)]$$

y por tanto, es única.

$\Leftarrow$  Sea la forma bilineal  $f$ , definida por  $f(x, y) = \frac{1}{2} [\Phi(x + y) - \Phi(x) - \Phi(y)]$ , vemos que es simétrica:

$$\begin{aligned} f(x, y) &= \frac{1}{2} [f(x + y, x + y) - f(x, x) - f(y, y)] = \\ &= \frac{1}{2} [f(\cancel{x}, \cancel{x}) + f(\cancel{y}, \cancel{y}) + f(x, y) + f(y, x) - f(\cancel{x}, \cancel{x}) - f(\cancel{y}, \cancel{y})] = \\ &= \frac{1}{2} [f(x, y) + f(y, x)] \end{aligned}$$

luego,

$$2f(x, y) = f(x, y) + f(y, x) \Rightarrow 2f(x, y) - f(\cancel{x}, \cancel{y}) = f(y, x) \Rightarrow f(x, y) = f(y, x)$$

luego,  $f$  es antisimétrica. Hemos demostrado que existe una única forma bilineal simétrica, cuya forma cuadrática asociada es  $\Phi$ . ■

**Corolario 1.1.36.** *Sea  $\Phi$  una forma cuadrática en  $V$  asociada a la forma bilineal  $g$ . La forma polar  $f_P$  de  $\Phi$  se puede obtener como:*

$$(I) \quad f_P(x, y) = \frac{1}{2} [\Phi(x + y) - \Phi(x) - \Phi(y)]$$

$$(II) \quad f_P(x, y) = \frac{1}{4} [\Phi(x + y) - \Phi(x - y)]$$

$$(III) \quad f_P(x, y) = \frac{1}{2} [g(x, y) + g(y, x)]$$

**Demostración.** (I) Ya lo hicimos antes. ■

- (II) Sabemos que  $\Phi(x+y) = \Phi(x) + \Phi(y) + 2f_P(x, y)$  y que  $\Phi(x-y) = f(x-y, x-y) = f(x, x) + f(y, y) - f(x, y) - f(y, x) = \Phi(x) + \Phi(y) - 2f_P(x, y)$ , luego, restando ambas expresiones:

$$\Phi(x+y) - \Phi(x-y) = 4f_P(x, y) \Rightarrow f_P(x, y) = \frac{1}{4} [\Phi(x+y) - \Phi(x-y)] \quad \blacksquare$$

- (III) Si  $\Phi$  es la forma cuadrática asociada a  $g$ , entonces

$$\Phi(x+y) = \cancel{\Phi(x)} + \cancel{\Phi(y)} + g(x, y) + g(y, x) = \cancel{\Phi(x)} + \cancel{\Phi(y)} + 2f_P(x, y)$$

luego,

$$f_P(x, y) = \frac{1}{2} [g(x, y) + g(y, x)]$$

### Matriz asociada a una forma cuadrática

Dada una forma cuadrática  $\Phi$  en  $V$  y dada una base  $B$  de  $V$ , llamaremos matriz asociada a  $\Phi$  respecto de la base  $B$  a la matriz asociada respecto de la base  $B$  de la forma polar de  $\Phi$ . En particular, la matriz asociada a una forma cuadrática es siempre una matriz simétrica. Llamaremos rango de  $\Phi$  al rango  $rg(\Phi)$  de su matriz asociada, que coincide con el rango de su forma polar. Luego,

$$\Phi(x) = X^t \cdot A \cdot X = \sum_{i,j=1}^n a_{ij} x_i x_j \stackrel{a_{ij}=a_{ji}}{\Rightarrow} \Phi(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j$$

Luego,

$$\text{Para } n = 2 : \Phi(x, y) = a_{11}x^2 + a_{22}y^2 + 2a_{12}xy$$

$$\text{Para } n = 3 : \Phi(x, y) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz$$

### Conjugación respecto de una forma cuadrática

Veremos ahora los vectores conjugados a una forma cuadrática, junto algunas propiedades que veremos que también usaremos en la definición de producto escalar.

**Definición 1.1.19.** Sea  $\Phi : V \rightarrow \mathbb{K}$  una forma cuadrática, y sea  $f_P : V \times V \rightarrow \mathbb{K}$  su forma polar. Dos vectores  $x, y \in V$  se dice que son conjugados respecto de  $\Phi$  si  $f_P(x, y) = 0$ . Se dice que el vector  $x$  es autoconjunto si es conjugado consigo mismo, es decir,  $\Phi(x) = 0$ .

Dado un conjunto  $S \subseteq V$ , consideremos el conjunto de los vectores de  $V$  conjugados con todos los vectores de  $S$ :

$$S^C = \{x \in V \mid f_P(x, y) = 0, \quad \forall x, y \in V\}$$

**Proposición 1.1.37.** Para cada subconjunto no vacío  $S$  de  $V$ , el conjunto  $S^C$  es un subespacio vectorial de  $V$ . Además,  $S^C = (L(S))^C$ .

**Demostración.** Sean  $a, b \in \mathbb{K}$  y  $u, v \in S^C$  arbitrarios, entonces para cada  $y \in S$  se tiene,

$$f_P(a \cdot u + b \cdot v, y) = a \cdot f_P(u, y) + b \cdot f_P(v, y) = 0$$

Luego,  $a \cdot u + b \cdot v \in S^C$ , y en consecuencia,  $S^C$  es subespacio vectorial de  $V$ . ■

Por otra parte, puesto que  $S \subseteq L(S)$ , se tiene que  $(L(S))^C \subseteq S^C$ . Para la otra inclusión, consideramos  $x \in S^C$  y sea  $y \in L(S)$  arbitrario. Entonces  $y$  se escribe como combinación lineal de vectores de  $S$ ,

$$y = a_1 s_1 + \dots + a_k s_k; \quad a_1, \dots, a_k \in \mathbb{K}; \quad s_1, \dots, s_k \in S$$

Así pues,

$$f_P(x, y) = f_P(x, a_1 s_1 + \dots + a_k s_k) = a_1 f_P(x, s_1) + \dots + a_k f_P(x, s_k) = 0$$

ya que  $x \in S^C$  y los  $s_i \in S$ . Por tanto,  $x \in (L(S))^C$  y así,  $S^C \subseteq (L(S))^C$ . Luego, hemos demostrado que  $S^C = (L(S))^C$  ■

**Definición 1.1.20.** Se llama núcleo de la forma cuadrática  $\Phi$  al subespacio  $V$  a

$$N(\Phi) = V^C = \{x \in V \mid f_P(x, y) = 0; \quad \forall y \in V\}$$

Se dice que  $\Phi$  es no degenerada si  $N(\Phi) = 0$ , es decir, si el único vector que es conjugado a todos los vectores de  $V$  es el vector 0.

#### 1.1.5.3. Signatura de una forma cuadrática real

**Teorema 1.1.21.** Sea  $V$  un  $\mathbb{R}$ -espacio vectorial de dimensión finita y sea  $\Phi : V \rightarrow \mathbb{R}$  una forma cuadrática. Existe una base de  $V$  para la cual, la matriz asociada a  $\Phi$  es diagonal.

**Teorema 1.1.22** (Ley de inercia de Sylvester). *Sea  $V$  un  $\mathbb{R}$ -espacio vectorial de dimensión finita y sea  $\Phi : V \rightarrow \mathbb{R}$  una forma cuadrática. Si  $D_1$  y  $D_2$  son matrices diagonales asociadas a  $\Phi$  respecto de distintas bases  $B_1$  y  $B_2$ , entonces el número de elementos positivos y elementos negativos en  $D_1$  y  $D_2$ , es el mismo.*

**Demostración.** Como el núcleo de elementos positivos más el de negativos es en cualquier caso igual al rango de  $\Phi$ , por lo que, bastará probar que el número de positivos será igual para todas las bases de  $V$  que proporcionen una forma diagonal.

Sean pues  $B_1$  y  $B_2$  bases de  $V$  para las cuales la matriz de  $\Phi$  es diagonal y de forma que tienen en la diagonal, y de forma que tienen en la diagonal  $p$  y  $t$  elementos positivos respectivamente, y veamos que  $p = t$ . Escribiendo,

$$B_1 = \{u_1, u_2, \dots, u_p, u_{p+1}, \dots, u_n\}$$

$$B_2 = \{v_1, v_2, \dots, v_t, v_{t+1}, \dots, v_n\}$$

las matrices asociadas a  $\Phi$  respecto de  $B_1$  y  $B_2$  serán,

$$\begin{pmatrix} \Phi(u_1) & 0 & \dots & 0 \\ 0 & \Phi(u_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Phi(u_n) \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \Phi(v_1) & 0 & \dots & 0 \\ 0 & \Phi(v_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & \Phi(v_n) \end{pmatrix}$$

Por hipótesis, tenemos que

$$\Phi(u_1) > 0, \Phi(u_2) > 0, \dots, \Phi(u_p) > 0, \Phi(u_{p+1}) \leq 0, \dots, \Phi(u_n) \leq 0$$

$$\Phi(v_1) > 0, \Phi(v_2) > 0, \dots, \Phi(v_t) > 0, \Phi(v_{t+1}) \leq 0, \dots, \Phi(v_n) \leq 0$$

Considerando los subespacios de  $V$  siguientes,

$$U = L(u_1, \dots, u_p), \quad W = L(v_{t+1}, \dots, v_n)$$

Para cada  $0 \neq x \in U$ , se verifica  $\Phi(x) > 0$ , y para cada  $y \in W$ , se verifica  $\Phi(y) \leq 0$ . Por tanto, es evidente que  $U \cup W = \{0\}$  y por tanto,  $\dim(U + W) = \dim U + \dim W = p + n - t$ . Como  $U + W \leq V$ , entonces  $p + n - t \leq n \Rightarrow p \leq t$ . Por simetría, llegamos también a que  $t \leq p$ , por tanto,  $p = t$ , luego, hay el mismo número de elementos negativos y positivos. ■

Llamaremos **signatura** de la forma cuadrática real  $\Phi$  al par  $sg(\Phi) = (p, q)$ , donde  $p$  es el número de elementos positivos y  $q$ , el de negativos en una forma diagonal de  $\Phi$ .

Notemos que  $p$  es igual al número de autovalores positivos de la matriz de  $\Phi$ , y  $q$  es igual al número de autovalores negativos.

Por otro lado, el rango de  $\Phi$  es igual al número de filas no nulas de su forma diagonal, y por tanto, como ya hemos mencionado,  $rg(\Phi) = p + q$ .

**Teorema 1.1.23.** *Sea  $\Phi : V \rightarrow \mathbb{R}$  una forma cuadrática real, y llamemos  $n = \dim V$ , entonces,*

(I)  $\Phi$  es definida positiva si y solo si  $sg(\Phi) = (n, 0)$

(II)  $\Phi$  es definida negativa si y solo si  $sg(\Phi) = (0, n)$

(III)  $\Phi$  es semidefinida positiva si y solo si  $sg(\Phi) = (r, 0)$  con  $r < n$

(IV)  $\Phi$  es semidefinida negativa si y solo si  $sg(\Phi) = (0, r)$  con  $r < n$

**Demostración.** (I) Supongamos que  $\Phi$  es definida positiva, es decir,  $\Phi(x) > 0; \forall 0 \neq x \in V$  y sea  $B = \{e_1, e_2, \dots, e_n\}$  una base de  $V$  para la cual la matriz asociada a  $\Phi$  es diagonal,

$$\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

y notemos que, por la propia definición de matriz asociada a una forma cuadrática, se tiene  $d_i = \Phi(e_i)$ . Así pues, siendo  $\Phi$  definida positiva, se obtiene  $d_1 > 0, \dots, d_n > 0$  y por tanto,  $sg(\Phi) = (n, 0)$  ✓

Supongamos ahora que  $sg(\Phi) = (n, 0)$ . Entonces con la misma notación de antes, se tiene  $d_1 > 0, \dots, d_n > 0$  y, en consecuencia, si un vector cualquiera no nulo  $x = (x_1, \dots, x_n)_B$  de  $V$ , se verifica  $\Phi(x) = d_1x_1^2 + \dots + d_nx_n^2 > 0$  por ser cada  $d_i > 0$  ■

- (II) Supongamos que  $\Phi$  es definida negativa, es decir,  $\Phi(x) < 0; \forall 0 \neq x \in V$  y sea  $B = \{e_1, e_2, \dots, e_n\}$  una base de  $V$  para la cual la matriz asociada a  $\Phi$  es diagonal,

$$\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

y notemos que, por la propia definición de matriz asociada a una forma cuadrática, se tiene  $d_i = \Phi(e_i)$ . Así pues, siendo  $\Phi$  definida negativa, se obtiene  $d_1 < 0, \dots, d_n < 0$  y por tanto,  $sg(\Phi) = (0, n)\checkmark$

Supongamos ahora que  $sg(\Phi) = (0, n)$ . Entonces con la misma notación de antes, se tiene  $d_1 < 0, \dots, d_n < 0$  y, en consecuencia, si un vector cualquiera no nulo  $x = (x_1, \dots, x_n)_B$  de  $V$ , se verifica  $\Phi(x) = d_1x_1^2 + \dots + d_nx_n^2 < 0$  por ser cada  $d_i < 0$  ■

- (III) Por el teorema de Sylvester, sabemos que existen  $r$  valores propios positivos y  $n - r$  valores propios que son cero (ya que no es definida positiva y no hay valores propios negativos). Esto da lugar a una signatura de  $(r, 0)$  con  $r < n$ , ya que si  $r$  fuera igual a  $n$ , sería definida positiva, lo cual contradice nuestra suposición inicial de que es solamente semidefinida positiva.✓

Inversamente, si  $sg(\Phi) = (r, 0)$  con  $r < n$ , entonces la matriz asociada  $A$  a la forma cuadrática  $\Phi$  tiene  $r$  valores propios positivos y  $n - r$  ceros. Esto significa que para cualquier vector  $v$ ,  $\Phi(v) = v^T A v$  será una suma de términos no negativos, dado que los valores propios negativos corresponden a términos negativos en esta suma y no hay ninguno. Por lo tanto,  $\Phi(v) \geq 0$  para todo  $v$ , lo que significa que es semidefinida positiva. ■

- (IV) Supongamos que  $\Phi$  es semidefinida negativa. Esto significa que para todo vector  $v \in V$ ,  $\Phi(v) \geq 0$ . Dado que  $\Phi$  no es definida negativa, no todos los valores propios pueden ser negativos (de lo contrario,  $\Phi(v) < 0$  para todo  $v \neq 0$ ). Entonces, algunos de los valores propios deben ser cero. La signatura  $sg(\Phi)$  cuenta el número de valores propios negativos y positivos de la matriz simétrica asociada a  $\Phi$ . Si  $\Phi$  es semidefinida negativa, entonces no tiene valores propios positivos, lo que implica que  $sg(\Phi) = (0, r)$  con  $r$  siendo el número de valores propios negativos, y  $r < n$  porque si  $r$  fuera igual a  $n$ , sería definida negativa.✓

Ahora supongamos que  $sg(\Phi) = (0, r)$  con  $r < n$ . Esto indica que hay  $r$  valores

propios negativos y  $n - r$  ceros (puesto que no hay valores propios positivos). Por lo tanto, para cualquier vector  $v$ , la forma cuadrática  $\Phi(v) = v^T A v$  será una suma de términos no positivos, debido a que los valores propios positivos resultarían en términos positivos en esta suma, y no hay ninguno. Esto significa que  $\Phi(v) \leq 0$  para todo  $v \in V$ , y por lo tanto  $\Phi$  es semidefinida negativa.

■

## 1.2. Álgebra de Tensores

Llegamos a lo groso del capítulo, el **Álgebra de Tensores**. En este apartado vamos a ver qué es un tensor de forma matemática y cómo trabajar con ellos. También se mencionará cómo trabajamos los físicos con los tensores.

### 1.2.1. Producto tensorial: caso de dos términos

Vamos a ver qué es el **producto tensorial** y cómo los tensores se definen a partir de este.

**Proposición 1.2.1.** *Sea  $V$  un  $\mathbb{K}$ -espacio vectorial,  $\langle \cdot, \cdot \rangle$  el producto escalar euclídeo y  $B = \{v_1, \dots, v_n\}$  base de  $V$ ,*

$$\begin{aligned} f_v : V &\rightarrow V^* \\ v &\mapsto f_v(v) = \langle v, \cdot \rangle \end{aligned}$$

$f_v$  es una aplicación lineal, concretamente es un isomorfismo.

**Demostración.** Vemos que  $f_v$  es aplicación lineal,

$$f_v(w_1 + w_2) = \langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle = f_v(w_1) + f_v(w_2) \checkmark$$

$$f_v(\lambda \cdot w) = \langle v, \lambda \cdot w \rangle = \lambda \langle v, w \rangle = \lambda f_v(w) \checkmark$$

para  $\forall \lambda \in \mathbb{K}$  y  $\forall w_1, w_2, w \in V$ . Luego, es aplicación lineal.

Veamos que es isomorfo demostrando que es biyectivo, pues ya hemos visto que es aplicación lineal.

Sabemos que  $\ker \{f_v\} = \{0\} \Leftrightarrow f_v$  es inyectiva. Luego, vemos si  $\ker \{f_v\} = \{0\}$ :

$$\ker \{f_v\} = \{w \in V, f_v(w) = 0\} = \{w \in V; \langle v, w \rangle = 0 \Leftrightarrow w = 0\}$$

Por tanto,  $\ker \{f_v\} = \{0\}$  y así,  $f_v$  es inyectiva. ✓

Usando el Primer Teorema de isomorfía, tenemos que  $\dim(V) = \dim(\ker \{f_v\}) + \dim(Im f_v)$ , pero como la  $\dim B = \dim B^*$ , siendo  $B$  base de  $V$  y  $B^*$  base de  $V^*$ , entonces  $\dim V = \dim V^*$ , y por tanto,  $\dim V = \dim Im f_v = \dim V^*$ , luego  $Im f_v$  es  $V^*$  y por tanto,  $f_v$  es sobreyectiva. ✓

Luego,  $f_v$  es un isomorfismo. ■

Veamos cómo se define el producto tensorial y sus propiedades.

**Definición 1.2.1.** Sea  $V$  un  $\mathbb{K}$ -espacio vectorial,  $V^*$  el dual de  $V$ , y  $g^1, g^2 \in V^*$  aplicaciones lineales, tal que  $g^1 : V \rightarrow \mathbb{K}$  y  $g^2 : V \rightarrow \mathbb{K}$ . Así, definimos el producto tensorial como,

(I) Producto tensorial entre dos formas  $g^1, g^2 \in V^*$ ,

$$\begin{aligned} g^1 \otimes g^2 : V \times V &\rightarrow \mathbb{K} \\ (v, w) &\mapsto g^1(v)g^2(w) \end{aligned}$$

(II) Producto tensorial entre dos vectores  $v_1, v_2 \in V$ ,

$$\begin{aligned} v_1 \otimes v_2 : V^* \times V^* &\rightarrow \mathbb{K} \\ (f, g) &\mapsto f(v_1)g(v_2) \end{aligned}$$

(III) Producto tensorial de una forma y un vector  $v_1 \in V, f^1 \in V^*$ ,

$$\begin{aligned} v_1 \otimes f^1 : V^* \times V &\rightarrow \mathbb{K} \\ (g, w) &\mapsto g(v_1)f^1(w) \end{aligned}$$

**Proposición 1.2.2.** Los productos tensoriales definidos anteriormente son formas bilineales.

**Demostración.** Usando  $\forall v_1, v_2, u_1, u_2, v, w, u \in V, \forall f^1, f^2, g, p, q \in V^*$  y  $\forall \lambda \in \mathbb{K}$ ,

(I)

$$\begin{aligned} f^1 \otimes f^2 : V \times V &\rightarrow \mathbb{K} \\ (v, w) &\mapsto f^1(v)f^2(w) \end{aligned}$$

siendo  $f^1, f^2 \in V^*$ . Veamos que es forma bilineal,

- )  $(f^1 \otimes f^2)(u_1 + u_2, v) = f^1(u_1 + u_2)f^2(v) = [f^1(u_1) + f^1(u_2)]f^2(v)$   
 $= f^1(u_1)f^2(v) + f^1(u_2)f^2(v) = (f^1 \otimes f^2)(u_1, v) + (f^1 \otimes f^2)(u_2, v), \checkmark$
- )  $(f^1 \otimes f^2)(v, u_1 + u_2) = f^1(v)f^2(u_1 + u_2) = f^1(v)[f^2(u_1) + f^2(u_2)]$   
 $= f^1(v)f^2(u_1) + f^1(v)f^2(u_2) = (f^1 \otimes f^2)(v, u_1) + (f^1 \otimes f^2)(v, u_2), \checkmark$
- )  $(f^1 \otimes f^2)(\lambda v, u) = f^1(\lambda v)f^2(u) = \lambda f^1(v)f^2(u) = \lambda(f^1 \otimes f^2)(v, u), \checkmark$
- )  $(f^1 \otimes f^2)(u, \lambda v) = f^1(u)f^2(\lambda v) = \lambda f^1(u)f^2(v) = \lambda(f^1 \otimes f^2)(u, v), \checkmark$

Luego,  $f^1 \otimes f^2$  es una forma bilineal. ■

(II)

$$\begin{aligned} v_1 \otimes v_2 : V^* \times V^* &\rightarrow \mathbb{K} \\ (f, g) &\mapsto f(v_1)g(v_2) \end{aligned}$$

- )  $(v_1 \otimes v_2)(f^1 + f^2, g) = (f^1 + f^2)(v_1)g(v_2) = [f^1(v_1) + f^2(v_1)]g(v_2)$   
 $= f^1(v_1)g(v_2) + f^2(v_1)g(v_2) = (v_1 \otimes v_2)(f^1, g) + (v_1 \otimes v_2)(f^2, g), \checkmark$
- )  $(v_1 \otimes v_2)(g, f^1 + f^2) = g(v_1)(f^1 + f^2)(v_2)g = g(v_1)[f^1(v_2) + f^2(v_2)]$   
 $= g(v_1)f^1(v_2) + g(v_1)f^2(v_2) = (v_1 \otimes v_2)(g, f^1) + (v_1 \otimes v_2)(g, f^2), \checkmark$
- )  $(v_1 \otimes v_2)(\lambda f, g) = (\lambda f)(v_1)g(v_2) = \lambda f(v_1)g(v_2) = \lambda(v_1 \otimes v_2)(f, g), \checkmark$
- )  $(v_1 \otimes v_2)(g, \lambda f) = g(v_1)(\lambda f)(v_2) = \lambda g(v_1)f(v_2) = \lambda(v_1 \otimes v_2)(g, f), \checkmark$

Luego,  $v_1 \otimes v_2$  es una forma bilineal. ■

(III)

$$\begin{aligned} v_1 \otimes f^1 : V^* \times V &\rightarrow \mathbb{K} \\ (g, w) &\mapsto g(v_1)f^1(w) \end{aligned}$$

- )  $(v_1 \otimes f^1)(p + q, w) = (p + q)(v_1)f^1(w) = [p(v_1) + q(v_1)]f^1(w) =$   
 $= p(v_1)f^1(w) + q(v_1)f^1(w) = (v_1 \otimes f^1)(p, w) + (v_1 \otimes f^1)(q, w), \checkmark$
- )  $(v_1 \otimes f^1)(g, u + w) = g(v_1)f^1(u + w) = g(v_1)[f^1(u) + f^1(w)] =$   
 $= g(v_1)f^1(u) + g(v_1)f^1(w) = (v_1 \otimes f^1)(g, u) + (v_1 \otimes f^1)(g, w), \checkmark$
- )  $(v_1 \otimes f^1)(\lambda g, w) = (\lambda g)(v_1)f^1(w) = \lambda g(v_1)f^1(w) = \lambda(v_1 \otimes f^1)(g, w), \checkmark$
- )  $(v_1 \otimes f^1)(g, \lambda w) = g(v_1)f^1(\lambda w) = \lambda g(v_1)f^1(w) = \lambda(v_1 \otimes f^1)(g, w), \checkmark$

Luego,  $v_1 \otimes f^1$  es una forma bilineal. ■

El producto tensorial no se da solo entre elementos de los espacios vectoriales o duales, sino que también se puede dar entre espacios, siendo el nuevo espacio generado un **espacio vectorial**.

**Proposición 1.2.3.** *El espacio  $V \otimes V$  tiene estructura de espacio vectorial.*

**Demostración.** 1. Vemos que  $(V \otimes V, +)$  es grupo abeliano:

(I) Vemos si la operación  $+$  es cerrada:

$\forall v, w, z \in V$  con  $v \otimes w, v \otimes z, w \otimes z \in V \otimes V$ , tenemos que ver si  $(v + w) \otimes z \in V \otimes V$ . Sabemos que,

$$\begin{aligned} v \otimes w : V^* \times V^* &\rightarrow \mathbb{R} \\ (f, g) &\mapsto f(v)g(w) \end{aligned}$$

luego,

$$\begin{aligned} (v + w) \otimes z : V^* \times V^* &\rightarrow \mathbb{R} \\ (f, p) &\mapsto f(v + w)p(z) \end{aligned}$$

Entonces,

$$\begin{aligned} ((v + w) \otimes z)(f, p) &= f(v + w)p(z) = [f(v) + f(w)]p(z) = \\ &= f(v)p(z) + f(w)p(z) = (v \otimes z)(f, p) + (w \otimes z)(f, p) \end{aligned}$$

Luego,  $(v + w) \otimes z \in V \otimes V$  y así, la operación  $+$  es cerrada. ✓

(II) Asociatividad:

Sean  $a \otimes b, c \otimes d, e \otimes f \in V \otimes V$ , tenemos que ver si  $a \otimes b + [c \otimes d + e \otimes f] = [a \otimes b + c \otimes d] + e \otimes f$ , tal que

$$\begin{aligned} (a \otimes b + [c \otimes d + e \otimes f])(p, q) &= p(a)q(b) + [p(c)q(d) + p(e)q(f)] = p(a)q(b) + p(c)q(d) + p(e)q(f) = \\ &= p(a + c + e)q(b + d + f) = p(a + c)q(b + d) + p(e)q(f) = [p(a)q(b) + p(c)q(d)] + p(e)q(f) = \\ &= ([a \otimes b + c \otimes d] + e \otimes f)(p, q) \end{aligned}$$

(III) Elemento neutro:

Sea  $e_1 \otimes e_2 \in V \otimes V$  el elemento neutro de  $V \otimes V$ , tal que

$$e_1 \otimes e_2 + v \otimes w = v \otimes w + e_1 \otimes e_2 = v \otimes w$$

Vemos el valor de este elemento neutro,

$$(e_1 \otimes e_2 + v \otimes w)(f, g) = (v \otimes w)(f, g)$$

$$f(e_1)g(e_2) + f(v)g(w) = f(v)g(w)$$

$$f(e_1 + v)g(e_w + w) = f(v)g(w) \Leftrightarrow \begin{cases} e_1 = 0 \\ e_2 = 0 \end{cases}$$

luego,  $e_1 \otimes e_2 = 0$ . ✓

(iv) Elemento simétrico:

$\forall v \otimes u \in V \otimes V, \exists \tilde{v} \otimes \tilde{u} \in V \otimes V$ , tal que

$$v \otimes u + \tilde{v} \otimes \tilde{u} = \tilde{v} \otimes \tilde{u} + v \otimes u = e_1 \otimes e_2 = 0$$

Veamos quién es  $\tilde{v} \otimes \tilde{u}$ ,

$$(v \otimes u + \tilde{v} \otimes \tilde{u})(f, g) = f(v)g(u) + f(\tilde{v})g(\tilde{u}) = (0 \otimes 0)(f, g) = f(0)g(0)$$

luego,

$$v + \tilde{v} = 0 \Rightarrow \tilde{v} = -v$$

$$u + \tilde{u} = 0 \Rightarrow \tilde{u} = -u$$

Por tanto, el elemento simétrico de  $v \otimes u$  es  $(-v) \otimes (-u)$ . ✓

(v) Conmutabilidad:

Sean  $v \otimes w, u \otimes z \in V \otimes V$ , entonces

$$(v \otimes w + u \otimes z)(f, g) = f(v)g(w) + f(u)g(z) = f(v+u)g(w+z) =$$

$$= f(u+v)g(z+w) = f(u)g(z) + f(v)g(w) = (u \otimes z + v \otimes w)(f, g) \checkmark$$

Luego, es grupo abeliano. ✓

2. Doble propiedad distributiva:

a)  $\forall \lambda, \mu \in \mathbb{R}, \forall v \otimes w \in V \otimes V$ ,

$$(\lambda + \mu) \cdot (v \otimes w)(f, g) = (\lambda + \mu)f(v)g(w) =$$

$$= \lambda f(v)g(w) + \mu f(v)g(w) = \lambda(v \otimes w)(f, g) + \mu(v \otimes w)(f, g) \checkmark$$

b)  $\forall \lambda \in \mathbb{R}, \forall v \otimes w, u \otimes z \in V \otimes V$ , tenemos que

$$\begin{aligned} \lambda(v \otimes w)(f, g) + \lambda(u \otimes z)(f, g) &= \lambda f(v)g(w) + \lambda f(u)g(z) = \\ &= \lambda [f(v)g(w) + f(u)g(z)] = \lambda(v \otimes w + u \otimes z)(f, g) \checkmark \end{aligned}$$

3. Propiedad pseudo-asociativa:

$\forall \lambda, \mu \in \mathbb{R}; \forall v \otimes w \in V \otimes V$ , tenemos que

$$\begin{aligned} \lambda \cdot [\mu \cdot (v \otimes w)(f, g)] &= \lambda [\mu f(v)g(w)] = \lambda f(\mu v)g(\mu w) = \\ &= f(\lambda \mu v)g(\lambda \mu w) = f(\mu \lambda v)g(\mu \lambda w) = \mu [f(\lambda v)g(\lambda w)] = (\mu \cdot \lambda) f(v)g(w) = (\mu \cdot \lambda)(v \otimes w)(f, g) \checkmark \end{aligned}$$

4. Elemento unitario del cuerpo:  $\forall v \otimes w \in V \otimes V; \tilde{\mu} \in \mathbb{R}$ , entonces  $\tilde{\mu} \cdot v \otimes w = v \otimes w \cdot \tilde{\mu} = v \otimes w$

$$(v \otimes w)(f, g) = f(v)g(w) \Rightarrow \frac{\tilde{\mu} \cdot v = v}{\tilde{\mu} \cdot w = w} \Leftrightarrow \tilde{\mu} = 1 \checkmark$$

Luego,  $(V \otimes V, +, \cdot)$  es un  $\mathbb{R}$ -espacio vectorial. ■

Al igual que cualquier otro espacio vectorial, el espacio  $V \otimes V$  deberá tener una **base**.

**Proposición 1.2.4.** Si tenemos un  $V$  espacio vectorial sobre  $\mathbb{K}$  con base  $B = \{v_1, \dots, v_n\}$ , entonces todo  $v \otimes w$  será combinación lineal de los elementos de la base de  $V \otimes V$  dada por  $B \otimes B = \{v_i \otimes v_j\}_{i,j=1}^n$

**Demostración.** Queremos ver que  $\{v_i \otimes v_j\}_{i,j=1}^n$  es base de  $V \otimes V$ . Para ello, tendremos que ver que esta base  $B \otimes B$  complete el espacio  $V \otimes V$  y que los vectores de la misma sean linealmente independientes.

Sabemos que  $v \otimes w \in V \otimes V$  y que

$$\begin{aligned} v \otimes w : V^* \times V^* &\rightarrow \mathbb{R} \\ (f, g) &\mapsto f(v)g(w) \end{aligned}$$

Luego, para que la base  $B \otimes B$  complete el espacio  $V \otimes V$ , se deberá poder expresar cualquier vector  $v \otimes w \in V \otimes V$  como combinación lineal de los vectores de  $B \otimes B$ .

Podemos usar  $B = \{v_i\}_{i=1}^n$  base de  $V$ , tal que

$$v = \sum_{i=1}^n \lambda^i v_i = \lambda^i v_i, \quad w = \sum_{j=1}^n \mu^j v_j = \mu^j v_j$$

Por tanto, usando  $f, g \in V^*$ , tenemos que

$$v \otimes w(f, g) = f(v)g(w) = f(\lambda^i v_i)g(\mu^j v_j) = \lambda^i f(v_i)\mu^j g(v_j) = \lambda^i \mu^j f(v_i)g(v_j) = \lambda^i \mu^j (v_i \otimes v_j)(f, g)$$

Luego, hemos expresado un vector del espacio  $V \otimes V$  como combinación lineal de los vectores de la base  $B \otimes B$ . ✓

Veamos que son linealmente independientes, para ello, se debe cumplir que,

$$\sum_{i,j=1}^n \lambda^{ij} (v_i \otimes v_j) = \lambda^{ij} (v_i \otimes v_j) = 0 \Leftrightarrow \lambda^{ij} = 0$$

Sabiendo que la base de  $V^*$  es  $B^* = \{f^1, f^2, \dots, f^n\}$ , tal que

$$f^i(v_i) = 1 \quad f^j(v_i) \stackrel{i \neq j}{=} 0 \Rightarrow f^i(v_j) = \delta_{ij}$$

Podemos evaluar lo anterior en dos elementos arbitrarios de  $B^*$ , tal que

$$0 = \lambda^{ij} (v_i \otimes v_j)(f^n, f^m) = \lambda^{ij} f^n(v_i) f^m(v_j) = \lambda_{ij} \delta_n^i \delta_m^j = \lambda^{nm}$$

luego,  $\lambda^{nm} = 0$  y por tanto, los vectores son linealmente independientes. ✓

Así, hemos demostrado que  $B \otimes B$  es base de  $V \otimes V$ . ■

**Nota 1.2.5.** Denotaremos  $v \otimes w \equiv h$ , tal que

$$\begin{aligned} h : V^* \times V^* &\rightarrow \mathbb{R} \\ (f^i, f^j) &\mapsto h(f^i, f^j) = h^{ij} \end{aligned}$$

siendo  $f^i, f^j \in B^*$ . Por tanto, para dos  $p, q \in V^*$  cualesquiera, escribiremos

$$(v \otimes w)(p, q) = h(p, q) = h\left(\sum_{i=1}^n p_i f^i, \sum_{j=1}^n q_j f^j\right) = p_i q_j (f^i, f^j) = h^{ij} p_i q_j$$

Veamos algunas **propiedades** del producto tensorial.

**Proposición 1.2.6.** *Sea  $V$  un  $\mathbb{R}$ -espacio vectorial,*

$$(I) \quad (v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w; \quad \forall v_1, v_2, w \in V.$$

$$(II) \quad w \otimes (v_1 + v_2) = w \otimes v_1 + w \otimes v_2, \quad \forall v_1, v_2, w \in V.$$

$$(III) \quad (\lambda v) \otimes w = \lambda v \otimes w, \quad \forall v, w \in V, \quad \forall \lambda \in \mathbb{R}.$$

$$(IV) \quad w \otimes (\lambda v) = \lambda w \otimes v, \quad \forall v, w \in V, \quad \forall \lambda \in \mathbb{R}.$$

$$(V) \quad v \otimes w \neq w \otimes v.$$

$$(VI) \quad v \otimes w \neq 0 \text{ si } v \neq 0 \text{ ó } w \neq 0.$$

$$(VII) \quad \text{Sea } a \otimes b \neq 0, \quad a \otimes b = a' \otimes b' \Leftrightarrow a' = \lambda a \text{ y } b' = \lambda^{-1}b.$$

$$(VIII) \quad V \otimes W \text{ es isomorfo con } W \otimes V.$$

**Demostración.** (I)  $\forall v_1, v_2, w \in V,$

$$((v_1 + v_2) \otimes w)(f, g) = f(v_1 + v_2)g(w) = [f(v_1) + f(v_2)]g(w) =$$

$$= f(v_1)g(w) + f(v_2)g(w) = (v_1 \otimes w)(f, g) + (v_2 \otimes w)(f, g) \quad \blacksquare$$

$$(II) \quad \forall v_1, v_2, w \in V,$$

$$(w \otimes (v_1 + v_2))(f, g) = f(w)g(v_1 + v_2) = f(w)[g(v_1) + g(v_2)] =$$

$$= f(w)g(v_1) + f(w)g(v_2) = (w \otimes v_1)(f, g) + (w \otimes v_2)(f, g) \quad \blacksquare$$

$$(III) \quad \forall v, w \in V \text{ y } \forall \lambda \in \mathbb{R},$$

$$((\lambda \cdot v) \otimes w)(f, g) = f(\lambda \cdot v)g(w) = \lambda \cdot f(v)g(w) = \lambda \cdot (v \otimes w)(f, g) \quad \blacksquare$$

$$(IV) \quad \forall v, w \in V \text{ y } \forall \mu \in \mathbb{R},$$

$$(w \otimes (\lambda \cdot v))(f, g) = f(w)g(\lambda \cdot v) = \lambda \cdot f(w)g(v) = \lambda \cdot (w \otimes v)(f, g) \quad \blacksquare$$

(V) Vemos que,  $(v \otimes w)(f, g) = f(v)g(w)$  y que  $(w \otimes v)(f, g) = f(w)g(v)$ , luego estos elementos serían iguales solo si  $f \equiv g$ .  $\blacksquare$

(vi) Sean  $v, w \in V$  y  $f, g \in V^*$ , tales que  $f \not\equiv 0$  y  $g \not\equiv 0$ , entonces

$$\begin{aligned} f(v) = 0 &\Leftrightarrow v = 0 \\ (v \otimes w)(f, g) = f(v)g(w) = 0 &\Leftrightarrow \quad \text{ó} \quad \blacksquare \\ g(w) = 0 &\Leftrightarrow w = 0 \end{aligned}$$

(vii)  $\Rightarrow$  Sea  $a \otimes b = a' \otimes b'$  entonces

$$(a \otimes b)(f, g) = f(a)g(b) = (a' \otimes b')(f, g) = f(a')g(b')$$

luego,

$$f(a)g(b) = f(a')g(b')$$

pero como  $a \neq a'$  y  $b \neq b'$ , debe haber una relación entre ambos, de tal forma que se cumpla la igualdad anterior. Supondremos que  $a$  y  $a'$  tienen una relación lineal (la más sencilla), tal que  $a' = \lambda a + c$ , luego

$$f(a)g(b) = f(a')g(b') = f(\lambda a + c)g(b') = f(\lambda a)g(b') + f(c)g(b') = \lambda f(a)g(b') + f(c)g(b')$$

Agrupamos términos de la igualdad, tal que,

$$0 : \quad 0 = f(c)g(b')$$

$$f(a) : \quad g(b) = \lambda g(b')$$

Por la propiedad (vi), como  $b' \neq 0$ , entonces  $c = 0$ . Además,

$$g(b) = \lambda g(b') \Rightarrow g(b') = \lambda^{-1}g(b) \Rightarrow g(b') = g(\lambda^{-1}b) \Rightarrow b' = \lambda^{-1}b$$

Luego,

$$\begin{aligned} a' &= \lambda a && \checkmark \\ b' &= \lambda^{-1}b \end{aligned}$$

$\Leftarrow$  Sea  $a' = \lambda a$  y  $b' = \lambda^{-1}b$ , entonces

$$(a' \otimes b')(f, g) = f(a')g(b') = f(\lambda a)g(\lambda^{-1}b) = \cancel{\lambda} \cancel{\lambda^{-1}} f(a)g(b) = (a \otimes b)(f, g) \checkmark$$

(viii) Sean  $V, W$  espacios vectoriales, tales que

$$\begin{aligned} V \otimes W &\rightarrow W \otimes V \\ v \otimes w &\mapsto w \otimes v \end{aligned}$$

Si suponemos que  $\dim V = n$  y  $\dim W = m$ , sabemos por tanto que  $\dim(V \otimes W) = n \cdot m$  y  $\dim(W \otimes V) = m \cdot n$ , luego tienen la misma dimensión y por tanto, son isomorfos. ✓

También podemos hacerlo sin usar la proposición de que  $\dim(V \otimes W) = n \cdot m$ . Es claro ver que la aplicación es inyectiva, pues no hay dos elementos con la misma imagen, ya que la imagen se forma al permutar los elementos. Luego, al ser inyectivo, tenemos que  $\dim \text{Ker} = 0$ . Por el Primer Teorema de Isomorfía,

$$\dim(V \otimes W) = \dim \text{Ker} + \dim \text{Im}^0 = \dim \text{Im} = \dim(W \otimes V)$$

Luego, como  $V \otimes W$  y  $W \otimes V$  tienen la misma dimensión, entonces son isomorfos.

■

### 1.2.2. El espacio de tensores $(r,s)$ : Definición, propiedades y ejemplos

Una vez visto el producto tensorial de dos elementos, vamos a ver una generalización.

**Definición 1.2.2.** Sea  $V$  un espacio vectorial y  $V^*$  un espacio dual, definimos

$$\Omega^{r,s}(V) = \{f \text{ aplicación multilínea; } f : V^* \times \dots \times V^* \times V \times \dots \times V\}$$

es decir,

$$\Omega^{r,s}(V) \equiv V \otimes \dots \otimes V \otimes V^* \otimes \dots \otimes V^*$$

**Nota 1.2.7.** Sea  $V$  un espacio vectorial. Las formas multilíneales cuyas variables están en  $V^*$  o  $V$ , se denominan **tensores sobre  $V$**  y los espacios vectoriales que forman, se denominan **espacios tensoriales sobre  $V$** .

El número de variables de  $V^*$  y  $V$  se denominan los **grados** de un tensor; al número de variables de  $V^*$  se les denomina **grados contravariantes** y al número de variables de

*V, grados covariantes.* Así, una forma multilínea del tipo  $V^* \times V \times V$  es un tensor de tipo  $(1, 2)$ , denotado como  $V \otimes V^* \otimes V^* = T_2^1$ .

**Nota 1.2.8.** Un tensor de tipo  $(0, 0)$  se define como un escalar, tal que  $T_0^0 = \lambda$ .

Un tensor de tipo  $(1, 0)$  se denomina **vector contravariante o vector** y a uno del tipo  $(0, 1)$ , **vector covariante o covector**.

Un tensor de tipo  $(r, 0)$  se denomina **tensor contravariante** y uno del tipo  $(0, s)$ , se denomina **tensor covariante**.

Al igual que  $V \otimes V$  es un espacio vectorial, el conjunto  $\Omega^{r,s}(V)$  también debe serlo.

**Proposición 1.2.9.**  $\Omega^{r,s}(V)$  es espacio vectorial.

**Demostración.** Tenemos que ver que  $(\Omega^{r,s}(V), +, \cdot)$  es un  $\mathbb{R}$ -espacio vectorial, siendo  $+$  una operación interna y  $\cdot$  una operación externa, tal que

$$+ : \Omega^{r,s}(V) \times \Omega^{r,s}(V) \rightarrow \Omega^{r,s}(V) \quad y \quad \cdot : \Omega^{r,s}(V) \times \mathbb{R} \rightarrow \Omega^{r,s}(V)$$

Veamos si verifica las condiciones de espacio vectorial:

1. ¿ $(\Omega^{r,s}(V), +)$  es un grupo abeliano?

(I) ¿ $+$  es una operación cerrada?

Sabiendo que  $f, g \in \Omega^{r,s}(V)$  son aplicaciones multilíneales, entonces  $h = f + g$  será otra aplicación multilínea, lo vemos,

$$\begin{aligned} h(v_1^*, \dots, \alpha v_i^* + \lambda w_i^*, \dots, v_s^*) &= (f + g)(v_1^*, \dots, \alpha v_i^* + \lambda w_i^*, \dots, v_s^*) = \\ &= f(v_1^*, \dots, \alpha v_i^* + \lambda w_i^*, \dots, v_s^*) + g(v_1^*, \dots, \alpha v_i^* + \lambda w_i^*, \dots, v_s^*) = \\ &= \alpha f(v_1^*, \dots, v_i^*, \dots, v_s^*) + \lambda f(v_1^*, \dots, w_i^*, \dots, v_s^*) + \\ &\quad + \alpha g(v_1^*, \dots, v_i^*, \dots, v_s^*) + \lambda g(v_1^*, \dots, w_i^*, \dots, v_s^*) = \\ &= \alpha [f(v_1^*, \dots, v_i^*, \dots, v_s^*) + g(v_1^*, \dots, v_i^*, v_s^*)] + \\ &\quad + \lambda [f(v_1^*, \dots, w_i^*, \dots, v_s^*) + g(v_1^*, \dots, w_i^*, \dots, v_s^*)] = \\ &= \alpha(f + g)(v_1^*, \dots, v_i^*, \dots, v_s^*) + \lambda(f + g)(v_1^*, \dots, w_i^*, \dots, v_s^*) = \\ &= \alpha h(v_1^*, \dots, v_i^*, \dots, v_s^*) + \lambda h(v_1^*, \dots, w_i^*, \dots, v_s^*) \checkmark \end{aligned}$$

Luego,  $h \in \Omega^{r,s}(V)$ , y por tanto, la operación es cerrada.  $\checkmark$

(II) ¿Asociatividad?

$$\forall f, g, h \in \Omega^{r,s}(V)$$

$$(f + (g+h))(v) = f(v) + (g+h)(v) = f(v) + (g(v) + h(v)) = ((f)(v) + g(v)) + h(v) = \\ = (f + g)(v) + h(v) = ((f + g) + h)(v) \checkmark$$

(III) ¿Elemento neutro?

$$\forall f \in \Omega^{r,s}(V), \exists f^0 \in \Omega^{r,s}(V) \text{ tal que } f^0 + f = f = f + f^0 = f$$

$$(f + f^0)(v) = f(v) + f^0(v) = f(v) \Rightarrow f^0(v) = 0 \Rightarrow f^0 \equiv 0 \checkmark$$

(IV) ¿Elemento simétrico?

$$\forall f \in \Omega^{r,s}(V), \exists \tilde{f} \in \Omega^{r,s}(V) \text{ tal que } f + \tilde{f} = \tilde{f} + f = f^0$$

$$(f + \tilde{f})(v) = f(v) + \tilde{f}(v) = f^0(v) = 0 \Rightarrow \tilde{f}(v) = -f(v) \Rightarrow \tilde{f} \equiv -f \checkmark$$

(V) ¿Commutabilidad?

$$\forall f, g \in \Omega^{r,s}(V),$$

$$(f + g)(v) = f(v) + g(v) = g(v) + f(v) = (g + f)(v) \checkmark$$

Luego,  $(\Omega^{r,s}(V), +)$  es grupo abeliano. ■

2. Doble propiedad distributiva:

a)  $\forall \lambda, \mu \in \mathbb{R}$  y  $\forall f \in \Omega^{r,s}(V)$ ,

$$(\lambda + \mu)f(v) = f((\lambda + \mu)v) = f(\lambda v) + f(\mu v) = \lambda f(v) + \mu f(v) \checkmark$$

b)  $\forall \lambda \in \mathbb{R}$  y  $\forall f, g \in \Omega^{r,s}(V)$ ,

$$\lambda(f + g)(v) = (f + g)(\lambda v) = f(\lambda v) + g(\lambda v) = \lambda f(v) + \lambda g(v) \checkmark$$

3. Propiedad pseudo-asociativa:

$$\forall \lambda, \mu \in \mathbb{R} \text{ y } \forall f \in \Omega^{r,s}(V),$$

$$\lambda(\mu f(v)) = \lambda f(\mu v) = f((\lambda\mu)v) = (\lambda\mu)f(v) \checkmark$$

4. Elemento unitario de  $\mathbb{R}$ :

$$\forall f \in \Omega^{r,s}(V), \exists \tilde{\lambda} \in \mathbb{R} \text{ tal que } \tilde{\lambda} \cdot f = f \cdot \tilde{\lambda} = f,$$

$$\tilde{\lambda} \cdot f(v) = f(\tilde{\lambda} \cdot v) = f(v) \Rightarrow \tilde{\lambda}v = v \Rightarrow \tilde{\lambda} = 1 \checkmark$$

Luego,  $(\Omega^{r,s}(V), +, \cdot)$  es un  $\mathbb{R}$ -espacio vectorial. ■

Como  $\Omega^{r,s}(V)$  es un espacio vectorial, deberá de tener una **base**.

**Proposición 1.2.10.** *Si tenemos un conjunto  $V$  que sea un  $\mathbb{R}$ -espacio vectorial con base  $B = \{v_1, \dots, v_n\}$  y  $V^*$  el espacio dual de  $V$  con base  $B^* = \{f^1, \dots, f^n\}$ , entonces todo  $h \in \Omega^{r,s}$  será combinación lineal de  $B^{r,s} = \{v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes \dots \otimes f^{j_s}\}$  tal que  $h \equiv (h_{j_1, \dots, j_s}^{i_1, \dots, i_r})_{i_1, \dots, i_r, j_1, \dots, j_s}^n$*

**Demostración.** Tenemos que ver que los elementos de la base son linealmente independientes y para ello, se debe cumplir que

$$\sum_{\substack{i_1, \dots, i_r \\ j_1, \dots, j_s}}^n \lambda_{j_1, \dots, j_s}^{i_1, \dots, i_r} v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes \dots \otimes f^{j_s} = 0 \Leftrightarrow \lambda_{j_1, \dots, j_s}^{i_1, \dots, i_r} = 0$$

Luego, dados  $i_1^0, \dots, i_r^0, j_1^0, \dots, j_s^0$  índices fijos, y vamos a tomar

$$f \equiv \sum_{\substack{i_1, \dots, i_r \\ j_1, \dots, j_s}}^n \lambda_{j_1, \dots, j_s}^{i_1, \dots, i_r} (v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes \dots \otimes f^{j_s}) = \lambda_{j_1, \dots, j_s}^{i_1, \dots, i_r} (v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes \dots \otimes f^{j_s})$$

luego,

$$0 = f(f^{i_1^0}, \dots, f^{i_r^0}, v_{j_1^0}, \dots, v_{j_s^0}) = \lambda_{j_1, \dots, j_s}^{i_1, \dots, i_r} \left( f^{i_1^0}(v_{i_1}) \dots f^{i_r^0}(v_{i_r}) f^{j_1}(v_{j_1^0}) \dots f^{j_s}(v_{j_s^0}) \right)$$

sabemos que un elemento de la base de  $V$  con un elemento de la base de  $V^*$  cumple que

$$\begin{cases} f^i(v_j) & \stackrel{j \neq i}{=} 0 \\ f^i(v_i) & = 1 \end{cases}$$

luego, esto es una delta de Kronecker  $\delta_{i,j}$ , y entonces,

$$0 = \lambda_{j_1, \dots, j_s}^{i_1, \dots, i_r} \delta_{i_1}^{i_1^0} \dots \delta_{i_r}^{i_r^0} \delta_{j_1}^{j_1^0} \dots \delta_{j_s}^{j_s^0} = \lambda_{j_1^0, \dots, j_s^0}^{i_1^0, \dots, i_r^0} \Rightarrow \lambda_{j_1^0, \dots, j_s^0}^{i_1^0, \dots, i_r^0} = 0$$

y por tanto, los elementos son linealmente independientes. ✓

Ahora tenemos que comprobar que un elemento  $h \in \Omega^{r,s}(V)$  se puede escribir como combinación lineal de los elementos de la base, es decir,

$$h(g^{i_1}, \dots, g^{i_r}, w_{j_1}, \dots, w_{j_s}) = h_{j_1, \dots, j_s}^{i_1, \dots, i_r} (v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes \dots \otimes f^{j_s}) (g^{i_1}, \dots, g^{i_r}, w_{j_1}, \dots, w_{j_s})$$

Vamos a verlo:

**Nota 1.2.11.** Vamos a hacer primero una aclaración acerca de los índices. Cuando decimos que  $i_1$  va desde 1 hasta  $n$ , estamos diciendo que tenemos la sucesión  $1_1, 2_1, \dots, n_1$ . Por tanto, aunque en los sumatorios pongamos  $\sum_{i_1=1}^n a_{i_1}$ , lo correcto sería poner  $\sum_{i=1}^n a_{i_1}$ , pero esto puede llevar a confusión o a problemas cuando, por ejemplo,  $i_1$  no tenga los mismos elementos que  $i_7$ . Luego, para referirnos a un elemento  $i_k$ -ésimo, escribiremos  $a_{i_k} = \sum_{i_k=1}^n \gamma_{i_k} b_{i_k}$ .

