

# Instantones



Universidad de Oviedo

**Autor:** Miguel Pérez Alonso

**Tutor:** Patrick Meessen

**Facultad:** Ciencias

**Año académico:** 2021/22

**Titulación:** Doble Grado en Matemáticas y Física (opción B)



# Índice general

<b>Introducción</b>	<b>1</b>
<b>1. Preliminares matemáticos</b>	<b>3</b>
1.1. Topología . . . . .	3
1.1.1. Definiciones básicas y grupo fundamental . . . . .	3
1.1.2. Grupos de homotopía de orden superior y propiedades . . . . .	6
1.1.3. Grupos de homotopía de la $n$ -esfera . . . . .	11
1.2. Geometría diferencial . . . . .	12
1.2.1. Variedades . . . . .	12
1.2.2. Formas diferenciales . . . . .	15
1.2.3. Grupos y álgebras de Lie . . . . .	23
1.2.4. Representaciones de $SU(2)$ . . . . .	28
<b>2. Efecto túnel y Yang-Mills</b>	<b>33</b>
2.1. Integrales de camino . . . . .	33
2.1.1. WKB y probabilidades de efecto túnel . . . . .	36
2.2. Conexiones en teorías gauge . . . . .	38
2.2.1. Introducción a los fibrados . . . . .	38
2.2.2. Conexiones . . . . .	41
2.3. Teorías de Yang-Mills . . . . .	45
<b>3. Instantones</b>	<b>51</b>
3.1. Winding number . . . . .	53
3.2. Vacíos en Yang-Mills y el vacío $\theta$ . . . . .	59

3.3. El instantón $n = 1$ . . . . .	62
3.4. Cuaterniones . . . . .	67
3.5. La construcción ADHM . . . . .	71
3.6. Fibrados en el espacio de twistores . . . . .	74
<b>Conclusiones</b>	<b>78</b>
<b>Bibliografía</b>	<b>80</b>

# Introducción

Los instantones son un tipo particular de soluciones de las ecuaciones de movimiento de una teoría gauge. En concreto, son soluciones que tienen acción finita y su campo asociado cumple una cierta relación de dualidad. Estas características dan lugar a una serie de propiedades topológicas que hacen que el estudio de los instantones sea de gran interés.

Su aplicación más relevante es en el cálculo de probabilidades de transición entre los diferentes vacíos (estados fundamentales) de una teoría gauge: el término correspondiente a los instantones es el que más contribuye al valor de esta probabilidad.

Otras aplicaciones se encuentran en la teoría electrodébil, donde la presencia de instantones introduce eventos que violan la conservación del número bariónico, o en la cromodinámica cuántica, donde los instantones permiten explicar la diferencia entre las masas de los mesones  $\eta$  y  $\eta'$ . [1]

En este trabajo se introducirá la mayor parte de la maquinaria matemática necesaria para entender más o menos en profundidad las teorías gauge, para luego estudiar con detalle los instantones en el caso de una teoría Yang-Mills con grupo gauge  $SU(2)$ .

En el capítulo 1, definiremos los grupos de homotopía de un espacio topológico y desarrollaremos lo básico de la geometría diferencial en variedades, para después dar con algo de detalle la teoría de representaciones del grupo de Lie  $SU(2)$ .

En el capítulo 2, introducimos el formalismo de integrales de camino de la mecánica cuántica en su forma más sencilla y su relación con el efecto túnel. Luego, tras dar unas pinceladas de teoría de fibrados, desarrollamos la acción y las ecuaciones del movimiento de una teoría de Yang-Mills.

Construyendo sobre todo lo anterior, dedicamos el último capítulo al estudio de los instantones: sus principales propiedades, el concepto de número instantónico o *winding number*, la construcción explícita de estas soluciones y su importancia a la hora de calcular probabilidades de efecto túnel.

El grupo  $SU(2)$  en concreto es especialmente relevante porque es el que se usa para describir la interacción débil en el modelo estándar de partículas, además de ser la base de la teoría del momento angular en mecánica cuántica y tener una muy estrecha relación con los cuaterniones.

# Capítulo 1

## Preliminares matemáticos

### 1.1. Topología

#### 1.1.1. Definiciones básicas y grupo fundamental

**Definición 1.1** Sea  $X$  un conjunto no vacío. Un conjunto  $\tau \subseteq P(X)$  de subconjuntos de  $X$  se llama *topología* si satisface las tres siguientes propiedades:

1.  $\emptyset$  y  $X$  son elementos de  $\tau$ .
2. La unión arbitraria de elementos de  $\tau$  está en  $\tau$ .
3. La intersección finita de elementos de  $\tau$  está en  $\tau$ .

Bajo estas condiciones, diremos que  $(X, \tau)$  es un **espacio topológico** [2] y a los elementos de  $\tau$  los llamaremos *abiertos*.

**Definición 1.2** Sean  $(X, \tau_X), (Y, \tau_Y)$  dos espacios topológicos. Una aplicación  $f : (X, \tau_X) \rightarrow (Y, \tau_Y)$  es *continua* si para cada  $U \in \tau_Y$  se tiene que  $f^{-1}(U) \in \tau_X$ .

Si  $f$  es continua, biyectiva y su inversa es continua, se dice que  $f$  es un **homeomorfismo**.

El concepto de homeomorfismo es el equivalente topológico de los isomorfismos de espacios vectoriales o grupos. Si dos espacios topológicos son homeomorfos, comparten propiedades como la compacidad o la conexión. Estas propiedades que se conservan bajo homeomorfismos se suelen llamar **invariantes topológicos**.

Ahora introducimos un concepto más débil que el ser homeomorfo, que nos será útil cuando hablemos de teorías gauge.

**Definición 1.3** Sean  $(X, \tau_X), (Y, \tau_Y)$  espacios topológicos y  $f, g : X \rightarrow Y$  dos aplicaciones continuas entre ellos. Se dice que  $f$  es **homotópica** a  $g$  (denotado  $f \sim g$ ) si existe una aplicación (que llamaremos **homotopía**)  $H: X \times [0, 1] \rightarrow Y$  que verifique [2]:

- $H$  es continua
- $H(x, 0) = f(x) \forall x \in X$
- $H(x, 1) = g(x) \forall x \in X$

Si  $g$  es una aplicación constante, se dice que  $f$  es homotópicamente nula.

**Definición 1.4** Sean  $(X, \tau_X), (Y, \tau_Y)$  espacios topológicos,  $f, g : X \rightarrow Y$  dos aplicaciones continuas entre ellos y  $A \subseteq X$ . Se dice que  $f$  es **homotópica a  $g$  respecto a  $A$**  (denotado  $f \sim_A g$ ) si existe una homotopía  $H: X \times [0, 1] \rightarrow Y$  que verifique  $H(a, t) = f(a) = g(a)$  para todo  $a \in A, t \in [0, 1]$ .

**Definición 1.5** Sean  $(X, \tau_X), (Y, \tau_Y)$  dos espacios topológicos. Se dice que  $X$  e  $Y$  son **homotópicamente equivalentes** si existen dos aplicaciones  $f, g : (X, \tau_X) \rightarrow (Y, \tau_Y)$  ambas continuas y que cumplan  $f \circ g \sim Id_Y, g \circ f \sim Id_X$ .

Para tener una noción más o menos visual de qué es una homotopía, podemos ver su acción como una deformación con continuidad. Si, por ejemplo, tomamos un círculo en  $\mathbb{R}^2$  (que podemos ver como la imagen de una cierta aplicación  $f$ ) y lo queremos deformar en una elipse (que sería la imagen de una aplicación  $g$ ), lo podemos hacer sin cambiar demasiado la forma del círculo. La homotopía nos detalla el proceso de deformar el círculo para convertirlo en la elipse a través del argumento  $t$ . En  $t = 0$  tenemos el círculo, en  $t = 1$  tenemos la elipse y en el resto de instantes tenemos algo intermedio.

De la misma forma, los espacios homotópicamente equivalentes se pueden deformar uno en el otro con continuidad.

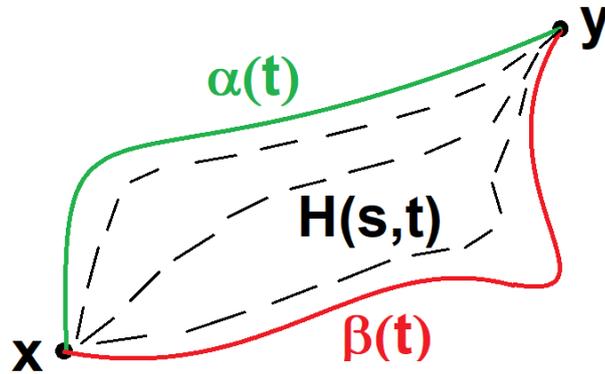


Figura 1.1: Homotopía entre dos caminos  $\alpha(t)$  y  $\beta(t)$  que empiezan y acaban en el mismo punto. Cada línea discontinua corresponde a un valor distinto de  $s$ .

Se puede demostrar [2] que la relación entre aplicaciones "ser homotópicas" (en general o respecto a un conjunto) y la relación entre espacios topológicos "ser homotópicamente equivalentes" son todas relaciones de equivalencia, de manera que todas las aplicaciones homotópicas entre sí las podemos agrupar en una sola clase.

Para definir los grupos de homotopía nos interesan las clases de equivalencia de un tipo de aplicaciones determinadas: los lazos.

**Definición 1.6** Sea  $(X, \tau)$  un espacio topológico. Se llama **camino** a cualquier aplicación continua  $\alpha : [0, 1] \rightarrow X$ . A los elementos de  $X$   $\alpha(0) = x_0$  y  $\alpha(1) = x_1$  se les llama **puntos inicial y final del camino respectivamente**. Si  $x_0 = \alpha(0) = \alpha(1)$ , llamaremos al camino  $\alpha$  **lazo** basado en  $x_0$ .

Usando caminos  $\alpha$  y  $\beta$  que cumplan  $\alpha(1) = \beta(0)$  (que incluyen a todos los lazos) se puede definir el siguiente producto:

$$(\alpha * \beta)(t) = \begin{cases} \alpha(2t) & 0 \leq t \leq \frac{1}{2} \\ \beta(2t - 1) & \frac{1}{2} \leq t \leq 1 \end{cases} \quad (1.1)$$

Y la operación inverso  $\alpha(t) \rightarrow \alpha(1 - t)$ .

Notemos que este producto no es necesariamente asociativo, pues los caminos se recorren en un orden concreto y, al hacer un producto  $\alpha * \beta * \gamma$ , dependiendo del producto que hagamos primero, vamos a dividir el intervalo  $[0, 1]$  de formas diferentes. Además,  $\alpha * \alpha^{-1}$  es un camino que recorre dos veces la imagen de  $\alpha$ , una vez en cada sentido. Esto claramente va a depender de quién sea  $\alpha$ , así que solo con este producto no podríamos definir una estructura de grupo.

Dentro de todos los lazos, consideremos los lazos basados en el mismo punto  $x_0$  y que son homotópicos respecto al conjunto  $A = \{0, 1\}$ , es decir, que existe una homotopía  $H(s, t)$  tal que  $H(0, t) = H(1, t) = x_0$  para todo  $t \in [0, 1]$ .

**Definición 1.7** Consideremos las clases de equivalencia

$$[\alpha] = \{\beta \text{ lazo basado en } x_0 \in X \mid \alpha \sim_{\{0,1\}} \beta\}. \quad (1.2)$$

El conjunto  $\pi_1(X, x_0) = \{[\alpha] \mid \alpha \text{ es lazo basado en } x_0\}$  junto con el producto

$$[\alpha] * [\beta] = [\alpha * \beta] \quad (1.3)$$

forma un grupo llamado **grupo fundamental** (o primer grupo de homotopía) de  $X$  respecto al punto  $x_0$ .

Este producto está bien definido: si tenemos dos representantes  $\alpha$  y  $\alpha'$  de la clase  $[\alpha]$  relacionados por una homotopía  $F$  y dos representantes  $\beta$  y  $\beta'$  de la clase  $[\beta]$  relacionados por una homotopía  $G$ , entonces

$$H(s, t) = \begin{cases} F(2s, t) & 0 \leq s \leq \frac{1}{2} \\ G(2s - 1, t) & \frac{1}{2} \leq s \leq 1 \end{cases} \quad (1.4)$$

es una homotopía de  $\alpha * \beta$  a  $\alpha' * \beta'$  [3], por lo que  $[\alpha * \beta] = [\alpha' * \beta']$ .

Un poco más adelante daremos la demostración de que el grupo fundamental es, en efecto, un grupo.

### 1.1.2. Grupos de homotopía de orden superior y propiedades

Ahora que hemos visto cuál es el primer grupo de homotopía, definamos el resto de grupos de homotopía. La extensión, conocido el grupo fundamental, es bastante directa.

Denotemos  $I^n = [0, 1]^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid 0 \leq x_1, \dots, x_n \leq 1\}$  al cubo unidad de  $\mathbb{R}^n$ . Definamos también su frontera

$$\partial I^n = \{(x_1, \dots, x_n) \in I^n \mid \text{al menos un } x_i \text{ es } 0 \text{ ó } 1 \ (i = 1, \dots, n)\}. \quad (1.5)$$

**Definición 1.8** Sea  $(X, \tau_X)$  un espacio topológico. Llamamos  **$n$ -camino** a cualquier aplicación continua  $\alpha : I^n \rightarrow X$ . Si  $\alpha(\partial I^n) = x_0$ , es decir, si  $\alpha$  permanece constante en la frontera de  $I^n$ , diremos que es un  **$n$ -lazo** basado en  $x_0$ .

Los  $n$ -lazos tienen una cualidad muy importante: como envían todo el conjunto  $\partial I_n$  al mismo punto, a efectos prácticos podemos ver los  $n$ -lazos como aplicaciones que, en vez de salir de  $I^n$ , salen de  $I^n/\partial I^n \cong S^n$  [3], donde

$$S^n = \{(x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \mid (x^1)^2 + \dots + (x^{n+1})^2 = 1\}. \quad (1.6)$$

De manera un poco rudimentaria, podemos ver esto en el caso  $n = 2$  como que, si dobláramos todo el borde de un cuadrado de papel para llevarlo al mismo punto, obtendríamos algo parecido a una bola.