Luego,  $\forall w_{j_1}, \dots, w_{j_s} \in V$  y  $\forall g^{i_1}, \dots, g^{i_r} \in V^*$ , tal que

$$\begin{aligned} w_{j_k} &= \sum_{j_k=1}^n \mu^{j_k} v_{j_k} = \mu^{j_k} v_{j_k} \\ g^{i_k}(v) &= \sum_{i_k=1}^n \mu_{i_k} f^{i_k}(v) = \mu_{i_k} f^{i_k}(v) \end{aligned}$$

Luego, tomando un  $h \in \Omega^{r,s}(V)$ , tal que

$$\begin{aligned} h(v_{j_p}) &= h_{j_p}; \quad h_{j_p} \in \mathbb{R} \quad p = 1, 2, \dots, s \\ h(f^{i_q}) &= h^{i_q}; \quad h^{i_q} \in \mathbb{R} \quad q = 1, 2, \dots, r \end{aligned}$$

Entonces,

$$\begin{aligned} h(g^{i_1}, \dots, g^{i_r}, w_{j_1}, \dots, w_{j_s}) &= h(g^{i_1}) \dots h(g^{i_r}) h(w_{j_1}) \dots h(w_{j_s}) = \\ &= h\left(\sum_{i_1=1}^n \mu^{i_1} f^{i_1}\right) \dots h\left(\sum_{i_r=1}^n \mu^{i_r} f^{i_r}\right) h\left(\sum_{j_1=1}^n \mu_{j_1} v_{j_1}\right) \dots h\left(\sum_{j_s=1}^n \mu_{j_s} v_{j_s}\right) = \\ &= \mu^{i_1} h(f^{i_1}) \dots \mu^{i_r} h(f^{i_r}) \mu_{j_1} h(v_{j_1}) \dots \mu_{j_s} h(v_{j_s}) = \\ &= \{\text{Podemos agrupar los escalares de tal forma que}\} = \\ &= \mu_{j_1, \dots, j_s}^{i_1, \dots, i_r} h(f^{i_1}) \dots h(f^{i_r}) h(v_{j_1}) \dots h(v_{j_s}) = \\ &= \mu_{j_1, \dots, j_s}^{i_1, \dots, i_r} h^{i_1} \dots h^{i_r} h_{j_1} \dots h_{j_s} = \mu_{j_1, \dots, j_s}^{i_1, \dots, i_r} h_{j_1, \dots, j_s}^{i_1, \dots, i_r} \end{aligned}$$

Usando que  $\mu_{j_1, \dots, j_s}^{i_1, \dots, i_r} = (v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes f^{j_s})(g^{i_1}, \dots, g^{i_r}, w_{j_1}, \dots, w_{j_s})$ , tenemos

$$h(g^{i_1}, \dots, g^{i_r}, w_{j_1}, \dots, w_{j_s}) = h_{j_1, \dots, j_s}^{i_1, \dots, i_r}(v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes f^{j_s})(g^{i_1}, \dots, g^{i_r}, w_{j_1}, \dots, w_{j_s})$$

Luego,  $h \in \Omega^{r,s}(V)$  es combinación lineal de los elementos de la base. ✓

Por tanto,  $B^{r,s} = \{v_{i_1} \otimes \dots \otimes v_{i_r} \otimes f^{j_1} \otimes \dots \otimes f^{j_s}\}$  es base de  $\Omega^{r,s}(V)$ . ■

**Corolario 1.2.12.** *Sea  $V$  un espacio vectorial con  $\dim(V) = n$ . Luego,  $\dim \Omega^{r,s}(V) = n^{r+s}$*

**Demostración.** Sabemos que

$$\Omega^{r,s}(V) \equiv V \otimes \dots \otimes V \otimes V^* \otimes \dots \otimes V^*$$

Por tanto, al ser  $\dim(V) = n$ , al tener  $V r$ -veces, tenemos que  $\dim(V \otimes \dots \otimes V) = n^r$  y como  $\dim(V) = \dim(V^*)$ , entonces tenemos que  $\dim(V^* \otimes \dots \otimes V^*) = n^s$ , por tanto,  $\dim \Omega^{r,s}(V) = n^r \cdot n^s = n^{r+s}$  ■

**Nota 1.2.13.** *El producto escalar es un tensor de tipo  $(0,2)$ , es decir,  $\langle \cdot, \cdot \rangle \in \Omega^{0,2}(V)$  tal que*

$$\langle \cdot, \cdot \rangle = \sum g_{ij} f^i \otimes f^j \equiv g_{ij} f^i \otimes f^j$$

Ahora vamos a ver el caso particular del producto escalar. El cuál es muy importante, pues nos permitirá escribir **tensores independientes de sus bases**.

**Nota 1.2.14.** *El producto escalar  $\langle \cdot, v \rangle$  es un tensor de tipo  $(0,1)$ , es decir,  $\langle \cdot, v \rangle \in \Omega^{0,1}(V)$  tal que*

$$\langle \cdot, v \rangle = g_{ij} (f^i \otimes f^j)(\cdot, v)$$

**Proposición 1.2.15.** *Dado  $\{v_1, v_2, \dots, v_n\}$  base de  $V$ , tenemos que  $\langle \cdot, v_1 \rangle, \dots, \langle \cdot, v_n \rangle$  es base de  $V^*$*

**Demostración.** Tenemos que comprobar que los elementos de la base son linealmente independientes, es decir,

$$\lambda_1 \langle \cdot, v_1 \rangle + \lambda_2 \langle \cdot, v_2 \rangle + \dots + \lambda_n \langle \cdot, v_n \rangle = 0 \Leftrightarrow \lambda_1 = \dots = \lambda_n = 0$$

Evaluando un  $v_k \in B$ , tenemos

$$0 = \lambda_1 \langle v_k, v_1 \rangle + \dots + \lambda_k \langle v_k, v_k \rangle + \dots + \lambda_n \langle v_k, v_n \rangle$$

Como los elementos de  $B$  son linealmente independientes, por la condición de base dual, se debe cumplir que

$$\langle v_i, v_j \rangle = \delta_{ij}$$

Entonces,

$$0 = \lambda_1 \langle v_k, v_1 \rangle^0 + \dots + \lambda_k \cancel{\langle v_k, v_k \rangle^1} + \dots + \lambda_n \langle v_k, v_n \rangle^0 = \lambda_k$$

Por tanto,  $\lambda_k = 0$ , luego son linealmente independientes. Además, como esta base es una base dual de  $V^*$ , tendrá la misma dimensión que  $V$ , por la Proposición 1.1.7 y así, el conjunto generador pasa a ser base. ■

Así, dado un espacio vectorial métrico no degenerado,  $(V, g)$  de dimensión finita, se puede establecer un isomorfismo entre  $V$  y su dual  $V^*$ , que **no depende de bases, sino solo de la métrica,  $g$** . Esto permite identificar de forma natural los vectores de  $V$  con las formas lineales sobre  $V$ .

**Nota 1.2.16.** *Vamos a identificar el producto escalar como  $\langle \cdot, v \rangle \equiv g(\cdot, v)$  para simplificar la notación.*

Ahora vamos a definir dos aplicaciones, **bemol** y **sostenido**, las cuáles nos permitirán 'bajar' o 'subir' índices, es decir, transformar el tipo del tensor.

**Definición 1.2.3.** *Definimos la aplicación **bemol** como,*

$$\begin{aligned} b : V &\rightarrow V^* \\ v &\mapsto v^b \end{aligned}$$

*la cuál nos sirve para 'bajar índices'*

**Proposición 1.2.17.** *Sea  $(V, g)$  un espacio vectorial dotado de una métrica no degenerada.*

1. *Para cada  $v \in V$  la aplicación*

$$v^b \equiv g(v, \cdot) : V \rightarrow \mathbb{R}$$

$$v^b(\omega) = g(v, \omega) \quad \forall \omega \in V$$

*es lineal, es decir,  $v^b \in V^*$ .*

2. La aplicación **bemol** ('bajar índices'), dada por

$$\begin{aligned}\flat : V &\rightarrow V^* \\ v &\mapsto v^\flat\end{aligned}$$

es un isomorfismo de espacios vectoriales.

**Demostración.** 1. Veamos que es aplicación lineal,

$$\forall v, w, u \in V \text{ y } \forall \lambda \in \mathbb{R},$$

$$(i) \quad v^\flat(w + u) \equiv g(v, w + u) = g(v, w) + g(v, u) = v^\flat(w) + v^\flat(u) \checkmark$$

$$(ii) \quad v^\flat(\lambda w) \equiv g(v, \lambda w) = \lambda g(v, w) = \lambda v^\flat(w) \checkmark$$

■

2. Debemos probar que es una aplicación lineal y biyectiva.

Vemos que es lineal por 1. ✓

Comprobamos que es biyectiva, es decir, es inyectiva y sobreyectiva. Comprobamos que es inyectiva viendo que el  $\text{Ker } \{\flat\} = \{0\}$ ,

$$\text{ker } \{\flat\} = \{\tilde{w} \in V; v^\flat(\tilde{w}) = 0\}$$

usando que  $v^\flat \equiv g(v, \cdot)$ , tenemos que  $0 = v^\flat(\tilde{w}) \equiv g(v, \tilde{w})$  y como  $v$  es cualesquiera  $v \in V$ , al ser  $g$  no degenerada, se cumple que  $g(v, \tilde{w}) = 0$  si y solo si  $\tilde{w} = 0$  ó  $v = 0$ , pero como  $v$  es arbitrario, nos quedamos con  $\tilde{w} = 0$  y por tanto,  $\text{ker } \{\flat\} = \{0\}$ , luego es inyectiva. ✓

Vemos que es sobreyectiva usando el Primer Teorema de Isomorfía, el cuál nos dice que

$$\dim(V) = \dim(\text{ker } \{\flat\}) + \dim(\text{Im } \{\flat\})$$

pero como  $\text{ker } \{\flat\} = \{0\}$ , entonces,  $\dim(\text{ker } \{\flat\}) = 0$  y además, vemos que  $\text{Im } \{\flat\} = V^*$  y sabemos que  $\dim(V) = \dim(V^*)$ , luego

$$\dim(V) = \underbrace{\dim(\text{ker } \{\flat\})}_0 + \dim(\text{Im } \{\flat\}) = \dim(V^*)$$

luego, al ser la imagen de la misma dimensión que el espacio de entrada, decimos que  $\flat$  es sobreyectiva. ✓

Por tanto, la aplicación  $\flat$  es un isomorfismo. ■

**Definición 1.2.4.** Definimos la aplicación **sostenido** como la inversa de la aplicación bemol (pues la aplicación bemol es biyectiva y por tanto tiene inversa), dada por

$$\begin{aligned}\sharp : V^* &\rightarrow V \\ v^\flat &\mapsto v\end{aligned}$$

la cual nos permite subir índices.

Una caracterización alternativa del sostenido es la siguiente.

**Proposición 1.2.18.** Sea  $(V, g)$  un espacio vectorial métrico no degenerado de dimensión finita y  $\phi \in V^*$ . Entonces  $\phi^\sharp$  es el único vector que verifica,

$$g(\phi^\sharp, v) = \phi(v), \quad \forall v \in V$$

**Demostración.** Aplicando las definiciones,

$$g(\phi^\sharp, v) = (\phi^\sharp)^\flat(v) = \phi(v)$$

Además, si otro vector  $u_\phi \in V$  verificara esa relación, se tendría

$$g(\phi^\sharp - u_\phi, v) = \phi(v) - \phi(v) = 0,$$

y, al ser no degenerada,  $\phi^\sharp - u_\phi = 0 \Rightarrow u_\phi = \phi^\sharp$ . ■

**Nota 1.2.19.** Sea  $w = w^i v_i \equiv w^i$  un vector de  $V$  y sea  $g(\cdot, w)$  una 1-forma métrica asociada, tal que  $g(\cdot, w) = w_j f^j \equiv w_j$ . Se tiene entonces que

$$w_j = g_{ij} w^i \quad y \quad w^i = g^{ij} w^i$$

A esto lo denominamos **subida** y **bajada** de índices (métrica).

Además, usaremos la base de productos escalares, pues

$$V \longleftrightarrow V^*$$

$$v \rightarrow v^* \text{ depende de la base } B^*$$

$$v \rightarrow g(\cdot, v) \text{ no depende de la base } B^*, \text{ sino de la métrica } g$$

**Nota 1.2.20.** Los nombres 'subir y bajar' índices provienen de la Relatividad General. En Mecánica Cuántica, Dirac introdujo una nomenclatura distinta. Considerando un espacio vectorial euclídeo (y, con más generalidad, un espacio de Hilbert) con producto escalar  $\langle \cdot, \cdot \rangle$ ; los vectores  $v$  y  $w$  se denotan como un 'ket',  $v \equiv |v\rangle$ ,  $w \equiv |w\rangle$  y sus bemoles como un 'bra',  $v^\dagger \equiv \langle v|$ ,  $w^\dagger \equiv \langle w|$ , de modo que  $v^\dagger(w)$  es el 'braket'  $\langle v|w\rangle$ .

### 1.2.3. Contracción de tensores

Una vez que hemos visto cómo subir y bajar índices, podemos definir una operación denominada **contracción** de tensores, la cuál encoge un tensor  $(r, s)$  a uno  $(r - 1, s - 1)$ . La definición general se obtiene a partir del siguiente caso especial.

**Lema 1.2.21.** Hay una única aplicación lineal  $C : \Omega_1^1(V) \rightarrow \mathbb{R}$  llamada contracción  $(1,1)$ , tal que

$$\begin{aligned} C : \quad \Omega_1^1(V) &\rightarrow \mathbb{R} \\ v \otimes f &\mapsto C(v \otimes f) = f(v) \end{aligned}$$

para todo  $v \in V$  y  $f \in V^*$ .

**Demostración.** (Esta demostración usa el concepto de matrices de cambio de base, por lo que se recomienda ver la sección 1.2.4)

Tomando  $B = \{v^1, v^2, \dots, v^n\}$  base de  $V$  y  $B^* = \{f_1, f_2, \dots, f_n\}$  base de  $V^*$ , podemos escribir un tensor de tipo  $(1, 1)$  como

$$A \equiv \sum A_j^i f_i \otimes v^j$$

Como  $C(f_i \otimes v^j) = f_i(v^j) = \delta_i^j$ , por la condición de base dual, no nos queda otra opción, más que definir,

$$C(A) = \sum A_i^i = \sum A(f_i, v^i)$$

Entonces,  $C$  tiene las propiedades requeridas en las bases  $B, B^*$ . Luego, para obtener la función general requerida es suficiente con mostrar que esta definición es independiente de la elección del sistema de coordenadas. Así, tomando una nueva base de  $V$ ,  $B' =$

$\{w^1, w^2, \dots, w^n\}$  y otra de  $V^*$ ,  $B^{*'} = \{q_1, q_2, \dots, q_n\}$ , tenemos

$$\begin{aligned} C(A) &= \sum_m A(q_m, w^m) = \sum_m A\left(\sum_i a_i^m f_m, \sum_j b_j^m v^m\right) \\ &= \sum_{i,j,m} a_i^m b_j^m A(f_i, v^j) = \sum_{i,j} \delta_i^j A(f_i, v^j) \\ &= \sum_i A(f_i, v^j) \end{aligned}$$

■

Para extender las contracciones  $(1,1)$ ,  $C$ , a un tensor de un tipo mayor, el esquema es especificar una componente covariante y otra contravariante y aplicar  $C$  a estos.

Suponemos un tensor  $A \in \Omega_r^s(V)$  y  $1 \leq r$  y  $1 \leq j \leq s$ . Fijamos las formas  $p_1, p_2, \dots, p_{r-1}$  y los vectores  $u_1, u_2, \dots, u_{s-1}$ . Entonces la función

$$(p, u) \rightarrow A(p_1, \dots, \underbrace{p_i, \dots, p_{r-1}}_{i\text{-ésima componente contravariante}}, u^1, \dots, \overbrace{u^j, \dots, u^{s-1}}^{j\text{-ésima componente covariante}})$$

es un tensor  $(1,1)$  que puede escribirse como

$$A(p_1, \dots, \cdot, \dots, p_{r-1}, u^1, \dots, \cdot, \dots, u^{s-1})$$

Aplicando la contracción  $(1,1)$  a este tensor, produce una función de valor real denotada por

$$(C_j^i A)(p_1, \dots, p_{r-1}, u^1, \dots, u^{s-1})$$

Siendo  $C_j^i A$  una función multilineal. Por tanto, esto es un tensor de tipo  $(r-1, s-1)$  llamado *la contracción de A sobre i, j*.

**Nota 1.2.22.** *Para poder contraer tensores, debemos tener superíndices y subíndices, así, podemos usar primero la métrica para subir o bajar índices y luego aplicar la contracción.*

**Ejemplo 1.2.23.** *Si tenemos un tensor de tipo  $(0,2)$ ,  $S \equiv S_{\alpha\beta}$ , podemos hacer,*

$$\begin{aligned} S_{\alpha\beta} &\rightarrow g^{\gamma\alpha} S_{\alpha\beta} = S_\beta^\gamma \rightarrow C_1^1 S_\beta^\gamma = S_\beta^\beta \\ \text{Tensor } (0,2) &\rightarrow \text{Tensor } (1,1) \rightarrow \text{Escalar} \end{aligned}$$

cosa que se puede simplificar simplemente usando,

$$S \equiv S_{\alpha\beta} \rightarrow g^{\beta\alpha} S_{\alpha\beta} = S_{\beta}^{\beta}$$

es decir, podemos contraer tensores con la propia métrica.

**Ejemplo 1.2.24.** Si

$$U_i^j = T_{ik}^{kj}$$

entonces

$$U_{i'}^{j'} = T_{i'k'}^{k'j'} = S_{i'}^i S_{k'}^l R_k^{k'} R_j^{j'} T_{il}^{kj} = S_{i'}^i \delta_k^l R_j^{j'} T_{il}^{kj} = S_{i'}^i R_j^{j'} T_{ik}^{kj} = S_{i'}^i R_j^{j'} U_i^j$$

donde hemos utilizado  $S_{k'}^l R_k^{k'} = \delta_k^l$ . Vemos que se transforma como un tensor (1,1).

Así, dado un tensor  $T_{kl}^{ij}$  de tipo (2,2), serán posibles las 4 contracciones

$$T_{ki}^{kj}, \quad T_{ik}^{jk}, \quad T_{ik}^{kj}, \quad T_{ki}^{jk}$$

que originan 4 tensores de tipo (1,1). Por otro lado, las dos posibles contracciones dobles que dan lugar a un escalar (tensor de tipo (0,0)) son

$$T_{kj}^{kj}, \quad T_{kj}^{jk}$$

**Nota 1.2.25.** El producto escalar  $(\mathbb{R}^n, g_{ij})$  también se puede contraer. Pues  $g_{ij}$  es un tensor de tipo (0,2), al cual le podemos aplicar una contracción 1,1, pero primero lo pasamos a un tensor de tipo (1,1), variando sus índices, tal que

$$C_1^1(g^{ki}g_{ij}) = C_1^1(g_j^k) = g_j^j = n$$

donde sabemos que vale  $n$ , pues al ser un espacio de dimensión  $n$ , la matriz asociada a  $g$  será  $G \in \mathcal{M}_{n \times n}$  y por tanto, la traza será la suma de  $n$ -elementos. Sabemos que estos elementos son el 1, porque la traza es invariante frente a los cambios de base (cosa que veremos más adelante), por tanto, si cogemos el producto escalar usual en la base usual, la matriz asociada es la matriz de Gram, cuyos elementos son todos nulos, salvo la diagonal que está formada por 1.

### 1.2.4. Leyes de transformación

Las leyes de transformación, también se conocen como *cambios de bases* o *cambios de variables*. Vamos a ver cómo son estos cambios de base en los tensores, pero primero veremos los casos particulares de **vectores, formas y tensores de tipo (1,1)**.

#### Cambio de base de vectores

Queremos construir una matriz que nos permita cambiar las coordenadas de un vector en una base por las coordenadas del mismo vector en otra base, donde tomamos la función identidad.

Sea un espacio vectorial  $V$  con bases  $B = \{v_1, \dots, v_n\}$  y  $\tilde{B} = \{\tilde{v}_1, \dots, \tilde{v}_n\}$ . Consideramos un vector  $w \in V$ , que podemos escribir en ambas bases como,

$$w = x^1 v_1 + \dots + x^n v_n$$

$$w = \tilde{x}^1 \tilde{v}_1 + \dots + \tilde{x}_n \tilde{v}_n$$

Como el vector es el mismo, podemos igualar ambas expresiones,

$$x^1 v_1 + \dots + x^n v_n = \tilde{x}^1 \tilde{v}_1 + \dots + \tilde{x}_n \tilde{v}_n$$

o bien,

$$x^i v_i = \tilde{x}^i \tilde{v}_i$$

Expresando los elementos de la base  $B$  en función de los de la base  $B'$  tenemos,

$$\left\{ \begin{array}{l} v_1 = a_1^1 \tilde{v}_1 + \dots + a_1^n \tilde{v}_n \\ v_2 = a_2^1 \tilde{v}_1 + \dots + a_2^n \tilde{v}_n \\ \vdots \\ v_n = a_n^1 \tilde{v}_1 + \dots + a_n^n \tilde{v}_n \end{array} \right.$$

o bien,

$$v_i = a_i^j \tilde{v}_j$$

donde hemos agrupado las constantes en los diferentes  $a_i^j$ .

Sustituyendo,

$$\begin{aligned} x^1 v_1 + \dots + x^n v_n &= x^1 a_1^1 \tilde{v}_1 + \dots + x^1 a_1^n \tilde{v}_n + \dots + x^n a_n^1 \tilde{v}_1 + \dots + x^n a_n^n \tilde{v}_n = \\ &= (x^1 a_1^1 + \dots + x^n a_n^1) \tilde{v}_1 + \dots + (x^1 a_1^n + \dots + x^n a_n^n) \tilde{v}_n = \\ &= \tilde{x}^1 \tilde{v}_1 + \dots + \tilde{x}^n \tilde{v}_n \end{aligned}$$

Luego, lo que acompaña a cada vector deberá ser igual, es decir,

$$\begin{cases} \tilde{x}^1 = x^1 a_1^1 + \dots + x^n a_n^1 \\ \vdots \\ \tilde{x}^n = x^1 a_1^n + \dots + x^n a_n^n \end{cases}$$

o escrito en forma matricial como,

$$\begin{pmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^n \end{pmatrix} = \begin{pmatrix} a_1^1 & \dots & a_1^n \\ \vdots & \ddots & \vdots \\ a_n^1 & \dots & a_n^n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$$

donde la matriz

$$a_i^j = \begin{pmatrix} a_1^1 & \dots & a_1^n \\ \vdots & \ddots & \vdots \\ a_n^1 & \dots & a_n^n \end{pmatrix}$$

es la denominada **matriz de cambio de base de  $B$  en  $B'$** , pues estamos transformando las coordenadas del vector  $w$  en la base  $B$ , que son  $w = (x^1, \dots, x^n)$ , en las coordenadas en la base  $B'$ , que son  $w = (\tilde{x}^1, \dots, \tilde{x}^n)$ .

Este caso lo hemos hecho detallando los pasos a seguir. Los siguientes cambios de base lo haremos con notación de Einstein.

### Cambio de base de formas

Vamos a usar ahora notación de Einstein para simplificar la notación.

Sea  $V^*$  el espacio dual de un espacio vectorial  $V$ , con bases duales  $B^* = \{f^1, \dots, f^n\}$  y  $\tilde{B}^* = \{\tilde{f}^1, \dots, \tilde{f}^n\}$ . Consideremos un elemento  $g \in V^*$ , que podemos escribir en ambas bases como,

$$g = x_i f^i$$

$$g = \tilde{x}_i \tilde{f}^i$$

que podemos igualar,

$$x_i f^i = \tilde{x}_i \tilde{f}^i$$

Expresando los elementos de la base  $B^*$  en función de los de la base  $B^{*'}$  tenemos,

$$\left\{ f^i = b_j^i \tilde{f}^j \right.$$

Sustituyendo,

$$x_i f^i = x_i b_j^i \tilde{f}^j = \tilde{x}_i \tilde{f}^i$$

Luego, lo que acompaña a cada vector deberá ser igual, es decir,

$$\left\{ \tilde{x}_i = \sum_j a_i^j x_j \equiv b_i^j x_j \right.$$

o escrito en forma matricial como,

$$\begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & \ddots & \vdots \\ b_1^n & \dots & b_n^n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

donde la matriz

$$b_j^i = \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & \ddots & \vdots \\ b_1^n & \dots & b_n^n \end{pmatrix}$$

es la denominada **matriz de cambio de base de  $B^*$  en  $B'$** , pues estamos transformando las coordenadas de la aplicación lineal  $g$  en la base  $B^*$ , que son  $\{x_i\}$ , en las coordenadas en la base  $B^{*'}$ , que son  $\{\tilde{x}_i\}$ .

**Observación 1.2.26.** Sabemos que por la condición de base dual tenemos,

$$f^i(v_j) = \delta_i^j, \quad \tilde{f}^i(\tilde{v}_j) = \delta_i^j$$

Luego,

$$\begin{aligned} \delta_i^j &= f^i(v_j) = f^i(a_j^k \tilde{v}_k) = a_j^k f^i(\tilde{v}_k) = a_j^k (b_l^i \tilde{f}^l)(\tilde{v}_k) \\ &= a_j^k b_l^i \tilde{f}^l(\tilde{v}_k) = a_j^k b_l^i \delta_k^l = a_j^k b_k^i \end{aligned}$$

Por tanto, tenemos que  $a_j^k b_k^i = \delta_i^j$ , luego, podemos decir que una es la inversa de la otra.

### Cambio de base en tensores de tipo (1,1)

Tomando los espacios vectoriales y bases anteriores. Sea  $\Omega^{1,1}(V)$  un espacio vectorial con bases  $B^{1,1} = \{v_i \otimes f^j\}$  y  $\tilde{B}^{1,1} = \{\tilde{v}_l \otimes \tilde{f}^k\}$ . Tomando un elemento  $w \in \Omega^{1,1}(V)$  que lo escribimos en función de ambas bases,

$$w = w_j^i v_i \otimes f^j$$

$$w = \tilde{w}_l^k \tilde{v}_l \otimes \tilde{f}^k$$

Vamos a partir de las matrices anteriores, tal que

$$w_j^i v_i \otimes f^j = w_j^i (a_i^r \tilde{v}_r) \otimes (b_l^i \tilde{f}^l) = w_j^i (a_i^r b_l^j) \tilde{v}_r \otimes \tilde{f}^l$$

Luego, si anulamos el  $w_j^i$ , tenemos

$$v_i \otimes f^j = (a_i^r b_l^j) \tilde{v}_r \otimes \tilde{f}^l \equiv w_j^i = (a_i^r b_l^j) \tilde{w}_l^r$$

donde  $a_i^r$  y  $b_l^j$  son matrices de cambio de base, tanto de vectores como de formas y  $A \equiv a_i^r b_k^j$  será la **matriz de cambio de base de los tensores de tipo (1,1)**.

### Cambio de base en tensores de tipo (r,s)

Tomando los espacios y bases anteriores. Sea  $\Omega^{r,s}(V)$  un espacio vectorial, con bases  $B^{r,s} = \{v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes f^{j_1} \otimes \cdots \otimes f^{j_s}\}$  y  $\tilde{B}^{r,s} = \{\tilde{v}_{l_1} \otimes \cdots \otimes \tilde{v}_{l_r} \otimes \tilde{f}^{k_1} \otimes \cdots \otimes \tilde{f}^{k_s}\}$ . Repitiendo todo lo anterior, podemos ponerlo en forma general, tal que

$$\tilde{\Omega}_{i_1, \dots, i_k}^{j_1, \dots, j_s} = \Omega_{r_1, \dots, r_k}^{l_1, \dots, l_s} a_{l_1}^{j_1} a_{l_2}^{j_2} \dots a_{l_s}^{j_s} b_{i_1}^{r_1} b_{i_2}^{r_2} \dots b_{i_k}^{r_k}$$

donde  $M \equiv a_{l_1}^{j_1} a_{l_2}^{j_2} \dots a_{l_s}^{j_s} b_{i_1}^{r_1} b_{i_2}^{r_2} \dots b_{i_k}^{r_k}$  es la **matriz de cambio de base de los tensores de tipo (r,s)**.

### 1.2.5. Invariantes

Dado que los tensores suelen describirse en términos respecto de ciertas bases, cuando estos términos no dependen de la base empleada, los tensores se llamarán **invariantes**. O en otras palabras, los tensores que no se transforman frente a un cambio de base, serán los que llamaremos **invariantes**.

Vamos a intentar ilustrar este concepto definiendo un tensor invariante de tipo (1,1), denominado *traza*, que es un invariante conocido de las matrices. Si tenemos un tensor  $A = A_j^i e_i \otimes f^j$  que definimos como

$$\text{traza de } A = \text{tr}A = A_i^i$$

siendo la suma de los elementos de la diagonal principal de la matriz  $(A_j^i)$ . No es a priori evidente que hayamos definido algo que depende únicamente de  $A$ , ya que los  $A_j^i$  dependen no solo de  $A$  sino también de la base  $\{e_i\}$ . Para mostrar que  $\text{tr}A$  es un número determinado enteramente por  $A$  mismo y no por los  $e_i$  también, debemos demostrar la invariancia; es decir, si  $A$  se expresa en términos de otra base  $\tilde{e}_i$ , entonces la fórmula correspondiente en los nuevos componentes da el mismo número que antes. Así, escribimos  $A = \tilde{A}_j^i \tilde{e}_i \otimes \tilde{f}^j = A_j^i e_i \otimes f^j$  y veremos que  $A_j^i = \tilde{A}_j^i$ . Usando la misma notación de cambios de base que hemos visto en el apartado anterior, tenemos la ley de transformación siguiente,

$$\tilde{A}_m^n = A_j^i a_m^j b_i^n$$

de lo cual se obtiene

$$\tilde{A}_i^i = A_j^p a_i^j b_p^i = A_j^p \delta_p^j = A_i^i$$

Queda demostrado. Luego, tenemos la proposición,

**Proposición 1.2.27.** *La traza de un tensor de tipo (1,1) es un invariante.*

Para ver que no todas las expresiones en términos de las componentes de un tensor necesariamente serán un invariante, veamos el siguiente ejemplo.

**Ejemplo 1.2.28.** *Supongamos  $d = 2$  y  $A = e_1 \otimes e_1 + e_1 \otimes e_2$ , un tensor de tipo (0,2). La expresión de  $A_{ii}$  en este caso será  $A_{11} + A_{22} = 1 + 0 = 1$ . Ahora consideramos una nueva base dada por  $e_1 = \tilde{e}_1 + \tilde{e}_2$  y  $e_2 = \tilde{e}_2$ , entonces*

$$\begin{aligned} A &= (\tilde{e}_1 + \tilde{e}_2) \otimes (\tilde{e}_1 + \tilde{e}_2) + (\tilde{e}_1 + \tilde{e}_2) \otimes \tilde{e}_2 \\ &= \tilde{e}_1 \otimes \tilde{e}_1 + 2\tilde{e}_1 \otimes \tilde{e}_2 + \tilde{e}_2 \otimes \tilde{e}_1 + 2\tilde{e}_2 \otimes \tilde{e}_2 \end{aligned}$$

de la cuál se obtiene que  $\tilde{A}_{ii} = \tilde{A}_{11} + \tilde{A}_{22} = 1 + 2 = 3$ . Por tanto es diferente a la base primera, luego no es un invariante.

### **Nota Final**

Finalmente diremos que un tensor es todo aquel objeto matemático que satisfaga los cambios de base, o en otras palabras: *Un tensor es todo objeto matemático que transforma como un tensor.*

## Capítulo 2

# Teoría de la Relatividad Especial

*“Hay una fuerza motriz más poderosa que el vapor, la electricidad y la energía atómica: la voluntad”.*

(A. Einstein)

## 2.1. Repaso histórico

La física clásica, del siglo XIX, era una física bien asentada. La cuál explica la mecánica con el libro de Sir Isaac Newton titulado *Philosophiae Naturalis Principia Mathematica* y el electromagnetismo se explica con el libro de Maxwell titulado *Electricity and Magnetism*.

En 1887, Michelson y Morley iniciaron una revolución en la física con un experimento para medir la velocidad de la luz. El experimento consistía en medir la velocidad de la luz de un rayo paralelo al eje de rotación de la Tierra y de otro rayo perpendicular a este, esperándose obtener resultados diferentes. En cambio, se observó que ambos rayos iban exactamente igual, cosa que no tenía sentido en la época., por tanto, determinaron que la velocidad de la luz no era instantánea, sino que debía ser finita, y llegaron a un resultado de ésta bastante próximo al valor actual de la velocidad de la luz.

### 2.1.1. Relatividad Galileana

El Principio de Relatividad de Galileo establece que,

*"Es imposible determinar a base de experimentos (mecánicos) si un sistema de referencia está en reposo o en movimiento uniforme y rectilíneo".*

Esto se derivó de que en la Relatividad Galileana hay un espacio absoluto en el que las leyes de Newton son ciertas. Definiremos un *sistema de referencia inercial* (SRI) como aquel sistema referencia que se mueve a velocidad constante respecto al espacio absoluto. Además, todos los sistemas de referencia iniciales comparten un tiempo absoluto. Pero con la definición de SRI, el Principio de Relatividad se debe reformular con este concepto, así tenemos el Principio de Relatividad en formulación de equivalencia, que dice que

*"Todos los sistemas inerciales son equivalentes, es decir, todos los observadores inerciales ven la misma física".*

### Leyes de Newton

La Ley de Newton por excelencia es  $\vec{F} = m\vec{a} = -\nabla V(\vec{r} - \vec{r}_0)$ , donde  $V$  es la función potencial. Esta ley (y las demás) transforman bien bajo el grupo de transformaciones de Galileo, que son:

**1. Traslaciones temporales:**

$$t \rightarrow t' = t + t_0$$

**2. Traslaciones espaciales:**

$$\vec{r} \rightarrow \vec{r}' = \vec{r} + \vec{r}_i + \vec{v}t$$

donde  $\vec{v}$  es la velocidad relativa de un SRI con respecto al otro, y  $\vec{r}_i$  es el vector de posición entre los orígenes de ambos SRI al inicio.

**3. Rotaciones espaciales:**

$$\vec{a}' = R(\theta)\vec{a}$$

donde  $R(\theta)$  es la matriz de rotación.

Se puede ver que las Leyes de Newton no son covariantes, pero sí transforman bien, pues la física se mantiene, esto quiere decir que *las Leyes de Newton de la física transforman de forma covariante*.

El grupo de transformaciones de Galileo son simetrías que dan lugar a cantidades conservadas. Por tanto, si tenemos un Lagrangiano que sea invariante bajo traslaciones temporales, tendremos que el sistema conserva energía; si es invariante bajo traslaciones espaciales, conserva momento lineal; y si es invariante bajo rotaciones espaciales; conserva momento angular.

El grupo de transformaciones de Galileo NO deja invariante las ecuaciones de Maxwell, que son

$$(i) \quad \nabla \cdot \vec{E} = \rho/\epsilon_0; \quad (iii) \quad \nabla \cdot \vec{B} = 0$$

$$(ii) \quad \nabla \times \vec{B} = \partial_t \vec{E}/c^2 + \mu_0 \vec{J}; \quad (iv) \quad \nabla \times \vec{E} = -\partial_t \vec{B}$$

Si  $\rho = 0$  y  $\vec{J} = 0$ , es decir, estamos en vacío, podemos combinar las ecuaciones de Maxwell en una sola ecuación de ondas que se propaga a velocidad  $c = 299792,458$  m/s, resultado muy próximo al valor obtenido por Michelson y Morley, que además es independiente del sistema de referencia.

### 2.1.2. Transformaciones de Lorentz

Las transformaciones de Lorentz hacen que las ecuaciones de Maxwell transformen bien (sean covariantes). Estas transformaciones son:

$$\begin{aligned} (i) \quad t' &= \gamma \left( t - \frac{v}{c^2} x \right); & (iii) \quad y' &= y \\ (ii) \quad x' &= \gamma \left( x - vt \right); & (iv) \quad z' &= z \end{aligned}$$

donde  $v$  es la velocidad relativa entre SRI (que suponemos que se mueven en el eje  $X$ ), y  $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ .

Como estas transformaciones hacen que las leyes de Maxwell sean covariantes, diremos que las transformaciones de Lorentz sean más fundamentales que las transformaciones de Galileo.

Además, vemos que por la transformación (i) el tiempo ya **no es absoluto**, sino que depende del SRI, por lo que diremos que el tiempo es **relativo**.

## 2.2. Postulado de la Relatividad Especial

El razonamiento de Einstein fue era que, si todos los sistemas inerciales son equivalentes y todos los observadores inerciales ven la misma física, entonces todos estos observadores deben llegar a las mismas leyes de la física, cuando apuntan los resultados de sus experimentos. En otras palabras,

*"Las leyes de la física deben tener la misma forma en todos los sistemas inerciales".*

Formulado de esta manera, el Principio de Relatividad impone ciertas condiciones sobre la forma de las leyes de la física. Afirma que existen unas transformaciones, llamadas *cambios de coordenadas*, que relacionan las cantidades físicas medidas por un observador con las de otro. Para no salir de los SRI, estos cambios de coordenadas tienen una forma específica y estas transformaciones tienen la estructura matemática de un grupo. La formulación de Einstein del Principio de Relatividad implica por tanto, que las leyes de la física tienen que ser tales que mantienen la misma forma tras hacer un cambio de coordenadas entre dos sistemas inerciales. En otras palabras, las leyes de la física deben ser invariantes y las cantidades físicas que aparecen en estas leyes se tienen que transformar

adecuadamente bajo las transformaciones de ese grupo. Así, llegamos a la formulación del Primer Postulado de la Relatividad Especial.

### Primer Postulado: Principio de Relatividad

*"Todas las leyes físicas, en ausencia de fuerzas de gravedad, son idénticas en todos los sistemas de referencia inerciales".*

Donde se excluye la gravedad porque si se tiene en cuenta, provocaría que existieran fuerzas de marea que impedirían la definición del SRI.

Por otro lado, Einstein también se dio cuenta de la invariancia de la velocidad de la luz, cosa que recogió también en el Segundo Postulado de la Relatividad Especial.

### Segundo Postulado

*"La velocidad de la luz en vacío,  $c$ , es la misma en todos los sistemas de referencia inerciales, es decir, es una constante universal".*

Esto lleva también a un grado de profundidad muy elevado sobre la naturaleza del propio espacio. Pues Newton postuló que el espacio y el tiempo son absolutos, pero hemos visto en las transformaciones de Lorentz que el tiempo ya no es absoluto, sino que es relativo; por lo que, usando que velocidad=espacio/tiempo; sabiendo que la velocidad de la luz es absoluta, tenemos que velocidad(absoluta)=espacio(?) / tiempo(relativo), por lo que no queda otra de que el espacio no sea absoluto, sino que también sea **relativo**.

La Mecánica Clásica se recupera en el límite de  $c \rightarrow \infty$ , es decir, en el límite de que las interacciones sean instantáneas.

#### 2.2.1. Aspectos a tener en cuenta

Definimos un **evento** como un instante en el tiempo  $t$  y una posición en el espacio  $\vec{x}$ .

Como ya hemos dicho, en el límite  $c \rightarrow \infty$  se recupera la mecánica clásica, pero además, este límite lleva a que  $v/c \ll 1$ , por lo que si introducimos este límite en las transformaciones de Lorentz, recuperaremos las transformaciones de Galileo.

La simultaneidad no es absoluta, cosa que puede verse con un experimento mental: imaginemos dos observadores, uno que está en reposo respecto al otro que está en un tren infinitamente largo que se mueve a una velocidad cercana a la de la luz. Ahora imaginemos que caen dos rayos distanciados considerablemente cerca de las vías del tren. Como es lógico, el observador en reposo observa que ambos rayos caen de forma simultánea. Ahora bien, el observador dentro del tren no los verá simultáneos, sino que verá que cae primero el rayo hacia donde se aproxima y después cae el rayo del que se aleja, pues suponiendo que la velocidad de la luz es finita, la onda del primer rayo le llegará antes que la onda del segundo rayo. Esto implica que ambos observadores vean distintos sucesos, cosa que implica que la simultaneidad se rompa y ya no sea absoluta.

Esto quiere decir, que si ponemos  $\Delta t = t_2 - t_1 = 0$ , entonces  $\Delta t' \neq 0$ , pues

$$t'_2 - t'_1 = \Delta t' = \gamma(\Delta t - (v/c^2)\Delta x) = \gamma(-(v/c^2)\Delta x) \neq 0$$

Esto da lugar a diversas paradojas.

### 2.2.2. Adición de velocidades

En la Relatividad Galileana, la adición de velocidades es

$$\vec{V}' = \vec{v} + \vec{V}$$

donde  $\vec{v}$  es la velocidad relativa a la que se mueve el SRI  $S'$  con respecto a  $S$ , y  $\vec{V}$  y  $\vec{V}'$  son las velocidades de, por ejemplo, una partícula medidas en  $S$  y  $S'$ , respectivamente. Pero en Relatividad Especial cambia, pues deberemos usar las transformaciones de Lorentz generalizadas, que son,

$$t' = \gamma \left( t - \frac{\vec{v} \cdot \vec{x}}{c^2} \right); \quad \vec{x}' = \vec{x} + (\gamma - 1)(\hat{v} \cdot \vec{x})\hat{v} - \gamma \vec{v}t$$

donde  $\hat{v} = \vec{v}/|\vec{v}|$ . Ahora supongamos que una partícula se mueve con velocidad  $\vec{V} = \frac{d\vec{x}}{dt}$  en un SRI  $S$ , y con velocidad  $\vec{V}' = \frac{d\vec{x}'}{dt'}$  en otro SRI  $S'$ , por tanto, la transformación será,

$$\vec{V}' = \frac{\vec{V} + (\gamma - 1)(\hat{v} \cdot \vec{V})\hat{v} - \gamma \vec{v}}{\gamma \left( 1 - \frac{\vec{v} \cdot \vec{V}}{c^2} \right)}$$

Además, si descomponemos las velocidades en la parte paralela y perpendicular al velocidad del SRI, es decir,

$$\vec{V} = V_{||} \hat{v} + \vec{V}_{\perp}; \quad \vec{V}' = V'_{||} \hat{v} + \vec{V}'_{\perp}$$

entonces tendremos que

$$V'_{||} = \frac{V_{||} - |\vec{v}|}{\gamma \left(1 - \frac{|\vec{v}|V_{||}}{c^2}\right)}, \quad \vec{V}'_{\perp} = \frac{\vec{V}_{\perp}}{\gamma \left(1 - \frac{|\vec{v}|V_{||}}{c^2}\right)}$$

## 2.3. Dilatación temporal

La dilatación temporal es una causa directa de los postulados de Einstein. Veámoslo con un esquema,

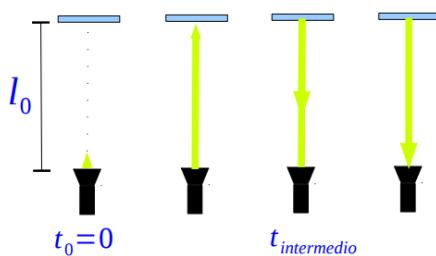


FIGURA 2.1: Espejos en reposo.

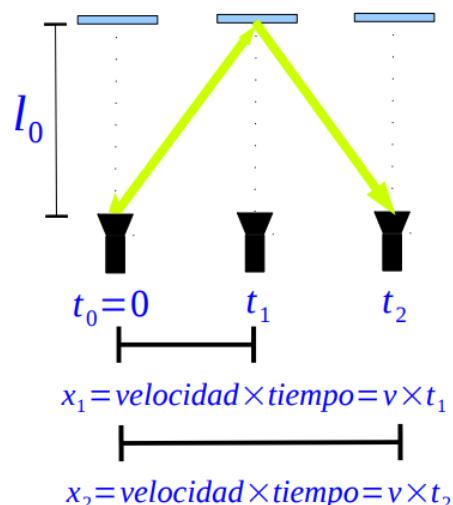


FIGURA 2.2: Espejos en movimiento.

Si nos fijamos en la Figura 2.1, al estar los espejos en reposo, el rayo de luz que sale de la linterna vuelve en un tiempo  $\Delta t = \frac{2l_0}{c}$ . En cambio, suponiendo que los espejos se mueven a velocidad  $\vec{v}$ , y que la distancia de los brazos del rayo es  $D$ , entonces ahora el tiempo que tarda el rayo en ir y volver es  $\Delta t' = \frac{2D}{c}$ . Usando el Teorema de Pitágoras podemos calcular  $D$ , tal que

$$D^2 = l_0^2 + \left(\frac{\Delta t' v}{2}\right)^2$$

Sustituyendo  $D$  y  $l_0$  de las ecuaciones de  $\Delta t$  y  $\Delta t'$ , tenemos

$$\left(\frac{\Delta t' c}{2}\right)^2 = \left(\frac{\delta t' v}{2}\right)^2 + \left(\frac{\Delta t c}{2}\right)^2$$

Por tanto, tenemos que el tiempo se dilata de la forma,

$$\Delta t' = \gamma \Delta t$$

y como  $\gamma > 1$  siempre, entonces  $\Delta t' > \Delta t$ , por eso se dilata el tiempo.

Vemos que en el SRI  $S'$  los relojes van más lento que en el SRI  $S$ , pues si consideramos como reloj el rebote de los fotones en los espejos, entonces en  $S$  los fotones van más rápido que los fotones en  $S'$ .

## 2.4. Contracción de longitudes

Tomamos dos eventos del espacio-tiempo, tal que

$$\Delta t' = \gamma \left( \Delta t - \frac{v}{c^2} \Delta x \right)$$

$$\Delta x' = \gamma (\Delta x - v \Delta t)$$

Asumimos que tomamos eventos que no están separados temporalmente, es decir, como si en  $S'$  tomásemos una foto, así,  $\Delta t' = 0$ . Por tanto, tendremos que  $\Delta x' = L'$  y  $\Delta x = L$ . Luego, sustituyendo tenemos que

$$\Delta x' = \frac{\Delta x}{\gamma} \implies L' = \frac{L}{\gamma}$$

Además, como  $\gamma > 1$ , tendremos que  $L > L'$ , por tanto, se habla de contracción de longitudes; donde  $L$  se conoce como **longitud propia**, que es la longitud del objeto respecto a un SRI en reposo respecto al objeto, es decir, el SRI centro de masas del objeto.

**Nota 2.4.1.** *Las transformaciones de Lorentz dejan invariante las distancias espaciotemporales, pues dados dos eventos  $(t_1, x_1)$  y  $(t_2, x_2)$  en  $S$ , y los eventos correspondientes  $(t'_1, x'_1)$  y  $(t'_2, x'_2)$  en  $S'$ , entonces*

$$-c^2(t_2 - t_1)^2 + (x_2 - x_1)^2 = -c^2(t'_2 - t'_1)^2 + (x'_2 - x'_1)^2$$

por tanto, tenemos una cantidad que es invariante al SRI.

## 2.5. Elemento de línea finito

Definimos el elemento de línea finito usando el concepto de la nota anterior, tal que

$$\Delta s^2 = -c^2 \Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 = -c^2 \Delta t^2 + \Delta \vec{x}^2$$

Así,

- si  $\Delta s^2 < 0$ , diremos que los eventos están separados temporalmente o que  $\Delta s^2$  es de género tiempo.
- si  $\Delta s^2 > 0$ , diremos que los eventos están separados espacialmente o que  $\Delta s^2$  es de género espacio.
- si  $\Delta s^2 = 0$ , diremos que los eventos están conectados por una señal luminosa o que  $\Delta s^2$  es de género luz o nulo.

Como  $\Delta s^2$  es un invariante de Lorentz, entonces esta clasificación es absoluta.

### 2.5.1. Elemento de línea infinitesimal

Consideramos dos eventos,  $(t, \vec{x})$  y  $(t + dt, \vec{x} + d\vec{x})$ , entonces definimos el elemento de línea infinitesimal como,

$$ds^2 = -c^2 dt^2 + d\vec{x}^2$$

donde

- si  $ds^2 < 0$ , diremos que los eventos están separados temporalmente o que  $ds^2$  es de género tiempo.
- si  $ds^2 > 0$ , diremos que los eventos están separados espacialmente o que  $ds^2$  es de género espacio.
- si  $ds^2 = 0$ , diremos que los eventos están conectados por una señal luminosa o que  $ds^2$  es de género luz o nulo.

### 2.5.2. Cono de luz

Cuando tenemos que  $ds^2 = 0$ , tenemos la situación del cono de luz, y nos permite representarlo en un eje de coordenadas, tal que

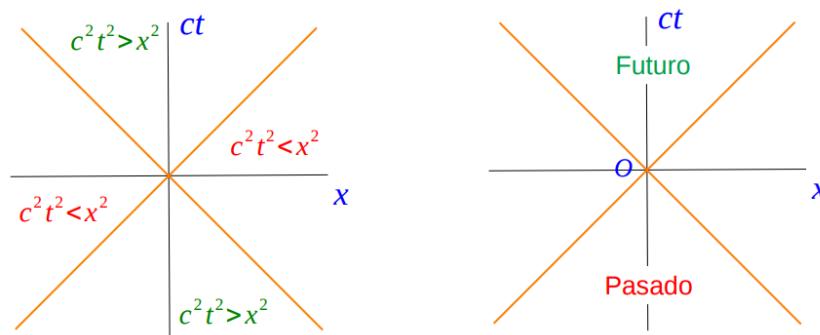


FIGURA 2.3: Cono de luz.

donde la distancia espacial entre los dos sucesos será igual a la distancia que recorre la luz en ese tiempo. Dentro del cono, los eventos están conectados temporalmente con  $(t, \vec{x})$ . Fuera del cono, tenemos los eventos desconectados temporalmente (causal) con  $(t, \vec{x})$ . En la propia superficie del cono se encuentran las señales luminosas emitidas desde  $(t, \vec{x})$ .

### 2.5.3. Sistema de referencia propio

Lo definimos como el sistema de referencia comóvil con la partícula, es decir, tiene la misma velocidad y dirección que la partícula, por lo que desde el punto de este sistema de referencia la partícula siempre está en reposo.

### 2.5.4. Tiempo propio

Se define como,

$$d\tau = \sqrt{-\frac{ds^2}{c^2}} = \sqrt{-\frac{-\epsilon^2 dt^2 + d\vec{x}^2}{\epsilon^2}} = \sqrt{dt^2 - \frac{d\vec{x}^2}{c^2}} = dt \sqrt{1 - \frac{1}{c^2} \frac{d\vec{x}^2}{dt^2}} = dt \sqrt{1 - \frac{\vec{v}^2}{c^2}} = \frac{dt}{\gamma}$$

Luego,

$$d\tau = dt/\gamma$$

Por tanto, el tiempo propio siempre pasa más lentamente que el tiempo del sistema laboratorio. Además, este tiempo será el tiempo del SR comóvil de la partícula.

## 2.6. Métrica de Minkowski y cuadrvectores

Definimos un cuadrvector posición como,

$$(ct, |vecx) \equiv x^\mu = (x^0, x^1, x^2, x^3)$$

donde  $\mu$  es un etiqueta con  $\mu = 0, 1, 2, 3$ .

Recordemos que  $ds^2 = -c^2 dt^2 + d\bar{x}^2$ , ahora tendremos,

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = -(dx^0)^2 + \sum_{i=1}^3 (dx^i)^2 = \sum_{\mu, \nu=0}^3 \eta_{\mu\nu} dx^\mu dx^\nu$$

Por tanto,  $\eta_{\mu\nu}$  será,

$$\eta_{\mu\nu} = diag[-1, 1, 1, 1] = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

siendo  $\eta_{\mu\nu}$  la métrica del espacio de Minkowski. Las propiedades son:

1. La inversa está bien definida, tal que

$$\eta^{\mu\nu} = diag[1, 1, 1, 1]$$

así,  $\eta_{\mu\nu}\eta^{\nu\rho} = \delta_\mu^\rho = diag[1, 1, 1, 1]$ .

2. La métrica con signo cambiado también se usa en diversos contextos, como en la física de partículas relativista.

### 2.6.1. Grupo de Poincaré

Recordemos que  $ds^2 = (ds')^2$  es un invariante; el grupo de simetrías que deja invariante a  $ds^2$  es el grupo de Poincaré, formado por:

- traslaciones espacio-temporales.
- reflexiones en el espacio y en el tiempo.

- transformaciones de Lorentz ortocronas (boosts y rotaciones espaciales) propias, es decir, aquellas que contengan a la identidad de forma continua.

## 2.7. Repaso de álgebra

Un vector  $\vec{v} \in V$ , siendo  $V$  un espacio vectorial, podemos escribir  $\vec{v}$  en función de una base de  $V$ , de la forma,

$$\vec{v} = v^\mu \hat{e}_\mu = v^0 \hat{e}_0 + v^1 \hat{e}_1 + \cdots + v^s \hat{e}_s$$

donde  $v^\mu$  son las componentes de  $\vec{v}$  en la base  $\{\hat{e}_\mu\}$ , siendo un vector columna.

Los vectores duales son aplicaciones lineales, tal que  $\vec{f} : V \rightarrow \mathbb{R}$ , formando un espacio vectorial dual  $V^*$ . Además,  $\vec{f} = f_\mu \hat{e}^\mu$ , siendo  $f_\mu$  un vector fila. Es decir, podemos representar  $\vec{f}$  en las componentes de la base dual  $\{\hat{e}^\mu\}$ , que es la única base que cumple que  $\langle \hat{e}^\mu, \hat{e}_\nu \rangle = \delta_\nu^\mu$ .

Dada una base dual, podemos extraer las componentes de  $\vec{v} \in V$  usando esta base, es decir,

$$\langle \hat{e}^\mu, \vec{v} \rangle = \langle \hat{e}^\mu, v^\nu \hat{e}_\nu \rangle = v^\nu \langle \hat{e}^\mu, \hat{e}_\nu \rangle = v^\nu \delta_\nu^\mu = v^\mu$$

Podemos hacer lo mismo con los vectores duales dada una base vectorial, tal que

$$\langle \hat{e}_\mu, \vec{f} \rangle = \langle \hat{e}_\mu, f_\nu \hat{e}^\nu \rangle = f_\nu \langle \hat{e}_\mu, \hat{e}^\nu \rangle = f_\nu \delta_\mu^\nu = f_\mu$$

También podemos hacer el producto escalar entre vectores y vectores duales, tal que

$$\langle \vec{f}, \vec{v} \rangle = \langle f_\mu \hat{e}^\mu, v^\nu \hat{e}_\nu \rangle = f_\mu v^\nu \langle \hat{e}^\mu, \hat{e}_\nu \rangle = f_\mu v^\nu \delta_\nu^\mu = f_\mu v^\mu$$

Si aplicamos una transformación pasiva  $M_\mu^\nu$ , es decir, dejamos el vector fijo y movemos el sistema de referencia, entonces vamos a decir que  $v'^\mu = M_\nu^\mu v^\nu$  y que  $\hat{e}'_\mu = (M^{-1})_\mu^\nu \hat{e}_\nu$ , pues

$$\vec{v}' = v'^\mu \hat{e}'_\mu = M_\nu^\mu v^\nu (M^{-1})_\mu^\rho \hat{e}_\rho = \delta_\nu^\rho v^\nu \hat{e}_\rho = v^\nu \hat{e}_\nu = \vec{v}$$

es decir, vemos que el vector no se ve afectado por las transformaciones pasivas, pero sus componentes sí se alteran.

Diremos que la métrica de Minkowski nos permite mapear vectores de  $V$  a vectores duales de  $V^*$ , pues dado un vector  $v^\mu \hat{e}_\mu \in V$ , vemos que

$$\begin{aligned}\eta_{\mu\nu} : \hat{e}_\rho &\rightarrow \hat{e}^\rho = \eta^{\rho\sigma} \hat{e}_\sigma \\ v^\rho &\mapsto v_\rho = \eta_{\rho\sigma} v^\sigma\end{aligned}$$

es decir, la métrica  $\eta_{\mu\nu} : \vec{v} \rightarrow \vec{v}^*$  transforma los vectores tal que,

$$\begin{aligned}v_0 &= -v^0; \quad v_i = v^i; \quad i = 1, 2, 3 \\ \hat{e}^0 &= -\hat{e}_0; \quad \hat{e}^i = \hat{e}_i\end{aligned}$$

También podemos usar  $\eta_{\mu\nu}$  para mapear vectores duales a vectores.

La métrica nos da un producto escalar, tal que

$$\langle \vec{v}^*, \vec{v} \rangle = \langle v_\mu \hat{e}^\mu, v^\nu \hat{e}_\nu \rangle = v_\mu v^\nu \langle \hat{e}^\mu, \hat{e}_\nu \rangle = v_\mu v^\nu \delta_\nu^\mu = v_\mu v^\mu = \eta_{\mu\nu} v^\mu v^\nu$$

La métrica  $\eta_{\mu\nu}$  nos permite subir y bajar índices, es decir, contraer índices.

La métrica transforma como un tensor de rango  $(0, 2)$ , es decir,

$$\eta_{\mu\nu}^{(M)} = (M^{-1})_\mu^\rho (M^{-1})_\nu^\sigma \eta_{\rho\sigma}$$

Denotaremos a  $\vec{v}$  y  $v^\mu$  indistintamente, y diremos que son **vectores contravariantes**. A los vectores del espacio dual los denotaremos por  $\vec{f}$  y  $f_\mu$  indistintamente, denominados **vectores covariantes**.

Un tensor de rango  $(r, s)$  es una aplicación multilineal, tal que

$$T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} : V_1^* \otimes V_2^* \otimes \dots \otimes V_r^* \otimes V_1 \otimes V_2 \otimes \dots \otimes V_s \rightarrow \mathbb{R}$$

Los tensores transforman como,

$$T'_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} = M_{\rho_1}^{\mu_1} M_{\rho_2}^{\mu_2} \dots M_{\rho_r}^{\mu_r} (M^{-1})_{\nu_1}^{\sigma_1} (M^{-1})_{\nu_2}^{\sigma_2} \dots (M^{-1})_{\nu_s}^{\sigma_s} T_{\sigma_1 \sigma_2 \dots \sigma_s}^{\rho_1 \rho_2 \dots \rho_r}$$

Con la métrica  $\eta_{\mu\nu}$  podemos calcular la traza de un tensor, tal que para un tensor  $T_{\mu\nu}$ , su traza será  $T = T_{\mu\nu} \eta^{\mu\nu}$ , es decir, hemos contraído todos los índices. Pero si tenemos  $T_\mu^\nu$ , su traza será  $T = T_\mu^\nu \delta_\nu^\mu$  y si tenemos  $T^{\mu\nu}$ , su traza será  $T = T^{\mu\nu} \eta_{\mu\nu}$ . O bien, podemos transformar los tensores y aplicar la primera definición, tal que  $T_\mu^\nu = T_{\mu\nu} \eta^{\sigma\nu}$  y  $T^{\mu\nu} = T_{\rho\sigma} \eta^{\rho\mu} \eta^{\sigma\nu}$ .

La métrica está relacionada con el elemento de línea, llegando a tener también la misma clasificación, tal que

- Si  $\eta_{\mu\nu}v^\mu v^\nu > 0$ , entonces diremos que  $v^\mu$  es un vector espacial.
- Si  $\eta_{\mu\nu}v^\mu v^\nu < 0$ , entonces diremos que  $v^\mu$  es un vector temporal.
- Si  $\eta_{\mu\nu}v^\mu v^\nu = 0$ , entonces diremos que  $v^\mu$  es un vector nulo.

Además, se cumple que un vector ortogonal a uno de tipo tiempo es espacial y un vector ortogonal a uno de tipo espacio o nulo, no tiene por qué ser temporal, es decir, es de cualquier género.

### 2.7.1. Tensor simétrico y antisimétrico

Un tensor simétrico será aquel que si se le permutan dos índices, permanece invariante, es decir,  $T_{\mu\nu} = T_{\nu\mu}$ . La parte simétrica de un tensor es

$$T_{(\mu\nu)} = \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu})$$

Un tensor antisimétrico es aquel que si se le permutan dos índices, cambia de signo, es decir,  $T_{\mu\nu} = -T_{\nu\mu}$ . La parte antisimétrica de un tensor es

$$T_{[\mu\nu]} = \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$$

Un tensor cualquiera siempre se puede descomponer en la suma de su parte simétrica y su parte antisimétrica, es decir,

$$R_{\mu\nu} = R_{(\mu\nu)} + R_{[\mu\nu]}$$

### 2.7.2. Transformaciones de Lorentz. Versión covariante

Recordemos que las transformaciones de Lorentz son,

$$\begin{aligned} (i) \quad dt' &= \gamma \left( dt - \frac{v}{c^2} dx \right); & (iii) \quad dy' &= dy \\ (ii) \quad dx' &= \gamma (dx - vdt); & (iv) \quad dz' &= dz \end{aligned}$$

Por tanto, usando la notación covariante, podemos agruparlas todas en una sola ecuación, tal que

$$dx'^\mu = \Lambda_\nu^\mu dx^\nu$$

donde

$$\Lambda_\nu^\mu = \begin{pmatrix} \gamma & -\gamma \frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

es la matriz de Lorentz.

Recordando que  $ds^2 = (ds')^2$ , vemos que la métrica de Minkowski y la matriz de Lorentz se pueden relacionar, tal que

$$\begin{aligned} (ds')^2 &= \eta_{\mu\nu} dx'^\mu dx'^\nu = \eta_{\mu\nu} \Lambda_\rho^\mu dx^\rho \Lambda_\sigma^\nu dx^\sigma \\ &\quad || \qquad \qquad \qquad || \\ ds^2 &= \eta_{\mu\nu} dx^\mu dx^\nu = \eta_{\mu\nu} \Lambda_\rho^\mu \Lambda_\sigma^\nu dx^\rho dx^\sigma \end{aligned}$$

Por tanto,

$$\eta_{\rho\sigma} = \eta_{\mu\nu} \Lambda_\rho^\mu \Lambda_\sigma^\nu$$

Es decir, la métrica de Minkowski permanece invariante bajo transformaciones de Lorentz.

Además, la inversa de la matriz de Lorentz también está definida,  $(\Lambda^{-1})_\nu^\mu$ , que también deja invariante la inversa de la métrica de Minkowski, y cumple que

$$\Lambda_\nu^\mu (\Lambda^{-1})_\rho^\nu = \delta_\rho^\mu$$

Para pasar de la matriz de Lorentz  $\Lambda_\mu^\nu$  a su inversa  $(\Lambda^{-1})_\nu^\mu$ , teniendo un movimiento con velocidad  $v$  a lo largo del eje  $X$ , solo debemos cambiar  $v$  por  $-v$ ; igual para rotaciones. Las propiedades de esta matriz son:

1. La traspuesta de la matriz de Lorentz es  $(\Lambda^{-1})_\nu^\mu = \Lambda_\nu^\mu$ .
2. Las componentes de la matriz de Lorentz para boosts de forma general son,

$$\Lambda_0^0 = \gamma; \quad \Lambda_i^0 = -\gamma \frac{v^i}{c}; \quad \Lambda_0^i = -\gamma \frac{v^i}{c}; \quad \Lambda_j^i = \delta_j^i + (\gamma - 1) \frac{v^i v^j}{\vec{v}^2}$$

3. Para las rotaciones de ángulo  $\theta$  alrededor de un eje  $\hat{\omega}^i$ , con  $\hat{\omega}^\mu = (0, \hat{\omega}^i)$  y  $\hat{\omega}^\mu \hat{\omega}_\mu = 1$ , las componentes de la matriz de Lorentz serán,

$$\Lambda_0^0 = 0; \quad \Lambda_0^i = 0 = \Lambda_i^0; \quad \Lambda_j^i = \cos \theta \delta_j^i + (1 - \cos \theta) \hat{\omega}^i \hat{\omega}_j - \sin \theta \mathcal{E}_{jk}^i \hat{\omega}^k$$

donde  $\mathcal{E}_{jk}^i$  es el tensor de Levi-Civita espacial, que actúa parecido a la delta de Kronecker, tal que

$$\mathcal{E}_{ijk} = \begin{cases} +1 & \text{si } i \neq j \neq k \text{ y la permutación es par} \\ -1 & \text{si } i \neq j \neq k \text{ y la permutación es impar} \\ 0 & \text{en otro caso} \end{cases}$$

Para la métrica de Minkowski,  $\eta_{\mu\nu} = \text{diag}[-1, 1, 1, 1]$ , tenemos que el tensor de Levi-Civita cumple que

$$\mathcal{E}_{ijk} = \mathcal{E}^{ijk} = \mathcal{E}_{jk}^i$$

## 2.8. Cinemática relativista

Para tratar la cinemática de forma relativista, debemos formalizar los conceptos de la cinemática clásica en formulación covariante.

### 2.8.1. Vector cuadrvélocidad

Representa la velocidad espacio-temporal de una partícula puntual y se obtiene derivando respecto al tiempo propio de la partícula su cuadrívector, tal que

$$u^\mu = \frac{dx^\mu}{d\tau} = \dot{x}^\mu$$

Si estamos en el sistema de referencia comóvil de la partícula, entonces la cuadrvélocidad será  $u^\mu = (c, 0, 0, 0)$ , pues en este sistema de referencia, la partícula se encuentra en reposo espacial relativo; no existe el reposo relativo temporal.

Si estamos en un SRI cualquiera donde la partícula se mueve a velocidad  $\vec{v} = \frac{d\vec{x}}{dt}$ , entonces la cuadrvélocidad será  $u^\mu = \gamma(c, \vec{v})$ , pues debemos reemplazar el  $dt$  por  $d\tau$ , recordando que  $dt = \gamma d\tau$ .

Además, sabemos que el producto escalar es un invariante, por lo que usando el sistema de referencia comóvil de la partícula, obtenemos que  $u^\mu u_\mu = -c^2$ .

### 2.8.2. Vector cuadrimomento

Representa el momento espacio-temporal de una partícula, y se obtiene multiplicando la masa en reposo de la partícula por su cuadrvélocidad. Así,

$$p^\mu = m_0 u^\mu \Rightarrow \begin{cases} p^0 = m_0 \gamma c \Rightarrow E = cp^0 = m_0 \gamma c^2 \\ p^i = m_0 \gamma v^i \Rightarrow \vec{p} = m_0 \gamma \vec{v} \end{cases}$$

Por tanto, el cuadrimomento es

$$p^\mu = (E/c, p^1, p^2, p^3)$$

Podemos calcular también el producto escalar de cuadrimomentos, tal que  $p^\mu p_\mu = -m^2 c^2$ .

### 2.8.3. Vector cuadriaceleración

Representa la aceleración espacio-temporal de una partícula puntual. Se calcula derivando respecto al tiempo propio el vector cuadrvélocidad de la partícula, o derivando dos veces el cuadrivector de la partícula respecto al tiempo propio, tal que

$$b^\mu = \frac{d^2 x^\mu}{d\tau^2} = \frac{du^\mu}{d\tau} = \dot{u}^\mu = \ddot{x}^\mu$$

La cuadriaceleración de una partícula puntual que se mueve a velocidad  $v$  y aceleración  $a$  en un SRI cualquiera es,

$$b^\mu = \left( \gamma^4 \frac{\vec{v} \cdot \vec{a}}{d\tau^2}, \gamma^4 \frac{(\vec{v} \cdot \vec{a}) \vec{v}}{c^2} + \gamma^2 \vec{a} \right)$$

donde  $\vec{a} = \frac{d\vec{v}}{dt}$  y  $\frac{d\gamma}{dt} = \gamma^3 \frac{\vec{v} \cdot \vec{a}}{c^2}$ .

Las propiedades de la cuadriaceleración son:

1.  $u^\mu b_\mu = 0$
2.  $b^\mu b_\mu = \gamma^4 \left( \gamma^2 \frac{\vec{a} \cdot \vec{v}}{c^2} + \vec{a} \cdot \vec{a} \right) \geq 0$ . Por tanto, esto implica que  $b^\mu$  es de género espacio.

### 2.8.4. Derivación

Al igual que tenemos un gradiente tridimensional, podemos construir un gradiente cuadridimensional, tal que

$$\partial_\mu f = f_{,\mu} = \frac{\partial f}{\partial x^\mu} = \left( \frac{1}{c} \frac{\partial f}{\partial t}, \nabla f \right)$$

siendo un vector covariante, por lo que podemos construir su versión contravariante, tal que

$$\partial^\mu f = \eta^{\mu\nu} \partial_\nu f = \left( -1 \frac{1}{c} \frac{\partial f}{\partial t}, \nabla f \right)$$

### 2.8.5. Operador D'Alembertiano

Se define como el Laplaciano cuadridimensional, tal que

$$\square f = \eta^{\mu\nu} \partial_\mu \partial_\nu f = -c^2 \partial_t^2 f + \nabla^2 f$$

Es invariante bajo transformaciones de Lorentz, pues si tenemos dos SRI  $S'$  y  $S$ , los operadores D'Alembertianos de cada SRI cumplirán que  $\square' = \square$ .

### 2.8.6. Tensor de Levi-Civita

Es un tensor totalmente antisimétrico. se define como,

$$\mathcal{E}^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{si } \mu \neq \nu \neq \rho \neq \sigma \text{ y la perturbación es par} \\ -1 & \text{si } \mu \neq \nu \neq \rho \neq \sigma \text{ y la perturbación es impar} \\ 0 & \text{en otro caso} \end{cases}$$

Este tensor cumple que,

1.  $\mathcal{E}_{0123} = -\mathcal{E}^{0123}$ .
2.  $\mathcal{E}^{\mu\nu\rho\sigma} \mathcal{E}_{\mu\nu\rho\sigma} = -4!$ .
3.  $\mathcal{E}^{\mu\nu\rho\sigma} \mathcal{E}_{\mu\nu\rho\gamma} = -3! \delta_\gamma^\sigma$ .
4.  $\mathcal{E}^{\mu\nu\rho\sigma} \mathcal{E}_{\alpha\beta\rho\sigma} = -2! \delta_\mu^{[\alpha} \delta_\beta^{\nu]}$ .
5.  $\mathcal{E}^{\mu\nu\sigma\rho} \mathcal{E}_{\alpha\beta\gamma\rho} = -1! \delta_{[\alpha}^\mu \delta_{\beta}^\nu \delta_{\gamma}^\sigma$ .
6.  $\mathcal{E}^{\mu\nu\sigma\rho} \mathcal{E}_{\alpha\beta\gamma\delta} = \delta_{[\alpha}^\mu \delta_{\beta}^\nu \delta_{\gamma}^\sigma \delta_{\delta]}^\rho$ .

## 2.9. Grupo de Poincaré

Este grupo es el grupo de transformaciones que deja invariante el elemento de línea. Tenemos,

- **Traslaciones espaciotemporales:** son del tipo  $x'^\mu = x^\mu + \alpha^\mu$ , donde  $\alpha^\mu = cte$ . El operador momento en mecánica cuántica se define como  $\hat{P}_\mu = -i\partial_\mu = -i\frac{\partial}{\partial x^\mu}$ , pero si definimos el operador  $\hat{U}_{\alpha^\mu} = \exp[i\alpha^\mu \hat{P}_\mu]$ , que podemos expandirlo en serie de potencias, tal que

$$\hat{U}_{\alpha^\mu} f(x^\mu) = \left[ 1 + \alpha^\mu (\partial_\mu) + \frac{1}{2!} (\alpha^\mu (\partial_\mu)^2) + \dots \right] f(x^\mu) = f(x^\mu + \alpha^\mu) = f(x'^\mu)$$

por tanto, este operador nos permite hacer que  $\hat{U}_{\alpha^\mu} x^\mu = x^\mu + \alpha^\mu = x'^\mu$ .

Al operador  $\hat{P}^\mu$  se le denomina *generador de traslaciones*. Podemos calcular el commutador, tal que  $[\hat{P}_\mu, \hat{P}_\nu] = 0$ , si y solo si el grupo de traslaciones es un grupo abeliano, por lo que las traslaciones comutan.

- **Transformaciones de Lorentz:** están formadas por las rotaciones espaciales y las transformaciones de Lorentz puras, los boosts.

Sabemos que  $\det(\Lambda_{nu}^\mu) = \pm 1$ , pero nos vamos a quedar con el subgrupo propio de las transformaciones con  $\det(\Lambda_\nu^\mu) = +1$  y dentro de este subgrupo, nos quedamos con el subgrupo ortocrono, es decir, el subgrupo donde se mantiene la dirección temporal,  $\Lambda_0^0 \geq 1$ . Si consideramos una transformación de Lorentz infinitesimal, tendremos que

$$\Lambda_\nu^\mu = \delta_\nu^\mu + \delta\omega_\nu^\mu$$

Recordamos que  $\eta_{\mu\nu} = \eta_\mu^\rho \Lambda_\nu^\sigma \eta_{\rho\sigma}$ , por lo que sustituyendo tenemos que,

$$\delta\omega_{\mu\nu} + \delta\omega_{\nu\mu} = 0$$

a primer orden. Es decir,  $\delta\omega_{\mu\nu}$  es un tensor antisimétrico.

Al ser un tensor antisimétrico, podemos hacer la siguiente clasificación:

$$\delta\omega_{0i} = -\delta\omega_{i0} = \frac{\delta v_i}{c} = \delta\xi_i$$

que se hace cargo de los boosts.

$$\delta\omega_{ij} = -\delta\omega_{ji} = -\mathcal{E}_{ijk}\delta\theta^k$$

que se hace cargo de las rotaciones.

Decimos que los  $\xi_i$  y  $\theta^k$  son los vectores de Killing de las traslaciones espaciales y rotaciones, respectivamente.

Como el grupo de Lorentz tiene 4 traslaciones espaciotemporales, 3 boosts y 3 rotaciones; decimos que este grupo tiene 10 grados de libertad.

Veamos cómo cambian los vectores,

$$x'^\mu = x^\mu + \delta x^\mu$$

con

$$\delta x^\mu = \delta\omega_\nu^\mu x^\nu = -\frac{i}{2}\delta\omega^{\rho\sigma}\hat{L}_{\rho\sigma}x^\mu$$

donde

$$\hat{L}_{\mu\nu} = \hat{X}_\mu(-i\partial_\nu) - \hat{X}_\nu(-i\partial_\mu) = \hat{X}_\mu\hat{P}_\nu - \hat{X}_\nu\hat{P}_\mu$$

Definimos el momento angular espacial tal que

$$\hat{L}^i = \frac{1}{2}\mathcal{E}^{ijk}\hat{L}_{jk}$$

Definimos  $\hat{K}_i = \hat{L}_{0i}$  como el generador de los boosts, tal que

$$\delta x^\mu = i\left(\delta\vec{\theta}\cdot\vec{L} + \delta\vec{\xi}\cdot\hat{\vec{K}}\right)x^\mu$$

Vemos que

$$[\hat{L}_i, \hat{L}_j]f(x) = i\mathcal{E}_{ij}^k\hat{L}_k(f(x))$$

$$[\hat{L}_i, \hat{K}_j] = i\mathcal{E}_{ij}^k\hat{K}_k$$

Luego, no comutan, pues no es lo mismo rotar y hacer un boost, que hacer un boost y luego rotar.

Además, si hacemos dos transformaciones de Lorentz en dos direcciones diferentes, tampoco comutan, pues  $[\hat{K}_i, \hat{K}_j] = -i\mathcal{E}_{ij}^k\hat{L}_k$ . Esto es lo que se conoce como *Rotación de*

Weigner, que se traduce en: 'boost + boost = boost + rotación'.

Para boosts finitos, es decir, velocidad finita, podemos identificar el vector  $\vec{\xi}$  con la velocidad, tal que

$$\vec{\xi} = \vec{v} \cdot \arctan\left(\frac{|\vec{v}|}{c}\right)$$

La versión matricial de  $\hat{L}_i$  y  $\hat{K}_i$  son,

$$(L_1)_{\nu}^{\mu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}; \quad (L_2)_{\nu}^{\mu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}; \quad (L_3)_{\nu}^{\mu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
  

$$(K_1)_{\nu}^{\mu} = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad (K_2)_{\nu}^{\mu} = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad (K_3)_{\nu}^{\mu} = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

## 2.10. Dinámica relativista

Sabemos que la dinámica viene descrita por el Lagrangiano y que las ecuaciones del movimiento se hallan minimizando la acción. La acción se define en términos del Lagrangiano como,

$$S = \int_{\tau_1}^{\tau_2} d\tau \mathcal{L}(\tau, x^{\mu}, \dot{x}^{\mu})$$

donde usamos el tiempo propio  $\tau$ .

Dada una trayectoria  $x^{\mu}(\tau)$ , consideramos variaciones de la trayectoria, de la forma  $x^{\mu}(\tau) + \delta x^{\mu}(\tau)$ , tal que  $\delta x^{\mu}|_{\tau_1, \tau_2} = 0$ .

### 2.10.1. Principio variacional de acción estacionaria

Las trayectorias físicas son aquellas que dejan invariante la acción estacionaria (minimizan o maximizan la acción). Esto se traduce en,

$$\delta S = \int_{\tau_1}^{\tau_2} \left( \frac{\partial \mathcal{L}}{\partial x^{\mu}} \delta x^{\mu} + \frac{\partial \mathcal{L}}{\partial \dot{x}^{\mu}} \delta \dot{x}^{\mu} \right) d\tau = \int_{\tau_2}^{\tau_1} \left[ \left( \frac{\partial \mathcal{L}}{\partial x^{\mu}} - \frac{d}{d\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^{\mu}} \right) \right) \delta x^{\mu} + \frac{d}{d\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^{\mu}} \delta x^{\mu} \right) \right] d\tau = 0$$

donde  $\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \delta x^\mu \Big|_{\tau_1, \tau_2} = 0$ , pues  $\delta x^\mu \Big|_{\tau_1, \tau_2} = 0$ . Usando que  $\delta S = 0$ , obtenemos las ecuaciones de Euler-Lagrange, tal que

$$\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{d}{d\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \right) = 0$$

## 2.10.2. Cantidades Conservadas

**Teorema 2.10.1.** *Teorema de Noether Si tomamos variaciones de las trayectorias que no se anulan en los extremos y resulta que  $\delta S = 0$ , entonces para esta variación existe una simetría, es decir, existe una cantidad conservada.*

Por tanto, a esta variación le podemos asociar una ley de conservación, tal que

$$\frac{d}{\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \delta x^\mu \right) = 0 \implies \delta Q_{\delta x} \equiv \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \delta x^\mu = cte$$

## 2.10.3. Partícula libre relativista

Construimos la acción usando el elemento de línea, tal que

$$S = -mc \int_{s_1}^{s_2} \sqrt{-ds^2} = -mc^2 \int_{\tau_1}^{\tau_2} \sqrt{\dot{t}^2 - \dot{\vec{x}}^2/c^2} d\tau = -mc^2 \int_{t_1}^{t_2} \sqrt{1 - \frac{\vec{v}^2}{c^2}} dt = -mc^2 \int_{t_1}^{t_2} \frac{dt}{\gamma}$$

Por tanto, el Lagrangiano relativista de una partícula libre será

$$\mathcal{L}(\tau, x^\mu, \dot{x}^\mu) = \sqrt{\dot{t}^2 - \frac{\dot{\vec{x}}^2}{c^2}}$$

Aplicamos las ecuaciones de Euler-Lagrange, tal que

$$\text{Eje 0: } \frac{\partial \mathcal{L}}{\partial x^0} - \frac{d}{d\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^0} \right) = 0 \implies \frac{d}{d\tau} \left( \frac{\dot{t}}{\sqrt{\dot{t}^2 - \frac{\dot{\vec{x}}^2}{c^2}}} \right) = 0$$

En resumen,

$$\frac{d}{d\tau} (m u^\mu) = 0 = \frac{d}{d\tau} (m \dot{x}^\mu)$$

La acción es invariante de Lorentz (de Poincaré) y a velocidades pequeñas ( $|\vec{v}| \ll c$ ), recuperando el límite no relativista del Lagrangiano,

$$S = -mc^2 \int \sqrt{1 - \frac{|\vec{v}|^2}{c^2}} dt \approx -mc^2 \int dt + \int dt \left( \frac{1}{2} m |\vec{v}|^2 \right) + \mathcal{O}(|\vec{v}|^4/c^4) \xrightarrow{0}$$

El cuadrimomento de la partícula será  $p_\mu = mu_\mu = m\gamma(-c, \vec{v})$ , siendo una cantidad conservada. La simetría asociada es la invariancia bajo traslaciones. Luego, dada una traslación infinitesimal  $\delta x^\mu = \delta\alpha^\mu = cte$ , tenemos que

$$\delta Q = \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \delta x^\mu = \frac{\partial \mathcal{L}}{\partial \dot{\alpha}^\mu} \delta \alpha^\mu = \frac{(mu_\mu)}{\sqrt{-u^\mu u_\mu}} \delta \alpha^\mu = p_\mu \delta \alpha^\mu$$

como  $\delta Q = cte$  y  $\delta \alpha^\mu = cte$ , entonces

$$\frac{d(\delta Q)}{d\tau} = 0 \Rightarrow \frac{d}{d\tau}(mu_\mu) = 0 \Rightarrow \frac{dp_\mu}{d\tau} = 0 \Rightarrow p_\mu = cte$$

siendo la cantidad conservada de las traslaciones.

Por otro lado, si hacemos una transformada de Lorentz,  $\delta x^\mu = \delta\omega_\nu^\mu x^\nu$ , la cantidad conservada será,

$$\delta Q = \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \delta x^\mu = p_\mu \delta\omega_\nu^\mu x^\nu = \frac{1}{2} \delta\omega_{\mu\nu} (x^\mu p^\nu - x^\nu p^\mu) = \frac{1}{2} M^{\mu\nu} \delta\omega_{\nu\mu}$$

donde  $M^{\mu\nu} = (x^\mu p^\nu - x^\nu p^\mu)$  representa el momento angular. Además, vemos que  $\dot{M}^{\mu\nu} = 0$ , por lo que  $M^{\mu\nu} = cte$ , entonces la cantidad conservada de las transformaciones de Lorentz es el momento angular.

#### 2.10.4. Sistema de N-partículas

Si tenemos N-partículas, de forma general, tendremos un Lagrangiano tal que

$$\mathcal{L}(\tau, x_1^\mu, x_2^\mu, \dots, x_N^\mu, \dot{x}_1^\mu, \dot{x}_2^\mu, \dots, \dot{x}_N^\mu) = \mathcal{L}(\tau, x_n^\mu, \dot{x}_n^\mu)$$

Análogamente al caso de una partícula, tendremos las ecuaciones de Euler-Lagrange,

$$\frac{\partial \mathcal{L}}{\partial x_n^\mu} - \frac{d}{d\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}_n^\mu} \right) = 0$$

que serán N-ecuaciones.

Para las cantidades conservadas, tendremos que

$$\delta Q_{\delta x} = \sum_{n=1}^N \frac{\partial \mathcal{L}}{\partial \dot{x}_n^\mu} \delta x_n^\mu$$

Para una transformación infinitesimal  $(\delta x_1^\mu, \dots, \delta x_N^\mu)$  que deje invariante el Lagrangiano, tendremos que  $\delta Q_{\delta x}$  es conservado.

En general tendremos,

- Para la invariancia bajo traslaciones, se conserva el cuadrimomento.
  - Si tenemos el Lagrangiano,  $\mathcal{L} = \mathcal{L}(\tau, \dot{x}_n^\mu)$ ; si hacemos traslaciones  $\delta x_n^\mu$ , el Lagrangiano queda invariante, luego se conserva el momento  $p_n^\mu = \eta^{\mu\nu} \frac{\partial \mathcal{L}}{\partial \dot{x}^\nu}$  de la partícula.
  - Si tenemos el Lagrangiano,  $\mathcal{L} = \mathcal{L}(\tau, x_n^\mu, \dot{x}_n^\mu)$  que sea invariante bajo una traslación global cuya perturbación sea  $\delta x_n^\mu = \delta \alpha^\mu = cte$ , entonces el momento total se conserva,

$$P^\mu = \sum_{n=1}^N p_n^\mu$$

pero no se conservan los momentos individuales. Donde  $P^0$  representa la energía total del sistema (masas + energía cinética + energía potencial) y  $P^i$  representa el momento lineal del sistema.

- Para la invariancia bajo transformaciones de Lorentz tendremos dos casos,
  - Si las partículas no interaccionan entre ellas, se conserva el cuadrimomento angular,

$$M_n^{\mu\nu} = x_n^\mu p_n^\nu - x_n^\nu p_n^\mu$$

- Si las partículas interaccionan, pero el sistema está aislado, entonces se conserva el cuadrimomento angular total,

$$M^{\mu\nu} = \sum_{n=1}^N M_n^{\mu\nu}$$

pero no los individuales.

## 2.11. Campos relativistas

Un campo escalar relativista será,

$$\begin{aligned}\phi(x) : \mathbb{R}^4 &\rightarrow \mathbb{R} \\ x^\mu &\mapsto \lambda\end{aligned}$$

Su acción es una función de la forma,

$$S = \int_{\mathcal{M}} d^4x \mathcal{L}(\phi, \partial_\mu \phi)$$

donde  $\mathcal{L}$  es la densidad lagrangiana. Aplicamos el principio variacional se acción estacionaria bajo variaciones del campo que se anulan en la frontera  $\partial\mathcal{M}$  de la variedad  $\mathcal{M}$ . Luego, tenemos la condición de que  $\delta\phi|_{\partial\mathcal{M}} = 0$ , tal que

$$\delta S = \int_{\mathcal{M}} d^4x \left( \frac{\partial \mathcal{L}}{\partial \phi} \delta \phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta (\partial_\mu \phi) \right)$$

Imponemos que las variaciones sean del tipo  $\delta(\partial_\mu \phi) = \partial_\mu(\delta\phi)$ , pues solo estamos variando el campo y no las coordenadas, así la acción se transforma en,

$$\delta S = \int_{\mathcal{M}} d^4x \left( \frac{\partial \mathcal{L}}{\partial \phi} \delta \phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\mu(\delta\phi) \right) = \int_{\mathcal{M}} d^4x \left[ \left( \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \right) \delta \phi + \underbrace{\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta \phi \right)}_{\int_{\partial\mathcal{M}} (...) \delta \phi = 0} \right] = 0$$

Por tanto, vemos que se cumplen las ecuaciones de Euler-Lagrange para un campo escalar relativista,

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0$$

**Ejemplo 2.11.1.** Si tenemos un campo escalar sin masa cuya densidad lagrangiana es  $\mathcal{L} = \partial_\mu \phi \partial^\mu \phi = \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi$ , tenemos que

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = \frac{\partial}{\partial (\partial_\mu \phi)} (\partial_\rho \phi \partial_\sigma \phi \eta^{\rho\sigma}) = \delta_\rho^\mu \partial_\sigma \phi \eta^{\rho\sigma} + \delta_\sigma^\mu \partial_\rho \phi \eta^{\rho\sigma} = 2\partial^\mu \phi$$

$$\frac{\partial \mathcal{L}}{\partial \phi} = 0$$

Luego, aplicando la ecuación de Euler-Lagrange, tenemos

$$-\partial_\mu (\partial^\mu \phi) = 0$$

es decir, tenemos  $\square\phi = 0$ .

### 2.11.1. Cantidades conservadas

Consideremos variaciones de la forma  $x'^\mu = x^\mu + \delta x^\mu$ , que induce  $\phi'(x') = \phi(x)$ . Entonces,

$$\phi'(x') = \phi'(x) + \delta x^\mu \partial_\mu \phi'(x) = \phi(x)$$

$$\phi'(x) - \phi(x) = -\delta x^\mu \partial_\mu \phi'(x) = -\delta x^\mu \partial_\mu \phi(x)$$

Si hacemos una transformación de coordenadas, vemos que

$$d^4x' = d^4x(1 + \partial_\mu \delta x^\mu)$$

y que

$$\mathcal{L}' = \mathcal{L} + \delta x^\mu \partial_\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \phi} (\phi'(x) - \phi(x)) + \frac{-\delta x^\mu \partial_\mu \phi}{\partial(\partial_\mu \phi)} (\partial_\mu \phi' - \partial_\mu \phi) - (\delta x^\nu \partial_\nu)(\partial_\mu \phi)$$

Si introducimos todo en la acción, tenemos

$$\begin{aligned} \delta S &= \int d^4x \left[ \mathcal{L}(\partial_\mu \delta x^\mu) + \delta x^\mu \partial_\mu \mathcal{L} - \frac{\partial \mathcal{L}}{\partial \phi} (\delta x^\nu \partial_\nu)(\partial_\mu \phi) - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} (\delta x^\nu \partial_\nu)(\partial_\mu \phi) \right] = \\ &= \int d^4x \left[ \mathcal{L}(\partial_\mu \delta x^\mu) + \delta x^\mu \partial_\mu \mathcal{L} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) (\delta x^\nu \partial_\nu)(\partial_\mu \phi) - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} (\delta x^\nu \partial_\nu)(\partial_\mu \phi) \right] = \\ &= \int d^4x \partial_\mu \left[ \left( \delta_\nu^\mu \mathcal{L} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi \right) \delta x^\nu \right] \end{aligned}$$

Si  $\delta x^\mu$  es una traslación constante,  $\delta x^\mu = \delta \alpha^\mu$ , entonces tenemos que

$$T_\nu^\mu = \delta_\nu^\mu \mathcal{L} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi$$

es conservado y se denomina *tensor de energía-impulso*, que tiene el contenido energético del campo de forma local.

Además, como  $\delta S = 0$ , entonces tendremos que  $\partial_\mu T_\nu^\mu = 0$ , por lo que es una ley de conservación. Este tensor es simétrico y toda densidad lagrangiana, con suficiente simetría, tiene asociado un tensor de energía-impulso que cumple que  $\partial_\mu T_\nu^\mu = 0$ .

# Capítulo 3

## Geometría Diferenciable

*“Primero tienes que aprender las reglas del juego, y después jugar mejor que nadie”.*  
(A. Einstein)

### 3.1. Breve repaso histórico

La Geometría Diferenciable surge de una generalización de la geometría euclíadiana que lleva a cabo Riemann, el cuál lo presentó en la Universidad de Gotinga, en su lección inaugural con el título "*Sobre las hipótesis en las que se basa la geometría*", teniendo en el tribunal al mismísimo Gauss.

Hasta entonces, la geometría se había estudiado siempre en el espacio tridimensional euclídeo, y el concepto de 'curvatura' se concebía únicamente para curvas y superficies en el espacio, mientras que Riemann planteó poder trabajar con 'espacios coordenados' sin suponerlos contenidos en el espacio euclídeo ni en ningún otro espacio. Puso en evidencia su relación con la curvatura que Gauss había definido para superficies en el espacio euclídeo, abriendo las puertas a definir un concepto general de curvatura que permitía afirmar, por ejemplo, que un espacio tridimensional fuera 'curvo', cosa inconcebible hasta entonces.

Una vez asentado los conceptos y demostrado la congruencia de la exposición de Riemann, pasamos al siglo XX, cuando Albert Einstein toma el tratado de Levi-Civita para estudiar el cálculo tensorial que usaría para desarrollar la teoría general de la relatividad. Por aquél entonces, el cálculo tensorial era un mar de subíndices y superíndices que subían y bajaban de una fórmula a otra, pero no tardaron muchos matemáticos en encontrar un enfoque más conceptual que permitiera llegar a una comprensión más profunda de la teoría, desarrollando una 'geometría sin índices' o 'intrínseca', en la que los conceptos fundamentales de la geometría diferenciable son objetos algebraicos abstractos globales, y las expresiones coordenadas (con índices), son solo representaciones auxiliares locales, que en ocasiones son convenientes de usar.

### 3.2. Introducción

Tal y como dice la frase de Albert Einstein, en este capítulo vamos a aprender las 'reglas del juego' de la Relatividad General; estas son la Geometría Diferenciable (sí, es Diferenciable, no Diferencial).

Para poder abordar este capítulo, es recomendable hacer una lectura del **Apéndice B**.

**Espacios topológicos**, pues usaremos bastantes conceptos de Topología que son recomendables recordar, de todas formas, volveremos a escribir las definiciones y Teoremas necesarios.

### 3.2.1. Prefacio

El espacio euclídeo y lorentziano tienen propiedades especiales. Dados  $p$  y  $q$ , entonces  $(p+q)$  también son puntos del espacio, luego, el espacio euclídeo y lorentziano son espacios vectoriales. Es decir, siempre podemos escribir

$$x_p^\mu \vec{e}_\mu + x_q^\mu \vec{e}_\mu = (x_p^\mu + x_q^\mu) \vec{e}_\mu$$

donde  $\vec{e}_\mu$  son los vectores de la base ortonormal del espacio vectorial,

En coordenadas curvilíneas esta propiedad se pierde.

**Ejemplo 3.2.1.** *Las coordenadas cilíndricas en dos dimensiones tenemos,*

$$\begin{aligned} x^1 &= r \cos \theta; & x^2 &= r \sin \theta \\ r &= \sqrt{x_1^2 + x_2^2}; & \theta &= \arctan \left( \frac{x_2}{x_1} \right) \end{aligned}$$

Vemos que  $(p+q)$  ya no pertenece al espacio 'padre' en general, con  $r_p + r_q$  ó  $\theta_p + \theta_q$ .

Por suerte, los desplazamientos infinitesimales sí forman un espacio vectorial. Es decir, el espacio tangente a cada punto de un espacio curvilíneo es un espacio vectorial.

Veamos cómo comparar dos puntos distintos de un espacio curvilíneo, pertenecientes a distintos espacios tangentes,

$$\begin{aligned} dx^1 &= \cos \theta dr - r \sin \theta d\theta \\ dx^2 &= \sin \theta dr + r \cos \theta d\theta \end{aligned}$$

Estos elementos forman un espacio vectorial. Tomando el diferencial de una función,

$$df = \frac{\partial f}{\partial x^1} dx^1 + \frac{\partial f}{\partial x^2} dx^2 = \frac{\partial f}{\partial r} dr + \frac{\partial f}{\partial \theta} d\theta$$

Por tanto, podemos hacer una identificación de las derivadas como los vectores y los diferenciables como los elementos de la base, tal que

$$\begin{aligned} df &= \frac{\partial f}{\partial x^1} dx^1 + \frac{\partial f}{\partial x^2} dx^2 = (\partial_\mu f) \vec{e}^\mu \\ &= \frac{\partial f}{\partial r} dr + \frac{\partial f}{\partial \theta} d\theta = (\tilde{\partial}_\mu f) \tilde{\vec{e}}^\mu \end{aligned}$$

Identificamos  $dx^\mu$  con una base de vectores covariantes y  $\partial_\mu f$  con las componentes de la base. Todos los  $d\vec{f}$  cumplen las propiedades de los elementos de un espacio vectorial.

Las leyes de transformación son,

$$x^\mu \equiv x^\mu(\tilde{x}); \quad dx^\mu \equiv \frac{\partial x^\mu}{\partial \tilde{x}^\mu} d\tilde{x}^\mu$$

por tanto,

$$\vec{e}^\mu \equiv \frac{\partial x^\mu}{\partial \tilde{x}^\mu} \tilde{\vec{e}}^\mu$$

Entonces, las coordenadas transformarán como,

$$\partial_\mu f = \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \tilde{\partial}_\nu f$$

Por tanto,

$$\partial_\mu f dx^\mu = \left( \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \tilde{\partial}_\nu f \right) \left( \frac{\partial x^\mu}{\partial \tilde{x}^\sigma} d\tilde{x}^\sigma \right) = \underbrace{\frac{\partial \tilde{x}^\nu}{\partial x^\mu} \frac{\partial x^\mu}{\partial \tilde{x}^\sigma}}_{\delta_\sigma^\nu} \tilde{\partial}_\nu f d\tilde{x}^\sigma f = \tilde{\partial}_\mu f d\tilde{x}^\mu$$

Por tanto, tenemos que  $\partial_\mu f dx^\mu = \tilde{\partial}_\mu f d\tilde{x}^\mu$  es covariante. Luego, para las coordenadas curvilíneas tenemos,

$$\begin{aligned} \partial_r &= \frac{\partial x^1}{\partial r} \partial x^1 + \frac{\partial x^2}{\partial r} \partial x^2 = \cos \theta \partial x^1 + \sin \theta \partial x^2 \\ \partial_\theta &= \frac{\partial x^1}{\partial \theta} \partial x^1 + \frac{\partial x^2}{\partial \theta} \partial x^2 = -r \sin \theta \partial x^1 + r \cos \theta \partial x^2 \end{aligned}$$

Así, el elemento de línea cambia como,

$$ds^2 = (dx^1)^2 + (dx^2)^2 = dr^2 + r^2 d\theta^2 = (d\tilde{x}^1)^2 + (\tilde{x}^1)^2 (d\tilde{x}^2)^2$$

Luego, antes teníamos que  $\delta_{\mu\nu} = \text{diag}(1, 1)$ , mientras que ahora tenemos  $g_{\mu\nu} = \text{diag}(1, r^2)$ . Así, la métrica transforma según,

$$\delta_{\mu\nu} = \frac{\partial x^\sigma}{\partial \tilde{x}^\mu} \frac{\partial x^\rho}{\partial \tilde{x}^\nu} \tilde{g}_{\sigma\rho}$$

Por tanto, la métrica es un tensor con dos índices covariantes. A partir de ahora, en coordenadas curvilíneas, identificaremos  $\partial_\mu$  con una base de vectores contravariantes, es decir,

$$\partial_\mu \longleftrightarrow \vec{e}_\mu$$

tal que  $\vec{v} = v^\mu \vec{e}_\mu = \partial_\mu v^\mu$ .

Ahora para derivar vectores deberemos tener cuidado. Notamos que si estamos en el espacio euclídeo con un vector  $\vec{v}$  en coordenadas cartesianas, entonces derivamos como,

$$d(\vec{v}) = \partial_\mu (v^\nu \vec{e}_\nu) dx^\mu = \left[ (\partial_\mu v^\nu) \vec{e}_\nu + v^\nu (\partial_\mu \vec{e}_\nu) \right] dx^\mu = (\partial_\mu v^\nu) \vec{e}_\nu dx^\mu$$

donde hemos tachado el segundo sumando, pues en coordenadas cartesianas los vectores de la base son constantes, por lo que la derivada es nula.