De la misma forma que para los 1-caminos, podemos definir un producto de  $n$ -caminos [4]

$$(\alpha * \beta)(t_1, \dots, t_n) = \begin{cases} \alpha(2t_1, t_2, \dots, t_n) & 0 \leq t_1 \leq \frac{1}{2} \\ \beta(2t_1 - 1, t_2, \dots, t_n) & \frac{1}{2} \leq t_1 \leq 1 \end{cases}. \quad (1.7)$$

También podemos definir el inverso  $\alpha \rightarrow \alpha^{-1} = \alpha(1 - t_1, t_2, \dots, t_n)$ .

**Definición 1.9** Sea  $(X, \tau_X)$  un espacio topológico,  $x_0 \in X$  y  $n \geq 1$ ,  $n \in \mathbb{N}$ . Consideramos las clases de equivalencia de la forma:

$$[\alpha] = \{\beta \text{ n-lazo basado en } x_0 \in X \mid \alpha \sim_{\partial I^n} \beta\} \quad (1.8)$$

El conjunto  $\pi_n(X, x_0) = \{[\alpha] \mid \alpha \text{ es un n-lazo basado en } x_0\}$  junto con el producto de  $n$ -caminos

$$[\alpha] * [\beta] = [\alpha * \beta] \quad (1.9)$$

forma un grupo llamado  **$n$ -ésimo grupo de homotopía** de  $X$  respecto al punto  $x_0$ . Si  $\pi_n(X, y)$  es el mismo para todo  $y \in X$ , denotaremos al grupo como  $\pi_n(X)$ .

El producto de clases de equivalencia está bien definido y se comprueba exactamente igual que en el caso del grupo fundamental. Comprobemos que los grupos de homotopía son grupos de verdad [3]. Sea  $X$  un espacio topológico y  $x_0 \in X$ .

- Propiedad asociativa: sean  $\alpha, \beta, \gamma$   $n$ -lazos arbitrarios basados en  $x_0$ . Para ver que

$$[(\alpha * \beta) * \gamma] = ([\alpha] * [\beta]) * [\gamma] = [\alpha] * ([\beta] * [\gamma]) = [\alpha * (\beta * \gamma)], \quad (1.10)$$

tenemos que encontrar una homotopía entre los lazos de una clase y los de la otra. Bastará hacerlo para los representantes  $(\alpha * \beta) * \gamma$  y  $\alpha * (\beta * \gamma)$ , ya que para el resto de elementos existirá una homotopía por transitividad de la relación que define  $\pi_n(X, x_0)$ .

Definamos la siguiente aplicación:

$$F(s_1, \dots, s_n, t) = \begin{cases} \alpha\left(\frac{4s_1}{1+t}, s_2, \dots, s_n\right) & 0 \leq s_1 \leq \frac{1+t}{4} \\ \beta(4s_1 - t - 1, s_2, \dots, s_n) & \frac{1+t}{4} \leq s_1 \leq \frac{2+t}{4} \\ \gamma\left(\frac{4s_1-t-2}{2-t}, s_2, \dots, s_n\right) & \frac{2+t}{4} \leq s_1 \leq 1 \end{cases} \quad (1.11)$$

$F$  es continua por un resultado conocido como lema del pegamento o *gluing lemma* [2], que dice que una función definida a trozos por funciones continuas que coinciden en la intersección de sus dominios es continua. Esto último se cumple porque que los tres lazos empiezan y acaban en  $x_0$ . En los puntos  $s_1 = \frac{1+t}{4}$  y  $s_1 = \frac{2+t}{4}$ :

$$\begin{aligned} F\left(\frac{1+t}{4}, s_2, \dots, s_n, t\right) &= \alpha\left(\frac{4}{1+t} \frac{1+t}{4}, s_2, \dots, s_n\right) = \alpha(1, s_2, \dots, s_n) = x_0 = \\ &= \beta\left(4\left(\frac{1+t}{4}\right) - t - 1, s_2, \dots, s_n\right) = \beta(0, s_2, \dots, s_n) \end{aligned} \quad (1.12)$$

$$\begin{aligned} F\left(\frac{2+t}{4}, s_2, \dots, s_n, t\right) &= \beta\left(4\left(\frac{2+t}{4}\right) - t - 1, s_2, \dots, s_n\right) = \beta(1, s_2, \dots, s_n) = x_0 = \\ &= \gamma\left(\frac{4}{2-t} \frac{2+t}{4} - \frac{2+t}{2-t}, s_2, \dots, s_n\right) = \gamma(0, s_2, \dots, s_n). \end{aligned} \quad (1.13)$$

Comprobamos las dos últimas condiciones,

$$F(s_1, s_2, \dots, s_n, 0) = \begin{cases} \alpha(4s_1, s_2, \dots, s_n) & 0 \leq s_1 \leq \frac{1}{4} \\ \beta(4s_1 - 1, s_2, \dots, s_n) & \frac{1}{4} \leq s_1 \leq \frac{1}{2} = (\alpha * \beta) * \gamma, \\ \gamma(2s_1 - 1, s_2, \dots, s_n) & \frac{1}{2} \leq s_1 \leq 1 \end{cases} \quad (1.14)$$

$$F(s_1, s_2, \dots, s_n, 1) = \begin{cases} \alpha(2s_1, s_2, \dots, s_n) & 0 \leq s_1 \leq \frac{1}{2} \\ \beta(4s_1, s_2, \dots, s_n) & \frac{1}{2} \leq s_1 \leq \frac{3}{4} = \alpha * (\beta * \gamma). \\ \gamma(4s_1 - 3, s_2, \dots, s_n) & \frac{3}{4} \leq s_1 \leq 1 \end{cases} \quad (1.15)$$

Por tanto,  $F$  es una homotopía.

- Elemento neutro: buscamos un lazo  $e$  tal que  $[e] * [\alpha] = [\alpha] * [e] = [\alpha]$  para todo lazo  $\alpha$  basado en  $x_0$ . El candidato más probable es el lazo constante  $e(s_1, \dots, s_n) = x_0 \forall (s_1, \dots, s_n) \in I^n$ . De manera similar a lo anterior, construimos una homotopía  $F_1$  entre  $\alpha * e$  y  $\alpha$ , así como una homotopía  $F_2$  entre  $e * \alpha$  y  $\alpha$ .

$$F_1(s_1, s_2, \dots, s_n, t) = \begin{cases} \alpha\left(\frac{2s_1}{1+t}, s_2, \dots, s_n\right) & 0 \leq s_1 \leq \frac{t+1}{2} \\ x_0 & \frac{t+1}{2} \leq s_1 \leq 1 \end{cases} \quad (1.16)$$

$$F_2(s_1, s_2, \dots, s_n, t) = \begin{cases} x_0 & 0 \leq s_1 \leq \frac{1-t}{2} \\ \alpha \left( \frac{2s_1+t-1}{t+1}, s_2, \dots, s_n \right) & \frac{1-t}{2} \leq s_1 \leq 1 \end{cases} \quad (1.17)$$

Se cumple que  $F_1(\frac{t+1}{2}, s_2, \dots, s_n, t) = x_0$ ,  $F_1(s_1, \dots, s_n, 0) = \alpha * e$  y  $F_1(s_1, \dots, s_n, 1) = \alpha$ , por lo que es la homotopía que buscamos. Se puede hacer la comprobación análoga para  $F_2$ . Por tanto, tenemos que en efecto  $e$  es el elemento neutro.

- Inverso: es bastante probable que  $[\alpha^{-1}]$  sea el inverso de  $[\alpha]$ . Veamos si podemos construir una homotopía entre  $\alpha * \alpha^{-1}$  y  $e$ :

$$F(s, t) = \begin{cases} \alpha(2s_1(1-t), s_2, \dots, s_n) & 0 \leq s_1 \leq \frac{1}{2} \\ \alpha(2(1-s_1)(1-t), s_2, \dots, s_n) & \frac{1}{2} \leq s_1 \leq 1 \end{cases} \quad (1.18)$$

Tenemos que  $F(s_1, s_2, \dots, s_n, 0) = \alpha * \alpha^{-1}$  y que  $F(s_1, s_2, \dots, s_n, 1)$  es la constante  $x_0$ , así que tenemos la existencia de inverso:  $[\alpha]^{-1} = [\alpha^{-1}]$ .

Una vez probada la estructura de grupo, podemos ver algunas de las propiedades más importantes de los grupos de homotopía.

**Proposición 1.1** [3] *Sea  $(X, \tau_X)$  un espacio topológico conexo por caminos (esto es, para cada par de elementos de  $X$  hay un camino cuyos extremos son esos elementos). Entonces  $\pi_n(X, x_0)$  y  $\pi_n(X, y)$  son grupos isomorfos para todo  $y \in X$  y para todo  $n \geq 1$ .*

**Definición 1.10** *Un espacio topológico  $(X, \tau_X)$  es **simplemente conexo** si para cada par de puntos de  $X$  existe un camino que los une (es decir,  $X$  es conexo por caminos) y su grupo fundamental en un punto es trivial.*

Esta definición de un conjunto simplemente conexo es la formalización de la intuición de que el conjunto "no tiene agujeros". En un conjunto simplemente conexo, todo lazo se puede contraer a un solo punto con continuidad (es decir, son homotópicos al camino constante). Si nuestro conjunto "tuviera agujeros", al tomar un lazo que rodee a algún agujero no podríamos contraerlo en un punto.

**Proposición 1.2** [3] *Sean  $(X, \tau_X)$  e  $(Y, \tau_Y)$  dos espacios topológicos tales que  $X$  e  $Y$  son homotópicamente equivalentes (por lo que existen dos aplicaciones  $h_1$  y  $h_2$  tales que  $h_1 \circ h_2 \sim Id_Y$ ).*

*Entonces  $\pi_n(X, x_0)$  es isomorfo a  $\pi_n(Y, h_1(x_0))$  para todo  $x_0 \in X$ , para todo  $n \geq 1$ .*

Como ser homeomorfo es un caso particular de tener el mismo tipo de homotopía, esto demuestra que los grupos de homotopía son invariantes topológicos. Si dos espacios son homeomorfos, necesariamente sus grupos de homotopía son los mismos.

**Teorema 1.1** [5] Sea  $(X, \tau_X)$  un espacio topológico.  $\pi_n(X, x_0)$  es un grupo abeliano si  $n \geq 2$ .

**Demostración:**

La demostración de este teorema se suele hacer de una forma bastante visual, apelando a que la frontera de  $I^n$  para  $n \geq 2$  es conexa y con infinitos puntos, de forma que al hacer el producto de dos clases de homotopía, puedes cambiarlas de orden sin ningún problema.

Aquí usaremos un truco algebraico bastante sencillo para verlo de forma rigurosa.

**Lema 1.1** (Argumento de Eckmann-Hilton) Sean  $*$  y  $\oplus$  dos aplicaciones binarias con elemento neutro sobre un mismo conjunto  $X$  que cumplan

$$(a * b) \oplus (c * d) = (a \oplus c) * (b \oplus d) \quad (1.19)$$

para todo  $a, b, c, d \in X$ . Entonces ambas operaciones son la misma y son conmutativas y asociativas.

**Demostración:**

Primero veamos que el elemento neutro de ambas operaciones  $1_*$  y  $1_\oplus$  (ambos elementos de  $X$ , por lo que se pueden introducir con la operación adecuada) son iguales:

$$1_* = 1_* * 1_* = (1_* \oplus 1_\oplus) * (1_\oplus \oplus 1_*) \underset{\text{hip.}}{=} (1_* * 1_\oplus) \oplus (1_\oplus * 1_*) = 1_\oplus \oplus 1_\oplus = 1_\oplus \equiv 1_\oplus \quad (1.20)$$

Ahora veamos que las operaciones son la misma:

$$a * b = (a \oplus 1) * (1 \oplus b) \underset{\text{hip.}}{=} (a * 1) \oplus (1 * b) = a \oplus b \quad (1.21)$$

Por último vemos la conmutatividad y la asociatividad:

$$a * b = (1 \otimes a) * (b \otimes 1) \underset{\text{hip.}}{=} (1 * b) \oplus (a \oplus 1) = b \oplus a = b * a \quad (1.22)$$

$$(a * b) * c = (a * b) * (1 * c) = (a * b) \oplus (1 * c) = (a * 1) * (b * c) = a * (b * c) \quad \blacksquare \quad (1.23)$$

Ahora volvamos a los grupos de homotopía. Supongamos que hubiéramos definido el producto de dos clases de homotopía de  $n$ -lazos  $[\alpha], [\beta] \in \pi_n(X, x_0)$  como la clase de un nuevo producto de  $n$ -caminos  $\oplus$ :

$$[\alpha] \oplus [\beta] = [\alpha \oplus \beta], \text{ con } \alpha \oplus \beta(t_1, \dots, t_n) = \begin{cases} \alpha(t_1, 2t_2, t_3, \dots, t_n) & 0 \leq t_2 \leq \frac{1}{2} \\ \alpha(t_1, 2t_2 - 1, t_3, \dots, t_n) & \frac{1}{2} \leq t_2 \leq 1 \end{cases} \quad (1.24)$$

Es claro que esta operación al igual que el producto usual  $*$  tiene al camino constante en  $x_0$  como elemento neutro. También podemos comprobar fácilmente con las definiciones que, si  $[\alpha], [\beta], [\gamma], [\delta] \in \pi_n(X, x_0)$ :

$$([\alpha] \oplus [\beta]) * ([\gamma] \oplus [\delta]) = ([\alpha] * [\gamma]) \oplus ([\beta] * [\delta]) \quad (1.25)$$

En particular, los representantes de ambos productos definen la misma aplicación, por lo que son homotópicas. Por tanto, usando el argumento de Eckmann-Hilton concluimos que  $\pi_n(X, x_0)$  es un grupo abeliano si  $n \geq 2$ . ■

Notemos que este método falla cuando  $n = 1$  porque no podemos definir el segundo producto de lazos de manera que "no interfiera" con el otro porque nos faltan variables.

### 1.1.3. Grupos de homotopía de la n-esfera

A lo largo del trabajo vamos a tratar con los grupos de homotopía de las esferas (en concreto de  $S^3$ ), así que no está de más saber cuáles son estos grupos. Al ser  $S^n$  un espacio topológico conexo por caminos para todo  $n \geq 1$ , todos los puntos tendrán el mismo grupo fundamental por la proposición 2.1. Las demostraciones rigurosas requieren resultados de topología que no son interesantes para lo que pretendemos con este trabajo y que no se usarían para nada más, por lo que las omitiremos.