Ahora vamos a derivar en coordenadas curvilíneas, tal que

$$d(\vec{v}) = \tilde{\partial}_\mu (\tilde{v}^\nu \vec{e}_\nu) d\tilde{x}^\mu = \left[ (\tilde{\partial}_\mu \tilde{v}^\nu) \vec{e}_\nu + \tilde{v}^\nu (\tilde{\partial}_\mu \vec{e}_\nu) \right] d\tilde{x}^\mu = \left[ (\tilde{\partial}_\mu \tilde{v}^\nu) \vec{e}_\nu + \tilde{v}^\nu \Gamma_{\mu\nu}^\sigma \vec{e}_\sigma \right] d\tilde{x}^\mu = (\nabla_\mu \tilde{v}^\nu) \vec{e}_\nu d\tilde{x}^\mu$$

donde  $\Gamma_{\mu\nu}^\sigma$  son los símbolos de Christoffel y  $\nabla_\mu$  es la derivada covariante que satisface:

1.  $\nabla_\mu g_{\nu\sigma} = 0$
2.  $\Gamma_{\nu\sigma}^\mu = \frac{1}{2} g^{\mu\sigma} [g_{\sigma\nu,\rho} + g_{\sigma\rho,\nu} - g_{\rho\nu,\sigma}]$ , con  $g_{\sigma\nu,\rho} = \tilde{\partial}_\rho g_{\sigma\nu}$

Por tanto, la derivada covariante se define como,

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma_{\mu\sigma}^\nu v^\sigma$$

donde usamos métricas planas en espacios curvilíneos, por lo que no requerimos del uso de variedades (esto lo veremos más adelante).

### 3.3. Variedad diferenciable

Una variedad diferenciable es un conjunto de puntos de dimensión  $n$  que localmente se parecen a  $\mathbb{R}^n$ .

### 3.3.1. Definiciones

**Definición 3.3.1.** *Un espacio topológico es un conjunto de puntos y sus entornos.*

**Definición 3.3.2.** *Un espacio topológico es de tipo Haussdorff si todo par de puntos poseen entornos abiertos disjuntos.*

Esto nos da una noción de 'separación' entre puntos.

**Definición 3.3.3.** *Un homeomorfismo (cartas coordenadas) es un aplicación entre dos espacios topológicos, tal que ella y su inversa son biyecciones continuas, es decir, mapean abiertos en abiertos.*

**Definición 3.3.4.** *Una variedad topológica es un espacio topológico de tipo Haussdorff, tal que cada punto tiene un entorno abierto homeomorfo a un abierto de  $\mathbb{R}^n$ . La denotaremos por  $\mathcal{M}$ .*

**Definición 3.3.5.** *Una carta (o sistema de coordenadas) es una dupla  $(U, \phi)$ , que viene dada por un abierto  $U \subset \mathcal{M}$  y un homeomorfismo*

$$\phi : u \rightarrow \phi(u) \in \mathbb{R}^n$$

que mapea un entorno abierto  $u \in U$  a  $\mathbb{R}^n$ .

**Definición 3.3.6.** *Un atlas  $\mathcal{C}^\infty$  es un conjunto de cartas  $(U_\alpha, \phi_\alpha)$  que cumplen:*

- *Cubren todo  $\mathcal{M}$ , es decir,  $\mathcal{M} = \bigcup_\alpha U_\alpha$ .*
- *Las cartas deben ser compatibles, es decir, si consideramos dos abiertos  $U_\alpha, U_\beta$  tal que  $U_\alpha \cap U_\beta \neq \emptyset$ , entonces la función de transición entre ambas cartas viene dada por*

$$F = \phi_\alpha \circ \phi_\beta^{-1} = \phi_\alpha(\phi_\beta^{-1})$$

*tal que*

$$F : \phi_\beta(U_\alpha \cap \phi_\beta) \rightarrow \phi_\alpha(U_\alpha \cup U_\beta)$$

*siendo  $F$  una función  $\mathcal{C}^\infty$  entre abiertos de  $\mathbb{R}^n$ .*

Además,  $\mathcal{M}$  tiene dimensión  $n$ , es decir,  $p \in \mathcal{M}$  es una colección de  $n$  números reales.

**Definición 3.3.7.** *Una variedad diferenciable es  $\mathcal{C}^\infty$  (suave) si es una variedad topológica con un atlas  $\mathcal{C}^\infty$ .*

Veamos de forma visual cómo actúan todas estas cosas,

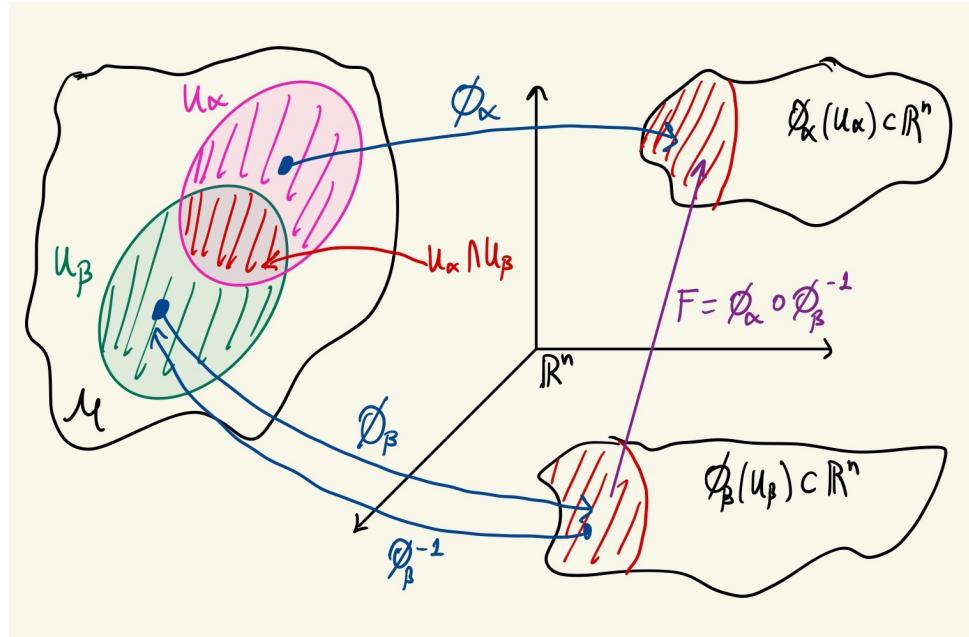


FIGURA 3.1: Esquema visual de cómo actúan los homeomorfismos sobre variedades.

**Nota 3.3.1.** Las aplicaciones  $\phi_\alpha$  pueden no ser  $\mathcal{C}^\infty$  (solo continuas), pero  $F = \phi_\alpha \circ \phi_\beta^{-1}$  sí debe ser  $\mathcal{C}^\infty$ .

Denotaremos por  $x^\mu$  a las coordenadas de un punto  $p \in \mathcal{M}$  definidas por  $\phi_\alpha(p)$ , es decir,  $x^\mu \equiv \phi^\mu(p)$ .

**Definición 3.3.8.** Una variedad es orientable si y solo si admite un atlas, tal que para cada par de abiertos  $U_\alpha$  y  $U_\beta$  no disjuntos, el Jacobiano entre coordenadas sea positivo,

$$\det \left( \frac{\partial x_\alpha}{\partial x_\beta} \right) > 0$$

Si la variedad tiene una frontera  $\partial\mathcal{M}$ , entonces sustituimos el mapeo a  $\mathbb{R}^n$  por el mapeo a  $\frac{1}{2}\mathbb{R}^n$ , siendo

$$\frac{1}{2}\mathbb{R}^n = \{x \in \mathbb{R}^n | x' \leq 0\}$$

de  $(n - 1)$ -dimensión.

### 3.3.2. Transformaciones de coordenadas

Definimos las transformaciones de coordenadas como

$$x' = \phi' \circ \phi^{-1}(x)$$

donde  $\phi$  puede ser un escalar, un vector, un campo vectorial, una uno-forma o un tensor. Veamos cada uno de ellos:

#### 3.3.2.1. Escalares

Sea una función cualquiera  $f : \mathcal{M} \rightarrow \mathbb{R}$ , esta induce a  $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$  mediante la regla

$$\tilde{f}(x) = (f \circ \phi^{-1})(x) = f(\phi^{-1}(x))$$

Diremos que  $f$  es suave o  $C^\infty$ , si  $\tilde{f} = f \circ \phi^{-1}$  lo es.

**Proposición 3.3.2.** *Si tenemos  $(U, \phi)$  y  $(U', \phi')$ , tal que  $U$  y  $U'$  son disjuntos, entonces se tiene que*

$$\tilde{f}(x') = \tilde{f}(x)$$

en  $p \in U \cap U'$ . Esta es la regla de transformación de los escalares.

En adelante vamos a referirnos a  $\tilde{f}$  como  $f$  y llamaremos  $F_{\mathcal{M}}$  al conjunto de funciones  $C^\infty$  sobre  $\mathcal{M}$ .

#### 3.3.2.2. Vectores

Como las variedades  $\mathcal{M}$  no son espacios vectoriales, no podemos definir vectores aquí. Entonces, lo que hacemos es definir el espacio tangente  $T_p$  en un punto  $p \in \mathcal{M}$ , definiéndose punto a punto. Consideramos para ello el conjunto de curvas  $\gamma(t)$  que pasan por  $p$ , es decir,

$$\gamma(t) : \mathbb{R} \rightarrow \mathcal{M}$$

tal que  $\gamma(t = t_0) = p$ .

Consideramos una función  $f$  arbitraria de  $F_{\mathcal{M}}$  y calculamos,

$$\begin{aligned} \frac{d}{dt} (f \circ \gamma(t))|_{t=t_0} &= \frac{d}{dt} \left( \underbrace{(f \circ \phi^{-1})}_{f(x)} \circ \underbrace{(\phi \circ \gamma(t))}_{\gamma^\mu(t)} \right) \Big|_{t=t_0} = \\ &= \frac{d}{dt} (f(x) \circ \gamma^\mu(t))|_{t=t_0} = \frac{d\gamma^\mu}{dt} \frac{\partial f}{\partial x^\mu} \Big|_{t=t_0} \equiv \vec{v}_\gamma(f)|_{t=t_0} \end{aligned}$$

donde  $\vec{v}_\gamma(f)|_{t=t_0}$  se define como un vector que vive en el espacio tangente (espacio vectorial)0 de un punto  $p \in \mathcal{M}$ .

Podemos interpretar  $\dot{\gamma}_\mu$  como las coordenadas del vector  $\vec{v}_\gamma$ , es decir,

$$v_\gamma^\mu = \dot{\gamma}^\mu$$

y además,  $\partial_\mu$  se interpreta como una base coordenada de vectores, es decir,  $\partial_\mu \equiv \vec{e}_\mu$ , ya que  $f$  es arbitraria.

La notación que usaremos para los vectores,  $\vec{v} : F_{\mathcal{M}} \rightarrow \mathbb{R}$ , será  $\vec{v} = v^\mu \vec{e}_\mu$  en  $T_p$ ; diremos que son vectores contravariantes.

La transformación de vectores se realizará usando dos bases coordenadas  $\{\vec{e}_\mu\}$  y  $\{\vec{e}'_\mu\}$  asociadas a las coordenadas  $x^\mu$  y  $x'^\mu$ , por lo que aplicando el Jacobiano tendremos

$$\vec{e}'_\mu = \frac{\partial x^\mu}{\partial x'^\nu} = \vec{e}_\nu \implies v'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} v^\nu$$

siendo esta la transformación de vectores. Además, los vectores de  $T_p$  satisfacen:

1.  $\vec{v}(\alpha f + \beta g) = \alpha \vec{v}(f) + \beta \vec{v}(g); \quad \forall f, g \in F_{\mathcal{M}}, \forall \alpha, \beta \in \mathbb{R}.$
2. La regla de Leibniz:

$$\vec{v}(fg) = g\vec{v}(f) + f\vec{v}(g); \quad \forall f, g \in F_{\mathcal{M}}$$

### 3.3.2.3. Campo vectorial en la variedad

Un campo vectorial en  $\mathcal{M}$  es una asignación de un vector  $v \in T_p$  en cada punto  $p \in \mathcal{M}$ , tal que sus componentes en cualquier base coordinada sean suaves ( $C^\infty$ ).

El conmutador, o corchete de Lie, se define como

$$[\vec{v}, \vec{w}] = \vec{v} \circ \vec{w} - \vec{w} \circ \vec{v} = v^\mu (\partial_\mu w^\nu) \partial_\nu - w^\mu (\partial_\mu v^\nu) \partial_\nu$$

donde  $[\vec{v}, \vec{w}] = \vec{u}$ , tal que  $u^\mu = v^\nu \partial_\nu w^\mu - w^\nu \partial_\nu v^\mu$ .

Sus propiedades son:

1. Antisimetría:

$$[\vec{v}, \vec{w}] = - [\vec{w}, \vec{v}]$$

2. Identidad de Jacobi:

$$[\vec{u}, [\vec{v}, \vec{w}]] + [\vec{v}, [\vec{w}, \vec{u}]] + [\vec{w}, [\vec{u}, \vec{v}]] = 0$$

3. Para una base coordenada,  $[\vec{e}_\mu, \vec{e}_\nu] = 0$ .

4. Para una base no coordinada,  $[\vec{e}_\mu, \vec{e}_\nu] \neq 0$

**Proposición 3.3.3.** *Sea  $\{\vec{e}_\mu\}$  un base coordinada, entonces su transformación*

$$\vec{e}'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} \vec{e}_\nu$$

*también es una base coordinada.*

Denotaremos las bases coordenadas con índices latinos, es decir, usaremos  $\vec{e}_a$ , con  $\vec{v} = v^a \vec{e}_a$ , y nos referiremos a  $v^a$  y  $\vec{v}$  indistintamente.

#### 3.3.2.4. Uno-formas

Las uno-formas son aplicaciones lineales reales en  $T_p$ , tal que

$$\vec{f} : T_p \rightarrow \mathbb{R}$$

donde el producto escalar  $\langle \vec{f}, \vec{v} \rangle \in \mathbb{R}$  con  $\vec{v} \in T_p$ .

Dada una base coordinada  $\{\vec{e}_a\}$  en  $T_p$ , existe un único conjunto de uno-formas  $\{\bar{e}^a\}$  que

cumplen  $\langle \vec{e}^a, \vec{e}_b \rangle = \delta_b^a$ , donde  $\delta_b^a$  es la delta de Kronecker. Diremos que  $\{\vec{e}^a\}$  es una base dual de  $T_p^*$ , que el espacio dual tangente a  $p \in \mathcal{M}$ , también llamado **espacio cotangente**.

Dados  $\vec{v} = v^a \vec{e}_a$  y  $\vec{f} = f_a \vec{e}^a$ , entonces  $\langle \vec{f}, \vec{v} \rangle = f_a v^a$ , por tanto, el producto escalar de vectores con uno-formas se puede interpretar como una contracción de índices.

Además, diremos que  $f_a$  son las componentes de un vector covariante.

Dada una función  $f : \mathcal{M} \rightarrow p$ , podemos construir una uno-forma de la siguiente manera:

-Dado un  $\vec{v}$ ,

$$\vec{v}(f) = v^\mu \partial_\mu f = v^\mu \delta_\mu^\nu \partial_\nu f = v^\mu \partial_\nu f \langle \vec{e}^\nu, \vec{e}_\mu \rangle = \langle \partial_\nu f \vec{e}^\nu, v^\mu \vec{e}_\mu \rangle = d\vec{f}, \vec{v} \rangle$$

por tanto,

$$d\vec{f} = \partial_\mu f \vec{e}^\mu = \partial_\mu f dx^\mu$$

De tal forma que  $\langle dx^\mu, \partial_\nu \rangle = \delta_\nu^\mu$ . Por tanto, las uno-formas transforman como

$$\vec{e}'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} \vec{e}^\nu \implies f'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} f_\nu$$

### 3.3.2.5. Tensores

Un tensor es una aplicación multilineal real de tipo  $(r, s)$ , denotada como  $T_{\mu_1 \mu_2 \dots \mu_r}^{\nu_1 \nu_2 \dots \nu_s}$  que actúa como,

$$T(\vec{v}_1, \vec{v}_2, \dots, \vec{f}_1, \vec{f}_2, \dots) = T_{\mu_1 \mu_2 \dots}^{\nu_1 \nu_2 \dots} v_1^{\mu_1} v_2^{\mu_2} \dots f_{1\nu_1} f_{2\nu_2} \dots$$

Los tensores transforman bien bajo transformaciones de coordenadas.

Dos tensores del mismo rango se pueden sumar (multiplicación por escalares).

**Definición 3.3.9.** Definimos la parte simétrica de un tensor de rango 2 como

$$T_{(\mu\nu)} = \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu})$$

**Definición 3.3.10.** Definimos la parte antisimétrica de un tensor de rango 2 como

$$T_{[\mu\nu]} = \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$$

Podemos cambiar el carácter tensorial, realizando una multiplicación tensorial o haciendo una contracción de índices.

**Ejemplo 3.3.4.** Dados  $R_{\mu\nu}$  y  $T_{\mu\nu}$ , podemos construir  $R \otimes T \rightarrow R_{\mu\nu}T_{\sigma\rho} \equiv G_{\mu\nu\sigma\rho}$ , siendo el producto tensorial

**Ejemplo 3.3.5.** Dados  $T_{\mu\nu}$  y  $v^\mu$ , podemos definir  $T_{\mu\nu}v^\nu = h_\mu$ , siendo una contracción de índices.

### Tensor métrico

Definimos el tensor métrico como la distancia cuadrática infinitesimal al desplazarnos en una dirección. Será por tanto un tensor de rango (0, 2), pues es cuadrática, y además asumimos que es un tensor simétrico.

En una base coordenada, las componentes del tensor métrico son

$$g_{\mu\nu} = g(\vec{e}_\mu, \vec{e}_\nu)$$

También lo escribiremos como

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu$$

donde  $ds^2$  es el elemento de línea.

Definimos la norma de un vector como

$$|\vec{v}| = \sqrt{|g_{\mu\nu}v^\mu v^\nu|}$$

Definimos el ángulo entre dos vectores  $\vec{v}$  y  $\vec{w}$  como

$$\cos \theta = \frac{g_{\mu\nu}v^\mu w^\nu}{\sqrt{|g_{\mu\nu}v^\mu v^\nu|}\sqrt{|g_{\mu\nu}w^\mu w^\nu|}}$$

con  $|\vec{v}| \neq 0$  y  $|\vec{w}| \neq 0$ .

Diremos que  $g_{\mu\nu}$  es una métrica **degenerada** en un punto  $p$  si es una matriz singular en cualquier base coordenada  $\{\vec{e}_\mu\}$ , es decir, que su determinante es igual a cero. En caso contrario, existirá un tensor de rango  $(2, 0)$  que cumple que  $g^{\mu\nu}g_{\nu\rho} = \delta_\rho^\mu$ . Es decir,  $g^{\mu\nu}$  es la **inversa** de  $g_{\mu\nu}$ .

La métrica (y su inversa) nos permiten 'subir' y 'bajar' índices, por tanto podremos transformar vectores en uno-formas,  $v_\mu = g_{\mu\nu}v^\nu$ , y viceversa,  $f^\mu = g^{\mu\nu}f_\nu$ .

### Signatura

La signatura de la métrica se obtiene diagonalizando la matriz  $g_{\mu\nu}$ . Si la signatura es del tipo  $(+++)$ , diremos que es una métrica riemanniana; si es del tipo  $(- + +)$ , diremos que es lorentziana.

Las métricas físicas son lorentzianas e introducen una estructura casual, tal que

- Si  $v^\mu v^\nu g_{\mu\nu} > 0$ , entonces  $v^\mu$  es un vector tipo espacial.
- Si  $v^\mu v^\nu g_{\mu\nu} < 0$ , entonces  $v^\mu$  es un vector tipo temporal.
- Si  $v^\mu v^\nu g_{\mu\nu} = 0$ , entonces  $v^\mu$  es un vector tipo luz o nulo.

También existe la métrica lorentziana  $(- - -)$  que se usa en física de partículas relativista, donde la estructura casual se invierte, es decir, el vector espacial pasa a temporal y viceversa.

## 3.4. Aplicaciones diferenciables y difeomorfismos

Si tenemos una solución relacionada por un difeomorfismo con la solución de las ecuaciones de Einstein, entonces la nueva solución es exactamente igual.

**Definición 3.4.1.** Una aplicación diferenciable  $U : \mathcal{M} \rightarrow \mathcal{M}'$  es  $\mathcal{C}^\infty$  si y solo si, dados dos atlas  $\{(U_\alpha, \phi_\alpha)\}, \{(U'_\beta, \phi'_\beta)\}$  de  $\mathcal{M}$  y  $\mathcal{M}'$  respectivamente, si las funciones

$$\phi'_\beta \circ \varphi \circ \phi'_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$$

son  $\mathcal{C}^\infty$ . Es decir, las coordenadas de  $\mathcal{M}$  mapean suavemente a  $\mathcal{M}'$ . (En principio,  $\dim \mathcal{M} \neq \dim \mathcal{M}'$ ).

La aplicación  $\varphi$  induce a una aplicación (natural) entre funciones de  $\mathcal{M}'$  a  $\mathcal{M}$ .

**Definición 3.4.2.** Definimos  $\varphi^*$  como el pull-back ('la aplicación hacia atrás') para  $f' : \mathcal{M}' \rightarrow \mathbb{R}$  tal que

$$(\varphi^* f')|_p = f' \circ \varphi(p) : \mathcal{M} \rightarrow \mathbb{R}$$

En el lenguaje de coordenadas, esto quiere decir que si  $f'(x') : \mathbb{R}^{n'} \rightarrow \mathbb{R}$ , entonces  $f'(x) = f'(x'(x))$ .

La aplicación inversa al pull-back se nombra como push-forward y se denota por  $\varphi_*$ . Para funciones, a priori, no está definida, ya que necesitamos  $\varphi^{-1}$ , que no está definido, al igual que  $x' = x'(x)$  sí está bien definido, pero su inversa no.

Debemos notar que  $\varphi$  sí induce una aplicación push-forward en el espacio tangente  $T_p$ . Pues si tenemos un vector  $\vec{v} \in T_p$ , entonces el vector  $(\varphi_* \vec{v}) \in T_{p'=\varphi(p)}$  sí está bien definido, y se define como

$$(\varphi_* \vec{v}) (f')|_{p'=\varphi(p)} = \vec{v} (\varphi^* f')|_p = \vec{v} (f' \circ \varphi)|_p$$

En coordenadas tenemos que

$$\vec{v} (f \circ \varphi)|_p = v^\mu \partial_\mu (f(x'(x))) = v^\mu \frac{\partial x'^\nu}{\partial x^\mu} \partial'_\nu f = ((\varphi_*)^\mu_\nu v^\nu) \partial'_\mu f = (\varphi_* \vec{v}) (f')|_{p'=\varphi(p)}$$

En resumen,

$$(\varphi_*)^\mu_\nu = \frac{\partial x'^\mu}{\partial x^\nu}$$

Normalmente, la transformación pull-back para vectores implica conocer  $\varphi^{-1}$ , que no está definido.

Para las uno-formas, el pull-back está bien definido, pues dado  $\vec{f}' \in T_{p'=\varphi(p)}^*$ , tenemos que

$$< \varphi^* \vec{f}', \vec{v} >|_p = < \vec{f}', \varphi_* \vec{v} >|_p$$

En el lenguaje de coordenadas tenemos que

$$(\varphi^* \vec{f}')_\mu v^\mu = f'_\mu (\varphi_*)^\mu_\nu v^\nu$$

por tanto, tenemos que

$$(\varphi^* \vec{f})_\mu = \frac{\partial x'^\nu}{\partial x^\mu} f'_\nu$$

Recordemos que el diferencial de  $f'$  es una uno-forma; y que el pull-back y el diferencial conmutan; por tanto,

$$\langle \varphi^*(f\vec{f'}), \vec{v} \rangle|_p = \langle d\vec{f'}, \varphi_*\vec{v} \rangle|_{\varphi(p)} = (\varphi_*\vec{v})(f')|_{\varphi(p)} = \vec{v}(\varphi^*f')|_p = \langle d(\varphi^*f'), \vec{v} \rangle|_p$$

**Definición 3.4.3.** *Un difeomorfismo es una aplicación diferenciable  $\varphi : \mathcal{M} \rightarrow \mathcal{M}'$ , tal que tanto ella como su inversa son biyectivas y de clase  $C^\infty$ , siempre que  $\dim \mathcal{M} = \dim \mathcal{M}'$ .*

Tanto las aplicaciones pull-back como las push-forward están bien definidas en el contexto de los difeomorfismos. Son isomorfos entre los espacios tangentes y cotangentes.

En resumen, diremos que si  $\mathcal{M}$  y  $\mathcal{M}'$  están relacionados mediante un difeomorfismo y tenemos  $(\mathcal{M}, g)$  y  $(\mathcal{M}', g')$ , entonces **son físicamente indistinguibles**.

Los difeomorfismos son una simetría de la relatividad general y tienen estructura de grupo, es decir,  $\varphi_! \circ \varphi_2 = \varphi_3$ .

Los difeomorfismos se pueden entender como transformaciones activas que mapean tensores en  $p$  a tensores en  $q = \varphi(p)$ , donde  $p, q \in \mathcal{M}$ . Pero admiten una perspectiva como transformaciones pasivas. Si tenemos una base coordenada con  $x^\mu$  en un entorno  $U$  de  $p$ , y la nueva base con  $x'^\mu$  en un entorno  $U'$  de  $q = \varphi(p)$ , siempre podemos definir unas nuevas coordenadas

$$y^\mu(p) = x'^\mu(\varphi(p))$$

un pull-back, tal que  $x^\mu = x^\mu(y)$ .

En resumen, transformamos tensores en  $p$  a tensores en  $p$ .

Desde un punto de vista físico, trabajar con transformaciones pasivas o activas es indistinguible, y como estamos acostumbrados a las pasivas, usaremos estas.

## 3.5. Derivada de Lie

La derivada de Lie nos dice cómo derivar funciones escalares, campos vectoriales y campos tensoriales de forma general. Además, nos dice cómo conectar espacios tangentes  $T_p$  y  $T_q$  de dos puntos  $p, q \in \mathcal{M}$  con  $p \neq q$ .

Para ello, necesitamos un campo vectorial  $\vec{\xi}$ , que define un conjunto de difeomorfismos. En concreto, por cada punto  $p \in \mathcal{M}$  pasa una curva  $\gamma_p(t)$  tal que

$$\dot{\gamma}_p(t) = \xi^\mu(p)$$

Esta curva define el difeomorfismo siguiente,

$$\varphi_t(p) = \gamma_p(t); \quad \varphi_t : U \rightarrow \varphi_t(U)$$

La derivada de Lie de una función escalar  $f$  a lo largo de  $\vec{\xi}$  se define como,

$$\mathcal{L}_{\vec{\xi}} f|_p = \vec{\xi}(f) \equiv \xi^\mu \partial_\mu f$$

La derivada de Lie de un campo vectorial se define como,

$$\mathcal{L}_{\vec{\xi}} \vec{v}|_p = \lim_{t \rightarrow 0} \frac{\varphi_t^*(\vec{v}(\varphi_t(p))) - \vec{v}|_p}{t}$$

Se puede demostrar que

$$\mathcal{L}_{\vec{\xi}} \vec{v}|_p = [\vec{\xi}, \vec{v}] = -\mathcal{L}_{\vec{v}} \vec{\xi}|_p$$

En coordenadas se escribe,

$$(\mathcal{L}_{\vec{\xi}} \vec{v})^\mu = \xi^\nu \partial_\nu v^\mu - v^\nu \partial_\nu \xi^\mu$$

Sus propiedades son:

1. La derivada de Lie preserva el tipo tensorial, las simetrías y las operaciones.
2. Es una aplicación lineal.
3. Satisface la regla de Leibniz,

$$\mathcal{L}_{\vec{\xi}}(\vec{u} \otimes \vec{v}) = (\mathcal{L}_{\vec{\xi}} \vec{u}) \otimes \vec{v} + \vec{u} \otimes (\mathcal{L}_{\vec{\xi}} \vec{v})$$

4. Cumple que

$$\langle \vec{f}, \mathcal{L}_{\vec{\xi}} \vec{v} \rangle = \mathcal{L}_{\vec{\xi}}(\langle \vec{f}, \vec{v} \rangle) - \langle \mathcal{L}_{\vec{\xi}} \vec{f}, \vec{v} \rangle$$

La derivada de un tensor se define como

$$\begin{aligned} \mathcal{L}_{\vec{\xi}}(T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r}) &= \xi^\mu \partial_\mu T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} - T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_2 \dots \mu_r} \partial_\nu \xi^{\mu_1} - T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \nu \dots \mu_r} \partial_\nu \xi^{\mu_2} - \dots - T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_{r-1} \nu} \partial_\nu \xi^{\mu_r} + \\ &+ T_{\nu \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} \partial_\nu \xi^\nu + \dots + T_{\nu_1 \nu_2 \dots \nu_{s-1} \nu}^{\mu_1 \mu_2 \dots \mu_r} \partial_\nu \xi^\nu \end{aligned}$$

La derivada de Lie nos permite saber cuándo hay una simetría, pues si  $\mathcal{L}_{\xi}\vec{v} = 0$ , entonces  $\vec{v}$  es simétrico bajo  $\varphi_t(p)$ .

La derivada de Lie necesita conocer los campos vectoriales y sus derivadas en las direcciones no tangenciales a las curvas que unen  $p$  y  $q$ . Como la derivada de Lie necesita mucha información de primeras, puesto que hay infinitos campos fuera de la curva, introduciremos la noción de **derivada covariante**. Esta derivada, por otro lado, solo necesitará información sobre la curva y las direcciones tangentes a ella.

### 3.5.1. Derivada covariante

La derivada covariante es una derivada que está enteramente definida en  $T_p$  y al actuar sobre un tensor,  $T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r}$ , nos devolverá otro tensor de tipo  $(r, s+1)$ , tal que

$$\nabla_{\nu_{s+1}} T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} \equiv T_{\nu_1 \nu_2 \dots \nu_s; \nu_{s+1}}^{\mu_1 \mu_2 \dots \mu_r}$$

Sus propiedades son las siguientes:

1. Es lineal,

$$\nabla_{\nu} (\alpha T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} + \beta S_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r}) = \alpha \nabla_{\nu} T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} + \beta \nabla_{\nu} S_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r}$$

con  $\alpha, \beta \in \mathbb{R}$ .

2. Satisface la regla de Leibniz (regla de la cadena),

$$\nabla_{\nu} \left( T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} S_{\sigma_1 \sigma_2 \dots \sigma_{s'}}^{\rho_1 \rho_2 \dots \rho_{r'}} \right) = S_{\sigma_1 \sigma_2 \dots \sigma_{s'}}^{\rho_1 \rho_2 \dots \rho_{r'}} \nabla_{\nu} T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} + T_{\nu_1 \nu_2 \dots \nu_s}^{\mu_1 \mu_2 \dots \mu_r} \nabla_{\nu} S_{\sigma_1 \sigma_2 \dots \sigma_{s'}}^{\rho_1 \rho_2 \dots \rho_{r'}}$$

3. Commuta con la contracción,

$$\nabla_{\nu} (T_{\mu \nu_2 \dots \nu_s}^{\mu \mu_2 \dots \mu_r}) = \nabla_{\nu} T_{\mu \nu_2 \dots \nu_s}^{\mu \mu_2 \dots \mu_r}$$

4. Sobre funciones actúa como,

$$\nabla_{\mu} f = (df)_{\mu}$$

En coordenadas,

$$(df)_{\mu} = \partial_{\mu} f$$

De la propiedad 4. tenemos que

$$\nabla_\mu f = (\vec{df})_\mu \partial_\mu f$$

Sobre  $\langle \vec{f}, \vec{v} \rangle = f_\mu v^\mu$  actúa  $\nabla_\mu(f_\nu v^\mu) = f_\nu \partial_\mu v^\nu$ .

### Observación 3.5.1.

- $\partial_\mu v^\nu$  no es un tensor, pues

$$\partial'_\mu v'^\nu = \frac{\partial x^\rho}{\partial x'^\mu} \partial_\rho \left( \frac{\partial x'^\nu}{\partial x^\sigma} v^\sigma \right) = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x'^\nu}{\partial x^\sigma} \partial_\rho v^\sigma + \underbrace{\frac{\partial x^\rho}{\partial x'^\mu} \partial_\rho \left( \frac{\partial x'^\nu}{\partial x^\sigma} \right) v^\sigma}$$

donde el término señalado hace que no sea un tensor, pues hace que no transforme como un tensor.

- $\nabla_\mu v^\nu$  es un tensor, pues

$$\begin{aligned} \nabla_\mu(f_\nu v^\nu) &= \partial_\mu(f_\nu v^\nu) = f_\nu \partial_\mu v^\nu + v^\nu \partial_\mu f^\nu \\ &\parallel \\ f_\nu \nabla_\mu v^\nu + v^\nu \nabla_\mu f_\nu &= f_\nu (\partial_\mu v^\nu + \Gamma_{\mu\rho}^\rho v^\rho) + v^\nu (\partial_\mu f_\nu + \bar{\Gamma}_{\mu\nu}^\rho f_\rho) = \\ &= f_\nu \partial_\mu v^\nu + v^\nu \partial_\mu f_\nu + \underbrace{\Gamma_{\mu\sigma}^\rho v^\sigma f_\rho + \bar{\Gamma}_{\mu\rho}^\sigma v^\rho f_\sigma}_{\parallel} \\ &\parallel \\ &0 \end{aligned}$$

donde el último término se anula porque  $\bar{\Gamma}_{\mu\rho}^\sigma = -\Gamma_{\mu\rho}^\sigma$ .

En resumen, tenemos que

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma_{\mu\rho}^\nu v^\rho$$

es la derivada covariante de vectores.

$$\nabla_\mu f_\nu = \partial_\mu f_\nu - \Gamma_{\mu\nu}^\rho f_\rho$$

es la derivada covariante de formas.

Notemos que  $\nabla_\mu v^\nu$  es un tensor, pero  $\partial_\mu v^\nu$  no es un tensor, por tanto  $\Gamma_{\mu\nu}^\rho$  no transforma como un tensor, pues

$$\Gamma'_{\mu\rho} = \frac{\partial x'^\nu}{\partial x^\gamma} \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\delta}{\partial x'^\rho} \Gamma_{\sigma\delta}^\gamma - \frac{\partial x^\gamma}{\partial x'^\nu} \frac{\partial x^\sigma}{\partial x'^\rho} \partial_\gamma \left( \frac{\partial x'^\nu}{\partial x^\sigma} \right)$$

En cambio, la combinación  $\partial_\mu v^\nu + \Gamma_{\mu\rho}^\nu v^\rho$  sí es un tensor.

La derivada covariante de un tensor tipo  $(r, s)$  viene dada por,

$$\begin{aligned}\nabla_\mu T_{\nu_1\nu_2\dots\nu_s}^{\mu_1\mu_2\dots\mu_r} &= \partial_\mu T_{\nu_1\nu_2\dots\nu_s}^{\mu_1\mu_2\dots\mu_r} + \Gamma_{\mu\rho}^{\mu_1} T_{\nu_1\nu_2\dots\nu_s}^{\rho\mu_2\dots\mu_r} + \dots + \Gamma_{\mu\rho}^{\mu_r} T_{\nu_1\nu_2\dots\nu_s}^{\mu_1\mu_2\dots\mu_{r-1}\rho} - \\ &\quad - \Gamma_{\mu\nu_1}^\rho T_{\rho\nu_2\dots\nu_s}^{\mu_1\mu_2\dots\mu_r} - \dots - \Gamma_{\mu\nu_s}^\rho T_{\nu_1\nu_2\dots\nu_{s-1}\rho}^{\mu_1\mu_2\dots\mu_r}\end{aligned}$$

En una base no coordenada cualquiera, tenemos que la diferencia de dos conexiones en un tensor de tipo  $(1, 2)$  se cumple que

$$\begin{cases} (\nabla_a - \bar{\nabla}_a) f_b = -C_{ab}^c f_c \\ (\nabla_a - \bar{\nabla}_a) v^b = C_{ac}^b v^c \end{cases} \quad \text{con } C_{bc}^a = \Gamma_{bc}^a - \bar{\Gamma}_{bc}^a, \text{ que es un tensor.}$$

De entre todas las conexiones posibles, vamos a tomar aquellas conexiones que son simétricas, es decir, que  $\nabla_\mu \nabla_\nu f = \nabla_\nu \nabla_\mu f$ . Por tanto, tendremos que los símbolos de Christoffel, en una base coordinada, son simétricos,

$$\Gamma_{\mu\nu}^\rho = \Gamma_{\nu\mu}^\rho = \Gamma_{(\mu\nu)}^\rho$$

Como consecuencia, tenemos que

$$\left( \mathcal{L}_\xi \vec{v} \right)^\mu = \xi^\nu \partial_\nu v^\mu - v^\nu \partial_\nu \xi^\mu = \xi^\nu \nabla_\nu v^\mu - v^\nu \nabla_\nu \xi^\mu$$

Por tanto, podemos sustituir  $\partial_\mu \rightarrow \nabla_\mu$  en la derivada de Lie.

### 3.6. Conexión de Levi-Civita

Es la única conexión simétrica y compatible con la métrica, es decir,

$$\nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma_{\mu\nu}^\sigma g_{\sigma\rho} - \Gamma_{\mu\rho}^\sigma g_{\nu\sigma} = 0$$

Por tanto, podemos llegar a una definición de los símbolos de Christoffel en la conexión de Levi-Civita que solo depende de la métrica, tal que

$$\Gamma_{\nu\rho}^\mu = \frac{1}{2} g^{\mu\sigma} (g_{\sigma\nu,\rho} + g_{\sigma\rho,\nu} - g_{\nu\rho,\sigma})$$

La conexión de Levi-Civita preserva las normas y los ángulos bajo el transporte paralelo (que se explicará en la siguiente sección).

Además, para la conexión de Levi-Civita se cumple que

$$(\mathcal{L}_{\vec{\xi}} g_{\mu\nu}) = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu$$

Podemos calcular las simetrías de nuestra variedad resolviendo  $(\mathcal{L}_{\vec{\xi}} g_{\mu\nu}) = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0$ , obteniendo así las cantidades simétricas de nuestra variedad con la métrica escogida.

### 3.7. Transporte paralelo

Es una forma 'barata' de movernos de un punto  $p$  a un punto  $q$  de la variedad. Decimos que es 'barata' porque solo necesitamos una curva  $\gamma_p(t)$  que pase por  $p$  y  $q$ , y una conexión  $\nabla_a$ .

Sea  $\vec{v} \in T_p$  y sea  $t^a$  el vector tangente a  $\gamma_p(t)$ . Dada una conexión  $\nabla_a$ , se define el transporte paralelo de  $\vec{v}$  sobre  $\gamma_p(t)$  como

$$t^a \nabla_a v^b = 0$$

teniendo una solución única. En coordenadas tenemos que  $t^\mu = \dot{x}^\mu = \frac{dx^\mu}{dt}$ , tal que

$$t^\mu \partial_\mu v^\nu + \Gamma_{\mu\rho}^\nu t^\mu v^\rho = 0$$

que equivale a

$$\dot{x}^\mu \partial_\mu v^\nu + \Gamma_{\mu\rho}^\nu \frac{dx^\mu}{dt} v^\rho = \frac{dv^\nu}{dt} + \Gamma_{\mu\rho}^\nu \frac{dx^\mu}{dt} v^\rho = 0$$

Tenemos un conjunto de  $n$ -ecuaciones diferenciales ordinarias de primer orden y lineales en  $v^\mu$ . Existe solución y es única. Además, cualquier combinación lineal de vectores  $v^\mu$  también es un vector de transporte paralelamente.

El transporte paralelo induce un isomorfismo entre los espacios tangentes  $T_p$  y  $T_q$ . Este isomorfismo depende de  $\nabla_\mu$  y de  $\gamma_p(t)$ .

**Nota 3.7.1.** Si la curvatura asociada a  $\nabla_a$  es cero, entonces el transporte paralelo no dependerá de  $\gamma_p(t)$ .

Para la conexión de Levi-Civita, dados  $v^\mu$  y  $u^\mu$ , que satisfacen  $t^\mu \nabla_\mu v^\nu = 0$  y  $t^\mu \nabla_\mu u^\nu = 0$ , con  $t^\mu$  tangente a  $\gamma_p(t)$ , entonces

$$t^\mu \nabla_\mu (v^\nu u^\rho g_{\nu\rho}) = u^\rho g_{\nu\rho} t^\mu \cancel{\nabla_\mu v^\nu} + v^\nu g_{\nu\rho} t^\mu \cancel{\nabla_\mu u^\rho} + v^\nu u^\rho t^\mu \cancel{\nabla_\mu g_{\nu\rho}} = 0$$

Por tanto, los ángulos y las normas de los vectores se preservan al ser transportados paralelamente con la conexión de Levi-Civita, independientemente de  $\gamma_p(t)$ .

### 3.8. Geodésicas

Una curva  $\gamma_p(s)$ , donde usamos el parámetro afín  $s$ , se dice que es geodésica si su vector tangente cumple que,

$$\frac{dv^\mu}{ds} = 0 \iff v^\mu \nabla_\mu v^\nu = 0$$

Este parámetro afín es único salvo multiplicación y adición de una constante. En coordenadas tenemos,

$$\begin{aligned} v^\mu \partial_\mu v^\nu + \Gamma_{\mu\rho}^\nu v^\mu v^\rho &= 0 \\ \left| \right| \quad & \\ \frac{dx^\mu}{ds} \partial_\mu \left( \frac{dx^\nu}{ds} \right) + \Gamma_{\mu\rho}^\nu \frac{dx^\mu}{ds} \frac{dx^\rho}{ds} &= 0 \end{aligned}$$

Por tanto, la ecuación de la geodésica queda

$$\frac{d^2x^\nu}{ds^2} + \Gamma_{\mu\rho}^\nu \frac{dx^\mu}{ds} \frac{dx^\rho}{ds} = 0$$

teniendo así un sistema de  $n$ -ecuaciones diferenciales ordinarias de segundo orden no lineales. Estas ecuaciones diferenciales tienen solución y es única, pues satisfacen teoremas de existencia y unicidad.

Localmente, podemos encontrar coordenadas normales  $x = x(x')$ , donde la ecuación de las geodésicas queda como,

$$\frac{d^2x'^\mu}{ds^2} = 0$$

### 3.8.1. Geodésicas como Principio Variacional

Postulamos una acción,

$$S = \int \sqrt{ds^2} = \int \sqrt{g_{\mu\nu}(x) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda = \int \mathcal{L} d\lambda$$

donde si  $\mathcal{L}^2 > 0$ , entonces tenemos vectores espaciales.

Si interpretamos  $\mathcal{L}$  como un Lagrangiano, podemos obtener las ecuaciones de Euler-Lagrange, tal que

$$\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{d}{d\lambda} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \right) = 0$$

o bien,

$$\frac{\partial(\mathcal{L}^2)}{\partial x^\mu} - \frac{d}{d\lambda} \left( \frac{\partial(\mathcal{L}^2)}{\partial \dot{x}^\mu} \right) = -2 \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \dot{\mathcal{L}}$$

donde  $\dot{x}^\mu = \frac{\partial x^\mu}{\partial \lambda}$ . Lo calculamos,

$$\begin{aligned} \frac{\partial(\mathcal{L}^2)}{\partial x^\mu} &= (\partial_\mu g_{\rho\nu}) \frac{dx^\rho}{d\lambda} \frac{dx^\nu}{d\lambda} \\ \frac{\partial(\mathcal{L}^2)}{\partial \dot{x}^\mu} &= 2g_{\mu\nu} \frac{dx^\nu}{d\lambda} \\ \frac{d}{d\lambda} \left( \frac{\partial(\mathcal{L}^2)}{\partial \dot{x}^\mu} \right) &= 2g_{\mu\nu} \frac{d^2 x^\nu}{d\lambda^2} + 2(\partial_\sigma g_{\mu\nu}) \frac{dx^\sigma}{d\lambda} \frac{dx^\nu}{d\lambda} \end{aligned}$$

Por tanto,

$$\frac{\partial(\mathcal{L}^2)}{\partial x^\mu} - \frac{d}{d\lambda} \left( \frac{\partial(\mathcal{L}^2)}{\partial \dot{x}^\mu} \right) = -2g_{\mu\nu} \left( \frac{d^2 x^\nu}{d\lambda^2} + \Gamma_{\rho\sigma}^\nu \dot{x}^\rho \dot{x}^\sigma \right) = -2 \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \dot{\mathcal{L}}$$

Entonces, la ecuación de la geodésica generalizada queda,

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\rho\sigma}^\mu \dot{x}^\rho \dot{x}^\sigma = \dot{x}^\mu \frac{\dot{\mathcal{L}}}{\mathcal{L}}$$

Para llegar a la ecuación de la geodésica debemos usar el parámetro afín, pues el término de la derecha de la igualdad no se anula debido a que  $\lambda$  no es el parámetro afín. Luego, redefinimos  $\mathcal{L} d\lambda \equiv ds$ , teniendo así la ecuación de la geodésica,

$$\frac{d^2 x^\mu}{ds^2} + \Gamma_{\rho\sigma}^\mu \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 0$$

### 3.8.2. Derivación de los términos de Christoffel mediante las geodésicas

Tomamos como Lagrangiano  $\tilde{\mathcal{L}} = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ . Veamos un ejemplo de cómo derivar los símbolos de Christoffel usando la ecuación de la geodésica.

**Ejemplo 3.8.1.** Tomamos la métrica en coordenadas cilíndricas,

$$ds^2 = dr^2 + r^2d\theta^2 + dz^2$$

por tanto, el Lagrangiano queda

$$\tilde{\mathcal{L}} = \dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2$$

resolvemos las ecuaciones de Euler-Lagrange,

para el eje  $z$ :

$$\frac{\partial \tilde{\mathcal{L}}}{\partial z} = 0; \quad \frac{\partial \tilde{\mathcal{L}}}{\partial \dot{z}} = 2\dot{z} \Rightarrow 0 - \frac{d}{ds}(2\dot{z}) = 0 \Rightarrow \ddot{z} = 0$$

luego, las ecuaciones de la geodésica para el eje  $z$  quedan,

$$\ddot{z} + \Gamma_{\mu\nu}^z \dot{x}^\mu \dot{x}^\nu = 0$$

por tanto  $\Gamma_{\mu\nu}^z = 0$ .

Para el eje  $r$ :

$$\frac{\partial \tilde{\mathcal{L}}}{\partial r} = 2r\dot{\theta}^2; \quad \frac{\partial \tilde{\mathcal{L}}}{\partial \dot{r}} = 2\dot{r} \Rightarrow 2r\dot{\theta}^2 - \frac{d}{ds}(2\dot{r}) = 0 \Rightarrow \ddot{r} = r\dot{\theta}$$

luego, las ecuaciones de la geodésica para el eje  $r$  quedan,

$$\ddot{r} + \Gamma_{\mu\nu}^r \dot{x}^\mu \dot{x}^\nu = 0$$

por tanto,

$$\Gamma_{\theta\theta}^r = -r; \quad \Gamma_{r\theta}^r = 0 = \Gamma_{rz}^r = \Gamma_{zr}^r = \Gamma_{\theta r}^r$$

Para el eje  $\theta$ :

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \theta} = 0; \quad \frac{\partial \tilde{\mathcal{L}}}{\partial \dot{\theta}} = 2r^2\dot{\theta} \Rightarrow 0 - \frac{d}{ds}(2r^2\dot{\theta}) = 0$$

luego, las ecuaciones de la geodésica para el eje  $\theta$  quedan,

$$\ddot{\theta} + 2\frac{1}{r}\dot{r}\dot{\theta} = 0$$

por tanto,

$$\Gamma_{r\theta}^\theta = \frac{1}{r} = \Gamma_{\theta r}^\theta$$

donde no aparece el 2 porque es simétrico.

### 3.8.3. Densidad tensorial

Sea un tensor  $T_{\rho\sigma\dots}^{\mu\nu\dots}$  multiplicando a  $(\sqrt{g})^\omega$ , con  $\omega = \dots, -2, -1, 0, 1, 2, \dots$ , su derivada covariante se define como

$$\nabla_\mu (\sqrt{g}^\omega T_{\nu_1\nu_2\dots}^{\mu_1\mu_2\dots}) = (\sqrt{-g})^\omega \nabla_\mu T_{\nu_1\nu_2\dots}^{\mu_1\mu_2\dots} - \frac{\omega}{2} \Gamma_{\nu\mu}^\nu (\sqrt{g})^\omega T_{\nu_1\nu_2\dots}^{\mu_1\mu_2\dots}$$

de forma que  $\nabla_\mu(\sqrt{g}) = 0$ .

## 3.9. Tensores de Curvatura

Dada una conexión  $\nabla_a$  y una uno-forma, el operador  $(\nabla_a \nabla_b - \nabla_b \nabla_a)f_c$  es lineal, es decir, sea  $h$  una función escalar, entonces

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)(hf_c) = h(\nabla_a \nabla_b - \nabla_b \nabla_a)f_c$$

Esto implica que

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)f_c = \mathcal{R}_{abc}^d f_d$$

donde  $\mathcal{R}_{abc}^d$  es el **tensor de Riemann**, se puede interpretar como que  $\nabla_a \nabla_b \equiv \rightarrow \uparrow$  y  $\nabla_b \nabla_a \equiv \uparrow \rightarrow$ , sacando así la curvatura de la variedad.

Para derivar  $(\nabla_a \nabla_b - \nabla_b \nabla_a)v^c$ , recordemos que la conexión es simétrica, por tanto  $(\nabla_a \nabla_b - \nabla_b \nabla_a)(f_c v^c) = 0$ , luego

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)v^c = \mathcal{R}_{abd}^c v^d$$

Las propiedades del tensor de Riemann son:

1. Es antisimétrico,  $\mathcal{R}_{abc}^d = -\mathcal{R}_{bac}^d$ .
2.  $\mathcal{R}_{[abc]}^d = 0$ , puesto que  $\nabla_{[a} \nabla_{b]} f_c = 0$ .

3. Cumple la identidad de Bianchi,

$$\nabla_{[a} \mathcal{R}_{bc]d}^e = 0$$

puesto que  $\nabla_{[a} \nabla_b \nabla_{c]} f_d = 0$ .