En primer lugar tenemos que  $\pi_1(S^1) = \mathbb{Z}$ . Esto se puede demostrar directamente definiendo la aplicación siguiente [2]:

$$\begin{aligned} \phi : \mathbb{Z} &\rightarrow \pi_1(S^1) \\ n &\rightarrow [\omega_n(\theta)] = [(\cos(2n\pi\theta), \sin(2n\pi\theta))] \end{aligned} \quad (1.26)$$

Se puede comprobar que  $\phi$  es un isomorfismo y se suele demostrar en la mayoría de libros de topología algebraica. Además, la aplicación  $\omega_n$  en todo un periodo  $\theta \in (0, 2\pi)$  da exactamente  $n$  vueltas alrededor del círculo cuyo sentido está determinado por el signo de  $n$ .

Pasando a las esferas en dimensión superior, resulta que se puede encontrar un isomorfismo entre  $\pi_k(S^n)$  y  $\pi_{k+1}(S^{n+1})$  si  $k < 2n - 1$  y  $n > 1$  [4].

Entonces, con calcular el grupo fundamental y el segundo grupo de homotopía de  $S^2$  nos bastaría para calcular  $\pi_k(S^n)$  para todo  $k \leq n$ .

Para ver  $\pi_1(S^2)$  notamos que la esfera es simplemente conexa ([2] tiene una demostración). Por definición y por la proposición 1.1, concluimos que  $\pi_1(S^2) = 0$ .

Para el segundo grupo de homotopía recordamos que los 2-lazos de  $S^2$  son aplicaciones  $g : S^2 \rightarrow S^2$ . Estas aplicaciones las podemos ver como  $g(\theta, \phi) = (\theta', \phi')$ , ya que los dos ángulos determinan completamente el punto de la esfera [3].

Si fijamos, por ejemplo,  $\phi' = 2\phi$ , tenemos que  $g$  daría dos "vueltas" alrededor de la esfera en el sentido de que recorre el dominio del ángulo  $\phi$  dos veces. Este lazo estaría en la clase del 2 de  $\pi_2(S^2)$ . Podemos hacer esto para cualquier entero con  $g_k(\theta, \phi) = (\theta, k\phi)$  y, aunque no de forma totalmente rigurosa, tenemos que  $\pi_2(S^2) = \mathbb{Z}$  [3].

Podemos resumir todo esto en un resultado:

**Teorema 1.2** *Para todo entero positivo  $n$ ,  $\pi_n(S^n) = \mathbb{Z}$  y  $\pi_k(S^n) = 0$  si  $k < n$ .*

## 1.2. Geometría diferencial

### 1.2.1. Variedades

Las variedades son generalizaciones de las superficies de  $\mathbb{R}^3$  en las que se puede definir un cálculo que extienda lo que conocemos de integración y derivación en  $\mathbb{R}^n$ .

En la sección anterior hemos trabajado con espacios topológicos arbitrarios. Para hablar de variedades, vamos a tener que suponer a partir de ahora que los espacios topológicos tienen una serie de propiedades, que se pueden resumir en que son localmente muy parecidos a  $\mathbb{R}^n$ : en cada punto  $p$ , supondremos que podemos encontrar un homeomorfismo entre un abierto de nuestro espacio topológico y un abierto de  $\mathbb{R}^n$ . Para más detalles, se puede consultar [3] o [6], por ejemplo.

Todas las teorías de campos usan una variedad como su espacio-tiempo, por lo que es imprescindible saber de variedades si queremos hablar mínimamente en serio de lo que pasa en estas teorías.

**Definición 1.11** *Sea  $(M, \tau)$  un espacio topológico,  $U \in \tau$ . Una aplicación  $\phi : U \subseteq M \rightarrow \phi(U) \subseteq \mathbb{R}^n$  se llama **carta sobre  $M$**  si es un homeomorfismo y  $\phi(U)$  es un abierto de  $\mathbb{R}^n$ . A la carta  $\phi$  se le suele denotar  $(U, \phi)$  para tener en cuenta su dominio.*

Las cartas nos permitirán dar una noción parecida a las coordenadas en  $\mathbb{R}^n$  en espacios topológicos más abstractos. Para cada punto  $p$  de  $M$ , tendremos una serie de funciones reales que nos definen el punto correspondiente de  $\mathbb{R}^n$ ,  $\phi(p) = (x^1(p), x^2(p), \dots, x^n(p))$ . Usando esto, se suele denotar también a la carta  $\phi$  como  $(U, x^1, \dots, x^n)$ .

**Definición 1.12** Sea  $(M, \tau)$  un espacio topológico. Llamaremos **atlas**  $C^\infty$  a un conjunto de cartas  $\{(\phi_i, U_i)\}_{i \in I}$  tales que  $\bigcup_{i \in I} U_i = M$  y tales que si  $U_i \cap U_j \neq \emptyset$  para algún par  $i \neq j$ , entonces la composición

$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (1.27)$$

es una función de clase  $C^\infty$  (en cuyo caso se dirá que  $\phi$  y  $\psi$  son **cartas compatibles**).

La condición de que el "cambio de coordenadas"  $\phi_i \circ \phi_j^{-1}$  sea  $C^\infty$  nos asegura que no haya problemas de definición a la hora de usar cartas cuyos dominios se intersequen.

**Definición 1.13** Sea  $(M, \tau)$  un espacio topológico. Al conjunto  $M$  junto con un atlas  $A$  de clase  $C^\infty$  se le llama **variedad diferenciable** (o simplemente **variedad**). Se impone que todas las cartas de  $A$  tengan como conjunto de llegada el mismo  $\mathbb{R}^n$ , a la dimensión  $n$  la llamaremos **dimensión de la variedad**.

Se puede alterar un poco esta definición de variedad para introducir un concepto de frontera: imponiendo que las cartas tengan como imagen abiertos del semiespacio superior

$$\mathbb{A}^n = \{(x^1, \dots, x^n) \in \mathbb{R}^n \mid x^n \geq 0\}, \quad (1.28)$$

podemos definir la frontera de  $M$  como los puntos que, tras aplicarles una carta, acaban en el hiperplano  $x^n = 0$ . También se define el interior de una variedad como el complementario (en  $M$ ) de la frontera.

Se cumple [6] que la frontera es una variedad de dimensión  $n - 1$  y que el interior es una variedad de dimensión  $n$ , ambas sin frontera.

**Definición 1.14** Una variedad  $M$  de dimensión  $n$  se dice **orientable** si, para todo par  $(U, \phi)$  y  $(V, \psi)$  de cartas sobre  $M$ , el determinante de la matriz jacobiana de  $\phi \circ \psi^{-1}$  tiene signo positivo en todo punto de  $U \cap V$ .

Las variedades orientables son de especial interés porque en ellas se puede definir una integración<sup>1</sup> consistente con la que conocemos en  $\mathbb{R}^n$ . Sin una orientación bien definida, resultados como el teorema de Stokes fallan en determinados casos [6].

**Definición 1.15** Sea  $M$  una variedad de dimensión  $n$  y  $(U, \phi)$  una carta sobre  $M$ . Se dice que una aplicación  $f : U \subset M \rightarrow \mathbb{R}$  es de clase  $C^\infty$  si la composición  $f \circ \phi^{-1} : \phi(U) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  es de clase  $C^\infty$ .

<sup>1</sup>Es posible introducir una integración en variedades no orientables, pero no la vamos a utilizar.

Este procedimiento de componer una función con las cartas para obtener funciones que sabemos derivar es un procedimiento estándar en variedades [6]. Usando esto, podemos definir, por ejemplo, las aplicaciones  $C^\infty$  entre variedades.

**Definición 1.16** Sean  $M$  una variedad de dimensión  $n$  y  $L$  una variedad de dimensión  $m$ . Una función  $f : M \rightarrow L$  se dice de clase  $C^\infty$  si la composición  $\psi \circ f \circ \phi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  es una función de clase  $C^\infty$  para toda carta  $\phi$  de  $M$  y para toda carta  $\psi$  de  $L$ .

Dentro de las variedades se pueden definir curvas de la misma forma que definíamos caminos en un espacio topológico cualquiera: dado un punto  $p$  de una variedad  $M$ , podemos definir una función  $\sigma(t)$  haciendo variar  $p$  a lo largo del "tiempo"  $t$ . A la curva se le puede asignar un vector "tangente" derivando con las coordenadas que proporciona una carta: si  $\phi(p) = (x^1(p), \dots, x^n(p))$  es una carta sobre  $M$ , entonces [4]

$$p \longrightarrow p(t) = \sigma(t) \xrightarrow{\phi} x^i(\sigma(t)) \longrightarrow \frac{dx^i(\sigma(t))}{dt} \quad (1.29)$$

Estos vectores tangentes  $\frac{dx^i(\sigma(t))}{dt}$  los podemos ver como operadores que actúan sobre funciones  $f$  definidas en un entorno de  $\sigma(t)$  y que nos devuelven cómo cambian a medida que avanzan por la curva.

$$\frac{d(f(\sigma(t)))}{dt} = \frac{d(f \circ \phi^{-1} \circ \phi \circ \sigma)(t)}{dt} = \frac{\partial f}{\partial x^i} \frac{dx^i}{dt} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i}(f) \equiv X(f) \quad (1.30)$$

Esto nos dice que todo vector tangente se puede expresar como una combinación lineal de los vectores  $\frac{\partial}{\partial x^i}$  puesto que las derivadas de  $X^i$  respecto a  $t$  siempre van a estar evaluadas y van a ser siempre números. A su vez, toda combinación de las derivadas parciales de la forma (1.30) es vector tangente a la curva  $\sigma(t) = p + \frac{dx^i}{dt}t$ .

**Definición 1.17** Sea  $M$  una variedad de dimensión  $n$ ,  $p \in M$  y  $\phi = (x^1, \dots, x^n)$  una carta sobre  $M$  que tenga a  $p$  en su dominio. El espacio vectorial generado por los vectores  $\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\}$  (evaluados en  $p$ ) se llama **espacio tangente a  $p$**  y se denota  $T_p M$ .

De esta definición se puede ver que la expresión de un elemento del espacio tangente en términos de las parciales depende la carta que usemos para definir las parciales. Veamos que, aun así, el espacio tangente como conjunto no depende de las coordenadas.

**Proposición 1.3** Sea  $M$  una variedad de dimensión  $n$ ,  $(U, \phi)$  y  $(V, \psi)$  dos cartas de  $M$  diferentes y  $p \in U \cap V$ . Entonces el espacio tangente a  $p$  dado por ambas cartas es el mismo.

### Demostración:

Denotemos  $\phi(p) = (x^1(p), \dots, x^n(p))$  y  $\psi(p) = (y^1(p), \dots, y^n(p))$  a las coordenadas de cada una de las cartas. Evaluando las derivadas parciales (los elementos de las bases del espacio tangente) en una función y usando que  $\phi^{-1}(\phi(p)) = p$ ,

$$\frac{\partial f}{\partial x^i}(p) = \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(p)) \quad \frac{\partial f}{\partial y^i}(p) = \frac{\partial(f \circ \psi^{-1})}{\partial y^i}(\psi(p)). \quad (1.31)$$

Ahora, teniendo en cuenta que  $\psi \circ \phi^{-1}$  nos da las coordenadas  $y^i$  en función de las  $x^i$ ,

$$\frac{\partial f}{\partial x^i}(p) = \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(p)) = \frac{\partial(f \circ \psi^{-1} \circ \psi \circ \phi^{-1})}{\partial x^i}(\phi(p)). \quad (1.32)$$

Aplicando la regla de la cadena a la primera composición, que son funciones de las coordenadas  $y^i(x^1, \dots, x^n)$ , y usando que  $\psi \circ \phi^{-1}(\phi(p))$  nos da las coordenadas  $y^i(p)$  tenemos

$$\frac{\partial(f \circ \psi^{-1} \circ \psi \circ \phi^{-1})}{\partial x^i}(\phi(p)) = \frac{\partial f}{\partial y^j}(p) \frac{\partial y^j}{\partial x^i}. \quad (1.33)$$

Como los  $\frac{\partial y^j}{\partial x^i}$  son números, esto es una combinación lineal de los operadores  $\{\frac{\partial}{\partial y^j}\}$  actuando sobre  $f$ , que es una función arbitraria. En concreto, los coeficientes de esta combinación lineal son las entradas de la matriz jacobiana del cambio de coordenadas  $\psi \circ \phi^{-1}$ . Por tanto, los operadores  $\{\frac{\partial}{\partial x^i}\}$  que generan el espacio tangente según la carta  $\phi$  se pueden poner como combinaciones lineales de los operadores que generan el espacio tangente según  $\psi$ . Concluimos que los espacios que generan ambas bases son iguales. ■

### 1.2.2. Formas diferenciales

A continuación introducimos el concepto de una forma diferencial, que es la base de todos los objetos presentes en las teorías gauge y donde se cimienta todo el cálculo diferencial e integral en variedades [6]. Para ello necesitamos recordar unos conceptos de álgebra lineal.

**Definición 1.18** Sea  $V$  un espacio vectorial sobre  $\mathbb{R}$ . Una función  $T : V^n \rightarrow \mathbb{R}$  se dice **multilineal** si para todo  $1 \leq i \leq n$  se tiene:

$$T(v_1, \dots, v_i + w_i, \dots, v_n) = T(v_1, \dots, v_i, \dots, v_n) + T(v_1, \dots, w_i, \dots, v_n) \quad (1.34)$$

$$T(v_1, \dots, av_i, \dots, v_n) = aT(v_1, \dots, v_i, \dots, v_n) \text{ para todo } a \in \mathbb{R} \quad (1.35)$$

**Definición 1.19** Sea  $V$  un  $\mathbb{R}$ -espacio vectorial. Al espacio de todas las aplicaciones lineales

$f : V \rightarrow \mathbb{R}$  se le llama **espacio dual** de  $V$  y se denota  $V^*$ . El espacio dual tiene asociada una única base dual  $\{e^1, \dots, e^n\}$  tal que, si la base de  $V$  es  $\{e_1, \dots, e_n\}$ , entonces  $e^i(e_j) = \delta_j^i$ , con  $\delta_j^i$  la delta de Kronecker.

**Definición 1.20** Sea  $V$  un  $\mathbb{R}$ -espacio vectorial. Un **tensor  $(k, l)$**  o tensor  $k$  veces contravariante y  $l$  veces covariante es una aplicación de la forma

$$T : \underbrace{V^* \times \dots \times V^*}_{k \text{ copias}} \times \underbrace{V \times \dots \times V}_{l \text{ copias}} \rightarrow \mathbb{R} \quad (1.36)$$

tal que  $T$  es multilineal.

**Definición 1.21** Sea  $S$  un tensor  $(k, l)$  y  $T$  un tensor  $(p, q)$ . Se define el **producto tensorial** de  $S$  y  $T$ ,  $S \otimes T$  como

$$S \otimes T : \underbrace{V^* \times \dots \times V^*}_{k+p \text{ copias}} \times \underbrace{V \times \dots \times V}_{l+q \text{ copias}} \rightarrow \mathbb{R} \quad (1.37)$$

$$S \otimes T(v^1, \dots, v^{k+p}, v_1, \dots, v_{l+q}) = S(v^1, \dots, v^k, v_1, \dots, v_l)T(v^{k+1}, \dots, v^{k+p}, v_{l+1}, \dots, v_{l+q}).$$

Los tensores  $(k, l)$  forman un espacio vectorial de dimensión  $n^{k+l}$  [7]. Si denotamos a la base de  $V$  como  $\{e_1, \dots, e_n\}$  y a la base dual de  $V^*$  como  $\{e^1, \dots, e^n\}$  podemos escribir cualquier tensor  $T$  como

$$T = T^{i_1 \dots i_k}_{j_1 \dots j_l} e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_l} \quad (1.38)$$

A los números

$$T^{i_1 \dots i_k}_{j_1 \dots j_l} = T(e^{i_1}, \dots, e^{i_k}, e_{j_1}, \dots, e_{j_l}) \quad (1.39)$$

los llamaremos **componentes** del tensor  $T$ .

**Definición 1.22** Sea  $T$  un tensor  $(0, l)$  y  $P$  una permutación del grupo simétrico  $S_l$ .  $P$  se puede expresar como producto de  $k$  transposiciones y se define el signo de  $P$  como  $+1$  si  $k$  es par y  $-1$  si  $k$  es impar.

- $T$  se dice **totalmente simétrico** si

$$P(T) = T(v_{P(1)}, \dots, v_{P(l)}) = T(v_1, \dots, v_l) = T \quad \forall P \in S_l, \forall v_1, \dots, v_l \in V \quad (1.40)$$

- $T$  se dice **totalmente antisimétrico** si

$$P(T) = T(v_{P(1)}, \dots, v_{P(l)}) = \text{signo}(P) T(v_1, \dots, v_l) = \text{signo}(P) T \quad \forall P \in S_l, \forall v_1, \dots, v_l \in V \quad (1.41)$$

Estas propiedades de simetría se pueden definir de forma análoga para tensores  $(k, 0)$  y  $(k, l)$ , pero solo nos interesará el caso  $(0, l)$ .

Ahora estamos en condiciones de definir las formas diferenciales en una variedad.

**Definición 1.23** Sea  $M$  una variedad de dimensión  $n$  y  $p \in M$ . Llamaremos **espacio cotangente** a  $M$  en  $p$  al espacio dual de  $T_p M$ . Lo denotaremos  $T_p^* M$ . Dada una base  $\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\}$  de  $T_p M$  denotaremos a su base dual como  $\{dx^1, \dots, dx^n\}$ .

La acción de los elementos del espacio cotangente sobre los vectores tangentes se define de la siguiente manera: sea  $df \in T_p^* M$  y  $V = V^i \frac{\partial}{\partial x^i} \in T_p M$ . Entonces

$$df(V) \equiv V(f) = V^i \frac{\partial f}{\partial x^i} = \frac{\partial f}{\partial x^i} \delta_j^i V^j = \frac{\partial f}{\partial x^i} dx^i(V), \quad (1.42)$$

con lo cual  $df = \frac{\partial f}{\partial x^i} dx^i$ .

**Definición 1.24** Sea  $M$  una variedad de dimensión  $n$  y  $p \in M$ . Una  **$r$ -forma diferencial** o  **$r$ -forma** ( $r \geq 1$ ) es un tensor  $r$  veces covariante y totalmente antisimétrico que actúa sobre  $r$  copias del espacio tangente a  $M$  en  $p$  y cuyas componentes son funciones de clase  $C^\infty$ . Se define también una  $0$ -forma como una función  $C^\infty$  sobre  $M$ .

Llamaremos  $\Lambda_p^r(M)$  al conjunto de las  $r$ -formas que actúan sobre  $T_p M$ .

El ejemplo más típico de una 1-forma es la diferencial de una función de varias variables, como anticipamos en la definición del espacio cotangente. El resto de formas se pueden obtener de forma cómoda mediante la operación que definimos a continuación.

**Definición 1.25** Sea  $\alpha$  una  $r$ -forma y  $\beta$  una  $s$ -forma. Se define el **producto exterior** de  $\alpha$  y  $\beta$ , denotado  $\alpha \wedge \beta$ , como

$$\begin{aligned} \alpha \wedge \beta &= \frac{1}{r!s!} \sum_{P \in S_{r+s}} \text{signo}(P) \alpha(v_{P(1)}, \dots, v_{P(r)}) \beta(v_{P(r+1)}, \dots, v_{P(r+s)}) = \\ &= \frac{1}{r!s!} \sum_{P \in S_{r+s}} \text{signo}(P) P(\alpha \otimes \beta) \end{aligned} \quad (1.43)$$

Donde estamos sumando sobre todas las permutaciones posibles de los  $r + s$  vectores involucrados.

Se cumple que  $\alpha \wedge \beta$  es una  $r + s$ -forma. A partir de este hecho y de que los productos tensoriales son una base del espacio vectorial de los tensores, se puede probar [7] que toda  $r$ -forma se puede expresar como

$$\alpha = \alpha_{i_1 \dots i_r} dx^{i_1} \wedge \dots \wedge dx^{i_r} \quad (1.44)$$

y que  $\Lambda_p^r$  es un espacio vectorial de dimensión  $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ , con  $n$  la dimensión de la variedad.

El producto exterior cumple las siguientes propiedades.

**Proposición 1.4** [6] Sean  $\alpha, \beta, \eta, \omega$ , formas diferenciales sobre la misma variedad  $M$  con  $\alpha$  una  $r$ -forma y  $\beta$  una  $l$ -forma. Entonces:

1. Para todo  $a, b \in \mathbb{R}$ ,

$$(a\alpha + b\beta) \wedge \eta = a(\alpha \wedge \eta) + b(\beta \wedge \eta) \quad (1.45)$$

$$\eta \wedge (a\alpha + b\beta) = a(\eta \wedge \alpha) + b(\eta \wedge \beta) \quad (1.46)$$

2.  $\alpha \wedge \beta = (-1)^{rl} \beta \wedge \alpha$

3.  $dx^{i_1} \wedge \dots \wedge dx^{i_r} = 0$  si y solo si existe  $1 \leq j, k \leq r$  tal que  $i_j = i_k$ . Esto también implica que si  $r$  es mayor que la dimensión de la variedad, toda  $r$ -forma es nula.

4. Para todo conjunto de 1-formas  $\omega^1, \dots, \omega^r$  y vectores tangentes  $v_1, \dots, v_r$  ( $1 \leq r \leq n$ ) se cumple

$$\omega^1 \wedge \dots \wedge \omega^r(v_1, \dots, v_r) = \det(\omega^j(v_i)), \quad (1.47)$$

donde  $(\omega^j(v_i))$  es una matriz  $r \times r$ .

Estas propiedades nos dicen que en la práctica, a la hora de hacer el producto exterior de dos formas, podemos multiplicar las componentes de cada una y calcular por separado los productos exteriores del tipo  $dx^i \wedge dx^j$ . Por ejemplo, si  $\alpha = \sin(x^2) dx^1$  y  $\beta = e^{x^1} dx^1 \wedge dx^2 + x^3 dx^2 \wedge dx^3$ ,

$$\alpha \wedge \beta = \sin(x^2) e^{x^1} dx^1 \wedge dx^1 \wedge dx^2 + \sin(x^2) x^3 dx^1 \wedge dx^2 \wedge dx^3 = \sin(x^2) x^3 dx^1 \wedge dx^2 \wedge dx^3, \quad (1.48)$$

porque por antisimetría  $dx^1 \wedge dx^1 = -dx^1 \wedge dx^1 = 0$ .

Para acabar con el tema de las formas diferenciales vamos a ver algunos operadores que nos servirán como herramientas para construir nuevas formas.

**Definición 1.26** Sea  $M$  una variedad de dimensión  $n$ ,  $T_p M$  un espacio tangente cualquiera de  $M$ ,  $\alpha = \alpha_{i_1 \dots i_r} dx^{i_1} \wedge \dots \wedge dx^{i_r} \in \Lambda_p^r(M)$ . La **derivada exterior** se define como la aplicación  $d : \Lambda_p^r(M) \rightarrow \Lambda_p^{r+1}(M)$  dada por

$$d\alpha = \frac{1}{r!} \left( \frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^j} \right) dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}. \quad (1.49)$$

La derivada exterior generaliza lo que conocemos como la diferencial de una función. Nos van a interesar las siguientes propiedades:

**Proposición 1.5** [6] Sean  $\alpha$  y  $\beta$  dos formas diferenciales sobre una variedad  $M$ ,  $\alpha$  una  $r$ -forma. Entonces

1.  $d(\alpha + \beta) = d\alpha + d\beta$  si  $\beta \in \Lambda_p^r(M)$ .
2.  $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^r \alpha \wedge d\beta$ .
3.  $d(d(\alpha)) = d^2\alpha = 0$ .

**Demostración:**

1. Aplicando la definición

$$d(\alpha + \beta) = \frac{1}{r!} \left( \frac{\partial (\alpha + \beta)_{i_1 \dots i_r}}{\partial x^j} \right) dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}. \quad (1.50)$$

Ahora bien, usando la linealidad de los tensores,

$$(\alpha + \beta)_{i_1 \dots i_r} = (\alpha + \beta)(e_1, \dots, e_r) = \alpha(e_1, \dots, e_r) + \beta(e_1, \dots, e_r) = \alpha_{i_1 \dots i_r} + \beta_{i_1 \dots i_r}. \quad (1.51)$$

Por tanto

$$d(\alpha + \beta) = \frac{1}{r!} \left( \frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^j} \right) dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r} + \frac{1}{r!} \left( \frac{\partial \beta_{i_1 \dots i_r}}{\partial x^j} \right) dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r} = d\alpha + d\beta. \quad (1.52)$$

2. Como tenemos ya la linealidad probada, podemos considerar formas del tipo

$$\alpha = u dx^{i_1} \wedge \dots \wedge dx^{i_r} \quad \beta = v dx^{j_1} \wedge \dots \wedge dx^{j_s}, \quad (1.53)$$

con  $u(x_1, \dots, x_n), v(x_1, \dots, x_n)$  dos funciones  $C^\infty$  (es decir, 0-formas). Podemos calcular muy fácilmente la derivada exterior de este tipo de formas:

$$d\alpha = \frac{\partial u}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}. \quad (1.54)$$

Entonces, usando la regla del producto con  $u$  y  $v$ ,

$$\begin{aligned}
d(\alpha \wedge \beta) &= d((u \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}) \wedge (v dx^{j_1} \wedge \dots \wedge dx^{j_s})) = \\
&= d((uv \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}) \wedge (dx^{j_1} \wedge \dots \wedge dx^{j_s})) = \\
&= (vdu + u dv) \wedge ((dx^{i_1} \wedge \dots \wedge dx^{i_r}) \wedge (dx^{j_1} \wedge \dots \wedge dx^{j_s})) = \\
&= (du \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}) \wedge (v dx^{j_1} \wedge \dots \wedge dx^{j_s}) + \\
&+ (-1)^r (u dx^{i_1} \wedge \dots \wedge dx^{i_r}) \wedge (dv \wedge dx^{j_1} \wedge \dots \wedge dx^{j_s}) = \\
&= d\alpha \wedge \beta + (-1)^r \alpha \wedge d\beta,
\end{aligned} \tag{1.55}$$

donde el factor  $(-1)^r$  lo introducimos al permutar el  $dv$  con los  $dx^{i_k}$  por la antisimetría de las formas.

3. En esta propiedad estamos utilizando que las componentes de  $\alpha$  son funciones de clase  $C^\infty$ .

$$d^2\alpha = d\left(\frac{1}{r!} \left(\frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^j}\right) dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}\right) = \frac{1}{r!(r+1)!} \left(\frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^j \partial x^k}\right) dx^k \wedge dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}. \tag{1.56}$$

Esto es cero directamente, ya que, por la simetría de las derivadas segundas,

$$\frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^j \partial x^k} dx^k \wedge dx^j = -\frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^j \partial x^k} dx^j \wedge dx^k = -\frac{\partial \alpha_{i_1 \dots i_r}}{\partial x^k \partial x^j} dx^j \wedge dx^k. \tag{1.57}$$

Como las variables  $j$  y  $k$  son ambas mudas (estamos sumando sobre ellas) el último término es exactamente el mismo que el primero y por tanto  $d^2\alpha = 0$ . ■

El siguiente operador necesita de la introducción de una métrica en las variedades. Por simplicidad y para no indagar más de lo necesario en el tema de las métricas, supondremos que en todo momento la métrica es simplemente una matriz diagonal  $\eta_{\mu\nu}$  de unos y menos unos que será la métrica de Minkowski o la identidad cuando empecemos a hablar de teorías gauge.

**Definición 1.27** Sea  $M$  una variedad orientable de dimensión  $n$ ,  $T_p M$  un espacio tangente con base  $\{e_1, \dots, e_n\}$  y  $\eta$  una métrica. Se define el **operador estrella de Hodge** como:

$$\begin{aligned}
\star : \Lambda_p^r(M) &\rightarrow \Lambda_p^{n-r} \\
\star(e^{i_1} \wedge \dots \wedge e^{i_r}) &= \frac{1}{(n-r)!} \sum_{i_{r+1}, \dots, i_n=1}^n \epsilon_{i_1 \dots i_n} \eta^{i_1 i_1} \dots \eta^{i_r i_r} e^{i_{r+1}} \wedge \dots \wedge e^{i_n}
\end{aligned} \tag{1.58}$$

Donde  $\eta^{ij}$  es el elemento  $(i, j)$  de la inversa de  $\eta$  y  $\epsilon_{i_1 \dots i_n}$  son los símbolos de Levi-Civita.