4. Si la conexión es de Levi-Civita, entonces tenemos que

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) T_{b_1 b_2 \dots}^{a_1 a_2 \dots} = -\mathcal{R}_{abc}^{a_1} T_{b_1 b_2 \dots}^{ca_2 \dots} - \dots + \mathcal{R}_{abb_1}^c T_{cb_2 \dots}^{a_1 a_2 \dots} + \dots$$

Luego, en una base coordenada en la conexión de Levi-Civita, podemos definir el tensor de Riemann como,

$$\mathcal{R}_{\mu\nu\rho}^\sigma = \partial_\nu \Gamma_{\mu\rho}^\sigma - \partial_\mu \Gamma_{\nu\rho}^\sigma + \Gamma_{\mu\rho}^\lambda \Gamma_{\lambda\nu}^\sigma - \Gamma_{\nu\rho}^\lambda \Gamma_{\lambda\mu}^\sigma$$

El **tensor de Ricci** se define como,

$$\mathcal{R}_{ab} = \mathcal{R}_{acb}^c$$

El **escalar de Ricci** se define como,

$$\mathcal{R} = \mathcal{R}_{ab} g^{ab}$$

Para la conexión de Levi-Civita, tenemos que el tensor de Ricci es simétrico, es decir,  $\mathcal{R}_{ab} = \mathcal{R}_{ba}$ .

# **Capítulo 4**

## **Relatividad General**

*“La alegría de ver y entender es el más perfecto don de la naturaleza”.*

(A. Einstein)

## 4.1. Principio de Equivalencia

### 4.1.1. Principio de Equivalencia débil

Este Principio, formulado por Galileo, indica la *universalidad de la caída libre*, es decir, dos objetos con masa distinta caen a la misma vez (demostrado experimentalmente por el astronauta David Scott durante la misión Apollo 15 en la Luna en 1971). Este hecho indica que la velocidad de los objetos en caída libre es independiente de su masa, curiosa coincidencia de la mecánica newtoniana, pues

*"Las masas inerciales y las masas gravitatorias son indistinguibles".*

En la mecánica newtoniana aparecen dos tipos de masas, en distintos contextos y de distintos tipos. La primera es la *masa inercial*,  $m_i$ , que aparece en la segunda Ley de Newton como constante de proporcionalidad entre la fuerza  $\vec{F}$  ejercida en un cuerpo y su aceleración  $\vec{a}$ ,

$$\vec{F} = m_i \vec{a} \quad (4.1)$$

La masa inercial es por tanto una medida para la resistencia de un cuerpo a cambios de velocidad. Por otra parte, la *masa gravitacional*,  $m_g$ , es una medida de cómo interacciona un cuerpo gravitacionalmente con los demás del universo. El potencial gravitatorio causado por un cuerpo sobre otro viene dado por

$$\Phi = \frac{G_N m_g M_g}{r}$$

donde  $G_N$  es la constante de Newton y  $r$  la distancia entre las dos masas  $m_g$  y  $M_g$ . De forma que,

$$\vec{F} = -m_g \vec{\nabla} \Phi \quad (4.2)$$

Estas dos masas son indistinguibles, pues si no lo fueran, dos cuerpos con distinta masa acelerarían de distinta manera bajo la acción de un campo gravitatorio.

**Nota 4.1.1.** *No se conocen objetos con masa negativa o gravitacionalmente nulas. Todos los objetos se ven afectados por el campo gravitatorio y además, lo hacen de la misma forma.*

Como consecuencia de este Principio, igualando (4.1) y (4.2), vemos que  $\vec{a}$  y  $-\vec{\nabla}\Phi$  son también indistinguibles en una región local, es decir, lo suficientemente pequeña. Consecuencia que Einstein percibió, por lo que decidió reformular el Principio de Equivalencia de Galileo.

#### 4.1.2. Principio de Equivalencia de Einstein

*"Es imposible detectar la presencia de un campo gravitatorio en una región lo suficientemente pequeña mediante las leyes de la relatividad especial".*

Esto quiere decir que en regiones locales es imposible detectar campos gravitatorios usando cualquier ley física, entendiendo esto no como una aceleración, sino como la interacción gravitacional que caracteriza el campo.

Para probar esto, decide idear un experimento, llamado **efecto Doppler gravitacional**, que consiste en el sistema de un emisor y un receptor fijos, de forma que a un tiempo  $t_0$  el emisor emite una onda electromagnética con una longitud de onda  $\lambda_1$  y cuando ésta llega al receptor en  $t = t_0 + z/c$ , la longitud de onda cambia a  $\lambda_2$ . Ver Figura 4.1

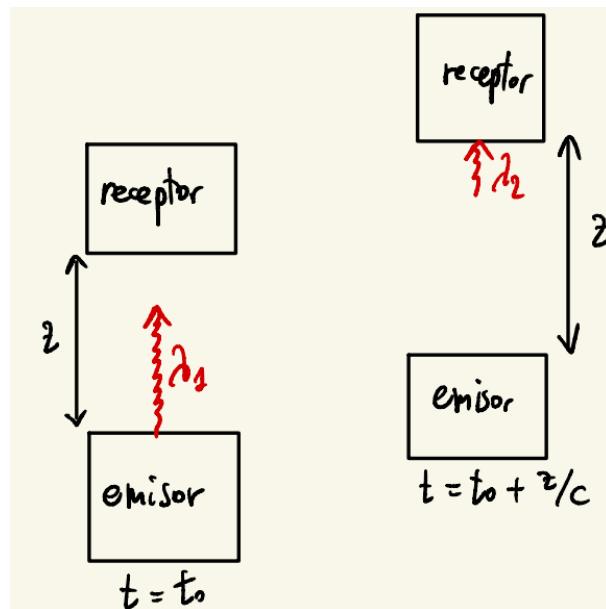


FIGURA 4.1: Esquema del efecto Doppler gravitacional

Así, usando el efecto Doppler relativista,

$$\nu_1 = \nu_2 \sqrt{\frac{1 + \Delta V/c}{1 - \Delta V/c}} \stackrel{\Delta V \ll c}{\approx} \nu_2 \left(1 + \frac{\Delta V}{c}\right)$$

y usando que  $\lambda_i \nu_i = c$ ,

$$\frac{\Delta\lambda}{\lambda_1} = \frac{\Delta V}{c} = \frac{a\Delta V}{c} = \frac{a \cdot z}{c^2}$$

Por tanto, como  $a$  y  $g$  son indistinguibles, tenemos

$$\frac{\Delta\lambda}{\lambda} = g \frac{z}{c^2} = \frac{1}{c^2} \int \partial_z \Phi dz = \frac{\Delta\Phi}{c^2}$$

Esto nos permite concluir que la variación en la longitud de onda  $\Delta\lambda$  de una señal electromagnética al propagarse en un campo gravitacional está directamente relacionada con la diferencia de potencial gravitatorio  $\Delta\Phi$ . En otras palabras, el efecto Doppler gravitacional es una manifestación experimental de cómo la gravedad afecta la propagación de la luz, lo que puede interpretarse como una consecuencia de la dilatación temporal gravitacional.

Por tanto, este resultado demuestra que un observador en reposo dentro de un campo gravitatorio no puede distinguir si el desplazamiento en la longitud de onda se debe a una aceleración (como en un sistema inercial no gravitatorio) o a la interacción gravitacional de un campo, validando así el principio de equivalencia propuesto por Einstein. Este principio establece que los efectos de un campo gravitacional local son indistinguibles de los efectos de un sistema acelerado, lo que constituye la base de la teoría de la relatividad general.

Además, este experimento mental se demostró empíricamente por Pound y Rebka en 1960.

Esta consecuencia permite volver a redefinir el Principio de Equivalencia.

#### 4.1.3. Principio de Equivalencia fuerte

*"Es imposible detectar la presencia de un campo gravitatorio en una región lo suficientemente pequeña mediante las leyes de la física, incluyendo la gravedad".*

Esto implica que los sistemas de referencia en caída libre (SRCL) son equivalentes a los sistemas de referencia inerciales (SRI) de la Relatividad Especial (son sistemas de referencia que se mueven libremente en un campo gravitacional). Estos sistemas de referencia solo se pueden definir localmente, pues estos sistemas se definen en los Espacios Tangentes, y como vimos en Geometría Diferenciable, estos se definen localmente sobre cada punto de

la Variedad Diferenciable. Por tanto, no podemos conectar dos SRCL de forma universal, pero podemos conectarlos mediante una curva  $\gamma_p(t)$ , usando el Transporte Paralelo. El resultado de hacer el transporte paralelo a lo largo de una curva dependerá de la propia curva  $\gamma_p(t)$ , exceptuando espacios planos (de curvatura cero), por lo que podremos usarlo.

Esto es una manifestación de la curvatura  $R_{abc}^d$  y su relación con la gravedad. Es decir, adoptamos una descripción geométrica de la gravedad.

## 4.2. Ecuaciones de Einstein

Como la ecuación de la geodésica describe las curvas, vamos a estudiar el límite de esta ecuación para llegar a una generalización. La ecuación que conocemos es

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma_{\nu\rho}^\mu \dot{x}^\nu \dot{x}^\rho = 0$$

donde  $\tau$  es el tiempo propio y  $\dot{x}^\mu = \partial_\tau x^\mu$ .

### Límite newtoniano (campo débil)

Tomamos este límite imponiendo que:

1. las velocidades son  $|\vec{v}| \ll c$ .
2. el campo gravitatorio débil

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

donde  $\eta_{\mu\nu}$  es la métrica de Minkowski y  $|h_{\mu\nu}| \ll 1$ .

3.  $h_{\mu\nu}$  es estático.

Entonces usando (1), tenemos que

$$\frac{dx^i}{d\tau} \frac{1}{c} \ll \frac{dt}{d\tau} \implies \frac{dx^i}{dt} \ll 1$$

implicando que

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma_{tt}^\mu \left( \frac{dt}{d\tau} \right)^2 \approx 0$$

Además, como  $\eta_{\mu\nu}$  y  $h_{\mu\nu}$  son estáticos, entonces  $g_{\mu\nu}$  es estático, así

$$\Gamma_{tt}^\mu = \frac{1}{2} g^{\mu\lambda} \left( \cancel{2g_{t\lambda,t}} - \cancel{g_{tt,\lambda}} \right)^0 = -\frac{1}{2} g^{\mu\lambda} g_{tt,\lambda}$$

Aplicando (2) queda que  $g_{\mu\nu}g^{\nu\rho} = \delta_\mu^\rho$  y por tanto  $g^{\mu\nu} = g^{\mu\nu} - h^{\mu\nu} + \mathcal{O}(h^2)$  con  $h^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}$ . De forma que  $\Gamma_{tt}^\mu = -\frac{1}{2}\eta^{\mu\nu}h_{tt,\nu} + \mathcal{O}(h^2)$ . En resumen,

$$\frac{d^2x^\mu}{d\tau^2} - \frac{1}{2}\eta^{\mu\nu}h_{tt,\nu} \left( \frac{dt}{d\tau} \right)^2 = 0$$

Entonces, para  $\mu = t$ , aplicando (3)  $\eta^{t\nu}h_{tt,\nu} = \eta^{tt}\cancel{h_{tt,\nu}}^0 = 0$ , tenemos que

$$\frac{d^2t}{d\tau^2} = 0 \implies t = a\tau + b$$

Para  $\mu = i = x, y, z$  tenemos

$$\frac{d^2x^i}{d\tau^2} - \frac{1}{2}\eta^{i\nu}h_{tt,\nu} \left( \frac{dt}{d\tau} \right)^2 = 0 \stackrel{\{t=\tau+b\}}{\implies} \frac{d^2x^i}{d\tau^2} - \frac{1}{2}h_{tt,i} = 0$$

Además, sabemos que la ecuación de Newton es

$$\frac{d^2x^i}{d\tau^2} + \Phi_{,i} = 0$$

y comparando ambas ecuaciones, vemos que  $g_{tt} = -1 - 2\Phi$ . Por tanto, la ecuación de las geodésicas nos permite obtener las ecuaciones de Newton.

Usando la ecuación de Poisson,  $\nabla^2\Phi = 4\pi G\rho$ , podemos reconstruir las ecuaciones dinámicas de  $g_{\mu\nu}$ , pero necesitamos las derivadas segundas de  $g_{\mu\nu}$ :

- **Opción A:**  $\nabla_\mu\nabla^\mu \rightarrow$  Problema con  $\nabla_\mu g_{\mu\nu} = 0$ .
- **Opción B:** El tensor de curvatura  $R_{\mu\nu\rho}^\sigma$  tiene derivadas segundas de  $g_{\mu\nu}$ . Notemos que  $\rho = T_{tt}$ , siendo  $T_{\mu\nu}$  es el tensor de energía-impulso del contenido material.

Sabemos que  $T_{\mu\nu}$  es un tensor simétrico de tipo  $(0,2)$  y la conservación de la energía implica que  $\nabla^\mu T_{\mu\nu} = 0$ . Por tanto, el primer tensor de curvatura candidato es  $R_{\mu\nu}$ , el tensor de Ricci, que también simétrico de tipo  $(0,2)$ , tal que  $R_{\mu\nu} = \kappa T_{\mu\nu}$ ; pero esto no es correcto, pues si tomamos la traza a ambos lados de la igualdad queda  $R = \kappa T$ , donde

$R = R_{\mu\nu}g^{\mu\nu}$  y  $T = T_{\mu\nu}g^{\mu\nu}$ , entonces

$$\left. \begin{array}{l} \nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R = \frac{1}{2}\kappa\nabla_\nu T \\ || \\ \kappa\nabla^\mu T_{\mu\nu} = 0 \end{array} \right\} \Rightarrow \nabla_\nu T = 0$$

cosa que solo ocurre en vacío, por lo que nos está restringiendo información importante, haciendo que este resultado no tenga sentido para cuando no estamos en vacío.

- **Opción C:** Podemos usar el tensor de Einstein, que es  $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ , y por tanto, las ecuaciones de Einstein quedan

$$G_{\mu\nu} = \kappa T_{\mu\nu} \implies R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}$$

Por tanto,  $\nabla^\mu G_{\mu\nu} = 0$  y es compatible con  $\nabla^\mu T_{\mu\nu} = 0$ .

Ahora debemos hallar el valor de  $\kappa$ , y para determinarlo, volvemos al límite newtoniano,

$$g_{tt} = -1 + h_{tt}; \quad g^{tt} = -1 - h_{tt} + \mathcal{O}(h^2)$$

Notemos que  $T = g^{\mu\nu}T_{\mu\nu} = g^{tt}T_{tt} = -T_{tt} + \mathcal{O}(h^2)$ , por lo que las ecuaciones de Einstein se pueden reescribir como,

$$R_{\mu\nu} = \kappa \left( T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T \right)$$

**Demostración.** (...)

■

Entonces,

$$R_{tt} = \frac{1}{2}\kappa \left( T_{tt} - \frac{1}{2}g_{tt}T \right) = \frac{1}{2}\kappa \left( T_{tt} + \frac{1}{2}(-T_{tt}) \right) = \frac{1}{2}\kappa T_{tt}$$

y también,

$$R_{tt} = R_{t\mu t}^\mu = R_{tit}^i = \partial_i \Gamma_{tt}^i - \partial_t \overset{0}{\cancel{\Gamma_{it}^i}} + \Gamma_{i\lambda}^i \Gamma_{tt}^\lambda - \Gamma_{t\lambda}^i \Gamma_{it}^\lambda \xrightarrow{\mathcal{O}(h^2)} = \partial_i \Gamma_{tt}^i = -\frac{1}{2} \partial_i \partial^i h_{tt} = -\frac{1}{2} \nabla^2 h_{tt}$$

Por tanto, igualando,

$$\nabla^2 \Phi = \frac{1}{2}\kappa T_{tt}$$

pues  $h_{tt} = -2\Phi$ . Por tanto, usando la ecuación de Poisson,

$$\nabla^2 h_{tt} = -2\nabla^2 \Phi = -2 \cdot (4\pi G\rho) = -\kappa T_{tt} = -\kappa\rho$$

Entonces,

$$\kappa = 8\pi G$$

con  $c = 1$ . Por tanto, las ecuaciones de Einstein quedan,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu}$$

Vemos que en el límite clásico obtenemos la ecuación de Poisson.

### 4.3. Principio de Covariancia General

La construcción de las ecuaciones de Einstein se conoce como *Principio de Covariancia General*, pues asumimos que no hay elementos universales más allá de la propia métrica del espacio-tiempo, es decir, que no existen regiones o direcciones privilegiadas. Es decir,

*"Las leyes de la física deben ser las mismas en cualquier sistema de coordenadas y no depender de un sistema de referencia particular".*

Con este principio pasamos de las derivadas parciales a la derivada covariante, y de la métrica de Minkowski a una métrica curva, tal que

$$\partial_\mu \longrightarrow \nabla_\mu; \quad y \quad \eta_{\mu\nu} \longrightarrow g_{\mu\nu}$$

La métrica  $g_{\mu\nu}$  determina la estructura del espacio-tiempo, es dinámica (satisface las ecuaciones de Einstein), es lorentziana y de dimensión 4. Asumimos que es  $\mathcal{C}^\infty$ . Además, al igual que en el espacio de Minkowski, esta métrica induce una estructura causal según

$$g_{\mu\nu}V^\mu V^\nu \begin{cases} < 0 & \rightarrow \text{temporal} \\ = 0 & \rightarrow \text{nula} \\ > 0 & \rightarrow \text{espacial} \end{cases}$$

### 4.4. Campos materiales

Los campos materiales son campos escalares o vectoriales o fluidos perfectos (lo detallaremos más adelante). Estos campos deben satisfacer los siguientes postulados:

- Causalidad local:** las ecuaciones del movimiento de los campos materiales permiten el intercambio de señales entre dos puntos de la variedad,  $p, q \in \mathcal{M}$ , tales que las curvas que unen estos puntos tienen vectores tangentes que no se hacen espaciales.

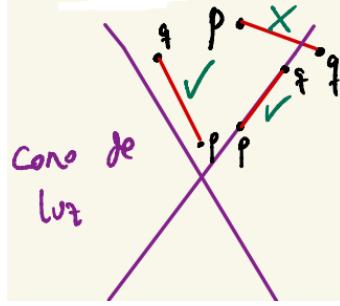


FIGURA 4.2: Cono de luz.

- Conservación de la energía y momento:** las ecuaciones de los campos materiales se obtienen a partir del tensor energía-impulso,  $T_{\mu\nu}$ , que se construye a partir de  $g_{\mu\nu}$ , los campos y sus derivadas. Además,  $\nabla^\mu T_{\mu\nu} = 0$ , es decir, se conserva.

#### 4.4.1. Fluido perfecto

Es un fluido que no conduce calor, ni es viscoso.

Para un observador comóvil (en caída libre con el fluido), verá un fluido isotrópico. Si  $u^\mu$  es la cuadrvelocidad del observador, entonces la cantidad  $T_{\mu\nu}u^\mu u^\nu = \rho$ , define la densidad de energía del fluido. Para cada dirección espacial,  $x_i^\mu$  con  $i = 1, 2, 3$ , las cantidades  $T_{\mu\nu}u^\mu x_i^\nu$  representan el flujo de energía en la dirección  $x_i^\mu$ ; pero esta cantidad no es nula, pues tenemos un fluido perfecto.

Además,  $T_{\mu\nu}x_i^\mu x_j^\nu$  representa el flujo de momento en la dirección ' $i$ ' a través de la superficie ortogonal a ' $j$ ', representando las fuerzas tangenciales o de cizalladura, de presión y viscosas del fluido. Pero para un fluido perfecto no hay flujos de momento paralelos a la superficie perpendicular a  $x_j^\mu$ , es decir, solo tendremos flujo de momento en las direcciones  $i = j$ . Además, como el fluido es isotrópico,

$$T_{\mu\nu}x_1^\mu x_1^\nu = T_{\mu\nu}x_2^\mu x_2^\nu = T_{\mu\nu}x_3^\mu x_3^\nu = P$$

es decir, representa la presión. En resumen, el tensor de energía-impulso puede escribirse como,

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu}$$

donde  $u^\mu$  representa el campo de velocidades del fluido.

Además, la presión puede reescribirse como,

$$P = \frac{1}{3} (u^\mu u^\nu + g^{\mu\nu}) T_{\mu\nu}; \quad \rho = T_{\mu\nu} u^\mu u^\nu$$

**Demostración.** (...)

■

Los fluidos perfectos tendrán una Ley de Conservación asociada,  $\nabla^\mu T_{\mu\nu} = 0$ , que podemos reescribir como

$$u^\mu \nabla_\mu \rho + (\rho + P) \nabla_\mu u^\mu = 0$$

obteniendo así la ecuación de continuidad,

$$\nabla_\mu (u^\mu \rho) + P \nabla_\mu u^\mu = 0$$

que asegura la conservación de la energía.

Por otro lado, tenemos

$$(u^\mu u^\nu + g^{\mu\nu}) \nabla^\rho T_{\rho\nu} = 0 \implies (\rho + P) u^\nu \nabla_\nu u^\mu + (g^{\mu\nu} + u^\mu u^\nu) \nabla_\nu P = 0$$

que equivale a la Segunda Ley de Newton.

Solo consideraremos fluidos barotrópicos, es decir,  $P = P(\rho)$ . Considerando  $P = 0$ , las ecuaciones del movimiento quedan,

$$\begin{aligned} \nabla_\mu (\rho u^\mu) &= 0; & \overbrace{u^\mu \nabla_\mu u^\nu}^{\text{ecuación geodésica}} &= 0 \\ &\parallel \\ \nabla_\mu (T_\nu^\mu u^\nu) &= 0 \end{aligned}$$

Por tanto, las ecuaciones de Einstein quedan,

$$G_{\mu\nu} = 8\pi G(\rho u_\mu u_\nu)$$

### 4.4.2. Campo escalar

Sea  $\phi$  un campo escalar de masa  $m$ , su tensor de energía-impulso es,

$$T_{\mu\nu} = \nabla_\mu \phi \nabla_\nu \phi - \frac{1}{2} g_{\mu\nu} (\nabla^\rho \phi \nabla_\rho \phi + m^2 \phi^2)$$

Usando la conservación del tensor de energía-impulso,  $\nabla^\mu T_{\mu\nu} = 0$ , obtenemos la ecuación de Clebsh-Gordan,

$$\nabla_\mu \nabla^\mu \phi - m^2 \phi = 0$$

**Demostración.** (...)

■

### 4.4.3. Campo electromagnético

Supondremos que no hay fuentes. Así, el tensor electromagnético es

$$F_{\mu\nu} = 2\nabla_{[\mu} A_{\nu]}$$

y el tensor de energía-impulso es,

$$T_{\mu\nu} = F_{\mu\rho} F_\nu^\rho - \frac{1}{4} g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma}$$

Usando la conservación del tensor de energía-impulso tenemos,

$$\nabla^\mu F_{\mu\nu} = 0$$

**Demostración.** (...)

■

## 4.5. Condiciones de energía

Estas condiciones se emplearán en el desarrollo de los teoremas de singularidad. Pues, al tensor energía-impulso se le deben imponer algunas propiedades para determinar si pueden formarse o no formarse singularidades en nuestro universo.

## Condición de energía débil

Dado el tensor energía-impulso  $T_{\mu\nu}$ , y cualquier vector temporal futuro  $u^\mu$ , entonces la condición de energía débil es  $T_{\mu\nu}u^\mu u^\nu \geq 0$ , es decir, la densidad de energía es positiva, cosa que implica que la masa es positiva y la gravedad es atractiva.

## Condición de energía fuerte

Dado el tensor de energía-impulso  $T_{\mu\nu}$ , y  $u^\mu$  cualquier vector temporal futuro, entonces la condición de energía fuerte es

$$\left( T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T \right) u^\mu u^\nu \geq 0$$

## Condición dominante de energía

Dado el tensor de energía-impulso  $T_{\mu\nu}$ , y  $u^\mu$  temporal futuro, entonces la condición dominante de energía es que el vector  $(-T^{\mu\nu}u_\nu)$  es temporal, nulo futuro o cero.

## Condición nula de energía

Dado el tensor de energía-impulso  $T_{\mu\nu}$ , y un vector  $k^\mu$  nulo futuro cualquiera, entonces la condición de energía es que  $T_{\mu\nu}k^\mu k^\nu \geq 0$ .

Estas condiciones se impican entre sí, pero no de cualquier forma, sino que la cadena de implicaciones es,

$$\text{dominante} \Rightarrow \text{débil} \Rightarrow \text{nula} \Leftarrow \text{fuerte}$$

La condición nula es la que nos garantiza los Teoremas de Singularidad, por tanto, bastará con ver alguna de las condiciones de energía para ver si la condición nula se satisface, pues en ocasiones la condición nula no es fácil de calcular.

### 4.5.1. Teoremas de Singularidad

Definiciones preliminares:

- **Geodésica inextensible incompleta:** Una trayectoria en el espacio-tiempo que termina abruptamente porque el espacio-tiempo deja de estar definido. Indica la presencia de una singularidad.
- **Superficie atrapada:** Una región del espacio-tiempo donde la gravedad es tan intensa que la luz y la materia no pueden escapar, típica de agujeros negros.
- **Condición de energía fuerte:** Supone que la energía y la presión de la materia cumplen ciertos límites, asegurando que la gravedad sea atractiva.
- **Espacio-tiempo globalmente hiperbólico:** Un espacio-tiempo que tiene una estructura causal bien definida, sin curvas cerradas de tipo tiempo (que implicarían paradojas causales).

A continuación, se presentan los enunciados de los teoremas de singularidad de Penrose y Hawking, traducidos de las publicaciones originales para reflejar fielmente sus formulaciones. Estos enunciados están extraídos y adaptados de los artículos citados y del libro *The Large Scale Structure of Space-Time* de Hawking y Ellis (1973), que formaliza estos resultados:

#### **Teorema 4.5.1. Teorema de Singularidad de Penrose (1965)**

*"Si una superficie atrapada existe en el espacio-tiempo, y si la relatividad general es válida, junto con la condición de energía fuerte, entonces hay al menos una geodésica de tipo tiempo o de tipo nulo inextensible incompleta en el espacio-tiempo."*

Es decir, una superficie atrapada es una región donde todos los rayos de luz convergen hacia adentro debido a la gravedad extrema. Este resultado implica que, en el colapso gravitacional (por ejemplo, en la formación de un agujero negro), el espacio-tiempo necesariamente desarrolla una singularidad donde las geodésicas terminan, es decir, el espacio-tiempo deja de ser definible en esa región.

#### **Teorema 4.5.2. Teorema de Singularidad de Hawking (1965)**

*"Si el espacio-tiempo es globalmente hiperbólico, satisface la condición de energía fuerte, y existe una superficie de Cauchy donde la expansión de las geodésicas de tipo tiempo es positiva en todo momento, entonces necesariamente existe una geodésica de tipo tiempo o de tipo nulo inextensible incompleta en el pasado."*

Es decir, este teorema se aplica al contexto cosmológico. Afirma que, si el universo está en expansión (como lo está nuestro universo observable), entonces debe haber existido una singularidad en el pasado (el Big Bang) que marcó el inicio del espacio-tiempo.

**Teorema 4.5.3. Teorema de Singularidad de Hawking-Penrose (1970)**

*"Si un espacio-tiempo contiene una trampa causal (como un horizonte de eventos), satisface la condición de energía media y es globalmente hiperbólico en el sentido de que no contiene curvas cerradas de tipo tiempo, entonces debe contener al menos una geodésica de tipo tiempo o de tipo nulo inextensible incompleta."*

Es decir, este teorema generaliza los anteriores y establece que, bajo condiciones físicas razonables (como la validez de la relatividad general y la ausencia de curvas cerradas de tipo tiempo), las singularidades son inevitables tanto en el colapso gravitacional como en la expansión inicial del universo.

## 4.6. Ecuaciones de Einstein y Principio de Mínima Acción

La acción de la formulación Lagrangiana para la relatividad general es formulada por Hilbert, pero éste dejó que Einstein lo publicara, y así, la acción de la relatividad general se denomina *acción de "Hilbert-Einstein"*. Se construye con cantidades invariantes bajo transformaciones generales de coordenadas (difeomorfismos). Se define como,

$$S_{HE} = \int d^4x \sqrt{|g|} R$$

donde  $\sqrt{|g|}$  es el Jacobiano y  $R$  el escalar de Ricci. que involucra derivadas segundas de  $g_{\mu\nu}$ , pero se pueden eliminar integrando por partes.

Consideramos variaciones  $\delta g_{\mu\nu}$  que se anulan en la frontera, es decir,  $\delta g_{\mu\nu}|_{\partial\mathcal{M}} = 0$ . Así, podemos variar la acción, tal que

$$\delta S_{HE} = \int d^4x \left[ \sqrt{|g|} g^{\mu\nu} \delta R_{\mu\nu} + \sqrt{|g|} (\delta g^{\mu\nu}) R_{\mu\nu} + (\delta \sqrt{|g|}) g^{\mu\nu} R_{\mu\nu} \right]$$

donde, asumiendo que la conexión es de Levi-Civita, tenemos

$$\delta \sqrt{|g|} = -\frac{1}{2} \sqrt{|g|} g_{\mu\nu} \delta g^{\mu\nu}; \quad \delta R_{\mu\nu\rho}^\sigma = \nabla_\nu (\delta \Gamma_{\mu\rho}^\sigma) - \nabla_\mu (\delta \Gamma_{\nu\rho}^\sigma)$$

recordando que  $\delta\partial() = \partial\delta()$  y que  $\Gamma$  no es un tensor, pero  $\delta\Gamma$  sí lo es.

**Demostración.** (...)

■

Por tanto,

$$g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\nu(g^{\mu\rho}\delta\Gamma_{\mu\rho}^\nu - g^{\nu\rho}\delta\Gamma_{\sigma\rho}^\sigma) = \nabla_\nu v^\nu$$

es decir, es igual a la divergencia de un vector. Por tanto,

$$\int_{\mathcal{M}} d^4x \sqrt{|g|} g^{\mu\nu} \delta R_{\mu\nu} = \int_{\mathcal{M}} d^4x \sqrt{|g|} \nabla_\nu v^\nu = \int_{\partial\mathcal{M}} d\sigma n_\nu v^\nu = 0$$

donde hemos usado el Teorema de Stokes.

Si acoplamos materia,

$$S = \frac{1}{\kappa} S_{HE} + S_{\mathcal{M}}$$

y tomamos variaciones con respecto a  $g_{\mu\nu}$ ,

$$\begin{aligned} \delta_g S &= \frac{1}{\kappa} \delta_g S_{HE} + \delta_g S_{\mathcal{M}} = 0 \\ &= \frac{1}{\kappa} \int d^4x \sqrt{|g|} (R_{\mu\nu} - \frac{1}{2}gR - \kappa T_{\mu\nu}) \delta g^{\mu\nu} = 0 \end{aligned}$$

donde hemos definido el tensor de energía-impulso como  $T_{\mu\nu} = -\frac{1}{\sqrt{g}} \frac{\delta S_{\mathcal{M}}}{\delta g^{\mu\nu}}$  y  $T^{\mu\nu} = \frac{1}{\sqrt{|g|}} \frac{\delta S_{\mathcal{M}}}{\delta g_{\mu\nu}}$ .

En términos de la densidad lagrangiana (o Lagrangiano), tenemos

$$S_{\mathcal{M}} = \int d^4x \sqrt{|g|} \mathcal{L}_{\mathcal{M}} = \int d^4x \mathcal{L}_{\mathcal{M}}$$

donde  $\mathcal{L}_{\mathcal{M}}$  es el Lagrangiano y  $\mathcal{L}_{\mathcal{M}}$  es la densidad lagrangiana. Así, el tensor de energía-impulso se reescribe como

$$T^{\mu\nu} = \frac{2}{\sqrt{|g|}} \frac{\partial \mathcal{L}_{\mathcal{M}}}{\partial g_{\mu\nu}} = 2 \frac{\partial \mathcal{L}_{\mathcal{M}}}{\partial g_{\mu\nu}} + g^{\mu\nu} \mathcal{L}_{\mathcal{M}}$$

$$T_{\mu\nu} = -\frac{2}{\sqrt{|g|}} \frac{\partial \mathcal{L}_{\mathcal{M}}}{\partial g^{\mu\nu}} = -2 \frac{\partial \mathcal{L}_{\mathcal{M}}}{\partial g^{\mu\nu}} + g_{\mu\nu} \mathcal{L}_{\mathcal{M}}$$

El tensor  $T_{\mu\nu}$  es conservado como consecuencia de la invariancia bajo transformaciones generales de coordenadas (difeomorfismos) de la acción,

$$S_{\mathcal{M}} = \int_{\mathcal{M}} d^4x \mathcal{L}(\phi, \nabla\phi, g)$$

Dado una densidad lagrangiana, aplicamos un difeomorfismo y vemos como transforman los campos. Consideramos un difeomorfismo  $\xi^\mu$  que se anula en frontera  $\partial\mathcal{M}$ . Tal que

$$\delta_\xi \phi = -\mathcal{L}_\xi \phi; \quad \delta_\xi g_{\mu\nu} = -\mathcal{L}_\xi g_{\mu\nu} = -2\nabla_{(\mu} \xi_{\nu)}$$

teniendo así una variación de la acción tal que

$$\delta_\xi S_{\mathcal{M}} = - \int_{\mathcal{M}} d^4x (\delta_\phi \mathcal{L}_{\mathcal{M}} \mathcal{L}_\xi \phi + \frac{\delta \mathcal{L}_{\mathcal{M}}}{\delta g_{\mu\nu}} \Delta_\mu \xi_\nu)$$

donde  $\delta_\phi \mathcal{L}_{\mathcal{M}} \mathcal{L}_\xi \phi = 0$ , pues se cumplen las ecuaciones del movimiento. Así,

$$\delta_\xi S_{\mathcal{M}} = - \int d^4x \sqrt{|g|} (T^{\mu\nu} \nabla_\mu \xi_\nu) = - \int d^4x \sqrt{|g|} \begin{bmatrix} \nabla_\mu (T^{\mu\nu} \xi_\nu) & -\xi_\nu \nabla_\mu T^{\mu\nu} \\ \int_{\partial\mathcal{M}} (...) = 0 \text{ ( } \xi^\mu \text{ es cero en } \partial\mathcal{M} \text{)} & \end{bmatrix} = 0$$

Derivando  $T_{\mu\nu}$  tenemos las acciones,

$$\text{Campo escalar : } S_\phi = -\frac{1}{2} \int d^4x \sqrt{|g|} (\nabla_\mu \phi \nabla_\nu \phi g^{\mu\nu} + m^2 \phi^2)$$

$$\text{Campo electromagnético : } S_A = -\frac{1}{16\pi} \int d^4x \sqrt{|g|} F_{\mu\nu} F^{\mu\nu}; \quad F_{\mu\nu} = 2\nabla_{[\mu} A_{\nu]}$$

## 4.7. Constante cosmológica

La constante cosmológica  $\Lambda$  es una consecuencia natural de las ecuaciones de Einstein, pues el tensor de Einstein  $G_{\mu\nu}$  no está únicamente definido,

$$G_{\mu\nu} \longrightarrow G_{\mu\nu} + \Lambda g_{\mu\nu} \equiv {}^\Lambda G_{\mu\nu}$$

Pero esta forma del tensor no es compatible con un universo estático, es decir, si  $\Lambda \neq 0$ , tendremos un universo dinámico (en expansión o contracción, dependiendo del signo de  $\Lambda$ ). También se cumple que

$$\nabla^\mu ({}^\Lambda G_{\mu\nu}) = 0$$

Originalmente, la constante cosmológica fue introducida por Einstein para describir un universo estático. Hoy forma parte del modelo estándar de cosmología  $\Lambda$ CDM. Además, esta constante puede relacionarse con el fluido que permea el universo, "Dark Energy".

La constante cosmológica se puede interpretar como un fluido perfecto, que las ecuaciones de Einstein

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}$$

podemos reescribirlas como,

$$G_{\mu\nu} = \kappa(T_{\mu\nu} + {}^{(\Lambda)}T_{\mu\nu}); \quad {}^{(\Lambda)}T_{\mu\nu} = \frac{-\Lambda}{8\pi G}g_{\mu\nu}$$

Teniendo una densidad de energía del fluido de la constante cosmológica tal que

$${}^{(\Lambda)}T_{\mu\nu}u^\mu u^\nu = \rho_\Lambda = \frac{\Lambda}{8\pi G}$$

y una presión 'cosmológica',

$$P_\Lambda = \frac{1}{3}(u^\mu u^\nu + g^{\mu\nu}){}^{(\Lambda)}T_{\mu\nu} = -\frac{\Lambda}{8\pi G}$$

En resumen, tenemos la ecuación de estado,

$$P_\Lambda = -\rho_\Lambda$$

La presión negativa hace que el universo se expanda y evita que colapse. Además, si el universo está colapsando, la constante cosmológica contribuye a este colapso.

El valor actual de la constante cosmológica es  $\Lambda = 10^{-52} \text{ m}^{-2}$ .

## 4.8. Soluciones dinámicas de las ecuaciones de Einstein

Las ecuaciones de Einstein

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = T_{\mu\nu}$$

son un conjunto de ecuaciones en derivadas parciales de segundo orden en  $g_{\mu\nu}$  no lineales. Debido a la covariancia general tenemos 4 ecuaciones estáticas con derivadas primeras en  $g_{\mu\nu}$  en la dirección temporal, teniendo 4 ligaduras.

Las ecuaciones de los campos materiales son similares y todo está acoplado.

$G_{\mu\nu}u^\nu$  no tiene derivadas segundas en la dirección  $u^\mu$ . Además,  $(G_{\mu\nu} - \kappa T_{\mu\nu})u^\nu = 0$ , inicialmente los datos iniciales tienen que cumplir 4 ecuaciones.

En vacío,  $G_{\mu\nu}n^\nu = 0$ , con  $\mu = 0, 1, 2, 3$ .

Teniendo en cuenta que  $g_{\mu\nu}$  y  $g_{\mu\nu} + \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu$  representan el mismo espacio-tiempo físico, tenemos otras 4 ecuaciones.

Si  $G_{\mu\nu}$  son 10 ecuaciones, quitamos 4 ligaduras y 4 componentes de  $g_{\mu\nu}$ , quedando 2 grados de libertad dinámicos, siendo 2 polarizaciones de ondas gravitacionales (similar al electromagnetismo).

No se conoce el espacio completo de soluciones.

Para gravedad en vacío existen teoremas de existencia y unicidad (Yvonne Choquet-Bruhat (1952)).

Hay 3 formas de resolver las ecuaciones de Einstein:

1. relatividad numérica.
2. soluciones aproximadas (ondas gravitacionales, límites newtonianos, etc.)
3. soluciones con simetría (agujeros negros con simetría esférica, modelos cosmológicos de Friedman-Roberson-Walker, etc.)

#### 4.8.1. Espacio-tiempo de Schwarzschild

Es una solución de las ecuaciones de Einstein en vacío esféricamente simétrica, siendo invariante bajo el grupo de rotaciones  $SO(3)$ , teniendo 3 campos vectoriales de Killing que satisfacen  $[\vec{v}^{(i)}, \vec{v}^{(j)}] = \epsilon^{ijk}\vec{v}^{(k)}$ , con  $i, j, k = 1, 2, 3$  (momentos angulares).

Asumiremos que la métrica es estática. Tenemos un CVK  $t^\mu = \delta_t^\mu$ . La métrica más general toma la forma,

$$ds^2 = -B(r)dt^2 + A(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$$

donde  $d\Omega^2 = (d\theta^2 + \sin^2 \theta d\varphi^2)$  es una métrica asociada a la 2-esfera unidad.

Vamos a determinar estas constantes usando el método de las geodésicas, con el lagrangiano

$$\tilde{\mathcal{L}} = -B(r)\dot{t}^2 + A(r)\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2$$

Derivamos los símbolos de Christoffel,

$$\begin{aligned}\Gamma_{tr}^t &= \frac{1}{2} \frac{1}{B} \frac{dB}{dr} = \frac{1}{2} \frac{B'}{B} = \Gamma_{rt}^t \\ \Gamma_{tt}^r &= \frac{1}{2} \frac{B'}{A}; \quad \Gamma_{rr}^r = \frac{1}{2} \frac{A'}{A}; \quad \Gamma_{\theta\theta}^r = -\frac{r}{A} \\ \Gamma_{\varphi\varphi}^r &= -\frac{r\sin^2\theta}{A}; \quad \Gamma_{\theta r}^\theta = \frac{1}{2} = \Gamma_{r\theta}^\theta; \quad \Gamma_{\varphi\varphi}^\theta = -\cos\theta\sin\theta \\ \Gamma_{r\varphi}^\varphi &= \frac{1}{2} = \Gamma_{\varphi r}^\varphi; \quad \Gamma_{\theta\varphi}^\varphi = \frac{\cos\theta}{\sin\theta} = \Gamma_{\varphi\theta}^\varphi\end{aligned}$$

Para el tensor de Riemann,

$$\begin{aligned}R_{trtr} &= \frac{1}{4}(2B'' - B'\frac{A'}{A} - \frac{(B')^2}{B}); \quad R_{t\theta t\theta} = \frac{rB'}{2A}; \quad R_{t\varphi t\varphi} = \sin^2\theta \frac{rB'}{2A} \\ R_{r\theta r\theta} &= \frac{rA'}{2A}; \quad R_{r\varphi r\varphi} = \sin^2\theta \frac{rA'}{2A}; \quad R_{\theta\varphi\theta\varphi} = \frac{(A-1)r^2\sin^2\theta}{A}\end{aligned}$$

Para el tensor de Ricci,

$$\begin{aligned}R_{tt} &= \frac{1}{4A} \left[ \left( \frac{4}{r} - \frac{A'}{A} \right) B' - \frac{(B')^2}{B} + 2B'' \right]; \quad R_{rr} = \frac{1}{r} \frac{A'}{A} + \frac{1}{4B} \left[ \frac{A'}{A} B' + \frac{(B')^2}{B} - 2B'' \right] \\ R_{\theta\theta} &= 1 + \frac{r}{2} \frac{A'}{A} - \frac{1}{A} - \frac{r}{2} \frac{B'}{AB}; \quad R_{\varphi\varphi} = \sin^2\theta R_{\theta\theta}\end{aligned}$$

Usando que  $G_{\mu\nu} = 0$ , entonces  $R_{\mu\nu} = 0$ , luego

$$-\frac{R_{tt}}{g_{tt}} + \frac{R_{rr}}{g_{rr}} = 0 \Rightarrow \frac{A'}{rA^2} + \frac{B'}{rAB} = 0 \Rightarrow \frac{A'}{A} = -\frac{B'}{B} \Rightarrow \begin{cases} A(r) = \frac{c_0}{B(r)} \\ B(r) = \frac{c_0}{A(r)} \end{cases}$$

Sustituyendo en  $R_{\theta\theta} = 0$  idénticamente y lo mismo para  $R_{\varphi\varphi} = 0$ .

Sustituyendo en

$$\frac{R_{rr}}{g_{rr}} = 0 \Rightarrow \frac{A'}{rA} - \frac{(A')^2}{rA^2} + \frac{A''}{2A^2} = 0 \Rightarrow \begin{cases} A(r) = \frac{c_1}{1+c_2/r} \\ B(r) = \tilde{c}_1(1 + \frac{c_2}{r}) \end{cases}$$

En el límite de campo débil,  $g_{tt} = -B(r) = -1 + \frac{2GM}{r}$ , por tanto  $\bar{c}_1 = 1$ ,  $c_2 = -2GM$  y  $c_0 = 1$ . Así, la métrica en el espacio-tiempo de Schwarzschild es

$$ds^2 = -\left(1 - \frac{2GM}{r}\right)dt^2 + \frac{1}{1 - \frac{2GM}{r}}dr^2 + r^2d\Omega^2$$

asumiendo que  $t^\mu = \delta_t^\mu$  CVK, es decir, que la métrica es estática.

**Teorema 4.8.1** (Teorema de Birkhoff). *Cualquier geometría de dimensión 4, esféricamente simétrica y que es Ricci plana ( $R_{\mu\nu} = 0 \Rightarrow G_{\mu\nu} = 0$ ), es decir, que se cumplan las ecuaciones de Einstein en vacío, es localmente difeomorfa a la métrica de Schwarzschild.*

En otras palabras, si tenemos un objeto esféricamente simétrico que rote o no, fuera de éste, la métrica será siempre de Schwarzschild.

**Demostración.** Sea  $ds^2 = -B(t, r)dt^2 + A(t, r)dr^2 + r^2d\Omega^2$ . El tensor de Ricci cumple que  $R_{\mu\nu} = 0$ , y en particular,

$$R_{tr} = \frac{\partial_t A}{rA} \implies A = A(r)$$

Con esto calculamos,

$$\frac{R_{tt}}{-g_{tt}} + \frac{R_{rr}}{g_{rr}} = 0 \implies \frac{\partial_r A}{rA^2} + \frac{\partial_r B}{ABr} = 0$$

como  $A = A(r)$ , solo nos queda que  $B = \tau(t)\tilde{B}(r)$ , es decir, debe haber separación de variables. La función la podemos reabsorber mediante una reparametrización temporal, tal que

$$B(r, t)dt^2 = \tilde{B}(r) \underbrace{\tau(t)dt^2}_{dT^2} = \tilde{B}(r)dT^2$$

Por tanto, tenemos la métrica,

$$ds^2 = -\tilde{B}(r)dT^2 + A(r)dr^2 + r^2d\Omega^2$$

El resto de la demostración es análoga a la demostración de la métrica de Schwarzschild. ■

Este Teorema nos dice que no existen ondas gravitacionales esféricamente simétricas, es decir, una estrella pulsante no puede emitir una onda gravitacional, pues fuera de éstas, la métrica es estática. Para poder producirlas, debemos romper la simetría esférica.

### 4.8.2. Geometría del interior de una estrella estática y esféricamente simétrica

Asumimos un fluido perfecto con esta simetría y estático, donde consideramos que la estrella tiene un radio  $R$ , tal que su tensor de energía-impulso es

$$\begin{aligned} T_{\mu\nu} &= (\rho + P)u_\mu u_\nu + Pg_{\mu\nu} && \text{si } r \leq R \\ T_{\mu\nu} &= 0 && \text{si } r > R \end{aligned}$$

donde  $\rho = \rho(r)$ ,  $P = P(r)$ ,  $u^\mu u^\nu g_{\mu\nu} = -1$  y  $u^\mu \propto \delta_t^\mu$ . Así, como  $\delta_t^\mu \delta_t^\nu g_{\mu\nu} = g_{tt} = -B(r)$ , entonces  $u^\mu = \delta_t^\mu / \sqrt{B(r)}$  y  $u_\mu = g_{\mu\nu} u^\nu = -\delta_\mu^t \sqrt{B(r)}$ .

Por tanto, para  $r \leq R$ ,

$$T_{tt} = (\rho + P)u_t u_t + Pg_{tt} = \rho B$$

Entonces, usando la definición del tensor de Einstein,

$$G_{tt} = 8\pi G T_{tt} = 8\pi G B \rho$$

Además,

$$T_{rr} = g_{rr}P = AP \Rightarrow G_{rr} = 8\pi G T_{rr} = 8\pi G A P$$

$$T_{\theta\theta} = g_{\theta\theta}P = r^2 P \Rightarrow G_{\theta\theta} = 8\pi G T_{\theta\theta} = 8\pi G r^2 P$$

$$T_{\varphi\varphi} = \sin^2\theta T_{\theta\theta} \Rightarrow G_{\varphi\varphi} = \sin^2\theta G_{\theta\theta}$$

Luego, sustituyendo  $g_{\mu\nu}$  en  $G_{\mu\nu}$  tenemos

$$\begin{aligned} G_{tt} &= \frac{B}{r^2} \left( 1 - \frac{1}{A} + \frac{rA'}{A^2} \right) \\ G_{rr} &= \frac{1}{r^2} (1 - A) + \frac{B'}{rB} \\ G_{\theta\theta} &= \frac{r^2}{2AB} \left( -\frac{A'}{rA} + \frac{B'}{r} - \frac{A'B'}{2A} - \frac{(B')^2}{2B} + B'' \right) \end{aligned}$$

Tenemos 3 ecuaciones de Einstein y necesitamos una cuarta ecuación, pero al tener un fluido perfecto, usamos también la ecuación de estado del fluido.

- Para  $tt$ :

$$\frac{\mathcal{B}}{r^2} = \left( 1 - \frac{1}{A} + \frac{rA'}{A^2} \right) = 8\pi G \rho \mathcal{B}$$

Si tomamos  $A(r) = \frac{1}{1-2Gm(r)/r}$ , entonces tenemos que

$$m(r) = 4\pi \int_0^r d\tilde{r} \tilde{r}^2 \rho(\tilde{r}) + m_0$$

siendo la masa que hay dentro del volumen de la estrella.

Si  $m_0 \neq 0$ , tenemos una singularidad cónica, pues si calculamos el volumen propio de la estrella, será distinto a  $4\pi R^2$ . Por tanto, tomamos  $m_0 = 0$ .

Notemos que  $A > 0$  siempre para que dentro de la estrella la métrica no haga cosas raras. Por tanto,  $r \geq 2Gm(r)$ . Además, la estrella en  $r = R$  debe tener una geometría continua y debe 'pegar' bien con la métrica de Schwarzschild, de forma que  $m(r = R) = M$ . Como  $A > 0$ , entonces  $\frac{R}{2} \geq GM$ , siendo  $GM$  el *radio de Schwarzschild*.

La masa propia de la estrella es

$$M_p = 4\pi \int_0^R d\tilde{r} \tilde{r}^2 \rho(\tilde{r}) \geq M$$

siendo mayor porque incluye la energía de enlace gravitacional.

■ **Para  $rr$ :**

Vemos que la parte radial queda,

$$\frac{(1-A)}{r^2} + \frac{B'}{rB} = 8\pi G P A \Rightarrow \frac{B'}{B} = \frac{2Gm(r) + 8\pi Gr^3 P(r)}{r(r - 2Gm(r))}$$

Tenemos un límite newtoniano si  $B = e^{2\phi}$  y  $r^3 P \ll Gm(r)$  ó  $r \gg Gm(r)$ , así queda,

$$\frac{d\phi}{dr} \approx \frac{m(r)}{r^2}$$

siendo el potencial gravitatorio.

■ **Para  $\theta\theta$ :**

No da información adicional.