Olvidémonos por un momento de la métrica en la ecuación (1.58). Lo que está haciendo el operador  $\star$  es tomar una  $r$ -forma y dar una  $(n - r)$ -forma cuyos elementos son los productos exteriores que no estaban en la forma original [8]. Ilustrémoslo un poco en dimensión dos usando la identidad como métrica.

$$\star(dx^1) = \frac{1}{(2-1)!} \sum_{j=1}^2 \epsilon_{1j} dx^j = \underbrace{\epsilon_{11}}_0 dx^1 + \epsilon_{12} dx^2 = dx^2 \quad (1.59)$$

$$\star(dx^2) = \frac{1}{(2-1)!} \sum_{j=1}^2 \epsilon_{2j} dx^j = \underbrace{\epsilon_{21}}_{-1} dx^1 + \underbrace{\epsilon_{22}}_0 dx^2 = -dx^1 \quad (1.60)$$

No necesitamos saber la acción de  $\star$  sobre las funciones (las 0-formas) porque  $\star$  conmuta con ellas: esto es consecuencia de que al evaluar en un punto

$$\star(f\omega)(p) = \star(\underbrace{f(p)}_{\in \mathbb{R}} \omega(p)) = f(p) \star \omega(p) \quad (1.61)$$

Por último, observemos que al aplicar dos veces el operador  $\star$  obtenemos una  $n - (n - k) = k$ -forma, lo que nos puede llevar a pensar que tal vez  $\star^2 = \star\star$  es la identidad o algo muy parecido a ella. Se puede demostrar [3] que, para una  $k$ -forma  $\alpha$ ,

$$\star\star\alpha = (-1)^{k(n-k)+s} \alpha, \quad (1.62)$$

donde  $n$  es la dimensión de la variedad y  $s$  es el número de autovalores negativos de la métrica  $\eta_{\mu\nu}$ .

**Definición 1.28** Sea  $M$  una variedad de dimensión  $n$  y  $T_p M$  un espacio tangente,  $\alpha$  una  $r$ -forma. La *codiferencial exterior* se define como la siguiente aplicación:

$$\begin{aligned} \delta_r : \Lambda_p^r(M) &\rightarrow \Lambda_p^{r-1}(M) \\ \delta_r(\alpha) &= (-1)^{n-r-1} \star(d\star\alpha) \end{aligned} \quad (1.63)$$

La secuencia que sigue la codiferencial exterior es la siguiente:

$$\Lambda_p^r(M) \xrightarrow{\star} \Lambda_p^{n-r}(M) \xrightarrow{d} \Lambda_p^{n-r+1}(M) \xrightarrow{\star} \Lambda_p^{n-(n-r+1)}(M) = \Lambda_p^{(r-1)}(M). \quad (1.64)$$

Teniendo en cuenta que  $d^2 = 0$ , tenemos que  $\delta^2 = \star d \star \star d \star \alpha = \star d^2 \star \alpha = 0$ .

Un último apunte que debemos hacer sobre la derivada exterior y la codiferencial exterior es que son operadores adjuntos. Esto significa, por definición, que dada una forma  $\alpha$  y un elemento del espacio dual de las formas  $\beta^*$ , se cumple que

$$(\delta\beta^*)(\alpha) = \beta^*(d\alpha). \quad (1.65)$$

Lo último que vamos a ver de geometría diferencial es la integración en variedades orientables.

En una variedad de dimensión  $n$  solo se puede definir una integral para las  $n$ -formas: estas formas las podemos expresar como  $\omega = g(x^1, \dots, x^n) dx^1 \wedge \dots \wedge dx^n$ , ya que  $\Lambda_p^n(M)$  tiene dimensión 1. También, por cuestiones de integrabilidad, se pide que las formas tengan soporte (el conjunto donde son no nulas) compacto.

Utilizando la carta  $(U, x^1, \dots, x^n)$  (en concreto, lo que se llama el *pullback* de la carta) se puede pasar de  $\omega$  a una  $n$ -forma de  $\mathbb{R}^n$ ,  $\tilde{\omega} = f(x^1, \dots, x^n) dx^1 \wedge \dots \wedge dx^n$ , y se define la integral como

$$\int_U \omega = \int_{\phi(U)} f(x^1, \dots, x^n) dx^1 \wedge \dots \wedge dx^n = \int_{\phi(U)} f(x^1, \dots, x^n) dx^1 \dots dx^n. \quad (1.66)$$

Si el atlas de  $M$  tiene más de una carta y queremos integrar sobre todo  $M$ , hay que tener más cuidado a la hora de escribir los dominios de integración y se ha de introducir el concepto de partición de la unidad [6]. En este trabajo no nos encontraremos con esta situación porque, aunque trabajemos en esferas, que no admiten atlas de una sola carta, las formas que integremos serán cero fuera del dominio de una carta, así que no profundizaremos en el tema.

Se puede demostrar [6], además, que la integral definida en (1.66) cumple las propiedades conocidas de la integral en  $\mathbb{R}^n$ .

Vamos a usar dos resultados muy importantes relacionados con la integración.

**Teorema 1.3** [6] (*de Stokes*) Sea  $M$  una variedad de dimensión  $n$  con frontera y  $\omega$  una  $(n-1)$ -forma con soporte compacto en  $M$ . Entonces

$$\int_M d\omega = \int_{\partial M} \omega. \quad (1.67)$$

Si  $\partial M = \emptyset$ , se impone  $\int_{\partial M} \omega = 0$ .

**Definición 1.29** Una forma diferencial  $\alpha$  se dice **cerrada** si  $d\alpha = 0$ .

**Definición 1.30** Una  $r$ -forma diferencial  $\alpha$  se dice **exacta** si existe una  $(r-1)$ -forma  $\beta$  tal que  $\alpha = d\beta$ .

**Teorema 1.4** [3] (**Lema de Poincaré**) Sea  $M$  una variedad y  $\omega$  una  $k$ -forma cerrada, es decir,  $d\omega = 0$ . Entonces  $\omega$  es localmente exacta, esto es, en cada punto existe una carta  $(U, \phi)$  donde  $\omega = d\alpha$  para cierta  $(k-1)$ -forma  $\alpha$ .

Es importante destacar que el lema de Poincaré es un resultado local:  $\omega = d\alpha$  solo es válido dentro del abierto  $U$ . Fuera de  $U$  no tiene por qué cumplirse.<sup>2</sup>

<sup>2</sup>De hecho, las formas cerradas pero no exactas definen los llamados grupos de cohomología de De Rham [3].

Un ejemplo de esto es la 1-forma  $\omega = \frac{-y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy = \omega_1(x, y)dx + \omega_2(x, y)dy$  sobre  $M = \mathbb{R}^2 \setminus \{(0, 0)\}$ .

Calculemos la derivada exterior de  $\omega$ :

$$\begin{aligned} d\omega &= \frac{\partial\omega_1}{\partial x}dx \wedge dx + \frac{\partial\omega_1}{\partial y}dy \wedge dx + \frac{\partial\omega_2}{\partial x}dx \wedge dy + \frac{\partial\omega_2}{\partial y}dy \wedge dy = \\ &= \frac{1}{(x^2+y^2)^2}(y^2-x^2) \underbrace{dy \wedge dx}_{-dx \wedge dy} + \frac{1}{(x^2+y^2)^2}(y^2-x^2)dx \wedge dy = 0. \end{aligned} \quad (1.68)$$

Si  $\omega$  fuera exacta existiría una función  $f$  tal que

$$\frac{\partial f}{\partial x} = \frac{-y}{x^2+y^2} \quad \frac{\partial f}{\partial y} = \frac{x}{x^2+y^2}. \quad (1.69)$$

Esto tiene como solución  $f(x, y) = \arctan(y/x) + C$ , que no es una función continua en  $M$ , pues tiene discontinuidades en toda la recta  $x = 0$ .

### 1.2.3. Grupos y álgebras de Lie

**Definición 1.31** Sea  $M$  una variedad diferenciable. Si en  $M$  se puede definir una operación binaria  $*$ :  $M \times M \rightarrow M$  tal que  $(M, *)$  es un grupo y tanto  $*$  como la operación que lleva cada elemento a su inverso son de clase  $C^\infty$ , diremos que  $M$  es un **grupo de Lie**.

Para lo que nos interesa, podemos pensar en los grupos de Lie simplemente como los grupos de matrices  $SU(n)$ ,  $U(n)$ ,  $SO(n)$ , etc.

Como posteriormente vamos a tratar los instantones de una teoría construida con  $SU(2)$  como base, vamos a pararnos un momento a ver alguna propiedad de este grupo de Lie. Como grupo de matrices, su definición es

$$SU(2) = \{A \in M_{2 \times 2}(\mathbb{C}) \mid A^\dagger A = \mathbb{1}_2, \det(A) = 1\}, \quad (1.70)$$

donde denotamos  $\mathbb{1}_n$  a la matriz identidad  $n \times n$ . Vamos a encontrar una forma mejor de expresar los elementos de  $SU(2)$  para darle una estructura de variedad. Sean  $a, b, c, d \in \mathbb{C}$ .

Entonces, si  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ,

$$A^\dagger A = \mathbb{1}_2 \Rightarrow A^\dagger = A^{-1} \Rightarrow \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}. \quad (1.71)$$

Ahora, calculando el determinante si  $a = \alpha + \beta i$ ,  $b = \gamma + \delta i$ ,

$$\det A = aa^* + bb^* = |a|^2 + |b|^2 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1. \quad (1.72)$$

Juntando esta última condición con que podemos escribir los elementos de  $\mathbb{C}^2$  como matrices  $2 \times 2$  y, a su vez,  $\mathbb{C}^2$  es homeomorfo a  $\mathbb{R}^4$ , concluimos que

$$SU(2) = \{(\alpha, \beta, \gamma, \delta) \in \mathbb{R}^4 \mid \alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1\} \quad (1.73)$$

es homeomorfo a la esfera en  $\mathbb{R}^4$  o 3-esfera  $S^3$ . A la 3-esfera se le puede dar una estructura de variedad de dimensión 3 con la ayuda de las llamadas proyecciones estereográficas, por lo que  $SU(2)$  también es una variedad de dimensión 3 y además comparte las propiedades topológicas de la 3-esfera: es una variedad compacta y simplemente conexa. Además, se cumple que  $\pi_1(SU(2)) = \pi_2(SU(2)) = 0$  y  $\pi_3(SU(2)) = \mathbb{Z}$  por el teorema 1.2.

Ahora vamos a introducir las álgebras de Lie. Todo grupo de Lie tiene asociada un álgebra que nos será de gran utilidad a la hora de definir los potenciales gauge, entre otras cosas.

**Definición 1.32** Sea  $M$  una variedad de dimensión  $n$ . Un **campo vectorial** sobre  $M$  es una aplicación que a cada punto  $p$  de  $M$  le asigna un vector tangente  $X_p = X^i(p) \frac{\partial}{\partial x^i} \Big|_p \in T_p M$ . Las componentes  $X^i : M \rightarrow \mathbb{R}$  son funciones  $C^\infty$ .

Al igual que con los vectores tangentes, podemos ver los campos vectoriales como operadores que actúan sobre funciones de  $M$  en  $\mathbb{R}$  de la siguiente forma:

$$X(f)(p) = X_p(f) = X^i(p) \frac{\partial f}{\partial x^i} \quad (1.74)$$

**Definición 1.33** Sean  $X$  e  $Y$  dos campos vectoriales sobre una variedad  $M$ . Se define el **corchete de Lie** de  $X$  e  $Y$  como

$$[X, Y](f) = X(Y(f)) - Y(X(f)) \quad (1.75)$$

Notemos que viendo cómo actúa  $XY$  sobre  $f$ , tenemos una combinación de derivadas segundas y por tanto  $XY$  no es un campo vectorial. Sin embargo, el corchete de Lie de dos campos vectoriales sí que es un campo vectorial.

El corchete de Lie satisface las siguientes propiedades [3]:

$$[X, Y] = -[Y, X] \quad (1.76)$$

$$[X, c_1Y_1 + c_2Y_2] = c_1[X, Y_1] + c_2[X, Y_2] \quad (1.77)$$

$$[c_1X_1 + c_2X_2, Y] = c_1[X_1, Y] + c_2[X_2, Y] \quad (1.78)$$

$$[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0 \quad (1.79)$$

Si la variedad  $M$  es un grupo de matrices, resulta que el corchete de Lie coincide exactamente con el conmutador de matrices:  $[A, B] = AB - BA$

**Definición 1.34** Sea  $G$  un grupo de Lie y  $a, g$  dos elementos de  $G$ . La traslación a la izquierda se define como  $L_g(a) = ga$ . Si  $e$  es el elemento neutro de  $G$  y  $A \in T_eG$ , al campo vectorial  $X = dL_g(A)$  se le llama **campo vectorial invariante por la izquierda** asociado a  $A$ .

A  $dL_g : T_eG \rightarrow T_gG$ , que se define tal que  $(dL_g(A))(f) = A(f \circ L_g)$ , se le suele llamar diferencial de  $L_g$ .

El concepto de campo vectorial invariante por la izquierda nos va a servir para definir el álgebra de Lie, pero en la práctica será más útil utilizar la alternativa con el espacio tangente que damos a continuación.

**Definición 1.35** Sea  $G$  un grupo de Lie. Al conjunto de todos los campos vectoriales sobre  $G$  invariantes por la izquierda junto con la operación del corchete de Lie se le llama **álgebra de Lie de  $G$** . Se denota  $(\mathfrak{g}, [\cdot, \cdot])$ .

De manera equivalente, se puede definir el álgebra de Lie como el espacio tangente del elemento neutro de  $G$ .

**Definición 1.36** Sea  $\mathfrak{g}$  un álgebra de Lie. Llamamos **complexificación** de  $\mathfrak{g}$  al álgebra de Lie formada por los elementos de la forma  $A_1 + iA_2$  con  $A_1, A_2 \in \mathfrak{g}$ . El corchete de Lie de la complexificación se define a partir de las propiedades del corchete de Lie de  $\mathfrak{g}$ :

$$[X_1 + iX_2, Y_1 + iY_2] = ([X_1, Y_1] - [X_2, Y_2]) + i([X_1, Y_2] + [X_2, Y_1]) \quad (1.80)$$

El concepto de complexificación nos va a ser útil cuando veamos las representaciones de  $SU(2)$ .

Ya que tenemos el ejemplo de  $SU(2)$  como grupo de Lie, podemos construir el álgebra de Lie de los grupos  $SU(n)$ , que denotamos  $\mathfrak{su}(n)$  [9].

En  $SU(n)$  el elemento neutro es la identidad. Recordemos que todas las matrices de este grupo satisfacen  $A^\dagger A = \mathbb{1}_n$  y  $\det A = 1$ .