Además tenemos,

$$0 = \nabla^\mu T_{\mu\nu} = u_\nu u^\mu \nabla_\mu (\rho + P) + (\rho + P)$$

donde  $u_\nu u^\mu$  depende de  $t$  y  $\partial_t(\rho + P) = 0$ , que depende solo de  $r$ . Entonces tenemos,

$$[y_\nu \nabla_\mu u^\mu + u^\mu \nabla_\mu y_\nu] + \nabla_\nu P = 0$$

Para  $\nu = t$ ,

$$\nabla_t P = \partial_t P = 0$$

$$\begin{aligned} & \underbrace{u_t \nabla_\mu u^\mu}_{= \partial_t u^t + \Gamma_{\mu t}^\mu u_t u^\mu} + \underbrace{u^t \partial_t u_t}_0 + \underbrace{u^t \Gamma_{tt}^t u_t}_0 \\ &= \underbrace{0}_0 + \underbrace{0}_0 + \underbrace{0}_0 \end{aligned}$$

Luego, no contribuye.

Para  $\nu = r$ ,

$$\nabla_r P = \partial_r P = P'$$

$$\begin{aligned} u_r \overset{0}{\nabla}_\mu u^{\mu+u^t \nabla_t u_r} &= u^t \partial_t u_r + u_r u^t \Gamma_{tr}^\mu u_\mu = \\ &= -u^t \Gamma_{tr}^t u_t = \Gamma_{tr}^t = \frac{1}{2} \frac{B'}{B} \end{aligned}$$

$$u^\mu u_\nu = u^t u_t = -1$$

Por tanto,

$$\frac{1}{2}(\rho + P) \frac{B'}{B} + P' = 0$$

Sustituyendo  $\frac{B'}{B}$ , tenemos

$$P' = -(\rho(r) + P(r)) \frac{Gm(r) + 4\pi r^3 P(r)}{r(r - 2Gm(r))} \quad (4.3)$$

siendo la **ecuación de Tolma-Oppenheimer-Volkof (TOV)**. Representa la variación de la presión de una estrella estática en términos del contenido energético y  $m(r)$ .

El límite newtoniano será para  $P \ll \rho$  y  $Gm(r) \ll r$ , tal que

$$\frac{dP}{dr} = -\rho \frac{m(r)}{r^2}$$

siendo la ecuación de equilibrio hidrostático.

Si introducimos la ecuación de estado  $P = \rho(r)$ , podemos resolver el sistema.

En cualquier caso, tenemos que  $\rho(r > R) = 0$  y  $P(r > R) = 0$ .

### 4.8.3. Estrellas de densidad uniforme

Como la densidad es uniforme, tendremos

$$\rho = \begin{cases} \rho_0 & \text{si } r \leq R \\ 0 & \text{si } r > R \end{cases}$$

En este caso, la masa queda,

$$m(r) = \begin{cases} \frac{4\pi}{3}\rho_0r^3 & \text{si } r \leq R \\ \frac{4\pi}{3}\rho_0R^3 & \text{si } r > R \end{cases}$$

Sustituyendo en (4.3), tenemos

$$P(r) = \rho_0 \frac{\left(1 - \frac{2GM}{R}\right)^{1/2} - \left(1 - \frac{2GMr^2}{R^3}\right)^{1/2}}{\left(1 - \frac{2GMr^2}{R^3}\right)^{1/2} - 3\left(1 - \frac{2GM}{R}\right)^{1/2}}$$

La presión en el centro de la estrella será

$$P(r=0) = \rho_0 \frac{1 - \left(1 - \frac{2GM}{R}\right)^{1/2}}{3\left(1 - \frac{2GM}{R}\right)^{1/2} - 1} = P_C$$

siendo la presión crítica.

Si  $R = \frac{9}{4}GM$ , entonces  $P_C \rightarrow \infty$ .

Según la Relatividad General, no podemos tener estrellas esféricas y estáticas de densidad uniforme con  $R \leq \frac{9}{4}GM$ . Las estrellas de neutrones alcanzan densidades de  $\rho \approx 6 \cdot 10^{17}$  kg/m<sup>3</sup> y tendrán  $M_{max} \leq 10M_0$ , siendo  $M_0$  la masa del Sol. En realidad, las estrellas de neutrones no pueden ser más de 2 o 3 veces la masa del Sol.

**(Ejercicio: Obtener  $A(r)$  y  $B(r)$  de las estrellas de neutrones).**

### 4.8.4. Colapso gravitatorio. Modelo de Oppenheimer-Snyder

Asumimos una estrella esférica, de radio  $R(t = t_0) = R_0$ , de densidad uniforme  $\rho = \rho(t)$  y sin presión  $P = 0$ . Por tanto, tenemos que  $T_{\mu\nu} = \rho u_\mu u_\nu$ .

Las ecuaciones de conservación nos dan

$$u^\mu \nabla_\mu u^\nu = 0$$

En la superficie de la estrella y teniendo en cuenta la continuidad de  $g_{\mu\nu}$ , las partículas seguirán las geodésicas,

$$-\left(1 - \frac{2GM}{r(s)}\right) \left(\frac{dt}{ds}\right)^2 + \frac{1}{1 - \frac{2GM}{r(s)}} \left(\frac{dr}{ds}\right)^2 = -1$$

En el límite  $r(s) \rightarrow 2GM$ , tenemos  $t| \rightarrow \infty$  y  $\dot{r} \rightarrow 0$ , es decir, en coordenadas  $(t, r)$  no veremos nunca a la estrella alcanzar  $r \rightarrow 2GM$ . En coordenadas de tiempo propio sí lo hace, teniendo que la cantidad  $u^\mu \xi_\mu = cte$ , con  $\xi^\mu = \delta^\mu_t$ . Esto implica que

$$\left(1 - \frac{2GM}{r}\right) t_0 cte \xrightarrow{r \rightarrow \infty} 1$$

Sustituyendo en  $u^\mu u_\nu g_{\mu\nu} = 1$ , obtenemos que  $\dot{r} = -\sqrt{\frac{2GM}{r}}$ . Así,

$$\Delta S = \int_{2GM}^R \frac{dr}{|\dot{r}|} = \frac{4GM}{3} \left[ \left( \frac{R_0}{2GM} \right)^{3/2} - 1 \right] < \infty$$

Además, podemos calcular lo que tarda desde  $r = 2GM$  hasta  $r = 0$ , tal que

$$\Delta S = \frac{4GM}{3} \Rightarrow \Delta S \approx 10^{-5} \frac{M}{M_0} \text{ (s)}$$

Además, cuando se produce el colapso gravitacional, se crea un **agujero negro**.

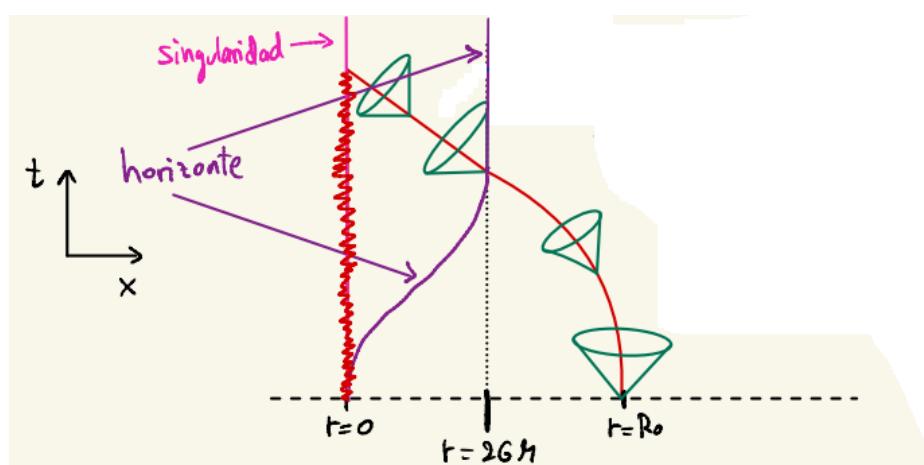


FIGURA 4.3: Esquema de un rayo de luz entrando en un agujero negro.

donde los conos verdes son los *conos de luz* y vemos que a partir del horizonte  $r = 2GM$  nada puede escapar del agujero negro, pues los conos de luz no tienen suficiente ángulo.

Además, dentro de un agujero negro nunca podremos llegar a ver la singularidad hasta que lleguemos a ella en un tiempo futuro, y también, usando la métrica de Schwarzschild, la coordenada radial se vuelve temporal y la temporal, radial, dentro del agujero negro.

Pues, como

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 + \frac{dr^2}{1 - \frac{2GM}{r}} + r^2 d\Omega^2$$

Entonces,

$$\nabla_\mu t \nabla_\nu t g^{\mu\nu} < 0 \quad \text{si } r > 2GM \quad \equiv \text{Vector temporal}$$

$$\nabla_\mu r \nabla_\nu r g^{\mu\nu} > 0 \quad \text{si } r > 2GM \quad \equiv \text{Vector espacial}$$

$$\nabla_\mu t \nabla_\nu t g^{\mu\nu} > 0 \quad \text{si } r < 2GM \quad \equiv \text{Vector espacial}$$

$$\nabla_\mu r \nabla_\nu r g^{\mu\nu} < 0 \quad \text{si } r < 2GM \quad \equiv \text{Vector temporal}$$

Roger Penrose demostró que las singularidades se forman tras el colapso de una estrella, de forma general.

#### 4.8.5. Propiedades de la métrica de Schwarzschild

- (I) La métrica es degenerada en  $r = 2GM$ ,  $r = 0$  y  $\theta = 0, \pi$ . Se puede demostrar que para  $r = 2GM$  y  $\theta = 0, \pi$  son singularidades de coordenadas, es decir, si hacemos un cambio de coordenadas solucionamos estas singularidades. Además, el parámetro afín se comporta bien en  $r = 2GM$ . Pero, para  $r = 0$ , sí tenemos una singularidad física, que se aprecia calculando los escalares de curvatura, pues como  $R_{\mu\nu} = 0$  es un tensor, será cero también en cualquier coordenada, teniendo así que  $R = 0$  y  $R_{\mu\nu} R^{\mu\nu} = 0$ , pero esto no da información.

Podemos usar el escalar de Krestchmann, que viene dado por

$$K = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} = \dots = \frac{48G^2 M^2}{r^6}$$

tal que  $\lim_{r \rightarrow 0} K = \infty$ , es decir, las fuerzas de marea son infinitas, teniendo así una singularidad física.

- (II) Si  $r = 2GM$  tenemos  $\nabla_\mu t \nabla_\nu t g^{\mu\nu} = 0$ , definiendo así el **horizonte de eventos** o de sucesos.
- (III) La métrica de Schwarzschild en  $r < 2GM$  es difeomorfa a una métrica de tipo Kantowski-Sachs, tal que

$$ds^2 = -\frac{dT^2}{\frac{2GM}{T} - 1} + \left( \frac{2GM}{T} - 1 \right) dx^2 + T^2 d\Omega^2$$

donde intercambiamos las coordenadas espaciales por las coordenadas temporales, denominándose homogénea, pues no depende de  $x$ , pero sí varía con  $T$ . Además, tiene singularidad en  $T = 0$ , tal que  $T \in (2GM, 0)$ .

#### 4.8.6. Métrica de Kerr. Agujero negro en rotación. 1963

Se define como,

$$\begin{aligned} ds^2 = & -\frac{\Delta(r) - a \sin^2 \theta}{\Sigma(r, \theta)} dt^2 - 2a \sin^2 \theta \frac{2GMr}{\Sigma(r, \theta)} d\varphi dt + \\ & + \frac{(r^2 + a^2)^2 - \Delta(r)a^2 \sin^2 \theta}{\Sigma(r, \theta)} \sin^2 \theta d\varphi^2 + \frac{\Sigma(r, \theta)}{\Delta(r)} dr^2 + \Sigma(r, \theta) d\theta^2 \end{aligned}$$

donde  $a = \frac{J}{M}$ , es decir, es el momento angular del agujero negro normalizado a su masa si  $M > 0$  y  $GM > a$ . Además,

$$\Sigma(r, \theta) = r^2 + a^2 \cos^2 \theta$$

$$\Delta(r) = r^2 + a^2 - 2GMr$$

#### Propiedades

- (I) Tiene dos campos vectoriales de killing, tales que

$$t^\mu = \delta_t^\mu; \quad \varphi^\mu = \delta_\varphi^\mu$$

- (II) Tiene dos horizontes en

$$r_\pm = GM \pm \sqrt{G^2 M^2 - a^2}$$

tal que  $\lim_{a \rightarrow 0} r_+ = 2GM$ .

- (III) Tiene dos ergosferas (zonas del espacio-tiempo donde todavía podemos compensar la rotación del espacio-tiempo provocada por el agujero negro). Tiene una ergosfera en el horizonte exterior,

$$r_+ < r < GM + (G^2 M^2 - a^2 \cos^2 \theta)^{1/2} < R_+$$

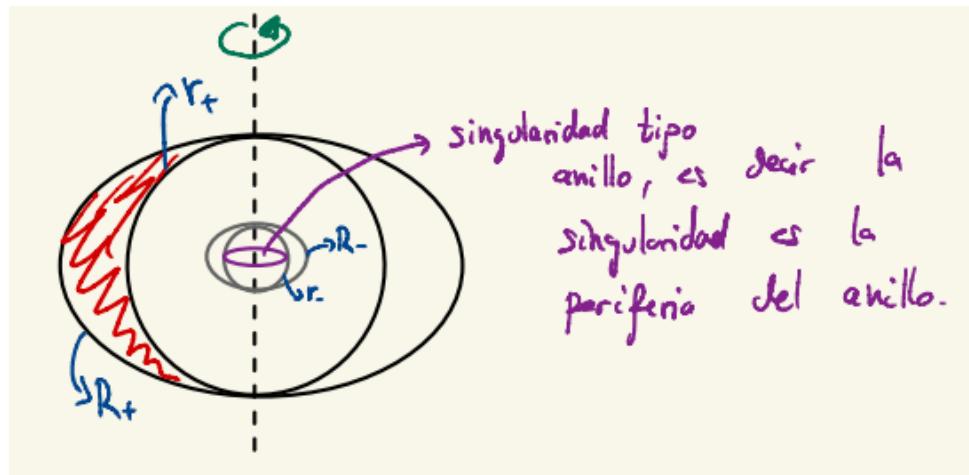


FIGURA 4.4: Esquema de un agujero negro de Kerr.

## 4.9. Testeos clásicos de la Relatividad General

#### 4.9.1. Geodésicas en una geometría de Schwarzschild

Sea  $u^\mu = \frac{dx^\mu}{d\tau} = \dot{x}^\mu$  el vector tangente a una geodésica y  $\tau$  es el parámetro afín. La norma es  $u^\mu u^\nu g_{\mu\nu} = \mathcal{E}$ , donde si  $\mathcal{E} = 0$ , tenemos vectores nulos y si  $\mathcal{E} = -1$ , tenemos vectores temporales. Usando la métrica de Schwarzschild tenemos,

$$u^\mu u^\nu g_{\mu\nu} = - \left(1 - \frac{2GM}{r}\right) \dot{t}^2 + \frac{\dot{r}^2}{1 - \frac{2GM}{r}} + r^2 \dot{\varphi}^2 = \mathcal{E}$$

donde  $\theta = \pi/2$ , representando el plano ecuatorial.

Recordamos que tenemos dos campos vectoriales de Killing,  $t^\mu = \delta_t^\mu$  y  $\Psi^\mu = \delta_\Psi^\mu$ , indicando que tenemos dos cantidades conservadas, que serán,

$$-u^\mu t^\nu g_{\mu\nu} = \left(1 - \frac{2GM}{r}\right) t = E$$

$$u^\mu \Psi^\nu g_{\mu\nu} = r^2 \dot{\varphi} = L$$

donde definimos las siguientes cantidades conservadas:

$$\left. \begin{array}{l} E : \text{ Energía por unidad de masa en reposo.} \\ L : \text{ Momento angular por unidad de masa en reposo.} \end{array} \right\} \text{ si } \mathcal{E} = -1$$

Si  $\mathcal{E} = 0$ , definimos  $\hbar E$  y  $\hbar L$ , que son la energía y el momento angular de un fotón.

Podemos resolver  $u^\mu u^\nu g_{\mu\nu} = \mathcal{E}$  como

$$\frac{1}{2} \dot{r}^2 + V(r) = \frac{E^2}{2}; \quad V(r) = \frac{1}{2} \left( 1 - \frac{2GM}{r} \right) \left( \frac{L^2}{r^2} - \mathcal{E} \right)$$

Tomando  $\dot{r} = 0$ , estudiamos órbitas circulares, tal que  $V'(r) = 0$  y  $V''$  nos dicen si son órbitas estables o no. Para  $\mathcal{E} = -1$ , se puede comprobar que si  $L^2 < 12G^2M^2$ , entonces  $V'(r) \neq 0$ , por lo que no hay órbitas circulares, sino que caen desde  $r = 2GM$  hasta  $r = 0$ .

Si  $L^2 > 12G^2M^2$ , entonces  $V'(r) = 0$  y obtenemos

$$R_\pm = \frac{L^2 \pm (L^4 - 12L^2G^2M^2)}{2GM}$$

teniendo con  $R_+$  una órbita estable alrededor del objeto de masa  $M$ , y con  $R_-$  tenemos una órbita inestable.

Notemos que  $R_+ > 6GM$  y que  $L \gg 2GM$ , entonces podemos aproximar  $R_+ \approx \frac{L^2}{GM}$ , siendo éste el límite newtoniano.

Existen órbitas circulares a radios menores, pero éstas son inestables, pues  $3GM < R_- < 6GM$ , siendo  $R_-$  es inestable, y como el mínimo del momento angular es  $L^2 = 12G^2M^2$ , entonces para  $3GM$ ,  $L^2 \rightarrow \infty$ .

Podemos expresar  $L = L(R_\pm, M)$  y  $E = E(R_\pm, M)$  resolviendo  $V'(R_\pm) = 0$ , tal que

$$E = \sqrt{2V(R_\pm)} = \frac{R_\pm - 2GM}{R_\pm^{1/2} (R_\pm - 3GM)^{1/2}}$$

de forma que

$$\lim_{R_+ \rightarrow \infty} E = 1; \quad \lim_{R_- \rightarrow 3GM} E = \infty; \quad \lim_{R_+ \rightarrow 6GM} E = \left(\frac{8}{2}\right)^{1/2} \lesssim 1$$

La energía radiada en forma de onda gravitacional desde  $R_+ >> 6GM$  hasta  $R_+ \sim 6GM$  será

$$\Delta E = 1 - \left(\frac{8}{9}\right)^{1/2} = 0.06$$

donde hemos comparado los dos límites de  $R_+$ . Este resultado nos dice que el cuerpo pierde un 6 % de su masa, emitiendo energía en forma de onda gravitacional.

#### 4.9.2. Variación del perihelio de Mercurio

La variación del perihelio de Mercurio no se puede explicar con Newton, siendo ésta de  $574.10 \pm 0.65$  arcsec/siglo. La corrección relativista es de 43 arcsec/siglo.

Dado una órbita circular, y desplazamos ligeramente con respecto a  $r = R_+$ , tal que

$$r(t) = R_+ + \delta r(t) \implies \delta \ddot{r} + \delta r V''(R_+) = 0$$

La frecuencia asociada al tiempo propio es

$$\omega_r^2 = \frac{d^2V}{dr^2} \Big|_{r=R_+} = \frac{GM(R_+ - 6GM)}{R_+^3(R_+ - 3GM)}$$

La frecuencia angular es  $\omega_\varphi = \dot{\varphi}$ , tal que

$$\omega_\varphi^2 = \frac{L^2}{R_+^4} = \frac{GM}{R_+^2(R_+ - 3GM)}$$

En el límite newtoniano tenemos que  $\omega_r \approx \omega_\varphi$ .

Las órbitas no circulares en la mecánica newtoniana son cerradas, pues  $\omega_r \approx \omega_\varphi$ , pero en Relatividad General, las órbitas no circulares son abiertas, pues  $\omega_r \neq \omega_\varphi$ .

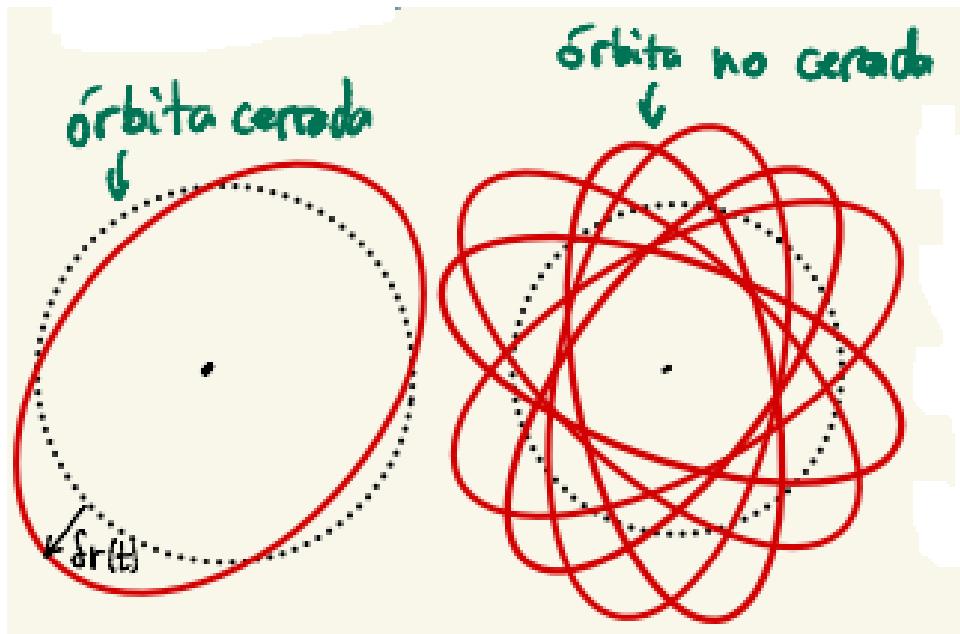


FIGURA 4.5: Comparación de una órbita cerrada y una abierta.

En las órbitas abiertas hay una precesión del ángulo en el que los máximos (o mínimos) de  $\delta_r$  son alcanzados.

La frecuencia de precesión es

$$\omega_p = \omega_\varphi - \omega_r = - \left[ \left( 1 - \frac{6GM}{R_+} \right)^{1/2} - 1 \right] \omega_\varphi$$

En el límite  $R_+ \gg 2GM$ , tenemos

$$\omega_p \approx \frac{3(GM)^{3/2}}{c^2 R_+^{5/2}}$$

Para una órbita elíptica tenemos

$$\omega_p \approx \frac{3(GM)^{3/2}}{c^2 a^{5/2} (1 - e)}$$

donde  $a$  es el semieje mayor y  $e$  es la excentricidad.

### 4.9.3. Corrimiento al rojo gravitacional

Supongamos dos observables estáticos,  $O_1$  y  $O_2$ , con  $R_1 < R_2$ . Al ser estáticos  $u_i^\mu \propto \delta_t^\mu$  tal que  $u_i^\mu u_i^\nu g_{\mu\nu} = -1$ , con  $i = 1, 2$ . Así,

$$u_1^\mu = \frac{\delta_t^\mu}{\sqrt{-g_{tt}}} \Big|_{r=R_1}; \quad u_2^\mu = \frac{\delta_t^\mu}{\sqrt{-g_{tt}}} \Big|_{r=R_2}$$

Suponemos que  $O_1$  manda una señal luminosa a  $O_2$ . El vector tangente  $k^\mu$  a la trayectoria geodésica (las trayectorias de los rayos de luz en cualquier geometría es siempre una geodésica) es nulo.

La cantidad  $k_\mu \xi^\mu$  es conservada, es decir,

$$k_\mu \xi^\mu \Big|_{r=R_1} = k_\mu \xi^\mu \Big|_{r=R_2}$$

Las frecuencias de emisión y recepción son,

$$\omega_1 = k_\mu u_1^\mu \Big|_{r=R_1}; \quad \omega_2 = k_\mu u_2^\mu \Big|_{r=R_2}$$

Entonces, si  $R_1 < R_2$  y  $\omega_2 < \omega_1$ , tenemos

$$\frac{\omega_1}{\omega_2} = \frac{(-g_{tt})^{1/2} \Big|_{r=R_2}}{(-g_{tt})^{1/2} \Big|_{r=R_1}} = \frac{\sqrt{1 - \frac{GM}{R_2}}}{\sqrt{1 - \frac{GM}{R_1}}} \Rightarrow \frac{\Delta\omega}{\omega} \begin{cases} R_1 \gg 2GM \\ R_2 \gg GM \end{cases} \approx \frac{GM}{R_2} - \frac{GM}{R_1} = \Delta\Phi$$

que coincide con lo medido en el experimento de Pond y Rebka en 1960.

En este caso, hay un corrimiento al rojo, pues está escapando la luz del pozo de potencial. (Puede interpretarse como un efecto Doppler, pero en vez de tener dos observadores en movimiento, los tenemos estáticos y, o bien el propio espacio se mueve, o bien hay un flujo de espacio entre ambos observadores.)

### 4.9.4. Deflexión de la luz

Vamos a estudiar geodésicas nulas,  $\mathcal{E} = 0$ . En este caso, el potencial es

$$V(r) = \frac{L^2}{2r^3} (r - 2GM)$$

y se puede ver que solo tenemos una órbita circular inestable en  $r = 3GM$  con  $E^2 = 2V(r = 3GM) = \frac{L^2}{27G^2M^2}$ .

Definimos el parámetro de impacto aparente como

$$b = \frac{L}{E}$$

si  $M = 0$ , entonces  $b$  sí es el parámetro de impacto real, definido como la distancia de máximo acercamiento, es decir,  $\dot{r} = 0$ .

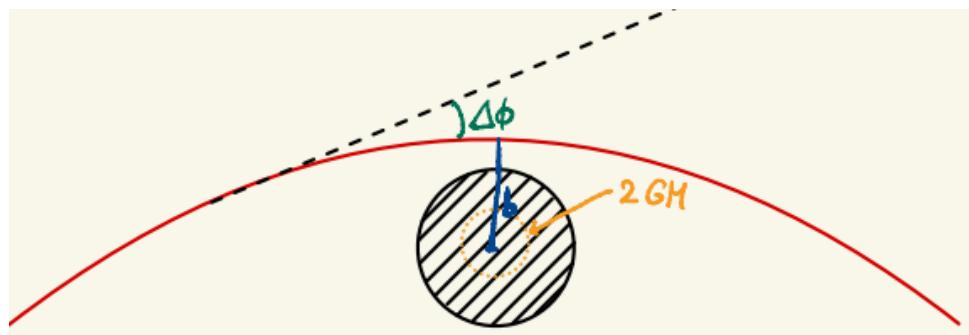


FIGURA 4.6: Esquema del parámetro de impacto.

Existe un radio de impacto crítico  $b_C = 3^{3/2}GM$ , siendo un parámetro de impacto para un objeto masivo, de forma que si  $r < b_C$ , todo cae dentro del agujero negro.

La sección eficaz del agujero negro vendrá dado por

$$\sigma = \pi b_C^2 = 27\pi G^2 M^2$$

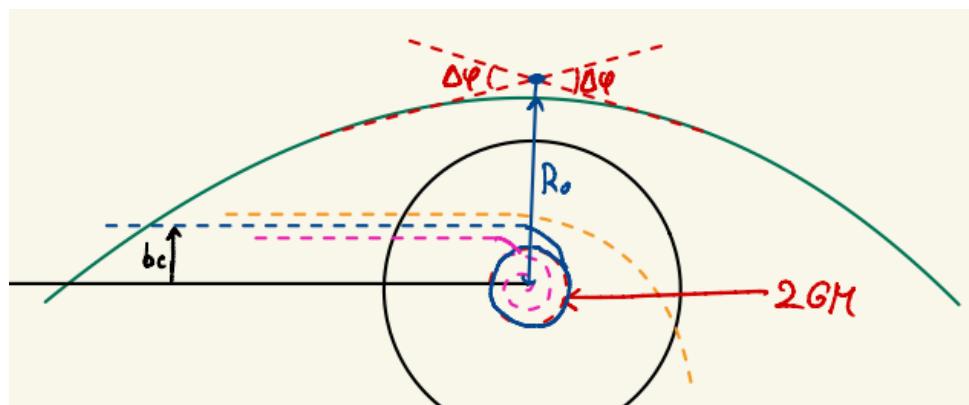


FIGURA 4.7: Esquema de un agujero negro.

Si estamos justo en  $r = b_C$ , comenzamos a orbitar el agujero negro sin caer; en este límite es donde se encuentra el **disco de acreción** de un agujero negro, siendo una órbita

inestable, pues con una pequeña perturbación terminas cayendo dentro.

$R_0$  es el punto de máximo acercamiento, que obtenemos resolviendo la ecuación

$$E = 2V(R_0) \Rightarrow R_0^3 - b^2(R_0 - 2GM) = 0$$

donde el  $R_0$  que nos interesa será la solución mayor de la ecuación y si  $M = 0$ , tenemos que  $R_0 = b$ . Si  $M \neq 0$ , tendremos

$$R_0 = \frac{2b}{\sqrt{3}} \cos \left( \frac{1}{3} \arccos \left[ -\frac{3^{3/2}GM}{b} \right] \right)$$

La trayectoria, se obtendrá resolviendo,

$$\frac{\dot{\varphi}}{r} = \frac{d\varphi}{dr} = \frac{L}{r^2} \left( E^2 - \frac{L^2}{r^3}(r - 2GM) \right)^{-1/2}$$

y como la trayectoria es simétrica en  $(-\infty, R_0]$  y  $[R_0, +\infty)$ , tenemos

$$\Delta\varphi = 2 \int_{R_0}^{\infty} \frac{dr}{\sqrt{r^4b^{-2} - r(r - 2GM)}}$$

Si  $M = 0$ , tenemos que  $\Delta\varphi = \pi$ , pues  $\Delta\varphi = 2 \arcsin(b/R_0) \xrightarrow{1} \pi$ .

Si  $M \neq 0$ , entonces  $\Delta\varphi = \pi + \delta\phi$ , donde si  $\delta\phi \ll 1$ , entonces  $R_0 \sim b$  y  $r \gg 2GM$ , obteniéndose que

$$\delta\phi = \frac{4GM}{bc^2}$$

siendo la ecuación de la deflexión de la luz relativista. Esta ecuación sirve para trayectorias fuera del límite fuerte (radio de la estrella).

Si se hace con Newton, sale que  $\delta\phi_N \propto \frac{1}{2}\delta\phi$ , debido a que con Newton no se tiene en cuenta las correcciones relativistas.

Para el Sol tenemos que  $M = M_O$  y  $b = R_O$ , luego  $\delta\phi_O = 1,75$  arcsec.

#### 4.9.5. Dilatación temporal gravitatoria

Tomamos

$$\frac{\dot{t}}{\dot{r}} = \frac{dt}{dr} = \left( 1 - \frac{2GM}{r} \right)^{-1} \left[ 1 - \left( 1 - \frac{2GM}{r} \right) \frac{b^2}{r^2} \right]^{-1/2}$$

Asumimos que mandamos un rayo desde la superficie de un cuerpo masivo y vuelve.

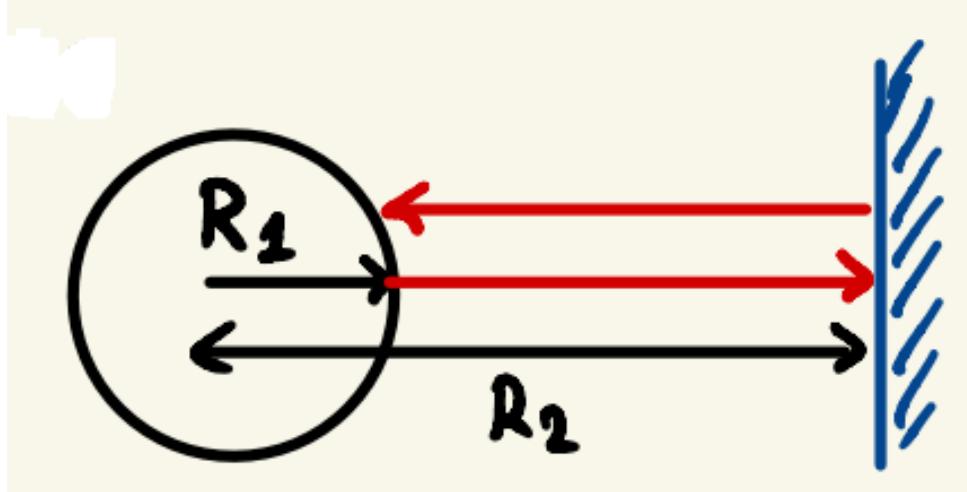


FIGURA 4.8: Esquema del experimento mental.

Vemos que resolviendo esta ecuación tenemos,

$$\Delta t = 2 \int_{R_1}^{R_2} \frac{dr}{1 - \frac{2GM}{r}} = \dots = 2r + 4GM \ln |r - 2GM| \Big|_{R_1}^{R_2}$$

Si  $M = 0$ , tenemos que  $\Delta t_0 = 2 \frac{R_2 - R_1}{c}$ .

Si  $M \neq 0$ , tenemos que  $\Delta t = \Delta t_0 + \delta t$ , con  $\delta t = \frac{4GM}{c^3} \ln \left( \frac{R_2}{R_1} \right)$ .

Pero lo tenemos en el espacio coordenado, y al ser un observador estático en  $r = R_1$ , debemos pasar al tiempo propio, tal que

$$\Delta\tau = \left( 1 - \frac{2GM}{c^2 R_1} \right) \Delta t \approx \Delta t_0 - \frac{2GM}{R_1 c^2} \Delta t_0 + \frac{4GM}{c^3} \ln \left( \frac{R_2}{R_1} \right)$$

Si tomamos un láser desde la superficie de la Tierra y rebota en un espejo en la superficie de la Luna, obtenemos,  $\Delta t_0 \approx 4 \cdot 10^{-6}$  s,  $\frac{-2GM}{R_1 c^2} \Delta t_0 \approx -6 \cdot 10^{-9}$  s y  $\delta t \approx 2,4 \cdot 10^{-1}$  s.

# Capítulo 5

## Ondas gravitacionales

*“La creatividad es la inteligencia divirtiéndose”.*

(A. Einstein)

## 5.1. Contexto histórico

(...)

Las ondas gravitacionales se forman en sistemas binarios de cuerpos masivos que orbitan. Como radian energía en forma de onda gravitacional, las órbitas se van reduciendo, pues pierden energía, haciendo que la orbitación se acelere, aumentando la frecuencia de la onda gravitacional, hasta que colisionen ambos cuerpos.

Las ondas gravitacionales viajan a la velocidad de la luz.

Los detectores de ondas gravitacionales funcionan como un interferómetro de Fabry-Perot, pero los espejos no están fijos, sino que están en péndulos, de forma que si pasa una onda gravitacional, los espejos se mueven y se mide la onda.

## 5.2. Ondas gravitacionales

Vamos a trabajar en el límite del campo débil, es decir, trabajamos con la métrica  $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$  y  $g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$ , donde  $\eta_{\mu\nu}$  es la métrica de Minkowski y  $|h_{\mu\nu}| \ll 1$  es una perturbación, tal que  $h^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}$ . Los símbolos de Christoffel serán,

$${}^{(1)}\Gamma_{\nu\rho}^\mu = \frac{1}{2}\eta^{\mu\sigma}(h_{\sigma\nu,\rho} + h_{\sigma\rho,\nu} - h_{\nu\rho,\sigma}) + \cancel{\text{termos de orden 0}}$$

El tensor de Ricci a orden 1 en  $h_{\mu\nu}$  será,

$${}^{(1)}R_{\mu\nu} = \partial_\sigma {}^{(1)}\Gamma_{\mu\nu}^\sigma - \partial_\mu {}^{(1)}\Gamma_{\sigma\nu}^\sigma + \cancel{\text{termos de orden 0}} \partial^\sigma \partial_{(\nu} h_{\mu)\sigma} - \frac{1}{2}\partial^\sigma \partial_\sigma h_{\mu\nu} - \partial_\mu \partial_\nu h$$

donde  $h = h_{\mu\nu}\eta^{\mu\nu}$ .

El tensor de Einstein será

$${}^{(1)}G_{\mu\nu} = {}^{(1)}R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}{}^{(1)}R$$

donde  ${}^{(1)}R = {}^{(1)}R_{\mu\nu}\eta^{\mu\nu} = \partial^\rho \partial^\sigma h_{\rho\sigma} - \partial^\rho \partial_\rho h$ .

Definimos,

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \Rightarrow \bar{h} = -h$$

donde  $\bar{h} = \bar{h}_{\mu\nu}\eta^{\mu\nu}$ .

En términos de  $\bar{h}_{\mu\nu}$ , tenemos

$${}^{(1)}G_{\mu\nu} = \partial^\sigma\partial_{(\nu}\bar{h}_{\mu)\sigma} - \frac{1}{2}\partial^\sigma\partial_\sigma\bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\partial^\sigma\partial^\rho\bar{h}_{\sigma\rho}$$

Recordemos que una métrica cualquiera  $g_{\mu\nu}$  está definida salvo difeomorfismos. Es decir, tenemos libertad para trabajar con  $g_{\mu\nu}$  o con  $g_{\mu\nu} + \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu$ .

Si tomamos  $|\xi_\mu| \ll 1$ , pero es comparable con  $h_{\mu\nu}$  ( $|\xi_\mu| \sim |h_{\mu\nu}|$ , entonces  $\eta_{\mu\nu} + h_{\mu\nu}$  y  $\eta_{\mu\nu} + \underbrace{h_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu}_{\mathcal{L}_\xi\eta_{\mu\nu}}$  representan la misma métrica.

Con esta libertad podemos simplificar nuestro problema, pues dado  $h_{\mu\nu}$ , podemos transformarlo como

$$\left. \begin{aligned} h'_{\mu\nu} &= h_{\mu\nu} + \partial_\nu\xi_\mu + \partial_\mu\xi_\nu \\ h' &= h + 2\partial_\mu\xi^\mu \end{aligned} \right\} \Rightarrow \bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu - \eta_{\mu\nu}\partial_\rho\xi^\rho$$

Entonces podemos transformar  $\partial^\mu\bar{h}_{\mu\nu}$  como

$$\partial^\mu\bar{h}'_{\mu\nu} = \partial^\mu\bar{h}_{\mu\nu} + \partial^\mu\partial_\mu\xi_\nu + \cancel{\partial_\nu\partial^\mu\xi_\mu} - \cancel{\eta_{\mu\nu}\partial^\mu\partial_\rho\xi^\rho}$$

Y si se cumpla que  $\partial^\mu\partial_\mu\xi_\nu = -\partial^\mu\bar{h}_{\mu\nu}$ , entonces tenemos que  $\partial^\mu\bar{h}'_{\mu\nu} = 0$ , que es una función de onda que podemos resolver. Podemos hallar  $\xi^\mu$  usando funciones de Green. Además, la nueva perturbación  $\bar{h}'_{\mu\nu}$  cumplirá el gauge de Lorentz, que es precisamente,  $\partial^\mu\bar{h}'_{\mu\nu} = 0$ . Así, el tensor de Einstein se reduce a

$${}^{(1)}G'_{\mu\nu} = -\frac{1}{2}\partial^\sigma\partial_\sigma\bar{h}'_{\mu\nu}$$

En el vacío sabemos que  ${}^{(1)}G'_{\mu\nu} = 0$ , por tanto en vacío se cumple que  $\partial^\sigma\partial_\sigma\bar{h}'_{\mu\nu} = 0$ , siendo una función de ondas electromagnéticas. Por tanto, vemos que las ondas gravitacionales se propagan a la velocidad de la luz,  $c$ .

Podemos incluso simplificar más los cálculos suponiendo ahora que

$$\xi^\mu = \bar{\xi}^\mu + \tilde{\xi}^\mu$$

donde  $\partial^\sigma \partial_\sigma \tilde{\xi}^\mu = -\partial^\mu \bar{h}_{\mu\nu}$  y  $\partial^\sigma \partial_\sigma \bar{\xi}_\nu = 0$ . Por lo que todavía tenemos libertad para simplificar  $\bar{h}'_{\mu\nu}$ . Notemos que

$$\left. \begin{array}{l} \bar{h}'_{\mu\nu} = \bar{H}'_{\mu\nu} e^{ik_\mu x^\mu} \Rightarrow \bar{\xi}_\mu = A_\mu e^{ik_\mu x^\mu} \\ \downarrow \\ \partial^\mu \bar{h}'_{\mu\nu} = 0 \end{array} \right\} \Rightarrow \bar{H}'_{\mu\nu} k^\nu = 0$$

Como  $\bar{h}'_{\mu\nu}$  es simétrico, entonces  $\bar{H}'_{\mu\nu}$  es simétrico.

Ahora vamos a tomar una nueva perturbación,

$$\bar{h}''_{\mu\nu} = \bar{h}'_{\mu\nu} + \partial_\mu \bar{\xi}_\nu + \partial_\nu \bar{\xi}_\mu - \eta_{\mu\nu} \partial_\sigma \bar{\xi}^\sigma$$

tal que

$$\bar{H}''_{\mu\nu} = \bar{H}'_{\mu\nu} + iA_\mu k_\nu + iA_\nu k_\mu - i\eta_{\mu\nu} A_\sigma k^\sigma$$

donde vamos a imponer que  $\bar{H}''_\mu^\mu = 0$  y  $\bar{H}''_{0i} = 0$  con  $i = 1, 2, 3$ . Así tenemos cuatro ecuaciones con cuatro incógnitas, que son las  $A_\mu$ . Notemos que  $\partial^\mu \bar{h}''_{\mu 0} = 0$ , por tanto  $\partial_0 \bar{h}''_{00} = 0$  y entonces  $\bar{H}''_{00} = 0$ . Además, como  $\bar{H}''_\mu^\mu = 0$ , entonces  $\bar{h}''_\mu^\mu = 0$  y por tanto,  $\bar{h}''_{\mu\nu} = h''_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h''^0 = h''_{\mu\nu}$ . Este gauge se conoce como *radiativo* o *TT-gauge* (traceless and transverse), pues la traza es nula.

Recapitulando tenemos que  $h_{ij} = h_{ji}$ , con  $i, j = 1, 2, 3$ , satisface que  $(-\partial_t^2 + \nabla^2)h_{ij} = 0$ , por lo que

$$h_{ij} = H_{ij} e^{ik_\mu x^\mu}$$

donde  $H_{ij}$  tiene 6 parámetros porque es simétrico y  $k_\mu = (\omega, k_i)$  con  $\omega^2 = |\vec{k}|^2 c^2$ , siendo  $\omega$  la velocidad angular y  $\vec{k}$  el vector número de onda. Por tanto, tenemos el sistema de ecuaciones siguiente,

$$A_{ij} k^j = 0; \quad A_i^i = 0$$

donde  $A_{ij} k^j = 0$  son 3 ecuaciones y  $A_i^i = 0$  es una ecuación, pero como  $H_{ij}$  tiene 6 parámetros, tendremos 6 parámetros - 4 ecuaciones, es decir, tenemos 2 grados de libertad, que se representan como 2 polarizaciones posibles de la onda gravitacional.

### 5.3. Ecuación de la desviación geodésica

Recordamos que  $T^\mu$  es el vector tangente a la geodésica, que satisface  $T^\mu \nabla_\mu T^\nu = 0$  y que  $\gamma_s(t) = \gamma_{p(s)}(t)$ , es decir, la familia de geodésicas  $\gamma_s(t)$  definen una superficie 2-D en la variedad  $\mathcal{M}$ , tal que  $t$  y  $s$  pueden definir coordenadas en la superficie (siempre que las geodésicas se crucen).

Si definimos  $S^\mu$  como el campo vectorial que define el desplazamiento infinitesimal entre geodésicas, entonces  $T^\mu$  y  $S^\mu$  son vectores de una base coordinada, es decir,  $\alpha_s T^\mu = 0$ , de forma que

$$S^\mu \nabla_\mu T^\nu = T^\mu \nabla_\mu S^\nu$$

Por tanto, definimos la velocidad relativa de separación entre geodésicas como  $v^\mu = T^\nu \nabla_\nu S^\mu$ .

Veamos ahora como escribir esto en función de la curvatura, y como  $T^\mu$  y  $S^\mu$  son vectores de una base coordinada, es decir, conmutan, así

$$\begin{aligned} a^\mu &= T^\nu \nabla_\nu (T^\rho \nabla_\rho S^\mu) = T^\nu \nabla_\nu (S^\rho \nabla_\rho T^\mu) = (T^\nu \nabla_\nu S^\rho) \nabla_\rho T^\mu + T^\nu S^\rho \nabla_\nu \nabla_\rho T^\mu = \\ &= \underbrace{(T^\nu \nabla_\nu S^\rho)}_{S^\nu \nabla_\nu T^\rho} \nabla_\rho T^\mu + T^\nu S^\rho \nabla_\rho \nabla_\nu T^\mu - \mathcal{R}_{\nu\rho\sigma}^\mu S^\rho T^\nu T^\sigma = \\ &= S^\nu \nabla_\nu \underbrace{(T^\rho \nabla_\rho T^\mu)}_{=0 \text{ (al ser geodésica)}} - \mathcal{R}_{\nu\rho\sigma}^\mu S^\rho T^\nu T^\sigma = -\mathcal{R}_{\nu\rho\sigma}^\mu S^\rho T^\nu T^\sigma = \frac{d^2 S^\mu}{dt^2} \end{aligned}$$

siendo esta la ecuación de la geodésica, donde  $t$  es el parámetro afín de la geodésica.

#### 5.3.1. Efecto de la una onda gravitacional que pasa cerca del observador

Tomemos un conjunto de objetos en reposo relativo inicialmente con cuadivelocidad  $u^\mu = \delta_t^\mu$ . Suponemos que dos de estos objetos están separados una distancia  $S^\mu$ . La ecuación de la geodésica es,

$$\frac{d^2 S^\mu}{dt^2} = \mathcal{R}_{\nu 00}^\mu S^\nu = \dots = \frac{1}{2} (\partial_t^2 h_\nu^\mu) S^\nu$$

donde  $h_\nu^\mu$  es una perturbación que cumple el gauge radiativo (realmente da igual en qué gauge trabajar, pues la curvatura y las ecuaciones de Einstein son independientes de los

gauges; usamos este gauge porque sabemos la forma de  $h_\nu^\mu$ ).

Supongamos que  $H_x = 0$  y  $H_t \neq 0$ , entonces  $S^t = 0 = S^z$  y tenemos dos ecuaciones geodésicas, tales que

$$\frac{d^2 S^x}{dt^2} = \frac{1}{2} S^x H_+ (-\omega^2 e^{ik_\mu x^\mu})$$

$$\frac{d^2 S^y}{dt^2} = -\frac{1}{2} S^y H_+ (-\omega^2 e^{ik_\mu x^\mu})$$

Integramos estas ecuaciones y obtenemos,

$$S^x = S^x(0) \left( 1 + \frac{1}{2} H_+ e^{ik_\mu x^\mu} \right)$$

$$S^y = S^y(0) \left( 1 - \frac{1}{2} H_+ e^{ik_\mu x^\mu} \right)$$

El efecto de la onda gravitacional sería el siguiente,

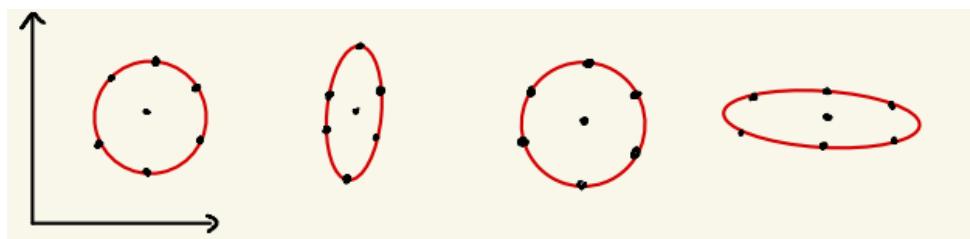


FIGURA 5.1: Efecto que sufren los objetos cuando pasa la onda gravitacional.

es decir, los objetos se contraen en la dirección vertical y horizontal.

Análogamente, para la polarización  $H_x \neq 0$  y  $H_t = 0$ , tenemos las soluciones,

$$S^x = S^x(0) + \frac{1}{2} H_x S^y(0) e^{ik_\mu x^\mu}$$

$$S^y = S^y(0) + \frac{1}{2} H_x S^x(0) e^{ik_\mu x^\mu}$$

Los efectos de la onda gravitacional son los siguientes,

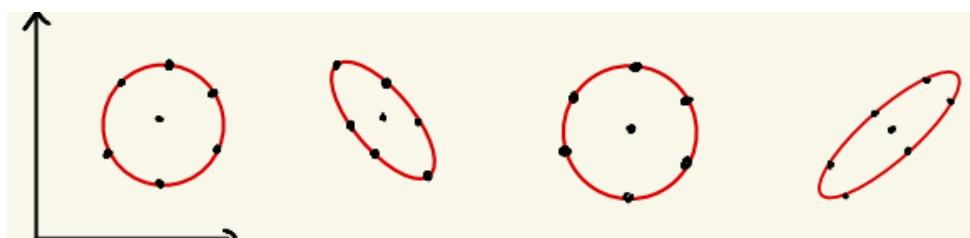


FIGURA 5.2: Efecto que sufren los objetos cuando pasa la onda gravitacional.

es decir, los objetos se contraen en las direcciones oblicuas.

No es difícil ver que si realizamos una rotación de 180°, el resultado es invariante, por tanto, las ondas gravitacionales **tienen espín 2**.

También existe la polarización circular, tal que

$$H_{+2} = \frac{1}{\sqrt{2}}(H_+ + iH_x)$$

$$H_{-2} = \frac{1}{\sqrt{2}}(H_+ - iH_x)$$

cuyo efecto es una rotación de los objetos sin contraerlos.

Si tenemos un interferómetro de Fabry-Perot, podemos llegar a medir estas ondas gravitacionales, pues dejamos los espejos en un péndulo, y cuando pase la onda hará que se muevan, tal que

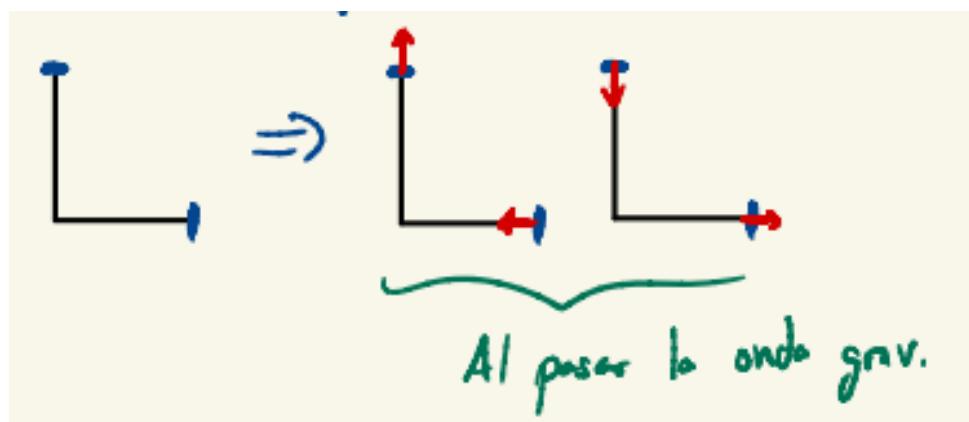


FIGURA 5.3: Efecto de la onda gravitacional sobre un interferómetro de Fabry-Perot.

### 5.3.2. Generación de las ondas gravitacionales

Las ecuaciones de Einstein en el gauge de Lorentz las podemos escribir como,

$$(1) 8\pi GT_{\mu\nu}; \quad -\frac{1}{2}\partial^\sigma\partial_\sigma\bar{h}_{\mu\nu} = 8\pi GT_{\mu\nu}; \quad \partial^\mu T_{\mu\nu} = 0 \Rightarrow \partial^\mu\bar{h}_{\mu\nu} = 0$$

Esta ecuación la podemos resolver utilizando las funciones de Green, tal que

$$\bar{h}_{\mu\nu}(x) = -16\pi G \int_{\mathcal{M}} G(x-y)T_{\mu\nu}(y)d^4y$$

donde  $x$  es la posición de la onda gravitacional,  $y$  es la posición de la fuente y  $\mathcal{G}$  es una función de Green. Por tanto, podemos decir que las funciones de Green son fuentes en el espacio.

De todas las funciones de Green usaremos la función de Green *retardada*, que representa la onda gravitacional de una fuente puntual que viaja hacia el futuro, siendo

$$\mathcal{G}(x - y) = -\frac{1}{4\pi|\vec{x} - \vec{y}|} \delta(|\vec{x} - \vec{y} - (x^0 - y^0)|) \cdot \Theta(x^0 - y^0)$$

donde

$$\Theta(x^0 - y^0) = \begin{cases} 1 & \text{si } x^0 > y^0 \\ 0 & \text{en otro caso} \end{cases}$$

Así, nos garantiza poder evaluar el cono de luz pasado de la onda gravitacional.

Notemos que  $t = x^0$  es el tiempo presente.

Si  $y^0 \geq x^0 = t$ , entonces  $\mathcal{G}(x - y) = 0$ , es decir, solo nos interesan las contribuciones de las fuentes en tiempos  $y^0 < t$ . Luego, integrando en  $y_0$  tenemos,

$$\bar{h}_{\mu\nu}(t, \vec{x}) = -4G \int \frac{1}{|\vec{x} - \vec{y}|} T_{\mu\nu} \delta(t - |\vec{x} - \vec{y}|, \vec{y}) d^3\vec{y}$$

Nos referimos a  $t_r = t - \frac{|\vec{x} - \vec{y}|}{c}$  como *tiempo retardado*. Así, dejando  $t_r$  fijo y tomando  $|\vec{x}| \gg |\vec{y}|$ , tenemos que  $|\vec{x} - \vec{y}| \approx R$ , tal que

$$\bar{h}_{\mu\nu}(t, \vec{x}) = -\frac{4G}{R} \int T_{\mu\nu}(t_r, \vec{y}) d^3\vec{y}$$

Como  $\partial_t \bar{h}_{0\nu} = \sum_{i=1}^3 \partial_i \bar{h}_{i\nu}$ , entonces  $\bar{h}_{i\nu}$  determina a  $\bar{h}_{0j}$  y  $\bar{h}_{00}$ . Por tanto, solo necesitamos estudiar la siguiente integral,

$$\int T^{ij}(t_r, \vec{y}) d^3y = \int d^3y \left[ \cancel{\frac{\partial}{\partial y^k}} (T^{ik} y^j) - \cancel{\frac{\partial T^{ik}}{\partial y^k}} y^j \right] = -\frac{\partial}{\partial t_r} \int d^3y T^{0i} y^j = \dots = \frac{1}{2} \frac{\partial^2}{\partial t_r^2} \int d^3y T^{00} y^i y^j$$

Además podemos escribirlo en términos del tensor de momento cuadrupolar, tal que

$$\int T^{ij}(t_r, \vec{y}) d^3y = \frac{1}{6} \partial_{t_r}^2 q^{ij}$$

donde

$$q^{ij}(t_r, R) = 3 \int d^3y T^{00}(t_r, y) y^i y^j$$

es el tensor de momento cuadrupolar. En resumen tenemos,

$$\bar{h}_{ij}(t, \vec{x}) = -\frac{2G}{3R} \partial_{t_r}^2 q_{ij}(t_r, R)$$

con  $t_r = t - R/c$ , siendo  $R$  la distancia a la fuente,  $t$  el tiempo de recepción y  $t_r$  el tiempo de emisión. Por tanto, si  $\partial_{t_r}^2 \neq 0$ , entonces se formarán ondas gravitacionales. No hay contribución dipolar debido a la conservación del tensor energía-impulso  $T_{\mu\nu}$ .

## 5.4. Energía de una onda gravitacional. Formalismo de Bondi

Este formalismo se desarrolla para explicar cómo es la energía de una onda gravitacional. Esta energía se denomina *potencia radiada* en ondas gravitacionales.

Debemos calcular el tensor de energía-impulso  $T_{\mu\nu}$  para ondas gravitacionales. Vamos a construir el tensor en el límite de campo débil, es decir,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}^{(1)} + h_{\mu\nu}^{(2)} + \dots$$

Podemos expandir el tensor de Einstein  $G_{\mu\nu} = 0$ , tal que

$$G_{\mu\nu}^{(0)} + G_{\mu\nu}^{(1)} + G_{\mu\nu}^{(2)} + \dots = 0$$

donde  $G_{\mu\nu}^{(1)}$  es lineal en  $h_{\mu\nu}^{(1)}$ , por tanto  $G_{\mu\nu}^{(1)} = 0$ . Y  $G_{\mu\nu}^{(2)}$  es lineal en  $h_{\mu\nu}^{(2)}$  y cuadrático en  $h_{\mu\nu}^{(1)}$ , por lo que  $G_{\mu\nu}^{(2)} = G_{\mu\nu}^{(2,1)} + G_{\mu\nu}^{(2,2)}$ , donde  $G_{\mu\nu}^{(2,2)}$  es la parte lineal en  $h_{\mu\nu}^{(2)}$  y  $G_{\mu\nu}^{(2,1)}$  es la parte cuadrática en  $h_{\mu\nu}^{(1)}$ , por lo que identificamos  $G_{\mu\nu}^{(2,1)} = 8\pi G t_{\mu\nu}$ , siendo  $t_{\mu\nu}$  cuadrático en  $h_{\mu\nu}^{(1)}$  y sus derivadas, es decir, será el tensor de energía-impulso para  $h_{\mu\nu}^{(1)}$ .

El problema de este tensor es que  $t_{\mu\nu}$  no es invariante bajo transformaciones del tipo  $g_{\mu\nu} \rightarrow g_{\mu\nu} + \mathcal{L}_{\xi}g_{\mu\nu}$ . Es decir, el tensor no está definido únicamente, por lo que la energía (potencia radiada) dependerá del sistema de coordenadas elegido. Pero si redefinimos la energía como

$$E = \int_{\Sigma} t_{00} d^3x$$

siendo  $\Sigma$  una hipersuperficie espacial, tenemos ahora sí definida la energía de forma única.

Si llevamos la energía radiada al infinito, tenemos

$$\delta E = \int_S t_{0\mu} n^\mu dt d\Omega$$

donde  $n^\mu n^\nu g_{\mu\nu} = +1$ . Por tanto, obtenemos

$$\delta E = \int \mathcal{P} dt$$

donde  $\mathcal{P}$  es la potencia radiada. Esta potencia puede calcularse y se obtiene,

$$\mathcal{P} = \frac{G}{45} \sum_{i,j=1}^3 \frac{d^3}{dt^3} \left( q_{ij} - \frac{1}{3} \delta_{ij} q \right)^2$$

donde  $q = \sum_{i,j=1}^3 q_{ij} \delta^{ij}$ .

# Capítulo 6

## Cosmología

*“Dos cosas son infinitas: la estupidez humana y el universo; y no estoy seguro de lo segundo”.*

(A. Einstein)

## 6.1. Conceptos básicos.

(...)

A escalas de más de 1000 Mpc, el universo es isotrópico y homogéneo.

$$\vec{v} = H_0 \vec{d}$$

donde  $H_0$  es la constante de Hubble.

El fondo cósmico de microondas nos indica que la potencia primordial es un invariante de escala.

Sabemos que a grandes escalas, el espacio-tiempo es homogéneo e isótropo. Definamos estos conceptos:

- **Isotropía:** una variedad diferenciable  $\mathcal{M}$  (espacio-tiempo) es isótropa alrededor de un punto  $p \in \mathcal{M}$ , si dados  $x^\mu, y^\mu \in T_p \mathcal{M}$ , existe una isometría tal que el push-forward de  $x^\mu$  es paralelo a  $y^\mu$  (sin transformar). Es decir, existen rotaciones que mapean  $x^\mu$  a  $y^\mu$ , tal que dejan la métrica invariante en  $p \in \mathcal{M}$ . La isotropía se puede definir punto a punto.
- **Homogeneidad:** dados  $p, q \in \mathcal{M}$  cualesquiera, existe una isometría que transforma  $p$  en  $q$ .

Un espacio-tiempo homogéneo no tiene por qué ser isótropo y un espacio-tiempo isótropo alrededor de un punto de la variedad no tiene por qué ser homogéneo. Pero si tenemos un espacio-tiempo isótropo en todos los  $p \in \mathcal{M}$ , entonces sí es homogéneo. Además, si es isótropo alrededor de un punto  $p \in \mathcal{M}$  y homogéneo, entonces será isótropo en todo  $p \in \mathcal{M}$ .

En adelante, aplicaremos homogeneidad e isotropía en hipersuperficies espaciales (no aplica en la dirección temporal, pues esta sí tiene una dirección 'privilegiada').

Nuestro espacio-tiempo tiene la topología  $\mathbb{R} \times \Sigma$ , donde  $\Sigma$  son secciones espaciales homogéneas e isótropas. En este caso,

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}(\vec{x})dx^i dx^j$$

donde  $a(t)$  es el factor de escala,  $\gamma_{ij}(\vec{x})$  es la métrica en las secciones espaciales que son máximamente simétricas, es decir, donde  $\mathcal{R}_{ijkl} = K[\gamma_{ik}\gamma_{jl} - \gamma_{il}\gamma_{jk}]$ , donde  $K$  podría depender de  $t$ , pero en todos los textos se asume que  $K = cte$  para evitar que la topología del espacio-tiempo cambie. El tensor de Ricci sobre las secciones espaciales  $\Sigma$  será,

$$\mathcal{R}_{jk} = \mathcal{R}_{ijkl}\gamma^{ik} = 2K\gamma_{jl}$$

Asumiendo que  $\gamma_{ij}(\vec{x})dx^i dx^j = A(r)dr^2 + r^2 d\Omega^2$ , siendo isótropa respecto a  $r = 0$ , y que  $\mathcal{R}_{ijkl}\gamma^{il} = 2K\gamma_{jl}$ , entonces

$$A(r) = \frac{1}{1 - Kr^2}$$

Por tanto, la métrica más general posible que sea homogénea e isótropa será la métrica de **Friedmann-Lemaître-Roberson-Walker**, tal que

$$ds^2 = -dt^2 + a(t)^2 \left( \frac{1}{1 - Kr^2} dr^2 + r^2 d\Omega^2 \right)$$

Notemos que si  $K \rightarrow \alpha K$ , podemos reabsorber  $\alpha$  con un cambio de coordenadas del tipo  $r \rightarrow \alpha^{-1}r$  y un reescalado del factor de escala,  $a(t) \rightarrow \alpha a(t)$ ; tendremos una métrica invariante. Por ello, solo nos interesarán los casos para  $K = -1, 0, +1$ .

## 1. Para $K=0$

Las secciones espaciales  $\Sigma$  son planas, pues la métrica queda

$$ds^2 = -dt^2 + a(t)^2 \underbrace{(dr^2 + r^2 d\Omega^2)}_{\text{Minkowski en } \mathbb{S}^2} = -dt^2 + a(t)^2 (dx^2 + dy^2 + dz^2)$$

Podemos identificar  $\Sigma$  con  $\mathbb{R}^3$ , si la topología es abierta, o con un 3-toro  $\mathbb{T}^3$ , si la topología es compacta.

## 2. Para $K=+1$

La métrica queda,

$$ds^2 = -dt^2 + a(t)^2 \left( \frac{1}{1 + r^2} dr^2 + r^2 d\Omega^2 \right) \stackrel{\{r=\sin\Psi\}}{=} -dt^2 + a(t)^2 (d\Psi^2 + \sin^2 \Psi d\Omega^2)$$

luego,  $\Sigma$  tiene topología de  $\mathbb{S}^3$ , es decir, es una 3-esfera de radio unidad, siendo cerrada y sin frontera.

### 3. Para $K=-1$

La métrica queda,

$$ds^2 = -dt^2 + a(t)^2 \left( \frac{1}{1-r^2} dr^2 + r^2 d\Omega^2 \right) \stackrel{\{r=\sinh \varphi\}}{=} -dt^2 + a(t)^2 (d\varphi^2 + \sinh^2 \varphi d\Omega^2)$$

luego,  $\Sigma$  tiene topología de un hiperboloide 3-dimensional de radio unidad; siendo abierto y con frontera.

La coordenada  $t$  que usamos en la métrica es el *tiempo comóvil*; asociado a observadores comóviles con  $u^\mu = \delta_t^\mu$ .

Sabemos que en la métrica de Schwarzschild en  $r = 0$  tenemos una singularidad física, cuyo radio viene dado por  $1 - \frac{2GM}{r} = 0$ , luego  $r_s = 2GM$ ; pero en la métrica de R-N, tendremos dos radios, pues vienen dados por  $1 - \frac{2GM}{r} - \frac{4\pi GQ^2}{r^2} = 0$ , luego  $r_\pm \neq 0$ . Para comprobarlo hacemos,

$$\mathcal{R}_{\mu\nu}\mathcal{R}^{\mu\nu} = 4 \left( \frac{4\pi GQ^2}{r^4} \right)^2 = (8\pi G)^2 T_{\mu\nu} T^{\mu\nu} = (8\pi G)^2 [(g^{tt}T_{tt})^2 + (g^{rr}T_{rr})^2 + (g^{\theta\theta}T_{\theta\theta})^2 + (g^{\varphi\varphi}T_{\varphi\varphi})^2]$$

#### 6.1.1. Observadores comóviles

Tomamos  $u^\mu = \delta_t^\mu$ , así, los símbolos de Christoffel quedan,

$$\Gamma_{ij}^0 = a\dot{a}\gamma_{ij}; \quad \Gamma_{0j}^i = \frac{\dot{a}}{a}\delta_j^i; \quad \Gamma_{\theta\theta}^r = r(1-Kr^2) = \frac{\Gamma_{\varphi\varphi}^r}{\sin^2 \theta}; \quad \Gamma_{r\theta}^\theta = \Gamma_{r\varphi}^\varphi = \frac{1}{4}; \quad \Gamma_{\varphi\varphi}^\theta = -\sin \theta \cos \theta = -\frac{\Gamma_{\theta\varphi}^\varphi}{\sin^2 \theta}$$

El tensor de Ricci queda,

$$\mathcal{R}_{tt} = -3\frac{\ddot{a}}{a}; \quad \mathcal{R}_{rr} = \frac{a\ddot{a} + 2\dot{a}^2 - 2K}{1 - Kr^2}; \quad \mathcal{R}_{\theta\theta} = r^2(a\ddot{a} + 2\dot{a}^2 + 2K) = \frac{\mathcal{R}_{\varphi\varphi}}{\sin^2 \theta}$$

El escalar de Ricci queda,

$$\mathcal{R} = \mathcal{R}_{\mu\nu}g^{\mu\nu} = \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + K) =$$

donde  $a = a(t)$ ,  $\dot{a} = \frac{da(t)}{dt}$  y  $\ddot{a} = \frac{d^2a(t)}{dt^2}$ .

Si  $T_{\mu\nu} = 0$ , tendremos varios casos,

- Si  $K = 0$ , entonces  $\dot{a} = 0$ , luego  $a(t) = cte$ .
- Si  $K = -1$ , entonces  $\dot{a} = +1$ .
- Si  $K = +1$ , entonces  $\dot{a}^2 = -1$ , pero este caso lo excluimos, pues no tiene sentido trabajar con métricas imaginarias.

### 6.1.2. Contenido de materia

Consideramos fluido perfecto, tal que  $T_{\mu\nu} = (\rho + P)u_\mu u_\nu + Pg_{\mu\nu}$ , con  $u^\mu = \delta_t^\mu$ . Usando la conservación de  $T_{\mu\nu}$  obtenemos que  $H = \dot{a}/a$ , denominado *parámetro de Hubble*. Tal que, si añadimos la ecuación de estado y  $\nabla_\mu T^{\mu\nu} = 0$ , tenemos que

$$\dot{\rho} + 3H(\rho + P) = 0$$

y

$$P = \omega\rho$$

donde consideramos  $\omega = cte$ , así podemos integrar y obtenemos,

$$\rho(t) = \rho_0 + \left( \frac{a_0}{a(t)} \right)^{-3(1+\omega)}$$

con  $a_0 = a(t_0)$  y  $\rho_0 = \rho(t_0)$ . Veamos la evolución del universo según el 'tiempo'  $\omega$ :

- Si  $\omega = 0$ , entonces  $\rho \sim a^{-3}$ , que modela un universo de materia fría que no interacciona;  $\rho$  está dominada por la masa en reposo.
- Si  $\omega = 1/3$ , entonces tenemos una etapa del universo dominada por partículas relativistas (fotones), donde  $T_{\mu\nu}g^{\mu\nu} = 0 = 3P - \rho$ , luego  $P = \rho/3$ , teniendo una etapa caliente del universo. En este caso,  $\rho \propto a^{-4}$ , es decir, los fotones se diluyen con un factor  $a^{-1}$  adicional, debido al desplazamiento al rojo de su frecuencia.
- Si  $\omega = -1$ , tenemos una constante cosmológica  $\Lambda$  que conocemos, donde  $\rho = cte$  y  $P = -\rho$ ; la densidad de energía no se diluye con  $a$ . Acaba dominando la expansión.

Las ecuaciones de Einstein nos dirán como  $a(t)$  depende con  $t$ , tal que

$$G_{\mu\nu} = 8\pi GT_{\mu\nu}; \quad \mathcal{R}_{\mu\nu} = 8\pi G \left( T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^{\mu\nu} \right)$$

donde

$$\begin{aligned}(0,0) \rightarrow -3\frac{\ddot{a}}{a} &= 4\pi G(\rho + 3P) \\ (i,i) \rightarrow \frac{\ddot{a}}{a} + 2\frac{(\dot{a})^2}{a^2} + 2\frac{K}{a^2} &= 4\pi G(\rho - P)\end{aligned}$$

Estas ecuaciones se combinan tal que

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P)$$

conocida como la **ecuación de Raychaudhuri** y también,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2}$$

conocida como la **ecuación de Friedmann**, que puede reescribirse como

$$\Omega - 1 = \frac{K}{a^2 H^2}$$

con  $\Omega = \frac{\rho}{\rho_{crit}}$ , siendo  $\rho_{crit} = \frac{3H^2}{8\pi G}$  denominada como densidad de energía crítica.

Midiendo  $\Omega$  se puede determinar  $K$ , tal que

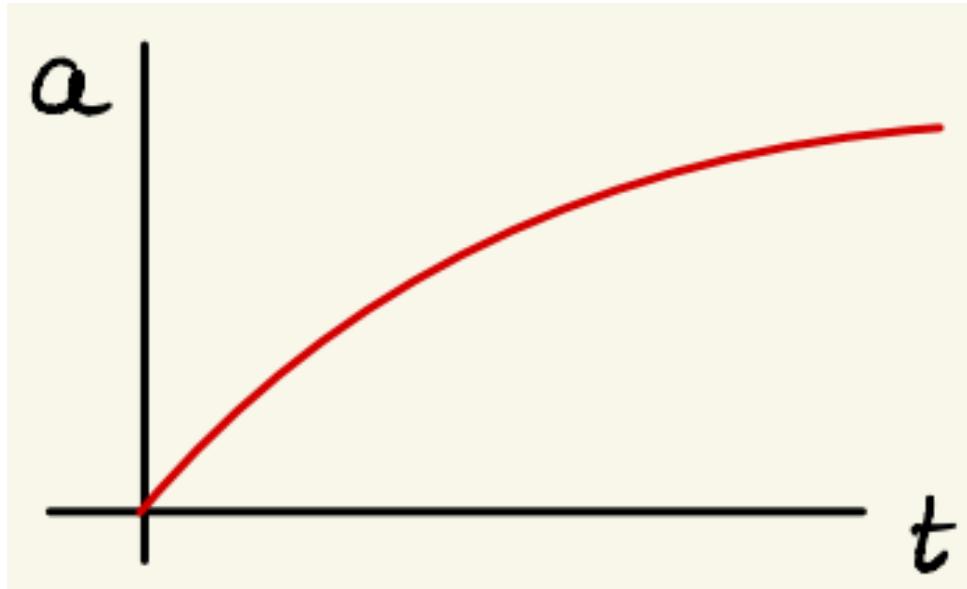
- Si  $\rho < \rho_{crit}$ , entonces  $\Omega < 1$ , por tanto  $K = -1$ , teniendo así un **universo abierto**.
- Si  $\rho > \rho_{crit}$ , entonces  $\Omega > 1$ , por tanto  $K = +1$ , teniendo así un **universo cerrado**.
- Si  $\rho \approx \rho_{crit}$ , entonces  $\Omega \approx 1$ , por tanto  $K \approx 0$ , teniendo así un **universo plano**.

Nuestras observaciones indican que  $\Omega \approx 1$ , por tanto nuestro universo, en la actualidad, es plano. También sabemos que  $\dot{a} > 0$ , por lo que el universo se expande.

Para  $\omega = 0, 1/3$  con  $\dot{a} < 0$ , tenemos un universo que desacelera. Pero gracias a supernovas se sabe que  $\dot{a} > 0$ , por lo que en la actualidad  $P < -\rho/3$ , luego la expansión se acelera. En el pasado remoto,  $\rho$  dominaba sobre  $K/a^2$ , pues  $\rho \sim a^{-3}$  y  $\rho \sim a^{-4}$  (en concreto  $\omega \rightarrow 0$  implica  $\rho \sim a^{-4}$ ). Luego,

$$\left(\frac{\dot{a}}{a}\right) \approx \frac{8\pi G}{3}\rho_{rad} = \frac{8\pi G}{3}\rho_0 \left(\frac{a_0}{a}\right)^4 \implies (\dot{a})^2 = \frac{8\pi G}{3}\rho_0 a^4 a^{-2}$$

en el límite de  $a \rightarrow 0$ ,  $\dot{a} \rightarrow \infty$ .

FIGURA 6.1: Representación de  $a$  frente a  $t$ .

En la Figura 6.1, vemos que  $a$  cuando  $a \rightarrow 0$ , llegamos a una singularidad de  $\rho \rightarrow \infty$ , por tanto esta singularidad será la **singularidad del Big Bang**.

Para  $K \leq 0$  tenemos que

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 + |k| > 0$$

El universo nunca rebota ni recolapsa.

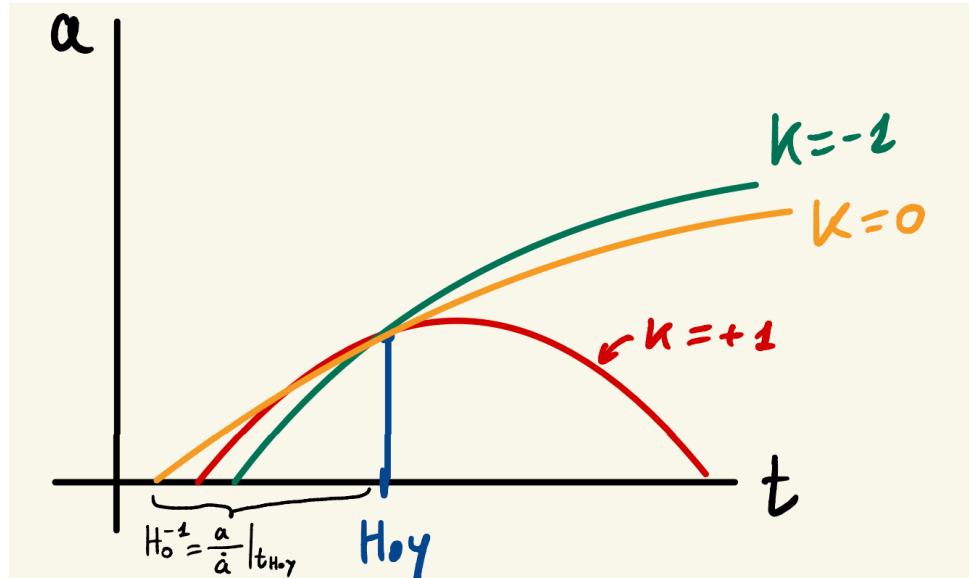
Recordando que  $P = \omega\rho$ , entonces  $\rho = \rho_0 \left(\frac{a_0}{a}\right)^{3(1+\omega)}$ . Volvamos a estudiar los distintos casos,

Para  $\omega = 0, 1/3$ , tenemos que  $\rho a^2 \xrightarrow[a \rightarrow \infty]{} 0$ . Entonces, para  $K = 0$  tenemos  $\dot{a}^2 \xrightarrow[a \rightarrow \infty]{} 0$ , es decir, el universo se expande indefinidamente, pero cada vez más rápido, de forma acelerada.

Para  $K = -1$ , tenemos que  $\dot{a}^2 = \frac{8\pi G}{3}a^2\rho - K$ . En algún instante  $t_{rec}$ , tendremos que  $\frac{8\pi G}{3}a^2\rho|_{t=t_{rec}} = K$ , por tanto tendremos que  $\dot{a} = 0$ .

-Si  $\dot{a} < 0$  inicialmente, entonces al final de la evolución tendremos que  $a \rightarrow 0$ .

-Si  $\dot{a} > 0$  inicialmente, entonces tendremos  $\dot{a} = 0$  en  $t = t_{rec}$  y para  $t > t_{rec}$  tendremos  $\dot{a} < 0$ , por lo que el universo comenzará a desacelerar, volviendo a contraerse y teniendo así un final del universo conocido como *Big Crunch*.

FIGURA 6.2: Evolución de  $a$  para distintos casos de  $K$ .

El factor de escala como función de  $t$  será,

- Para  $\omega = 0$  tendremos,

$$a(t) = \begin{cases} \frac{c}{2}(1 - \cos \phi(t)); & t = \frac{c}{2}(\phi(t) - \sin \phi(t)) \quad \text{si } K = +1 \\ & \qquad \qquad \qquad \left(\frac{9}{4}c\right)^{1/3} t^{2/3} \quad \text{si } K = 0 \\ \frac{c}{2}(\cosh \phi(t) - 1); & t = \frac{c}{2}(\sinh \phi(t) - \phi(t)) \quad \text{si } K = -1 \end{cases}$$

$$\text{con } c = \frac{8\pi G}{3}\rho a^2 = \text{cte.}$$

- Para  $\omega = 1/3$  tendremos,

$$a(t) = \begin{cases} \sqrt{\tilde{c}} \left[ 1 - \left( 1 - \frac{t}{\sqrt{\tilde{c}}} \right)^2 \right]^{1/2} & \text{si } K = +1 \\ (4\tilde{c})^{1/4} t^{1/2} & \text{si } K = 0 \\ \sqrt{\tilde{c}} \left[ \left( 1 + \frac{t}{\sqrt{\tilde{c}}} \right)^2 - 1 \right]^{1/2} & \text{si } K = -1 \end{cases}$$

$$\text{con } \tilde{c} = \frac{8\pi G}{3}\rho a^4 = \text{cte.}$$

- Para  $\omega = 1$ , si tenemos una constante cosmológica  $\Lambda < 0$ , por tanto  $\rho = -|\Lambda|$  y  $P = |\Lambda|$ . Solo podemos tener  $K = -1$ , por lo que

$$a(t) = \sqrt{\frac{3}{|\Lambda|}} \sin \left( \sqrt{\frac{|\Lambda|}{3}} t \right)$$

pero si tenemos una constante cosmológica  $\Lambda > 0$ , entonces sí podemos tener todos los valores de  $K$ , por tanto queda

$$a(t) = \begin{cases} \sqrt{\frac{3}{\Lambda}} \cosh\left(\sqrt{\frac{\Lambda}{3}}t\right) & \text{si } K = +1 \\ \sqrt{\frac{3}{\Lambda}} \exp\left(\sqrt{\frac{\Lambda}{3}}t\right) & \text{si } K = 0 \\ \sqrt{\frac{3}{\Lambda}} \sinh\left(\sqrt{\frac{\Lambda}{3}}t\right) & \text{si } K = -1 \end{cases}$$

### 6.1.3. Tensores de Killing

Los tensores de Killing solo se asocian con las geodésicas ( $K = 0$ ). Se denotan por  $K_{\mu\nu}$  y se cumple que  $\nabla_{(\rho}K_{\mu\nu)} = 0$ .

Para una geodésica  $T^\mu \nabla_\mu T^\nu = 0$ , tenemos que  $K^2 = T^\mu T^\nu K_{\mu\nu}$  es conservada sobre la geodésica. Para nuestra geometría tenemos el tensor de Killing siguiente,

$$K_{\mu\nu} = a^2(t)(u_\mu u_\nu + g_{\mu\nu})$$

con  $u^\mu = \delta_t^\mu$ . Si  $v^\mu$  es el vector tangente de una geodésica temporal, entonces  $v^\mu v^\nu g_{\mu\nu} = -1$  y  $v^\mu \nabla_\mu v^\nu = 0$ ; si  $v^\mu = v^t \delta_t^\mu + v^i \delta_i^\mu$ , entonces tenemos que

$$v^\mu v_\mu = -(v^t)^2 + v^i v^j g_{ij} = -(v^t)^2 + v^i v^j a^2 \gamma_{ij} = -(v^t)^2 + |\vec{v}|^2 = -1 \implies (v^t)^2 = 1 + |\vec{v}|^2$$

Luego,

$$K^2 = a^2(v_\mu v^\mu + (u_\mu v^\mu)^2) = a^2(-1 + (v^t)^2) = a^2 + |\vec{v}|^2$$

por tanto,  $|\vec{v}| = \frac{K}{a}$ , luego, por la propia expansión del universo, la velocidad de las partículas libres disminuye, debido a que la energía cinética de las partículas se va 'diluyendo'.

Si  $v^\mu v_\mu = 0$ , es decir, tenemos fotones, entonces  $K^2 = a^2(u_\mu v^\mu)^2$ , por tanto,  $u_\mu v^\mu = \frac{K}{a} = \omega$ . Por tanto, los fotones pierden su energía con la expansión del universo. Luego, para un observador comóvil, tendremos que  $K = \omega_0 a_0 = \omega_1 a_1$ .

Si  $a_1 > a_0$ , entonces  $\frac{\omega_0}{\omega_1} = \frac{a_1}{a_0} > 1$ , por tanto, la expansión del universo produce un desplazamiento al rojo de los fotones. Definimos el desplazamiento al rojo como,

$$z = \frac{\lambda_0 - \lambda_1}{\lambda_1} = \frac{a_0}{a_1} - 1$$

### 6.1.4. Ley de Hubble

Solamente se aplica a galaxias próximas a nosotros, con desplazamientos al rojo despreciables, sin considerar la curvatura del espacio-tiempo.

Los fotones siguen geodésicas nulas, tal que

$$-t^2 + \frac{a^2(t)\dot{r}^2}{1-Kr^2} = 0 \iff \int_{t_1}^{t_0} \frac{1}{a(t)} dt = \int_{r_1}^{r_0} \frac{1}{\sqrt{1-Kr^2}} dr$$

donde  $(t_0, r_0)$  y  $(t_1, r_1)$  son tiempos/posiciones de emisión y recepción.

Tomamos galaxias no muy lejanas, tal que

$$a(t_1) = a_0 + \dot{a}_0(t_0)(t_1 - t_0) + \frac{1}{2}\ddot{a}(t_0)(t_1 - t_0)^2 + \dots$$

recordando que  $z + 1 = \frac{a_0}{a_1}$  tenemos que

$$\frac{1}{1+z} = 1 + H_0(t_1 - t_0) - q_0 H_0^2(t_1 - t_0)^2 + \dots$$

donde  $H_0 = \frac{\dot{a}_0}{a_0}$  y  $q_0 = -\frac{a_0\ddot{a}_0}{\dot{a}_0^2}$  parámetro de desaceleración.

Si  $H_0(t_1 - t_0)$  es pequeño, entonces

$$t_0 - t_1 = H_0^{-1} \left[ z - \left( 1 + \frac{q_0}{2} \right) z^2 + \dots \right]$$

Además,

$$r_0 - r_1 = a_0^{-1} \left[ (t_0 - t_1) + \frac{1}{2} H_0 (t_0 - t_1)^2 + \dots \right] = \frac{1}{a_0 H_0} \left[ z - \frac{1}{2} (1 + q_0) z^2 + \dots \right]$$

No conocemos  $r_0 - r_1$ , pero sabemos  $L(W)$ , que es la luminosidad absoluta, y  $F(W/m^2)$ , que es el flujo de fotones que nos llega. Por tanto,

- en el espacio plano podemos definir  $4\pi d_L^2 = \frac{L}{F}$ , donde  $d_L$  es la distancia lumínica.
- en el espacio de Robinson-Walker no podemos definirlo. Los fotones emitidos y recibidos tienen  $\hbar\omega_0$  y  $\hbar\omega_1$ . Si fueron emitidos durante un tiempo  $\delta t_1$ , los recibimos

en  $\delta t_0 = \frac{a_0}{a_1} \delta t_1$ . Por tanto, la potencia recibida será,

$$\mathcal{P}_0 = \frac{\hbar\omega_0}{\delta t_0} = \frac{\hbar\omega_1}{\delta t_1} \frac{a_1^2}{a_0^2}$$

entonces,

$$F = \frac{\mathcal{P}_0}{A} = \frac{\hbar\omega_1}{\delta t_1} \frac{a_1^2}{a_0^2} \frac{1}{4\pi r^2 a_0^2} = L \frac{a_1^2}{a_0^2} \frac{1}{4\pi r^2 a_0^2}$$

Por tanto, la distancia lumínica del espacio curvo será,

$$d_L = a_0 r (1 + z)$$

Usando  $r = r_0 - r_1$ , obtenemos la **Ley de Hubble**, tal que

$$d_L = H_0^{-1} \left[ z + \frac{1}{2} (1 - q_0) z^2 + \dots \right]$$

El telescopio Planck nos ha dado que la constante de Hubble del fondo cósmico es  $H_0 = 68$  km/s/Mpc, mientras que el telescopio Webb-Hubble, nos da que la constante de Hubble para galaxias cercanas es  $H_0 = 73$  km/s/Mpc, cuya diferencia es entre  $4\sigma - 6\sigma$ . Se tienen diversas hipótesis para explicar esto, pero ninguna ha sido demostrada. Una de estas hipótesis es que la constante de Hubble no es realmente una constante, sino que depende de la densidad de energía del universo, pero esta densidad varía según las zonas del universo, pues no es homogéneo en todas las escalas, y entonces, nuestra zona del universo estaría más 'vacía' que la zona del fondo cósmico, por eso esta diferencia.

# Apéndice A

## Espacios topológicos

El principal interés de explicar los Espacios Topológicos surge del hecho de que el espacio-tiempo en Relatividad General tiene la estructura de un espacio topológico. En este apéndice recogemos varias definiciones y teoremas clave relativos a los espacios topológicos.

**Definición A.0.1.** *Un espacio topológico  $(X, \mathcal{T})$  se trata de un conjunto  $X$  con una colección  $\mathcal{T}$  de subconjuntos de  $X$  que satisface las siguientes propiedades:*

1. *La unión de un colección arbitraria finita o infinita de subespacios pertenecientes a  $\mathcal{T}$ , también pertenece a  $\mathcal{T}$ , es decir, si  $O_\alpha \in \mathcal{T}$  para todo  $\alpha$ , entonces  $\bigcup_\alpha O_\alpha \in \mathcal{T}$ .*
2. *La intersección de un número finito de subespacios de  $\mathcal{T}$  pertenece también a  $\mathcal{T}$ , es decir, si  $O_1, \dots, O_n \in \mathcal{T}$ , entonces  $\bigcap_{i=1}^n O_i \in \mathcal{T}$ .*
3. *El conjunto completo  $X$  y el conjunto vacío  $\emptyset$ , pertenecen a  $\mathcal{T}$ .*

**Nota A.0.1.**  *$\mathcal{T}$  es referido a una topología sobre  $X$ , y los subconjuntos de  $X$ , que se enumeran en la colección  $\mathcal{T}$ , se denominan conjuntos abiertos.*

Cualquier conjunto  $X$  se puede convertir fácilmente en un espacio topológico tomando  $\mathcal{T} = \{\text{todos los subconjuntos de } X\}$ , denominado *topología discreta*, o tomando  $\mathcal{T} = \{X, \emptyset\}$ , denominado *topología indiscreta*.

Un ejemplo mucho más interesante es el espacio topológico que se obtiene tomando  $\mathcal{T} = \mathbb{R}$ , el conjunto de los números reales, y definiendo  $\mathcal{T}$  para que esté formado por todos los subconjuntos de  $\mathbb{R}$ , que puede ser expresado como la unión de intervalos abiertos  $(a, b)$ . Así, tomando  $\mathcal{T}$  de esta forma sobre  $\mathbb{R}$ , un intervalo abierto es un conjunto

abierto; históricamente, este ejemplo es la razón por la que la terminología de 'conjunto abierto' se usa en la discusión de un espacio topológico abstracto.

También podemos definir topologías inducidas, pues la definición de espacio topológico nos permite jugar con los subconjuntos de subconjuntos.

**Definición A.0.2.** *Si  $(X, \mathcal{T})$  es un espacio topológico y  $A$  es un subconjunto cualquiera de  $X$ , podemos convertir  $A$  en un espacio topológico definiendo la topología  $\mathcal{S}$  en  $A$ , que consiste en todos los subconjuntos de  $A$  que pueden expresarse como intersecciones de elementos de  $\mathcal{T}$  con  $A$ , es decir,  $\mathcal{S} = \{U \mid U = A \cap O, O \in \mathcal{T}\}$ .  $\mathcal{S}$  se llama la '**topología inducida**' (o '**topología relativa**').*

En estos espacios topológicos también podemos definir un producto cartesiano, pues al trabajar con conjuntos está bien definido. De hecho, si  $(X_1, \mathcal{T}_1)$  y  $(X_2, \mathcal{T}_2)$  son espacios topológicos, entonces podemos introducir el producto cartesiano,

$$X_1 \times X_2 = \{(x_1, x_2) \mid x_1 \in X_1, x_2 \in X_2\}$$

dentro de un espacio topológico  $(X_1 \times X_2, \mathcal{T})$ , definiendo  $\mathcal{T}$  para que esté compuesto por todos los subconjuntos de  $X_1 \times X_2$ , que pueden ser expresados como uniones de la forma  $O_1 \times O_2$  con  $O_1 \in \mathcal{T}_1$  y  $O_2 \in \mathcal{T}_2$ .  $\mathcal{T}$  se denomina **producto topológico**, y usando esta definición de topología en  $\mathbb{R}$ , por construcción de topologías producto, podemos definir una topología en  $\mathbb{R}^n$ . La topología que obtenemos es la misma que se obtendría directamente definiendo  $\mathcal{T}$  para que esté formado por todos los subconjuntos de  $\mathbb{R}^n$ , que pueden ser expresados por uniones de bolas abiertas.

También podemos definir la continuidad de las funciones de los conjuntos de los espacios topológicos,

**Definición A.0.3.** *Sean  $(X, \mathcal{T})$  un espacio topológico con topología  $\mathcal{T}$ , y  $(Y, \mathcal{S})$  un espacio topológico con topología  $\mathcal{S}$ . Decimos que una función  $f : X \rightarrow Y$  es continua si para todo conjunto abierto  $O \in \mathcal{S}$  (es decir, cualquier conjunto abierto en  $Y$ ), la imagen inversa de  $O$  bajo  $f$ , denotada como*

$$f^{-1}[O] = \{x \in X \mid f(x) \in O\},$$

*es un conjunto abierto en  $X$  (es decir,  $f^{-1}[O] \in \mathcal{T}$ ).*

Para funciones de  $\mathbb{R}$  en  $\mathbb{R}$ , es fácil verificarlo usando la definición de topología sobre  $\mathbb{R}$ , esta definición de continuidad es equivalente a la definición usual  $\epsilon - \delta$ .

**Definición A.0.4.** *Si  $f$  es continua, inyectiva, sobreyectiva y su inversa es continua, entonces  $f$  se denomina **homeomorfismo**, y  $(X, \mathcal{T})$  y  $(Y, \mathcal{S})$  se dice que son **homeomorfos**. Los espacios topológicos homeomorfos tienen las mismas propiedades que los espacios topológicos.*

Antes hemos usado el concepto de 'conjunto abierto', pero también podemos definir los 'conjuntos cerrados', de forma que si  $(X, \mathcal{T})$  es un espacio topológico, un subconjunto  $C$  de  $X$  se dice que es *cerrado* si su complemento  $X - C \equiv \{x \in X \mid x \notin C\}$  es abierto. Así, por ejemplo, un intervalo cerrado  $[a, b]$  de  $\mathbb{R}$  (con la topología estándar sobre  $\mathbb{R}$ ) es un conjunto cerrado. A partir de los axiomas de los espacios topológicos, es inmediato ver que la intersección de cualquier colección arbitraria de conjuntos cerrados es cerrada y que la unión de un número finito de conjuntos cerrados es también cerrada. Cabe resaltar que un conjunto puede ser ni abierto y ni cerrado, por ejemplo, el intervalo medio-abierto  $[a, b)$  en  $\mathbb{R}$ ; o puede ser abierto y cerrado al mismo tiempo, como son todos los subconjuntos de la topología discreta. De hecho, la posibilidad de tener subconjuntos abiertos y cerrados a la vez da origen a la definición de **conectividad**,

**Definición A.0.5.** *Un espacio topológico  $(X, \mathcal{T})$  se dice que es **conexo** (o **conectado**) si el único subconjunto que es abierto y cerrado al mismo tiempo es el conjunto completo  $X$  y el conjunto vacío  $\emptyset$ .*

Tenemos que  $\mathbb{R}^n$ , con la topología estándar definida, es conexo.

En topología, uno de los conceptos fundamentales asociados a un subconjunto de un espacio es el de su adhesión. Este concepto permite formalizar la idea de los puntos donde el conjunto 'se acumula' dentro del espacio, incluyendo tanto los puntos que pertenecen al conjunto como aquellos que se encuentran arbitrariamente cerca de él. Formalmente, podemos definir la adhesión de la siguiente manera,

**Definición A.0.6.** *Sea  $(X, \mathcal{T})$  es un espacio topológico y  $A$  es un subconjunto arbitrario de  $X$ , la adhesión (o cierre),  $\overline{A}$ , de  $A$  se define como la intersección de todos los conjuntos cerrados que contienen a  $A$ .*

Claramente,  $\overline{A}$  es cerrado, contiene a  $A$ , y es igual a  $A$  si y solo si  $A$  es cerrado.

**Definición A.0.7.** *Sea  $(X, \mathcal{T})$  es un espacio topológico y  $A$  es un subconjunto arbitrario de  $X$ , el interior de  $A$  se define como la unión todos los conjuntos abiertos contenidos dentro de  $A$ .*

Claramente, el interior de  $A$  es abierto, está contenido en  $A$ , y es igual a  $A$  si y solo si  $A$  es abierto.

**Definición A.0.8.**  *$(X, \mathcal{T})$  es un espacio topológico y  $A$  es un subconjunto arbitrario de  $X$ , la frontera de  $A$ , denotada como  $\bar{A}$ , se define como todos los puntos que se encuentran en  $\bar{A}$ , pero no están en el interior de  $A$ .*

**Definición A.0.9.** *Sea  $(X, \mathcal{T})$  un espacio topológico, un entorno de  $x \in X$  es cualquier  $A \subset X$  tal que  $x$  pertenece al interior de  $A$ .*

En particular, cualquier conjunto abierto conteniendo  $x$  es un entorno de  $x$ .

**Definición A.0.10.** *Una base de entornos en  $x$  es una colección de entornos de  $x$  tal que todo entorno de  $x$  contiene a algún elemento de esta colección.*

En particular, la colección de todos los conjuntos abiertos conteniendo  $x$  es una base de entornos sobre  $x$ , aunque generalmente hay muchas otras posibilidades para bases de entornos.

**Definición A.0.11.** *Una base de entornos de  $X$  es una especificación de una base de entornos para cada  $x \in X$*

Las topologías suelen definirse especificando una base de entornos. El procedimiento es el siguiente:

1. Un entorno de  $x$  es cualquier conjunto que contiene a algún entorno base de  $x$ .
2. Un conjunto es abierto si es un entorno de cada uno de sus puntos.

Este proceso asegura que todos los conjuntos abiertos se generan a partir de la base de entornos, cumpliendo con las propiedades necesarias para definir una topología en el espacio. También es interesante ver que los conjuntos cerrados, los puntos de adherencia y de frontera pueden definirse directamente en términos de la base de entornos,

**Definición A.0.12.** Un conjunto  $G$  es cerrado si y solo si, para cada punto  $x$  que no pertenece a  $G$ , existe una base de entornos de  $x$  que no interseca a  $G$ . La adherencia de un conjunto  $A$  consiste en aquellos puntos  $x$  tales que cada base de entornos de  $x$  interseca a  $A$ . La frontera de  $A$  consiste en aquellos puntos  $x$  tales que cada base de entornos de  $x$  interseca tanto a  $A$  como a  $X - A$ .

## Espacios métricos

Las bases de entornos, y por tanto las topologías, frecuentemente se definen en términos de una *métrica* o *función distancia*, que es la función  $d : X \times X \rightarrow \mathbb{R}$ , que verifica:

1. Para todo  $x, y \in X$ ,  $d(x, y) \geq 0$  (positividad).
2. Si  $d(x, y) = 0$ , entonces  $x = y$  (no degeneración).
3. Para todo  $x, y \in X$ ,  $d(x, y) = d(y, x)$  (simetría).
4. Para todo  $x, y, z \in X$ ,  $d(x, y) + d(y, z) \geq d(x, z)$  (desigualdad triangular).

No hay ningún cambio esencial si también añadimos  $+\infty$  como un valor de  $d$ . Un conjunto con función métrica se denomina **espacio métrico**.

**Definición A.0.13.** La bola abierta con centro en  $x$  y radio  $r > 0$  con respecto a  $d$  se define como

$$B(x, r) = \{y \mid d(x, y) < r\}$$

Entonces, puede demostrarse que cada bola abierta serviría como base de entornos para una topología  $X$ , la *topología métrica* de  $d$ .

De forma más general: **Para cualquier espacio métrico, la colección de todos los subconjuntos que pueden expresarse como uniones de bolas abiertas define una topología.**

## Espacio Hausdorff

Vamos a introducir los espacios Hausdorff, pues nos permiten definir las variedades y son muy útiles en Geometría Diferencial.

**Definición A.0.14.** *Un espacio topológico  $(X, \mathcal{T})$  se dice que es Hausdorff si para cada par de puntos distintos  $p, q \in X$ ,  $p \neq q$ , existen conjuntos abiertos  $O_p, O_q \in \mathcal{T}$  tal que  $p \in O_p$ ,  $q \in O_q$ , y  $O_p \cap O_q = \emptyset$ .*

Es fácil comprobar que  $\mathbb{R}^n$ , con la topología estándar, es Hausdorff. También tenemos que una topología métrica siempre es Hausdorff.

## Compacidad

Una de las nociones más importantes en topología es la de la *compacidad*, que se define como,

**Definición A.0.15.** *Sea  $(X, \mathcal{T})$  un espacio topológico y  $A$  un subconjunto de  $X$ , una colección  $\{O_\alpha\}$  de conjuntos abiertos se denomina recubrimiento abierto de  $A$  si la unión de estos conjuntos contiene a  $A$ , i.e.  $A \subseteq \bigcup_\alpha O_\alpha$ . Una subcolección de estos conjuntos  $\{O_\alpha\}$  que también recubren  $A$  es referida como un sub-recubrimiento. El conjunto  $A$  se denomina compacto si cada recubrimiento abierto de  $A$  tiene un sub-recubrimiento finito (i.e., un sub-recubrimiento que consiste solo en un número finito de conjuntos).*

Así, por ejemplo, en cualquier espacio topológico, un conjunto formado por un solo punto, es compacto. Por otro lado, el intervalo abierto  $(0, 1)$  en  $\mathbb{R}$  (con la topología estándar) no es compacto ya que los conjuntos  $O_n = (1/n, 1)$  para  $n = 2, 3, \dots$ , ceden un recubrimiento abierto de  $(0, 1)$  el cuál admite sub-recubrimientos no finitos.

Los siguientes teoremas describen las implicaciones de la compacidad y muestran la utilidad de esta noción. Las demostraciones pueden encontrarse en cualquier texto de topología (e.g., Hocking and Young 1961; Kelley 1955).