Tomemos una curva en  $U(n)$ ,  $\sigma(t) = A(t)$ , con  $A \in U(n)$  tal que  $\sigma(0) = \mathbb{1}_n$ . Entonces, derivando en  $t = 0$ ,

$$0 = \left. \frac{d}{dt} \right|_{t=0} (\mathbb{1}_n) = \left. \frac{d}{dt} \right|_{t=0} (A^\dagger(t)A(t)) = \dot{A}(0)^\dagger A(0) + A^\dagger \dot{A}(0) = \dot{A}(0) + \dot{A}(0)^\dagger, \quad (1.81)$$

Como la derivada en 0 es un vector tangente, tenemos que todos los elementos de  $\mathfrak{u}(n)$  deben cumplir  $B + B^\dagger = 0$ . Veamos si estos son todos considerando la exponencial (de matrices) de una matriz de este tipo:

$$\exp(B)(\exp(B))^\dagger = \exp(B)(\exp(B^\dagger)) = \exp(B + B^\dagger) = \exp(0_{n \times n}) = \mathbb{1}_n \quad (1.82)$$

Por tanto  $\exp(B) \in U(n)$ , por lo que podemos tomar la curva  $\exp(tB)$  y derivarla en  $t = 0$  para obtener  $B = \left. \frac{d}{dt} \right|_{t=0} \exp(tB) \in \mathfrak{u}(n)$ . Este cálculo, de hecho, es un resultado general para cualquier grupo de Lie de matrices, si una matriz está en el álgebra de Lie, su exponencial debe estar en su grupo correspondiente<sup>3</sup>.

Entonces  $\mathfrak{u}(n) = \{B \in M_{n \times n}(\mathbb{C}) | B + B^\dagger = 0\}$ . Usemos ahora que el determinante debe ser siempre 1. Usando que si  $B \in \mathfrak{su}(n)$  entonces  $\exp(B) \in SU(n)$  y la fórmula  $\det(\exp(B)) = e^{\text{tr}(B)}$  tenemos

$$1 = \det(\exp(B)) = e^{\text{tr}(B)} = e^0 \Rightarrow \text{tr}(B) = 0 \quad (1.83)$$

Por tanto  $\mathfrak{su}(n) = \{B \in M_{n \times n}(\mathbb{C}) | B + B^\dagger = 0, \text{tr}(B) = 0\}$  es el conjunto de matrices antihermíticas y con traza 0.

La dimensión de  $\mathfrak{su}(n)$  es  $n^2 - 1$  [8]. En el caso  $n = 2$ , resulta que las matrices de Pauli multiplicadas por  $i$

$$i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad (1.84)$$

forman una base de  $\mathfrak{su}(2)$ . Es muy importante notar que, aunque las matrices de  $\mathfrak{su}(2)$  puedan tener entradas complejas, como espacio vectorial este álgebra de Lie es real porque  $S^3$  es una variedad real. Notemos, por ejemplo, que  $\sigma_3$  no cumple  $\sigma_3 + \sigma_3^\dagger = 0$ , por lo que no está en  $\mathfrak{su}(2)$ .

También se puede ver que  $\mathfrak{su}(2)$  es isomorfa al álgebra de Lie  $\mathfrak{so}(3)$ , correspondiente al grupo de rotaciones en el espacio tridimensional

$$SO(3) = \{A \in M_{3 \times 3}(\mathbb{R}) | AA^T = \mathbb{1}_n, \det(A) = 1\} \quad (1.85)$$

<sup>3</sup>En concreto, en la componente conexas del elemento neutro.

Análogamente al ejemplo que acabamos de hacer, se puede comprobar que es [8]

$$\mathfrak{so}(3) = \{X \in M_{3 \times 3}(\mathbb{R}) \mid X + X^T = 0, \text{tr}(X) = 0\}. \quad (1.86)$$

Para verlo se toma la siguiente base de  $\mathfrak{so}(3)$

$$L_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad (1.87)$$

y se construye el siguiente la siguiente aplicación entre las álgebras de Lie

$$\begin{aligned} \varphi : \mathfrak{so}(3) &\rightarrow \mathfrak{su}(2) \\ L_j &\rightarrow -\frac{i}{2}\sigma_j = t_j \end{aligned} \quad (1.88)$$

$\varphi$  es claramente lineal y biyectivo. Para ver que es homomorfismo (y, por tanto, isomorfismo) tenemos que comprobar que funciona bien con el corchete de Lie, es decir,  $\psi[a, b] = [\psi(a), \psi(b)]$ . Como

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \Rightarrow [t_i, t_j] = -\frac{1}{4}2i\epsilon_{ijk}\sigma_k = \epsilon_{ijk}t_k \quad (1.89)$$

$$[L_i, L_j] = \epsilon_{ijk}L_k, \quad (1.90)$$

podemos comprobar fácilmente que, efectivamente,

$$\psi([L_i, L_j]) = \psi(\epsilon_{ijk}L_k) = \epsilon_{ijk}t_k = [t_i, t_j] = [\psi(L_i), \psi(L_j)]. \quad (1.91)$$

Como último detalle vamos a dar una correspondencia entre  $SU(2)$  y  $SO(3)$ . Construimos el siguiente homomorfismo de grupos con  $X \in \mathfrak{su}(2)$ :

$$\begin{aligned} \varphi : SU(2) &\rightarrow SO(3) \\ U &\rightarrow UXU^\dagger \end{aligned} \quad (1.92)$$

Como  $\varphi(-U) = (-U)X(-U)^\dagger = \varphi(U)$  y  $\varphi(\mathbb{1}_2) = X$ , por lo que la identidad está en el núcleo, tenemos  $\ker(\varphi) = \{\mathbb{1}_2, -\mathbb{1}_2\} \cong \mathbb{Z}_2$ . Resulta que  $\varphi$  también es suprayectivo, por lo que aplicando el primer teorema de isomorfía para homomorfismos de grupos

$$SO(3) \cong SU(2)/\mathbb{Z}_2. \quad (1.93)$$

Esta relación se suele expresar como que  $SU(2)$  "cubre dos veces" a  $SO(3)$ .

Otra forma de ver esta correspondencia que es más familiar a alguien que estudia física es la siguiente: considerando las matrices de Pauli  $\sigma_1, \sigma_2$ , y  $\sigma_3$  (que son una base de  $SU(2)$ ), podemos definir un operador  $R$  de la siguiente manera:

$$R_{ij}(U) = \frac{1}{2} \text{Tr}(\sigma_j U \sigma_i U^{-1}) \quad (1.94)$$

La matriz  $R(U)$  es un elemento de  $SO(3)$  y de la definición de la entrada  $(i,j)$  podemos ver que  $R(U) = R(-U)$ , de forma que podemos volver a argumentar que  $SU(2)/\mathbb{Z}_2 \cong SO(3)$ .

#### 1.2.4. Representaciones de $SU(2)$

El objetivo de esta sección es ver cómo toda la teoría del momento angular que se enseña en los cursos de mecánica cuántica está cimentada en la teoría de representaciones del grupo  $SU(2)$  [10].

**Definición 1.37** Sea  $G$  un grupo de Lie y  $V$  un  $\mathbb{C}$ -espacio vectorial. Llamaremos **representación de  $G$  en  $V$**  a una aplicación de la forma

$$\begin{aligned} \rho : G &\rightarrow \text{Aut}(V) \\ g &\rightarrow \rho_g : V \rightarrow V \end{aligned}$$

Tal que  $\rho$  sea continua y un homomorfismo de grupos.

**Definición 1.38** Sea  $G$  un grupo de Lie,  $\mathfrak{g}$  su álgebra de Lie asociada y  $V$  un  $\mathbb{C}$ -espacio vectorial. Llamaremos **representación de  $\mathfrak{g}$  en  $V$**  a una aplicación de la forma

$$\begin{aligned} \rho : \mathfrak{g} &\rightarrow \text{End}(V) \\ g &\rightarrow \rho_g : V \rightarrow V \end{aligned}$$

Tal que  $\rho$  sea continua y un homomorfismo de álgebras de Lie.

Se puede probar [10] que el grupo de automorfismos  $\text{Aut}(V)$  de un espacio vectorial  $V$  de dimensión  $n < \infty$  es isomorfo al grupo lineal  $GL(n, \mathbb{C})$ , que es un grupo de Lie, y el álgebra de Lie de  $\text{Aut}(V)$  es el espacio  $\text{End}(V)$ , que es isomorfo a  $\mathfrak{gl}(n, \mathbb{C})$ . En la práctica, lo que estamos haciendo con una representación es asociar una matriz cuadrada a cada elemento del grupo o álgebra de Lie.

A simple vista puede parecer poco útil representar lo que en nuestro caso será un grupo de matrices como otro grupo de matrices, pero lo interesante va a ser encontrar las formas diferentes que actúa  $G$  sobre los espacios vectoriales.

Por ejemplo, sabemos que la ecuación de Schrödinger del átomo de hidrógeno es invariante bajo rotaciones en el espacio, y por tanto el espacio de sus soluciones también. Entonces, podemos construir un homomorfismo entre  $SO(3)$  (el grupo de Lie de las rotaciones) y el espacio vectorial que forman los autoestados del átomo de hidrógeno [10], dando lugar a una representación.

Como una representación  $\rho$  es un homomorfismo, se tiene que dados  $g_1, g_2 \in G$ ,  $\rho_{g_1 g_2} = \rho_{g_1} \circ \rho_{g_2}$ . A cada elemento  $\rho_g(v) \in V$  también lo denotaremos  $(\rho(g))(v)$  o  $gv$  como si fuera un producto, ya que podemos ver una representación como una acción de  $G$  sobre  $V$  de la forma  $\phi : G \times V \rightarrow V$  con  $\phi(g, v) = gv$ . Comprobamos que se cumplen las dos condiciones que definen una acción de un grupo [8]:

- $\phi(e_G, x) = e_g x = \rho_{e_G}(x) = Id_V(x) = x$
- $\phi(g, \phi(h, x)) = \phi(g, hx) = \rho_g(\rho_h(x)) = \rho_{gh}(x) = \phi(gh, x)$

Notemos también que el espacio vectorial  $V$  sobre el que definimos la representación no es ninguno en concreto, por lo que elegir una representación es equivalente a elegir el espacio vectorial sobre el que actúan los operadores (matrices) del grupo de Lie.

**Definición 1.39** Sea  $\rho$  una representación de un grupo de Lie  $G$  actuando sobre un espacio vectorial  $V$ . Un subespacio  $W \subseteq V$  se dice **invariante** si para todo  $g \in G$  y para todo  $w \in W$  se tiene  $\rho_g(w) = gw \in W$ .

Si los únicos subespacios invariantes de  $V$  son  $\{0\}$  y  $V$ ,  $\rho$  es una **representación irreducible**. De manera análoga se define una representación irreducible de un álgebra de Lie.

Ahora que tenemos los conceptos básicos de representaciones, vayamos al ejemplo concreto de  $SU(2)$ .

Consideremos el espacio vectorial  $V_m$  ( $m \in \mathbb{N}$ ) formado por polinomios homogéneos (los que cumplen que todos sus sumandos tienen el mismo grado total, por ejemplo  $xy^2 + x^3$ ) de grado  $m$  en dos variables complejas, es decir, polinomios de la forma [10]

$$f(z_1, z_2) = a_0 z_1^m + a_1 z_1^{m-1} z_2 + \dots + a_{m-1} z_1 z_2^{m-1} + a_m z_2^m, \quad a_j \in \mathbb{C} \forall j = 0, \dots, m. \quad (1.95)$$

Notemos que la dimensión de  $V_m$  es  $m + 1$ , pues son las posibles combinaciones que tenemos de hacer que los grados de las dos variables sumen  $m$ .

Vamos a considerar representaciones  $\Pi_m : SU(2) \rightarrow V_m$ . Para cada  $U \in SU(2)$  definimos

$$(\Pi_m(U)f)(z)Uf(z) = f(U^{-1}z) \quad \text{con } z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}. \quad (1.96)$$

Explícitamente, si  $f$  viene dado por (1.95),

$$(\Pi_m(U)f)(z_1, z_2) = \sum_{k=0}^m a_k (U_{11}^{-1}z_1 + U_{12}^{-1}z_2)^{m-k} (U_{21}^{-1}z_1 + U_{22}^{-1}z_2)^k. \quad (1.97)$$

Se comprueba fácilmente que esto de verdad define una representación de  $SU(2)$ :

$$(\Pi_m(U_1)\pi_m(U_2)f)(z) = (\pi_m(U_2)f)(U_1^{-1}z) = f(U_2^{-1}U_1^{-1}z) = (\pi_m(U_1U_2)f)(z). \quad (1.98)$$

Usando la relación entre los elementos de  $SU(2)$  y  $\mathfrak{su}(2)$  con la exponencial de matrices y el hecho de que  $SU(2)$  es simplemente conexo, se puede demostrar [10] que hay una correspondencia entre las representaciones de  $SU(2)$  y las representaciones de su álgebras de Lie. En este caso, la representación  $\pi_m$  tiene asociada la siguiente representación de  $\mathfrak{su}(2)$ :

$$(\pi_m(X)f)(z) = \frac{d}{dt}(f(e^{-tX}z))\Big|_{t=0} \quad (1.99)$$

Definiendo la curva  $z(t) = (z_1(t), z_2(t)) = e^{-tX}z$ , donde  $X \in \mathfrak{su}(2)$  (recordemos que  $e^{-tX}$  es una matriz de  $SU(2)$ ), podemos usar la regla de la cadena para ver la forma explícita de esta representación:

$$\pi_m(X)f = \frac{\partial f}{\partial z_1} \frac{\partial z_1(t)}{\partial t}\Big|_{t=0} + \frac{\partial f}{\partial z_2} \frac{\partial z_2(t)}{\partial t}\Big|_{t=0} \quad (1.100)$$

Ahora bien, en forma vectorial  $\frac{dz(t)}{dt}\Big|_{t=0} = -Xz$ , de forma que sustituyendo obtenemos

$$\pi_m(X)f = -\frac{\partial f}{\partial z_1}(X_{11}z_1 + X_{12}z_2) - \frac{\partial f}{\partial z_2}(X_{21}z_1 + X_{22}z_2) \quad (1.101)$$

Ahora es conveniente pasar a la complexificación de  $\mathfrak{su}(2)$ , que resulta ser isomorfa a  $\mathfrak{sl}(2; \mathbb{C})$ , el conjunto de todas las matrices complejas 2x2 con traza 0. Esto lo hacemos por dos razones:

1. No nos crea problemas porque el espacio  $V_m$  que estamos tratando es complejo y porque en  $\mathfrak{sl}(2; \mathbb{C})$  podemos extender las representaciones  $\pi_m$  como  $\pi_m(X + iY) = \pi_m(X) + i\pi_m(Y)$  sin ningún tipo de problema [10], de forma que solo tenemos que cambiar las matrices de  $\mathfrak{su}(2)$  por matrices de  $\mathfrak{sl}(2; \mathbb{C})$ .
2. En la complexificación existe una base de matrices muy conveniente:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (1.102)$$

Notemos que ninguna de estas tres matrices está en  $\mathfrak{su}(2)$ , pero podemos construirlas a partir de una combinación compleja de las matrices definidas en (1.84). Calculemos ahora cuánto vale  $\pi_m$  en cada elemento de esta base usando (1.101):

$$\begin{aligned}\pi_m(H) &= -z_1 \frac{\partial}{\partial z_1} + z_2 \frac{\partial}{\partial z_2} \\ \pi_m(X) &= -z_2 \frac{\partial}{\partial z_1}, \quad \pi_m(Y) = -z_1 \frac{\partial}{\partial z_2}\end{aligned}\tag{1.103}$$

Por último, aplicamos estos operadores a un polinomio base de  $V_m$ ,  $z_1^{m-k} z_2^k$  con  $k = 0, 1, \dots, m$ .

$$\pi_m(H)(z_1^{m-k} z_2^k) = -z_1((m-k)z_1^{m-k-1} z_2^k) - z_2(kz_1^{m-k} z_2^{k-1}) = (-m+2k)z_1^{m-k} z_2^k\tag{1.104}$$

$$\pi_m(X)(z_1^{m-k} z_2^k) = -z_2(m-k)z_1^{m-k-1} z_2^k = -(m-k)z_1^{m-k-1} z_2^{k+1}\tag{1.105}$$

$$\pi_m(Y)(z_1^{m-k} z_2^k) = -z_1 k z_1^{m-k} z_2^{k-1} = -k z_1^{m-k+1} z_2^{k-1}\tag{1.106}$$

Vemos que  $z_1^{m-k} z_2^k$  es un autovector de  $\pi_m(H)$  con autovalor  $(-m+2k)$  y  $\pi_m(X)$  y  $\pi_m(Y)$  son operadores que suben o bajan el exponente de  $z_2$ , respectivamente. Esto debería recordarnos a los operadores  $J_z, J_+$  y  $J_-$  de la teoría del momento angular que se da en mecánica cuántica.

Veamos que  $X$  e  $Y$  realmente dan los operadores escalera. En primer lugar, tenemos (salvo un factor 2) las mismas relaciones de conmutación que los operadores de momento angular

$$[H, X] = 2X \quad [H, Y] = -2Y \quad [X, Y] = H\tag{1.107}$$

Comparemos con  $[J_z, J_\pm] = \pm \hbar J_\pm$  y  $[J_+, J_-] = \hbar J_z$ . Ahora, dado un autovector de  $\pi_m(H)$ , llamémoslo  $v$ , con autovalor  $\alpha \in \mathbb{C}$ , veamos qué le pasa al vector  $\pi_m(X)(v)$ . Usando que los  $\pi_m$  son homomorfismos de álgebras de Lie,

$$[\pi_m(H), \pi_m(X)] = \pi_m([X, H]) = 2\pi_m[X] \Rightarrow\tag{1.108}$$

$$\Rightarrow \pi_m(H)\pi_m(X)(v) = \pi_m(X)\pi_m(H)(v) + 2\pi_m(X)(v) = \pi_m(X)(\alpha v) + 2\pi_m(X)v = (\alpha+2)\pi_m(X)v\tag{1.109}$$

Por tanto, o bien  $\pi_m(X)(v) = 0$  o bien es otro autovector de  $\pi_m(H)$  con autovalor  $\alpha+2$ . Análogamente,  $\pi_m(Y)(v)$  tendría autovalor  $\alpha-2$ . Notemos que, como las representaciones son de dimensión finita y  $H$  es una matriz hermítica, sus autovectores correspondientes a autovalores distintos tienen que ser linealmente independientes, por lo que en algún momento aplicar  $\pi_m(X)$  o  $\pi_m(Y)$  tiene que dar cero necesariamente, al igual que los operadores escalera.

Por tanto, podemos ver una muy clara similitud entre los operadores de momento angular y las representaciones  $\pi_m$ , así como entre los estados que se suelen denotar  $|j, m\rangle$  y los elementos de los

espacios  $V_m$ . De hecho, el número  $m$  lo podemos ver como el doble del "espín" de una partícula ( $m = 2j$ ), de forma que tengamos espacios de dimensión  $2j + 1$ .

Damos ahora los siguientes resultados sin demostración (se pueden consultar en [10], secciones 4.2, 4.6 y 4.7)

**Teorema 1.5** *Todas las representaciones  $\pi_m$  para  $m \geq 0$  son irreducibles (y por tanto las  $\pi_m$  también). Además, todos los autovalores de  $\pi_m(H)$  son números enteros.*

**Teorema 1.6** *Para todo natural  $m \geq 0$  existe una representación irreducible  $\pi$  de  $\mathfrak{sl}(2; \mathbb{C})$  de dimensión  $m + 1$  y dicha representación  $\pi$  es isomorfa a la representación  $\pi_m$  de  $\mathfrak{su}(2)$ .*

**Teorema 1.7** *Sea  $\phi$  un isomorfismo entre  $\mathfrak{su}(2)$  y  $\mathfrak{so}(3)$  y  $\sigma_m = \pi_m \circ \phi^{-1}$  una representación irreducible de  $\mathfrak{so}(3)$ . Definamos  $j = \frac{m}{2}$ . Si  $j$  es un número entero, entonces existe una representación  $\Sigma_m$  del grupo  $SO(3)$  tal que  $\Sigma_m(e^X) = e^{\sigma_m(X)}$  para todo  $X \in \mathfrak{so}(3)$ . Si  $j$  es un número semientero, no existe tal representación.*

Estos resultados nos dicen que cualquier representación irreducible de  $SU(2)$  se puede ver como una sobre los espacios  $V_m$  y además tiene que jugar bajo las mismas normas (relaciones de conmutación de la base, autovalores) que las  $\pi_m$ .

El teorema 1.7 es el resultado clave para la física: si el espín de una partícula es entero (es decir, si es un bosón) entonces podemos asociar sus estados de momento angular a representaciones de  $SO(3)$  (el grupo de las rotaciones en el espacio) y la propia partícula sigue las leyes de las rotaciones en el espacio. En cambio, para los fermiones esto no pasa. Como no podemos asociarles una representación de  $SO(3)$ , nos encontramos con la estructura de  $SU(2)$  que, si recordamos, se puede ver como dos "recubrimientos" de  $SO(3)$ . Esto explica, por ejemplo, que necesitemos rotar un ángulo de  $4\pi$  a un fermión para que vuelva a su estado inicial.

La representación con  $j = 1/2$  (o  $m = 1$ ), que está generada por las matrices de Pauli, tiene la particularidad de que a partir de productos tensoriales de sí misma puede generar todas las demás representaciones irreducibles de  $SU(2)$ . Esta característica define lo que se llama una **representación fundamental** del grupo o álgebra de Lie [3].

También destacamos que la representación fundamental de  $SO(3)$ , que se corresponde [10] con la de  $j = 1$  (o  $m = 2$ ), se suele llamar representación vectorial en física.

## Capítulo 2

# Efecto túnel y Yang-Mills

### 2.1. Integrales de camino

En mecánica cuántica, el propagador se usa para calcular la probabilidad de que un sistema pase de un estado inicial  $|x_i\rangle$  en un tiempo inicial  $t_i$  a un estado  $|x_f\rangle$  en un tiempo final  $t_f$ . Se define como:

$$\mathcal{K}(x_f, t_f; x_i, t_i) = \langle x_f | U(t_f, t_i) | x_i \rangle \quad (2.1)$$

Donde  $U(t_f, t_i) = e^{-\frac{i}{\hbar}(t_f - t_i)H}$  es el operador de evolución temporal cuando el hamiltoniano no depende del tiempo, es decir, el hamiltoniano es de la forma

$$H = \frac{P^2}{2m} + V(x) \quad (2.2)$$

El formalismo de integrales de camino o *path integrals* separa el camino entre  $x_i$  y  $x_f$  en  $N \rightarrow \infty$  trozos. A continuación damos un cálculo detallado de la integral de camino para el caso de un bosón. La integral de camino de un fermión necesita muchas más consideraciones matemáticas, como la introducción de los números de Grassmann [3], y no la necesitaremos para hablar de instantones. Por comodidad, usaremos unidades naturales de forma que  $\hbar = 1 = c$ .

En primer lugar, calcularemos el propagador para un intervalo de tiempo infinitesimal,

$$t_f - t_i = \epsilon \ll 1. \quad (2.3)$$

Como primer paso, introducimos la expresión del operador identidad en términos de la base de momentos en (2.1),

$$\langle x_f | e^{-i\epsilon H} | x_i \rangle = \int dk \langle x_f | e^{-i\epsilon H} | k \rangle \langle k | x_i \rangle = \int \frac{dk}{2\pi} e^{-ikx_i} e^{-i\epsilon H_x} e^{ikx_f}, \quad (2.4)$$

donde hemos usado que  $|k\rangle$  es un autoestado del momento lineal ( $P|k\rangle = i\partial_x|k\rangle$ ) para obtener  $H_x = \frac{-1}{2m}\partial_x^2 + V(x)$  y que  $\langle k|x\rangle = \frac{1}{\sqrt{2\pi}} e^{ikx}$ .

Ahora, utilizamos la relación de conmutación entre  $\partial_x$  y  $e^{ikx}$  [3]

$$\partial_x e^{ikx} = ik e^{ikx} + e^{ikx} \partial_x = e^{ikx} (ik + \partial_x) \Rightarrow \partial_x^n e^{ikx} = e^{ikx} (ik + \partial_x)^n, \quad (2.5)$$

de forma que, aplicando el caso  $n=2$  y conmutando con el operador de evolución temporal,

$$e^{-i\epsilon\left(\frac{-1}{2m}\partial_x^2 + V(x)\right)} e^{ikx_f} = e^{ikx_f} e^{-i\epsilon\left(\frac{-1}{2m}(ik+\partial_x)^2 + V(x)\right)}. \quad (2.6)$$

Desarrollando el cuadrado y volviendo a la ecuación del propagador,

$$\langle x_f | e^{-i\epsilon H} | x_i \rangle = \int \frac{dk}{2\pi} e^{i\left(\frac{\epsilon k^2}{2m} - k(x_f - x_i)\right)} e^{-i\epsilon\left(-i\frac{k}{m}\partial_x - \frac{\partial_x^2}{2m} + V(x)\right)} \quad (2.7)$$

Para aligerar la notación, podemos hacer el cambio de variable  $p = \sqrt{\epsilon/(2m)} k$ . Si completamos el cuadrado de la primera exponencial con un término  $\frac{m}{2\epsilon}(x_f - x_i)^2$  (que puede salir fuera de la integral), obtenemos

$$\langle x_f | e^{-i\epsilon H} | x_i \rangle = \sqrt{\frac{2m}{\epsilon}} e^{\frac{im}{2\epsilon}(x_f - x_i)^2} \int \frac{dp}{2\pi} e^{-i\left(p + \sqrt{\frac{m}{2\epsilon}}(x_f - x_i)\right)^2} e^{-i\epsilon\left(\sqrt{\frac{2}{\epsilon m}} p \partial_x - \frac{\partial_x^2}{2m} + V(x)\right)}. \quad (2.8)$$

Ahora podemos hacer un desarrollo de Taylor a primer orden en la segunda exponencial y definir un segundo cambio de variable  $q = p + \sqrt{\frac{m}{2\epsilon}}(x_f - x_i)$ . Teniendo cuidado con que el primer término lleva un  $\sqrt{\epsilon}$  dividiendo el propagador queda

$$\langle x_f | e^{-i\epsilon H} | x_i \rangle = \sqrt{\frac{2m}{\epsilon}} e^{\frac{im}{2\epsilon}(x_f - x_i)^2} \int \frac{dq}{2\pi} e^{-iq^2} \cdot \left(1 + (-i\epsilon)V(x) + \frac{i\epsilon^2}{2\epsilon}(x_f - x_i)\partial_x V(x)\right) + O(\epsilon^2) + O(\epsilon|x_f - x_i|^2). \quad (2.9)$$

Aquí, además de los términos de orden  $\epsilon^2$ , estamos despreciando los términos de orden  $\epsilon|x_f - x_i|^2$  porque al ser  $\epsilon$  un tiempo muy pequeño, la distancia entre  $x_f$  y  $x_i$  también será muy pequeña. Por último, notando que

$$V\left(\frac{x_f + x_i}{2}\right) = V\left(x_f - \left(\frac{x_f - x_i}{2}\right)\right) = V(x) + \frac{1}{2}(x_f - x_i)\partial_x V(x) + O(|x_f - x_i|^2), \quad (2.10)$$

haciendo la integral gaussiana de (2.9) y reconstruyendo la exponencial en el desarrollo de Taylor, tenemos la expresión final para el propagador en el tiempo infinitesimal  $\epsilon$ :

$$\langle x_f | e^{-i\epsilon H} | x_i \rangle = \sqrt{\frac{m}{2\pi i \epsilon}} e^{\frac{i\epsilon m}{2}\left(\frac{x_f - x_i}{\epsilon}\right)^2} e^{-i\epsilon V\left(\frac{x_f + x_i}{2}\right)}. \quad (2.11)$$

Notemos que  $\frac{x_f - x_i}{\epsilon}$  es una distancia dividida por un tiempo, por lo que en el límite  $\epsilon \rightarrow 0$  (donde también tenemos  $x_f \approx x_i$ ) podemos verla como una velocidad, de forma que el exponente de la exponencial se puede ver aproximadamente como la acción clásica [3]

$$\frac{i\epsilon m}{2} \left( \frac{x_f - x_i}{\epsilon} \right)^2 - i\epsilon V \left( \frac{x+y}{2} \right) \approx i\epsilon \left( \frac{1}{2} m v^2 + V(x) \right) \approx i \int_0^\epsilon L(x, v) dt = i\Delta S. \quad (2.12)$$