Quizás, el teorema más importante relativo a subconjuntos compactos de  $\mathbb{R}$  es el Teorema de Heine-Borel,

**Teorema A.0.16 (Heine-Borel).** *Un intervalo cerrado  $[a, b]$  de números reales es compacto (con la topología estándar sobre  $\mathbb{R}$ ).*

La relación general entre conjuntos compactos y cerrados se describe con los siguientes dos teoremas, las demostraciones son directas,

**Teorema A.0.17.** *Sea  $(X, \mathcal{T})$  un espacio topológico de Hausdorff y sea  $A \subset X$  compacto. Entonces  $A$  es cerrado.*

**Teorema A.0.18.** *Sea  $(X, \mathcal{T})$  compacto y  $A \subset X$  cerrado. Entonces  $A$  es compacto.*

Combinando los tres teoremas anteriores, llegamos al siguiente enunciado importante sobre la compacidad de subconjuntos de  $\mathbb{R}$ ,

**Teorema A.0.19.** *Un subconjunto  $A$  de los números reales es compacto si y solo si es cerrado y acotado.*

Se demuestra fácilmente que la propiedad de compacidad se conserva en mapas continuos. Tenemos,

**Teorema A.0.20.** *Una función continua de un espacio topológico compacto en  $\mathbb{R}$  es acotada y alcanza sus valores máximo y mínimo.*

El siguiente teorema proporciona una extensión inmediata de los resultados de compacidad de  $\mathbb{R}$  para  $\mathbb{R}^n$ .

**Teorema A.0.21 (Teorema de Tychonoff).** *Sea  $(X_1, \mathcal{T}_1)$  y  $(X_2, \mathcal{T}_2)$  espacios topológicos compactos. Entonces, el producto cartesiano  $X_1 \times X_2$  es compacto en el producto topológico.*

**Teorema A.0.22.** *El teorema anterior puede generalizarse para aplicar el producto de infinitos productos topológicos, pero el axioma de elección es necesario para esta generalización.*

Un corolario de este teorema y el anterior es

**Corolario A.0.2.** *Un subconjunto,  $A$ , de  $\mathbb{R}^n$  es compacto si y solo si es cerrado y acotado.*

Así, por ejemplo, la esfera  $n$ -dimensional  $S^n$  (definida como el conjunto de puntos en  $\mathbb{R}^{n+1}$  satisfaciendo  $x_1^2 + \dots + x_n^2 + 1^2 = 1$ ) en la topología inducida es compacta, así que es fácil ver que es un conjunto cerrado y acotado de  $\mathbb{R}^{n+1}$ .

## Convergencia de sucesiones

Otra noción que necesitaremos es la de convergencia de sucesiones.

**Definición A.0.23.** Una sucesión  $\{x_n\}$  de puntos en un espacio topológico  $(X, \mathcal{T})$  se dice que converge a un punto  $x$  si en cualquier entorno abierto  $O$  de  $x$  (i.e., un conjunto abierto  $O$  que contiene  $x$ ), hay un  $N$  tal que  $x_n \in O$ , para todo  $n > N$ .

El punto  $x$  se dice que es el *límite* de la sucesión. Es fácil comprobar que para  $\mathbb{R}$  (con la topología estándar) esto lleva a la definición usual de convergencia.

**Definición A.0.24.** Un punto  $y \in X$  se dice que es un punto de acumulación (o punto límite) de  $\{x_n\}$  si cada entorno abierto de  $y$  contiene infinitos números de la sucesión.

Sin embargo, en un espacio topológico general, si  $y$  es un punto de acumulación de  $\{x_n\}$ , podría no ser posible encontrar una sucesión  $\{y_n\}$  de puntos de la sucesión  $\{x_n\}$  tal que  $\{y_n\}$  converge hacia  $y$ . No obstante, el sentido de la convergencia de sucesiones hacia  $y$  siempre será posible si  $(X, \mathcal{T})$  es *primero numerable*, esto es, si para cada  $p \in X$  hay una colección numerable  $\{O_n\}$  de conjuntos abiertos tal que cada entorno abierto,  $O$ , de  $p$  contiene al menos un miembro de esta colección. Para  $\mathbb{R}^n$ , las bolas abiertas con radio racional centradas en puntos con coordenadas racionales, componen una colección numerable de conjuntos abiertos.

Una relación importante entre compacidad y convergencia de sucesiones está expresada en el Teorema de Bolzano-Weiestrass,

**Teorema A.0.25 (Teorema de Bolzano-Weiestrass).** Sea  $(X, \mathcal{T})$  un espacio topológico y sea  $A \subset X$ . Si  $A$  es compacto, entonces cada sucesión  $\{x_n\}$  de puntos de  $A$  tiene un punto de acumulación en  $A$ . Inversamente, si  $(X, \mathcal{T})$  es segundo numerable y cada sucesión en  $A$  tiene un punto de acumulación en  $A$ , entonces  $A$  es compacto. Así, en particular, si  $(X, \mathcal{T})$  es segundo numerable,  $A$  es compacto si y solo si cada sucesión en  $A$  tiene una convergencia de sucesiones cuyo límite está en  $A$ .

## Para-compacidad

Finalmente, definimos la noción de *para-compacidad*, una propiedad que las variedades deben satisfacer para evitar que sean "demasiado grandes".

**Definición A.0.26.** -Sea  $(X, \mathcal{T})$  un espacio topológico y sea  $\{O_\alpha\}$  un recubrimiento abierto de  $X$ . Un recubrimiento abierto  $\{V_\beta\}$  se dice que es un refinamiento de  $\{O_\alpha\}$  si para cada  $V_\beta$  existe un  $O_\alpha$  tal que  $V_\beta \subset O_\alpha$ .

-El recubrimiento  $\{V_\beta\}$  se dice que es localmente finito si cada  $x \in X$  tiene un entorno abierto  $W$  tal que solo un número finito de  $V_\beta$  satisfacen que  $W \cap V_\beta \neq \emptyset$ .

-El espacio topológico  $(X, \mathcal{T})$  se dice que es para-compacto si cada recubrimiento abierto  $\{O_\alpha\}$  de  $X$  tiene un recubrimiento localmente finito  $\{V_\beta\}$ .

Los conceptos de variedades se explican en el Apéndice B.

No es difícil ver que (verlo en e.g., Hocking and Young 1961) que cualquier espacio topológico de Hausdorff que sea localmente compacto (i.e., tal que cada punto tiene un entorno abierto con adherencia compacta) y que pueda ser expresado como una unión numerable de subconjuntos, es para-compacto. Así,  $\mathbb{R}^n$ ,  $S^m$  y sus productos verifican fácilmente ser para-compactos. En efecto, no es fácil construir ejemplos de espacios topológicos que satisfagan todos los requisitos de una variedad pero que no sean para-compactos; la 'recta larga' (ver Hocking and Young 1961) es quizás el ejemplo más simple, aunque para definirla se requiere el axioma de elección.

# Apéndice B

## Variedades

### B.1. Definición de Variedad

Una variedad, aproximadamente, es un espacio topológico en el cuál algunos entornos de cada punto admiten un sistema de coordenadas, siendo en funciones coordenadas reales en los puntos del entorno, que determina la posición de los puntos y la topología de ese entorno; esto es, el espacio es localmente cartesiano. Además, el paso de un sistema de coordenadas a otro es suave en la región superpuesta, por tanto, el significado de curva, función o mapa 'diferenciable' es consistente cuando se refiere a cualquier sistema. Una definición detallada se dará más adelante.

Los modelos matemáticos que se usan para describir sistemas físicos, usan las variedades como objeto básico de estudio, sobre el cual se puede definir una *estructura* adicional para obtener cualquier sistema en cuestión. El concepto generaliza e incluye los casos especiales de línea cartesiana, plano, espacio, y las superficies que se estudian en cálculo avanzado. La teoría de estos espacios se generaliza con las variedades, incluyendo las ideas de funciones diferenciables, curvas suaves, vectores tangentes, y campos vectoriales. Sin embargo, las nociones de distancia entre puntos y líneas rectas (o caminos más cortos) no son parte de la idea de una variedad, pero surgen como consecuencias de estructuras adicionales, las cuales pueden ser o no ser asumidas y en cualquier caso, no es único.

Una variedad tiene dimensión tal que, al igual que en los modelos físicos, su valor será igual al número de grados de libertad. Nos limitaremos a estudiar las variedades de dimensión finita.

Algunas definiciones preliminares podrán facilitar la definición de variedad:

**Definición B.1.1 (Carta).** *Sea  $(X, \mathcal{T})$  un espacio topológico, siendo  $\mathcal{T}$  una topología sobre el conjunto  $X$ , decimos que una **carta** en  $p \in X$  es una función  $\mu : O \rightarrow \mathbb{R}^d$ , donde  $O$  es un conjunto abierto que contiene  $p$  y  $\mu$  es un homeomorfismo sobre un conjunto abierto de  $\mathbb{R}^d$ .*

- La *dimensión* de una carta  $\mu : O \rightarrow \mathbb{R}^d$  es  $d$ .
- Las *funciones de coordenadas* de una carta son las funciones de valor real sobre  $O$ , dadas por las entradas de valores de  $\mu$ ; esto es, son las funciones  $x^i = u^i \circ \mu : O \rightarrow \mathbb{R}$ , donde  $u^i : \mathbb{R}^d \rightarrow \mathbb{R}$  son las coordenadas estándar sobre  $\mathbb{R}^d$ .
- Los  $u^i$  están definidos por  $u^i(a^1, \dots, a^d) = a^i$ . Los superíndices no son potencias, claro está, pero son simplemente la indexación tensorial habitual de coordenadas. Si se necesitan potencias, usaremos paréntesis extra,  $(x)^2$  en lugar de  $x^2$  para el cuadrado de  $x$ , pero normalmente el contexto tendrá información suficiente para saber la distinción y quitar el uso de estos paréntesis. Por tanto, para cada  $q \in O$ , tenemos que  $\mu q = (x^1 q, \dots, x^d q)$ , y entonces, podremos escribir también  $\mu = (x^1, \dots, x^d)$ . En otra terminología, podemos denominar  $\mu$  como un *mapa coordenado*,  $O$  como un *entorno coordenado*, y la colección  $(x^1, \dots, x^d)$  como *coordenadas o sistema coordenado sobre  $p$* .

Restringiremos los símbolos ' $u^i$ ' a este uso como coordenadas estándar de  $\mathbb{R}^d$ . Para  $\mathbb{R}^2$  y  $\mathbb{R}^3$  también usaremos  $x, y, z$  como coordenadas, como es habitual, exceptuando que las usaremos como funciones.

**Definición B.1.2.** -*Una función de valor real  $f : V \rightarrow \mathbb{R}$  es  $\mathcal{C}^\infty$  (continua a orden  $\infty$ ) si  $V$  es un conjunto abierto en  $\mathbb{R}^d$  y  $f$  tiene derivadas parciales continuas para todos los órdenes y tipos (cruzadas o no).*

-*Una función  $\varphi : V \rightarrow \mathbb{R}^e$  es un **mapa**  $\mathcal{C}^\infty$  si las componentes de  $u^i \circ \varphi : V \rightarrow \mathbb{R}$  son  $\mathcal{C}^\infty$ ,  $i = 1, \dots, e$ .*

De forma más general,  $\varphi$  es  $\mathcal{C}^k$ , siendo  $k$  un entero no negativo, si todas las derivadas parciales, hasta e incluyendo estas de orden  $k$ , existen y son continuas.  $\mathcal{C}^0$  implica simplemente continuidad. Un mapa  $\varphi$  es *analítico* si  $u^i \circ \varphi$  es real-analítico, esto es, puede ser expresado en un entorno para cada punto por medio de una serie de potencias convergente en coordenadas cartesianas, teniendo su origen en el punto. Los mapas analíticos son  $\mathcal{C}^\infty$ , pero no al revés.

**Ejemplo B.1.1.** Sea  $z = x + iy$ , una variable compleja, definimos  $u(x, y)$  como  $u + iv = e^{-1/z^4}$ ,  $u(0, 0) = 0$ . Entonces, vemos que  $u$  no es  $\mathcal{C}^\infty$ , y de hecho, no es continuo en el  $(0, 0)$ , pero las derivadas parciales de  $u$  existen en todos los órdenes y en todos los puntos, incluyendo el  $(0, 0)$ . De este modo, los requisitos para continuidad en la definición de  $\mathcal{C}^\infty$  no son triviales. Para funciones de una variable, es cierto que las funciones diferenciables son continuas.

Dos cartas  $\mu : U \rightarrow \mathbb{R}^d$  y  $\tau : V \rightarrow \mathbb{R}^e$  en un espacio topológico  $(X, \mathcal{T})$  son  $\mathcal{C}^\infty$ -relacionadas si  $d = e$  y cualquier  $U \cap V = \emptyset$  o  $\mu \circ \tau^{-1}$  y  $\tau \circ \mu^{-1}$  son mapas  $\mathcal{C}^\infty$ . El dominio de  $\mu \circ \tau^{-1}$  es  $\tau(U \cap V)$ , un conjunto abierto es  $\mathbb{R}^d$ . Ver Figura B.1

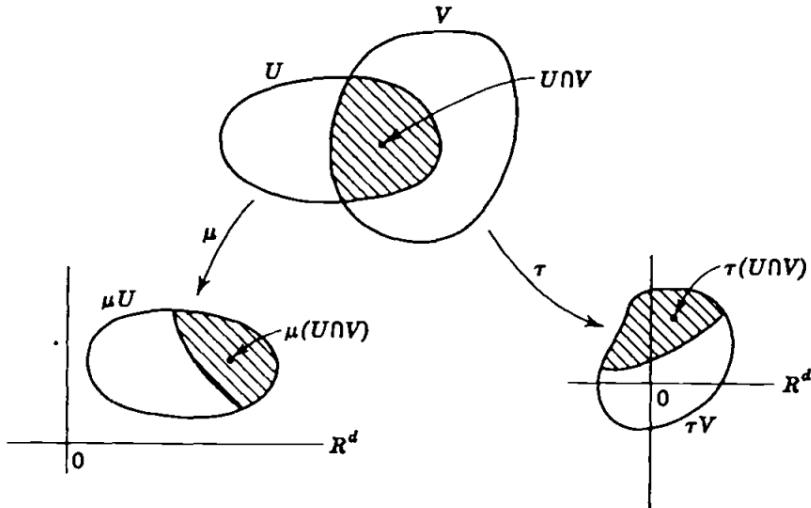


FIGURA B.1: Interpretación de ambas cartas. (donde  $R^d \equiv \mathbb{R}^d$ ) [6.1.4]

Otros grados de relación se definen reemplazando ' $\mathcal{C}^\infty$ ' por ' $\mathcal{C}^k$ ' o por 'analítico'. Dos cartas de la misma dimensión siempre son  $\mathcal{C}^0$ -relacionadas porque los mapas de coordenadas son continuos.

**Definición B.1.3.** Una variedad topológica  $\mathcal{C}^0$  es un espacio separable de Hausdorff tal que haya una carta de dimensión  $d$  en cada punto.

La dimensión de la variedad es la misma que la de las cartas. De esta forma, existe una colección de cartas  $\{\mu_\alpha : U_\alpha \rightarrow \mathbb{R}^\alpha \mid \alpha \in I\}$ , tal que  $\{U_\alpha \mid \alpha \in I\}$  es una cubierta del espacio. Dicha colección se denomina un *atlas*. Un atlas  $\mathcal{C}^\infty$  es aquel para el cual cada par de cartas es  $\mathcal{C}^\infty$ -relacionada. Una carta es *admisible* para un atlas  $\mathcal{C}^\infty$  si es  $\mathcal{C}^\infty$ -relacionado para cada carta en el atlas. En particular, los miembros de un atlas  $\mathcal{C}^\infty$  son admisibles.

**Definición B.1.4.** Una variedad  $\mathcal{C}^\infty$  es una variedad topológica junto con todas las cartas admisibles con algún atlas  $\mathcal{C}^\infty$ .

Nos referiremos a las 'variedades  $\mathcal{C}^\infty$ ' como 'variedades'. La razón de incluir todas las cartas admisibles en vez de simplemente aquellas que están en algún atlas dado es para transmitir la idea que ningún sistema de coordenadas particular es preferible sobre otro y también para resolver el problema lógico de decir qué *es* una variedad. El origen de este problema lógico es el hecho de que dos atlases diferentes pueden tener la misma colección de cartas admisibles, en cuyo caso, nos gustaría decir que tenemos una sola variedad, no dos variedades diferentes para cada atlas. Por otro lado, casi invariablemente ocurre que una variedad es especificada dando solo un atlas, no toda la colección de cartas admisibles.

**Definición B.1.5.** *Las variedades  $\mathcal{C}^\infty$  y las variedades analíticas reales se definen reemplazando ' $\mathcal{C}^\infty$ ' por ' $\mathcal{C}^k$ ' y 'analítica', respectivamente a lo largo de la cadena de definiciones anterior.*

Debe quedar claro que una variedad  $\mathcal{C}^\infty$  se convierte en una variedad  $\mathcal{C}^k$  simplemente ampliando la colección de cartas admisibles para incluir todas las relacionadas con  $\mathcal{C}^k$ , y, de forma similar, una variedad analítica real se convierte en una variedad  $\mathcal{C}^\infty$ . En cambio, una variedad  $\mathcal{C}^1$  se convierte en una variedad analítica real (y por lo tanto  $\mathcal{C}^\infty$ ), de muchas formas, descartando una colección adecuada de cartas admisibles de  $\mathcal{C}^1$  para dejar solo las cartas que están mutuamente relacionados analíticamente, pero este resultado no es del todo obvio, siendo un teorema muy difícil de Whitney. Se sabe que una variedad  $\mathcal{C}^0$  puede no llegar a convertirse en una variedad  $\mathcal{C}^1$ , y es aún más difícil de demostrar.

**Observación B.1.2.** *En la definición de sistema de coordenadas requerimos que el entorno coordenado y el rango en  $\mathbb{R}^d$  sean conjuntos abiertos. Esto es contrario al uso popular, o al menos más específico que el uso de coordenadas curvilíneas en cálculo avanzado.*

Por ejemplo, las coordenadas esféricas se utilizan incluso a lo largo de puntos del eje  $z$  donde ni siquiera son 1 a 1. Las razones de la restricción a conjuntos abiertos son que fuerzan uniformidad en la estructura local que simplifica el análisis en variedades (no hay 'puntos de frontera') e, incluso si la uniformidad local fuera forzada en otros aspectos, esto evita el problema de explicar qué entendemos por diferenciabilidad en puntos de frontera del entorno coordenado; esto es, las derivadas laterales no necesitan ser mencionadas. Por otro lado, en aplicaciones, con frecuencia surgen problemas de valores de frontera, cuya configuración es una *variedad con borde*. Estos espacios son más generales que las variedades y la generalidad extra surge de permitir una *variedad de borde* en una dimensión inferior. Los puntos de la variedad de borde tienen un entorno coordinado en la variedad

de borde que está unido a un entorno coordenado del interior de manera muy similar a como una cara de un cubo está unida al interior.

## B.2. Ejemplos de Variedades

### B.2.1. Espacio Cartesiano

Definimos una estructura de variedad en  $\mathbb{R}^d$ , de la manera más obvia, tomando como un atlas la única carta  $\mathbb{I} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , el mapa identidad. Las funciones coordenadas de esta carta son por lo tanto las coordenadas estándar (cartesianas)  $u^i$ . Cuando hablamos de  $\mathbb{R}^d$  como una variedad, nuestra intención es esta estructura estándar, a menos que se indique lo contrario.

Un mapa coordenado admisible  $\mathcal{C}^\infty$  sobre  $\mathbb{R}^d$  es un mapa  $\mathcal{C}^\infty$  1 a 1  $\mu : U \rightarrow \mathbb{R}^d$ , donde  $U$  es un conjunto abierto y el determinante del jacobiano es  $|\partial x^i / \partial u^i| \neq 0$ , donde  $x^i = u^i \circ \mu$  son las funciones coordenadas. Que el determinante del jacobiano no sea nulo solo es otra forma de requerir que el mapa  $\mu^{-1}$  sea  $\mathcal{C}^\infty$ .

Si  $f^i$ ,  $i = 1, 2, \dots, d$ , son funciones reales  $\mathcal{C}^\infty$  en algún conjunto abierto de  $\mathbb{R}^d$  y para algún  $p \in \mathbb{R}^d$  tenemos  $|\partial f^i / \partial u^i| \neq 0$ , entonces el teorema de la *función inversa* establece que hay un entorno  $U$  de  $p$  y un entorno  $V$  de  $(f^1 p, f^2 p, \dots, f^d p)$  tal que el mapa  $\mu = (f^1, f^2, \dots, f^d)$  lleva  $U$  a  $V$ , es 1 a 1, y tiene una inversa  $\mathcal{C}^\infty$ . Esto proporciona un medio eficaz de obtener coordenadas admisibles. En particular, coordenadas polares, coordenadas cilíndricas, coordenadas esféricas, y cualquier otras coordenadas curvilíneas personalizadas son coordenadas admisibles para  $\mathbb{R}^2$  y  $\mathbb{R}^3$  siempre que estén adecuadamente restringidas para ser 1 a 1 y tener un determinante jacobiano distinto de cero.

### B.2.2. Subvariedad Abierta

Si  $M$  es una variedad y  $N$  es cualquier conjunto abierto de  $M$ , entonces  $N$  hereda una estructura de variedad restringiendo la topología y mapas coordinados de  $M$  a  $N$ . Llamamos a  $N$  una *subvariedad abierta* de  $M$ . En particular, cualquier subconjunto de  $\mathbb{R}^d$  es una variedad de dimensión  $d$ .

### B.2.3. Producto de Variedades

Si  $M$  y  $N$  son variedades de dimensión  $d$  y  $e$  respectivamente, entonces a  $M \times N$  se le da una estructura de variedad tomando el producto topológico como su topología (básicamente, los entornos son productos de los entornos de  $M$  y  $N$ ) y como atlas se toma el producto de cartas de los atlases de  $M$  y  $N$ . Si  $\mu : U \rightarrow \mathbb{R}^d$  es una carta sobre  $M$ , y  $\varphi : V \rightarrow \mathbb{R}^e$  es una carta sobre  $N$ , su producto es  $(\mu, \varphi) : U \times V \rightarrow \mathbb{R}^{d+e}$ , que se define por  $(\mu, \varphi)(m, n) = (\mu m, \varphi n)$ . Si  $x^i$  son las funciones coordenadas de  $\mu$  e  $y^i$  son las funciones coordenadas de  $\varphi$ , entonces las coordenadas de  $(m, n)$  en el producto de cartas son  $(x^1 m, \dots, x^d m, y^1 n, \dots, y^e n)$ . Así, si  $p : M \times N \rightarrow M$  y  $q : M \times N \rightarrow N$  son las proyecciones,  $p(m, n) = m$ ,  $q(m, n) = n$ , las funciones coordenadas sobre  $U \times V$  son  $z^1 = x^1 \circ p, \dots, z^d = x^d \circ p, z^{d+1} = y^1 \circ q, \dots, z^{d+e} = y^e \circ q$ .

Esta operación de producto puede ser iterada evidentemente, y podemos tomar diferentes copias de la misma variedad como elementos. Así, incluso podemos descomponer una variedad  $\mathbb{R}^d = \mathbb{R} \overset{d-\text{veces}}{\times} \mathbb{R}$  ( $d$ -elementos). Es fácil ver que un círculo  $S^1$  (la curva) es una variedad unidimensional. Dibujando  $S^1$  como una parte de  $\mathbb{R}^2$ , vemos que un cilindro (la superficie) es la variedad  $S^1 \times \mathbb{R}$  y puede dibujarse en  $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$ .

Podemos considerar  $S^1 \times S^1$  como una unión,  $\{\{p\} \times S^1 | p \in S^1\}$ , de círculos  $\{p\} \times S^1$ , para cada  $p \in S^1$ . Ahora, si dibujamos el primer elemento como si estuviera en el plano  $XY$  de  $\mathbb{R}^3$ , satisfaciendo las ecuaciones  $x^2 + y^2 = 1$ ,  $z = 0$ , y para cada  $p$  en el primer elemento, representamos  $\{p\} \times S^1$  como un círculo más pequeño con centro en  $p$  y diámetros perpendiculares al primer círculo en  $p$ , entonces la unión  $S^1 \times S^1$  es la superficie de revolución del círculo pequeño sobre el eje  $Z$ , un toroide (Ver Figura B.2).

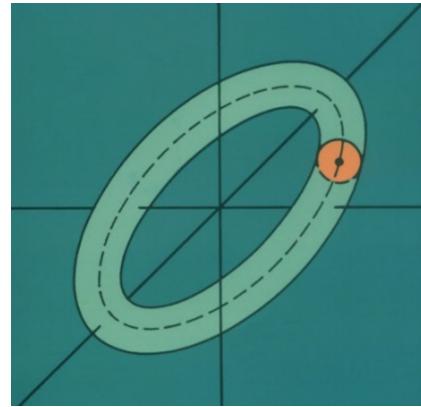
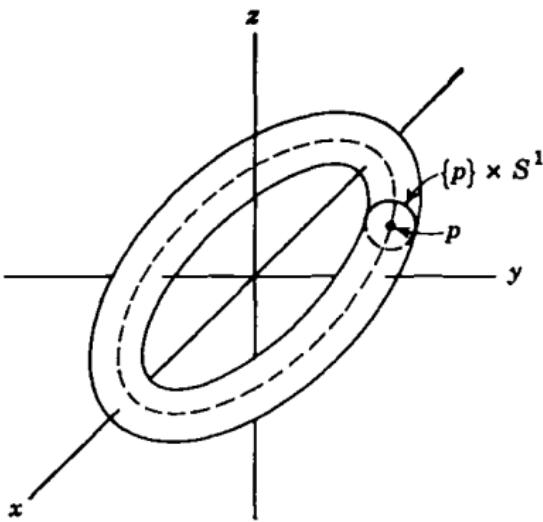


FIGURA B.3: Portada de [1]

FIGURA B.2: Representación de un toroide. [6.1.4]

No es difícil ver que la topología inducida de  $\mathbb{R}^3$  en el toroide es el producto topológico.

El toroide es la variedad subyacente que modela el conjunto de posiciones (el *espacio de configuraciones*) de un péndulo doble. Estamos pensando en un sistema mecánico que consiste en dos varillas, la primera que es libre de rotar en un plano sobre un eje fijo, y la segunda que rota sobre un eje en el plano que se fija relativo a la primera varilla, usualmente, pero no necesariamente es el plano de la primera varilla. Los ángulos que estas varillas hacen con un eje de coordenadas en sus planos pueden ser identificados con los ángulos  $u, v$  que se encuentran en la parametrización del toroide dado a continuación, dando una correspondencia 1 a 1 entre las posiciones del péndulo doble y del toroide. La articulación debe estar fija de modo que cada varilla es libre de hacer un giro completo sobre su eje, o solo una parte del toroide en este modelo. De hecho, si se bloquea la segunda varilla sobre el eje de la primera, de forma que  $v$  se restringe a  $0 < \epsilon < v < 2\pi$ , entonces el modelo es de un cilindro en vez de un toroide.

Añadiendo más varillas obtenemos sistemas físicos para los cuáles, el modelo es el producto de las copias de  $S^1$ . Si la articulación está fija de forma que la varilla es libre de moverse en el espacio en vez de en un plano, entonces serán necesarios algunos elementos de  $S^2$ . Finalmente, si un extremo de la primera varilla no está fijo del todo, pero permite moverse libremente en el espacio (o en un plano), entonces se necesitarán un elemento de  $\mathbb{R}^3$  (o de  $\mathbb{R}^2$ ).

De forma más general, si un sistema físico está compuesto por dos sistemas, para los cuales se pueden asumir todas sus posiciones independientemente del otro, entonces el sistema compuesto tiene como sus variedades de posición el producto de las variedades de posición de las dos componentes del sistema. Esto es así a pesar de que hay alguna ligadura dinámica (por ejemplo, gravitacional o elástica) entre las componentes.

Para ver más ejemplos ver la referencia [1, Chapter 1, pages 26-35]

### B.3. Mapas diferenciables

Si  $F : M \rightarrow N$ , donde  $M$  y  $N$  son variedades  $\mathcal{C}^\infty$ , entonces llamamos a  $F$  un mapa  $\mathcal{C}^\infty$  si la expresión de coordenadas para  $F$  consiste en mapas  $\mathcal{C}^\infty$  en espacios cartesianos. Ahora

elaboramos esta afirmación en una definición completa, en particular dejando claro qué significa 'expresión de coordenadas'.

**Definición B.3.1.** Sea  $\mu_1 : U \rightarrow \mathbb{R}^d$  y  $\mu_2 : V \rightarrow \mathbb{R}^e$  cartas  $\mathcal{C}^\infty$  sobre  $M$  y  $N$  variedades  $\mathcal{C}^\infty$ , de forma que  $U$  y  $V$  son subconjuntos abiertos de  $M$  y  $N$  respectivamente. Asumimos que  $F : M \rightarrow N$  es un mapa continuo, tal que  $W = F^{-1}V$  es un subconjunto abierto de  $M$ , ver Figura B.4.

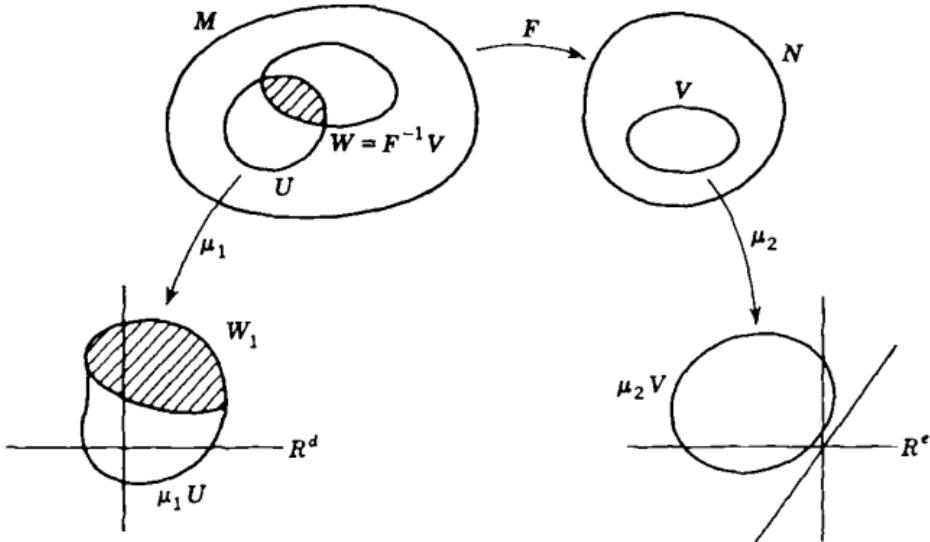


FIGURA B.4: Representación de los subconjuntos. [6.1.4]

Sea  $W_1 = \mu_1 W$ , tal que  $W_1$  es un conjunto abierto en  $\mathbb{R}^d$ . La expresión de coordenadas  $\mu_1 - \mu_2$  para  $F$  es el mapa  $\mu_2 \circ F \circ \mu_1^{-1} : W_1 \rightarrow \mathbb{R}^e$ . El mapa  $F$  es  $\mathcal{C}^\infty$  si todas las expresiones de coordenadas, para todas las cartas admisibles  $\mu_1, \mu_2$ , son mapas cartesianos  $\mathcal{C}^\infty$ .

**Proposición B.3.1.** Un mapa  $F : M \rightarrow N$  es  $\mathcal{C}^\infty$  si las expresiones de coordenadas  $\mu_\alpha - \mu_\beta$  de  $F$  son  $\mathcal{C}^\infty$  para aquellos  $\mu_\alpha$  en algún atlas de  $M$  y aquellos  $\mu_\beta$  en algún atlas de  $N$ .

**Demostración.** Sean  $\{\mu_\alpha : U_\alpha \rightarrow \mathbb{R}^d | \alpha \in I\}$  y  $\{\mu_\beta : V_\beta \rightarrow \mathbb{R}^e | \beta \in J\}$  los atlases de  $M$  y  $N$ , respectivamente, tal que para cada  $\alpha \in I$ ,  $\beta \in J$ ,  $\mu_\beta \circ F \circ \mu_\alpha^{-1}$  es un mapa  $\mathcal{C}^\infty$ . Supongamos que  $\mu_1, \mu_2$  son cualquier otras cartas como en la definición, así  $\mu_2 \circ F \circ \mu_1^{-1} : W_1 \rightarrow \mathbb{R}^e$ . Debemos demostrar que esto es  $\mathcal{C}^\infty$ , pero como ser  $\mathcal{C}^\infty$  es una propiedad local, basta con demostrarlo en un entorno de cada punto de  $W_1$ . Si  $m_1 \in W_1$ , entonces hay un  $\alpha \in I$  y un  $\beta \in J$  tal que  $\mu_1^{-1}m_1 = m \in U_\alpha$  y  $n = Fm \in V_\beta$ . Por hipótesis,  $\mu_\beta \circ F \circ \mu_\alpha^{-1}$  es un mapa cartesiano  $\mathcal{C}^\infty$ . Pero  $\mu_1$  y  $\mu_2$  están relacionadas de forma  $\mathcal{C}^\infty$  con  $\mu_\alpha$  y  $\mu_\beta$ , respectivamente. Por lo tanto,  $\mu_\beta \circ \mu_\alpha^{-1}$  está definida y es  $\mathcal{C}^\infty$  en algún entorno de  $m_1$ , y

$\mu_2 \circ \mu_1^{-1}$  está definida y es  $\mathcal{C}^\infty$  en algún entorno de  $n_\beta = \mu_\beta n$ .

La composición de mapas cartesianos  $\mathcal{C}^\infty$  es  $\mathcal{C}^\infty$ , por tanto  $\mu_2 \circ \mu_\beta^{-1} \circ \mu_\beta \circ F \circ \mu_\alpha^{-1} \circ \mu_\alpha \circ \mu_1^{-1}$  es un mapa  $\mathcal{C}^\infty$ . Sin embargo, está definido en algún entorno de  $m_1$  y coincide con la restricción de  $\mu_2 \circ F \circ \mu_1^{-1}$  en ese entorno, por tanto  $\mu_2 \circ F \circ \mu_1^{-1}$  es  $\mathcal{C}^\infty$  en un entorno de  $m_1$ . ■

En la práctica, para verificar que los mapas son  $\mathcal{C}^\infty$  se debe hacer mostrando que las componentes individuales de las expresiones de coordenadas tienen derivadas parciales continuas en todos los órdenes. Estas componentes son funciones  $u^i \circ \mu_2 \circ F \circ \mu_1^{-1} = f^i$ ,  $i = 1, 2, \dots, e$ , que son funciones reales con  $d$  variables reales definidas en un subconjunto abierto de  $W_1$  o  $\mathbb{R}^d$ .

Si tomamos  $y^i = u^i \circ \mu_2$  las funciones de coordenadas de  $\mu_2$  y  $x^j = u^j \circ \mu_1$ , las de  $\mu_1$ , entonces tenemos que  $y^i \circ F \circ \mu_1^{-1} = f^i$ , o  $y^i \circ F = f^i \circ \mu_1$ . Aplicando esto a  $m \in W$ , tenemos

$$y^i F m = f^i \mu_1 m = f^i(x^1 m, x^2 m, \dots, x^d m)$$

Suele escribirse esto como una ecuación entre funciones de la forma

$$y^i = f^i(x^1, x^2, \dots, x^d)$$

pero como esto no nos indica el papel del mapa  $F$  por sí mismo, preferimos usar la versión más precisa, que es

$$y^i \circ F = f^i(x^1, x^2, \dots, x^d)$$

Estas ecuaciones también se llaman *expresión de coordenadas de  $F$* .

En particular, podemos considerar el caso  $N = \mathbb{R}$  de funciones reales evaluadas sobre  $M$ . Es interesante que las funciones  $\mathcal{C}^\infty$  necesitan ser definidas directamente solo en este caso, y la definición general de un mapa  $\mathcal{C}^\infty$  se sigue con la siguiente proposición.

**Proposición B.3.2.** *Si  $F : M \rightarrow N$ , entonces  $F$  es  $\mathcal{C}^\infty$  si para cada función real  $\mathcal{C}^\infty$   $y : V \rightarrow \mathbb{R}$ , donde  $V$  es una subvariedad abierta de  $N$ ,  $y \circ F$  es una función real  $\mathcal{C}^\infty$  sobre la subvariedad abierta  $F^{-1}V$  de  $M$ .*

**Demostración.** Esto se deduce de manera trivial al tomar, sucesivamente, como  $y$  las funciones coordenadas  $y^i$  definidas en  $V \subset N$ . ■

Un **difeomorfismo** de  $M$  sobre  $N$  es una función  $\mathcal{C}^\infty$  biyectiva,  $F : M \rightarrow N$ , tal que el mapa inverso  $F^{-1} : N \rightarrow M$  también es  $\mathcal{C}^\infty$ . Dos variedades son **difeomorfas** si existe un difeomorfismo entre ellas. Esta es la noción natural de isomorfismo, o equivalencia, para variedades. Es una relación de equivalencia. Dos variedades difeomorfas son iguales en todas las propiedades que conciernen únicamente a su estructura como variedades. En particular, son topológicamente iguales, es decir, **homeomorfas**.

Una variedad topológica puede tener dos atlas  $\mathcal{C}^\infty$  diferentes que estén  $\mathcal{C}^\infty$ -relacionados, pero las dos variedades  $\mathcal{C}^\infty$  determinadas por estos atlas pueden seguir siendo difeomorfas. El problema es que el mapa identidad no es un difeomorfismo. De hecho, dos estructuras de variedad  $\mathcal{C}^\infty$  en una variedad de dimensión  $\leq 4$  son invariablemente difeomorfas. Por otro lado, cualquier variedad compacta de dimensión  $\geq 7$  admite varias estructuras de variedad  $\mathcal{C}^\infty$  no difeomorfas; esto es, puede haber un homeomorfismo entre dos variedades, pero no un difeomorfismo.

Para un simple ejemplo de diferentes estructuras  $\mathcal{C}^\infty$  que son difeomorfas, consideramos  $\mathbb{R}$  con la estructura estándar  $\{u : \mathbb{R} \rightarrow \mathbb{R}\}$  como atlas (con una carta), y  $M = \mathbb{R}$  con la estructura  $\{u^3 : \mathbb{R} \rightarrow \mathbb{R}\}$  como atlas (de nuevo, una carta). Ya que una carta admisible siempre es un difeomorfismo en su dominio de subvariedad abierta, el difeomorfismo de  $M$  sobre  $\mathbb{R}$  es un mapa coordenado de  $M$ ,  $u^3 : M \rightarrow \mathbb{R}$ . El difeomorfismo inverso es  $u^{1/3} : \mathbb{R} \rightarrow M$ . El mapa identidad  $u : \mathbb{R} \rightarrow M$  es  $\mathcal{C}^\infty$ , ya que la expresión de coordenadas es  $u^3 \circ u \circ u = u^3 : \mathbb{R} \rightarrow \mathbb{R}$ . El mapa identidad  $u : M \rightarrow \mathbb{R}$  no es  $\mathcal{C}^\infty$ , ya que la expresión de coordenadas es  $u \circ u \circ u^{1/3} = u^{1/3} : \mathbb{R} \rightarrow \mathbb{R}$ , que no es  $\mathcal{C}^\infty$ . Por tanto, el mapa identidad no es un difeomorfismo.

Existen ejemplos de estructuras de variedad  $\mathcal{C}^\infty$  no difeomorfas de dimensión  $\geq 7$ . No son fácil de describir.

## B.4. Subvariedades

Una variedad  $M$  está **incrustada** en una variedad  $N$  si existe un mapa  $\mathcal{C}^\infty$  biyectivo,  $F : M \rightarrow N$ , tal que para cada punto  $m \in M$ , existe un entorno  $U$  de  $m$  y una carta de  $N$  en  $Fm$ ,  $\mu : V \rightarrow \mathbb{R}^e$ ,  $\mu = (y^1, \dots, y^e)$ , tal que  $x^i = y^i \circ F|_U$ ,  $i = 1, 2, \dots, d$ , son coordenadas en  $U$  para  $M$ . El mapa  $F$  se denomina entonces una **incrustación** de  $M$  en  $N$ .

Si el requerimiento de que  $F$  sea biyectiva se omite, pero el requerimiento de obtener coordenadas para  $M$  de las de  $N$  por composición con  $F$  todavía se sostiene, entonces  $M$  se denomina **inmersa** en  $N$  y  $F$  se denomina una **inmersión**. Otra forma de indicar esto es requerir que cada punto  $m$  de  $M$  esté contenido en una subvariedad abierta  $U$  de  $M$ , la cuál es incrustada en  $N$  usando  $F$ . Por tanto, una inmersión es una incrustación local. Ver Figura B.5

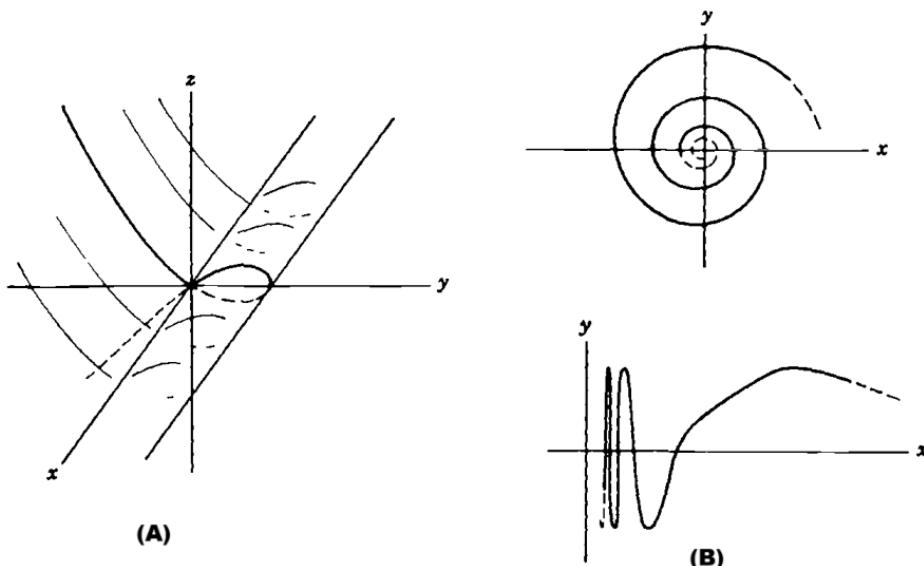


FIGURA B.5: Esquemas de una incrustación de  $\mathbb{R}^2$  en  $\mathbb{R}^3$  (A) y una inmersión (B).  
[6.1.4]

**Definición B.4.1.** Una **subvariedad** de  $N$  es un subconjunto  $FM$ , donde  $F : M \rightarrow N$  es una incrustación, dotada de estructura de variedad para la cuál  $F : M \rightarrow FM$  es un difeomorfismo.

La dimensión de una subvariedad obviamente no es mayor que la dimensión de la variedad que la contiene. Si fuera igual a la dimensión de la variedad que la contiene, entonces la subvariedad no es más que una subvariedad abierta, que definimos previamente.

La topología de una subvariedad no tiene por qué ser la topología inducida de la variedad más grande. Por supuesto, el mapa de inclusión es  $C^\infty$ , en particular continuo, de modo que los conjuntos abiertos de la topología inducida son conjuntos abiertos en la topología de la subvariedad. Sin embargo, la topología de la subvariedad puede tener muchos más conjuntos abiertos.

Una subvariedad debe estar metida en la variedad que la contiene de una forma especial. Por ejemplo, cosas como cúspides y esquinas están descartados, incluso si pueden aparecer en la imagen de un mapa inyectivo  $\mathcal{C}^\infty$ , que no es una incrustación. Para describir cuidadosamente la naturaleza especial de una subvariedad, definimos una *porción de coordenadas* de dimensión  $d$  en una variedad  $N$  de dimensión  $e$ , siendo un conjunto de puntos en un entorno coordinado  $U$  con coordenadas  $y^1, \dots, y^e$  de la forma  $\{m | m \in U, y^{d+1}m = c^{d+1}, \dots, y^em = c^e\}$ , donde  $c^i$  son constantes que determinan la porción. En otras palabras, una porción de coordenadas es la imagen bajo la inversa de un mapa coordinado de la parte  $d$ -dimensional del plano  $\mathbb{R}^e$  que yace dentro del rango de coordenadas.

**Proposición B.4.1.** *Si  $M$  es una subvariedad de  $N$ , entonces para cada  $m \in M$  existen coordenadas  $y^1, \dots, y^e$  para  $N$  en un entorno de  $m$  en  $N$ , tal que la porción de coordenadas correspondientes a las constantes  $c^{d+1} = y^{d+1}m, c^{d+2} = y^{d+2}m, \dots, c^e = y^em$  es un entorno de  $m$  en  $M$ , y la restricción de  $y^1, \dots, y^d$  a esa porción son coordenadas para  $M$ .*

**Demostración.** Sea  $F : P \rightarrow N$  una incrustación tal que  $FP = M$ . Tomamos las coordenadas  $z^1, z^2, \dots, z^e$  para  $N$  en un entorno de  $m$  en  $N$  tal que  $x^1 = z^1 \circ F|_U, \dots, x^d = z^d \circ F|_U$  son coordenadas en  $p = F^{-1}m$  en un entorno coordinado  $U \subset P$ . Ya que  $F$  es  $\mathcal{C}^\infty$ , podemos escribir

$$z^i \circ F = f^i(x^1, \dots, x^d); \quad i = 1, \dots, e$$

donde las  $f^i$  son funciones  $\mathcal{C}^\infty$  sobre un conjunto abierto en  $\mathbb{R}^d$ . Es claro que  $f^i(x^1, \dots, x^d) = x^i$  para  $i = 1, \dots, d$ , pero los demás  $f^i$ ,  $i > d$ , no es tan simple.

Definimos

$$\begin{aligned} y^i &= z^i, & i \leq d, \\ y^i &= z^i - f^i(z^1, \dots, z^d), & i > d \end{aligned}$$

Por tanto, es claro que

$$\begin{aligned} z^i &= y^i, & i \leq d, \\ z^i &= y^i + f^i(y^1, \dots, y^d), & i > d \end{aligned}$$

así que los mapas  $\mu_1 = (z^1, \dots, z^e)$  y  $\mu_2 = (y^1, \dots, y^e)$  están  $\mathcal{C}^\infty$  relacionados en ambos sentidos. El dominio de los  $y^i$  se incluye en el de los  $z^i$ , por lo que podemos afirmar que  $\mu_2$  es una carta admisible sin comprobar relaciones adicionales con otras coordenadas. Además,  $FU$  es la porción de coordenadas  $y^{d+1} = 0, \dots, y^e = 0$ , y las restricciones de  $y^1, \dots, y^d$  para  $FU$  corresponden a  $x^i$  bajo  $F$ , por lo que hay coordenadas para  $M$  sobre  $F$ . ■

**Observación B.4.2.** *No se afirma que obtengamos todos los puntos de  $M$  que yacen en  $N$ -entorno de  $m$  como miembros de una sola porción de coordenadas. De hecho, esto no es posible en algunos casos.*

Lo contrario a la Proposición B.4.1 es obvio usando la definición; esto es, si un subconjunto tiene estructura de variedad que está localmente determinada por una porción de coordenadas, donde coordenadas no constantes que proporcionan las coordenadas en la porción para la estructura de variedad del subconjunto, entonces el subconjunto es una subvariedad.

Whitney demostró que cualquier variedad es difeomorfa a una subvariedad de  $\mathbb{R}^e$ ; si  $d$  es la dimensión de la variedad, entonces necesitamos tomar un  $e$  no mayor a  $2d + 1$ .

**Teorema B.4.2** (Teorema de Whitney). *Cualquier variedad suave de dimensión  $d$  puede ser inmersa en  $\mathbb{R}^{2d}$  e incorporada en  $\mathbb{R}^{2d+1}$*

**Demostración.** Ver la demostración en [2, Chapter 6, page 134] ■

Por tanto, la teoría de variedades puede considerarse el estudio de subconjuntos especiales de espacios cartesianos, si se desea.

## B.5. Curvas diferenciables

En algunos contextos una curva es casi lo mismo que una subvariedad unidimensional, pero vamos a preferir tratar solo las curvas que tengan una parametrización específica. Técnicamente, entonces, cambiando la parametrización de una curva tendremos una curva diferente, pero normalmente ignoraremos la distinción y hablaremos de una curva como si fuera un conjunto de puntos. Generalmente nuestras curvas tendrán un punto inicial y final pero también consideraremos curvas con finales abiertos.

## Bibliografía

- [1] Richard L. Bishop and Samuel I. Goldberg. *Tensor Analysis on Manifolds*. Garden City, New York, Dover edition, 1968-1980-2023. ISBN 978-0-486-64039-6. [188](#), [189](#)
- [2] John M. Lee. *Introduction to Smooth Manifolds*. Springer New York Heidelberg dordrecht London, second edition edition, 2003,2013. ISBN 978-1-4419-9981-8. [195](#)
- [3] Luis Miguel. Merino González and Evangelina Santos Aláez. *Álgebra lineal : con métodos elementales*. Thomson, Madrid, [1<sup>a</sup> ed., 4<sup>a</sup> reimp.] edition, 2006 - 2010. ISBN 9788497324816. [15](#)
- [4] F. Zaldívar. *Introducción a la teoría de grupos*. Aportaciones matemáticas: Textos. Sociedad Matemática Mexicana, 2006. ISBN 9789686708660. URL <https://books.google.es/books?id=mNaJ1LC5WkUC>. [11](#)

## Bibliografía de Figuras

- [1] Richard L. Bishop and Samuel I. Goldberg, *Tensor Analysis on Manifolds*. Garden City, New York, Dover edition, 1968-1980-2023. ISBN 978-0-486-64039-6. Pag. 21
- [2] Richard L. Bishop and Samuel I. Goldberg, *Tensor Analysis on Manifolds*. Garden City, New York, Dover edition, 1968-1980-2023. ISBN 978-0-486-64039-6. Pag. 25
- [3] Richard L. Bishop and Samuel I. Goldberg, *Tensor Analysis on Manifolds*. Garden City, New York, Dover edition, 1968-1980-2023. ISBN 978-0-486-64039-6. Pag. 36
- [4] Richard L. Bishop and Samuel I. Goldberg, *Tensor Analysis on Manifolds*. Garden City, New York, Dover edition, 1968-1980-2023. ISBN 978-0-486-64039-6. Pag. 40