Ya tenemos el propagador para un instante de tiempo infinitesimal. Extendamos este resultado al caso general dividiendo el intervalo  $(t_i, t_f)$  en  $N$  subintervalos de longitud  $\epsilon$ :

$$t_k = t_i + k\epsilon, \quad \epsilon = \frac{t_f - t_i}{N}. \quad (2.13)$$

De la misma manera, dividimos el intervalo  $(x_i, x_f)$  en otros  $N$  subintervalos. Introduciendo las representaciones de la identidad según los autoestados de posición en la expresión del propagador,

$$\begin{aligned} \langle x_N | e^{-i(t_N - t_0)H} | x_0 \rangle &= \langle x_N | e^{-i(t_N - t_{N-1})H} \dots e^{-i(t_1 - t_0)H} | x_0 \rangle = \\ &= \int dx_{N-1} \dots dx_1 \langle x_N | e^{-i(t_N - t_{N-1})H} | x_{N-1} \rangle \langle x_{N-1} | e^{-i(t_{N-1} - t_{N-2})H} | x_{N-2} \rangle \dots \langle x_1 | e^{-i(t_1 - t_0)H} | x_0 \rangle. \end{aligned} \quad (2.14)$$

Acabamos de calcular cuánto valen los términos del integrando, por lo que si definimos

$$\Delta S_k = \epsilon \left( \frac{1}{2} m \left( \frac{x_k - x_{k-1}}{\epsilon} \right)^2 - V \left( \frac{x_k + x_{k-1}}{2} \right) \right), \quad (2.15)$$

en el límite  $N \rightarrow \infty$  tenemos

$$\langle x_N | e^{-i(t_N - t_0)H} | x_0 \rangle = \lim_{N \rightarrow \infty} \left( \frac{m}{2\pi i \epsilon} \right)^{N/2} \int \prod_{j=1}^{N-1} dx_j e^{i \sum_{k=1}^N \Delta S_k}, \quad (2.16)$$

Al estar fijando los puntos  $x_k$  y estar a orden uno en el desarrollo de Taylor, estamos construyendo un camino lineal a trozos de  $x_i \equiv x_0$  a  $x_f \equiv x_N$ . Esto nos permite cambiar la suma de las acciones por una integral del lagrangiano desde  $t_i$  a  $t_f$ , de forma similar a como lo hicimos en (2.12). Si ahora definimos

$$\mathcal{D}[x] = \lim_{N \rightarrow \infty} \left( \frac{m}{2\pi i \epsilon} \right)^{N/2} \prod_{j=1}^{N-1} dx_j. \quad (2.17)$$

obtenemos la expresión de la integral de caminos:

$$\langle x_N | e^{-i(t_N - t_0)H} | x_0 \rangle = \int \mathcal{D}[x] e^{i \int_{t_i}^{t_f} dt \left( \frac{1}{2} m \dot{x}^2 - V(x) \right)} = \int \mathcal{D}[x] e^{iS[x]}. \quad (2.18)$$

Aunque esta expresión venga de dos términos que en principio podrían divergir al tomar el límite  $N \rightarrow \infty$ , la combinación de ambos al calcularse en los problemas concretos sí es una función bien definida [3].

### 2.1.1. WKB y probabilidades de efecto túnel

El formalismo de integrales de camino nos será muy útil a la hora de calcular probabilidades de efecto túnel. Por ejemplo, el resultado de la conocida aproximación WKB para calcular este tipo de probabilidades se deduce de forma bastante sencilla usando la integral de camino.

Podemos separar cualquier trayectoria  $x(t)$  en una parte que satisface las ecuaciones del movimiento clásicas  $x_c(t)$  y otra que cuantifique los efectos cuánticos  $r(t)$ . Imponiendo que los extremos de la trayectoria clásica y cuántica coincidan ( $r(t_i) = r(t_f) = 0$ ), podemos calcular la acción

$$\dot{x}^2 = \frac{d}{dt}(x_c + r)^2 = \dot{x}_c^2 + 2\dot{x}_c\dot{r} + \dot{r}^2 = \dot{x}_c^2 + \partial_t(r\dot{x}_c) - r\ddot{x}_c + \dot{r}^2 \quad (2.19)$$

El segundo término es una derivada total, por lo que a efectos de integrarlo, no tendrá ninguna contribución porque los términos de frontera de  $r$  son 0 y podemos despreciarlo. Expandimos ahora el potencial con una serie de Taylor

$$V(x) = V(x_c) + V(r) = V(x_c) + \left. \frac{\partial V}{\partial x} \right|_{x_c} r + \sum_{n=2}^{\infty} \frac{r^n}{n!} \left. \frac{\partial^n V}{\partial x^n} \right|_{x_c}. \quad (2.20)$$

Juntándolo todo:

$$L(x, t) = \frac{1}{2}m\dot{x}_c^2 - V(x_c) - r(m\ddot{x}_c + \left. \frac{\partial V}{\partial x} \right|_{x_c}) + \frac{1}{2}m\dot{r}^2 - \sum_{n=2}^{\infty} \frac{r^n}{n!} \left. \frac{\partial^n V}{\partial x^n} \right|_{x_c}. \quad (2.21)$$

El tercer sumando son las ecuaciones de Euler-Lagrange evaluadas en la trayectoria clásica, así que es 0. Los dos últimos resultan ser despreciables en el régimen semiclásico [11] (que es donde se aplica la aproximación WKB), por lo que el lagrangiano es el de una partícula con una trayectoria clásica. El coeficiente que modulará la probabilidad de efecto túnel será, entonces,

$$K = e^{\frac{i}{\hbar}S(x_c)} = e^{\frac{i}{\hbar} \int_{t_i}^{t_f} dt L(x_c, t)} \quad (2.22)$$

Haciendo la transformada de Legendre que relaciona hamiltoniano y lagrangiano,

$$S(x_c) = \int dt L(x, \dot{x}, t) = \int dt (p\dot{x}_c - H) = \int dx p - E(t_f - t_i). \quad (2.23)$$

En la última igualdad hemos usado que el hamiltoniano es la energía total si no depende del tiempo, y que es constante.

Para admitir energías menores que el potencial (y que por tanto pueda haber efecto túnel) podemos expresar  $p(x) = i\sqrt{2m(V(x) - E)} = i|p(x)|$  de forma que sustituyendo en (2.22) obtenemos

$$K = e^{-\int_{x_c} dx |p(x)|} \quad (2.24)$$

El cuadrado de esto es la fórmula que da la aproximación WKB para la probabilidad de efecto túnel.

El paso  $p \rightarrow i|p|$  para modelar lo que sería una energía cinética negativa es la parte esencial de todo esto. Equivalentemente, como  $p = m\dot{x}$  podemos decir que estamos cambiando el tiempo como  $t \rightarrow -i\tau$ . A esto se le llama rotación de Wick y aunque pueda parecer inocua, tiene efectos físicos muy relevantes. En primer lugar, el signo menos en  $\tau$  cambia la métrica de Minkowski a la identidad, de forma que el espacio-tiempo pasa a ser  $\mathbb{R}^4$  (el espacio euclídeo) en vez del espacio-tiempo de Minkowski.

También hay cambios en el lagrangiano y la acción:

$$L(x) = \frac{m}{2} \left( \frac{\partial x}{\partial t} \right)^2 - V = \frac{1}{(-i)^2} \frac{m}{2} \left( \frac{\partial x}{\partial \tau} \right)^2 \equiv -L_E(x) \quad (2.25)$$

$$S[x] = \int dt L(x) = \int (-id\tau)(-L_E(x)) \equiv iS_E[x] \quad (2.26)$$

El subíndice  $E$  es de euclídeo. Haciendo la rotación de Wick a la path integral (2.18) tenemos

$$\mathcal{K} = \int \mathcal{D}_E[x] e^{-S_E[x]}. \quad (2.27)$$

Si queremos extender esto a una teoría de campos más general, en vez de un lagrangiano  $L(x, \dot{x})$  tendremos una densidad lagrangiana  $\mathcal{L}(\phi, \partial_\mu \phi)$ , de manera que nuestra acción será

$$S[\phi] = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi). \quad (2.28)$$

Bajo la rotación de Wick, la densidad lagrangiana y la acción siguen cambiando como en las ecuaciones (2.12) y (2.26), por lo que el propagador (2.27) será

$$\mathcal{K} = \int \mathcal{D}_E[\phi] e^{-S_E[\phi]}. \quad (2.29)$$

Podemos concluir que si encontramos soluciones con acción euclídea finita, podemos calcular el propagador y obtendremos una probabilidad distinta de 0 de que el efecto túnel entre ellas ocurra (y, por tanto, ocurrirá en algún momento). Los instantones son soluciones de este tipo.

Además, este efecto túnel se puede ver como una trayectoria clásica porque, como el lagrangiano (y por tanto el potencial) cambia de signo, lo que antes eran barreras prohibidas clásicamente, se convierten en caminos de energía cinética positiva tras hacer la rotación de Wick [12].

## 2.2. Conexiones en teorías gauge

### 2.2.1. Introducción a los fibrados

En las teorías de campos es muy común introducir una derivada covariante en vez de usar la derivada ordinaria para que el lagrangiano cambie bajo transformaciones locales (transformaciones gauge) igual que los campos de la teoría. Todo esto lleva a sus espaldas la geometría diferencial de fibrados. Con la introducción a las variedades que hemos visto en la sección 1.2 podemos dar lo básico que necesitaremos para entender las teorías gauge.

**Definición 2.1** *Un fibrado (diferenciable) es una tupla  $(E, \pi, M, F, G)$ , normalmente denotada como  $E \xrightarrow{\pi} M$ , formada por los siguientes elementos:*

- Una variedad  $E$  llamada espacio total.
- Una variedad  $M$  llamada espacio base o variedad base.
- Una variedad  $F$  llamada fibra.
- Una aplicación suprayectiva  $\pi : E \rightarrow M$  llamada proyección tal que para todo punto  $p \in M$ ,  $\pi^{-1}(p) = F_p \cong F$ .  $F_p$  es homeomorfa a  $F$  y se llama fibra en  $p$ .
- Un grupo de Lie  $G$  que actúa sobre  $F$  multiplicando a la izquierda, normalmente llamado grupo de estructura.

La estructura de fibrado también está fijada por dos tipos de aplicaciones:

- Dado un conjunto  $\{U_i\}_{i \in I}$  de abiertos de  $M$  tal que  $\bigcup_{i \in I} U_i = M$ , se puede definir un conjunto de homeomorfismos  $\phi_i : U_i \times F \rightarrow \pi^{-1}(U_i)$  de clase  $C^\infty$  llamados **trivializaciones locales**. Estas aplicaciones cumplen que si  $p \in U_i$  y  $q \in F$ ,  $\pi \circ \phi_i(p, q) = \pi \circ \phi_{i,p}(q) = p$ .
- Dado un par de abiertos tales que  $U_i \cap U_j \neq \emptyset$  y  $p \in U_i \cap U_j$ , impondremos que la aplicación  $t_{ij}(p) = \phi_{i,p}^{-1} \circ \phi_{j,p}$  sea un elemento del grupo de Lie  $G$ , de manera que existe la siguiente relación entre  $\phi_i$  y  $\phi_j$ :

$$\phi_j(p, q) = \phi_i(p, t_{ij}(p)f). \quad (2.30)$$

A las  $t_{ij}$  se les llama **funciones de transición**.

Si  $E = M \times F$ , se dice que el fibrado es trivial. El nombre de trivialización local viene de que la inversa  $\phi_i^{-1}$  identifica el conjunto  $\pi^{-1}(U_i) \subseteq E$  con el fibrado trivial  $U_i \times F$ .

Aclaremos un poco lo que son las funciones de transición. De la misma forma que hicimos para los espacios vectoriales con las representaciones, se puede definir una acción de un grupo de Lie en una variedad cualquiera. Tomemos dos abiertos de  $M$ ,  $U_i$  y  $U_j$ , tales que  $U_i \cap U_j \neq \emptyset$ . Entonces tenemos dos trivializaciones locales  $\phi_i$  y  $\phi_j$  definidas en  $U_i \cap U_j$ . Si tomamos un punto  $u \in E$  tal que al hacer la proyección  $\pi(u) = p \in U_i \cap U_j$ , usando las inversas de las trivializaciones tenemos que

$$\phi_i^{-1}(u) = (p, f_i), \quad \phi_j^{-1}(u) = (p, f_j), \quad (2.31)$$

donde  $f_i$  y  $f_j$  son dos elementos de la fibra  $F$ . La función de transición nos da la relación entre  $f_i$  y  $f_j$  a partir de la acción de  $G$  sobre las fibras:

$$\begin{aligned} t_{ij} : U_i \cap U_j &\rightarrow G \\ p &\rightarrow t_{ij}(p) : F \rightarrow F \\ f_i &\rightarrow f_j = t_{ij}(p)f_j \end{aligned} \quad (2.32)$$

Por consistencia, las funciones de transición deben cumplir

$$t_{ii}(p) = Id : F \rightarrow F \quad (2.33)$$

$$t_{ij}^{-1}(p) = t_{ji}(p) \quad (2.34)$$

$$t_{ij}(p)t_{jk}(p) = t_{ik}(p) \quad (2.35)$$

**Definición 2.2** Un fibrado  $(E, \pi, M, F, G)$  cuyas fibras sean espacios vectoriales se denomina **fibrado vectorial**.

Un fibrado  $(E, \pi, M, F, G)$  cuyas fibras sean isomorfas al grupo de Lie  $G$  se denomina **fibrado principal**.

**Definición 2.3** Sea  $M$  una variedad de dimensión  $n$ . Llamamos **fibrado tangente de  $M$**  (denotado  $TM$ ) a la unión disjunta de todos los espacios tangentes de  $M$ :

$$TM = \bigsqcup_{p \in M} T_p M = \bigcup_{p \in M} \{(p, x) | x \in T_p M\} \quad (2.36)$$

Equivalentemente, el fibrado tangente es un fibrado vectorial  $TM \xrightarrow{\pi} M$  tal que para todo  $p \in M$ ,  $\pi^{-1}(p) = T_p M$ .

**Definición 2.4** Sea  $E \xrightarrow{\pi} M$  un fibrado. Se llama **sección** del fibrado a una aplicación  $s : M \rightarrow E$  tal que  $\pi \circ s = Id_M$ .

Las secciones las vamos a usar en el contexto del fibrado tangente. La condición  $\pi \circ s = Id_M$  nos dice que para cada  $p \in M$ ,  $s(p)$  es un elemento de  $\pi^{-1}(p) = F_p = T_pM$ , el espacio tangente. Por tanto los campos vectoriales no son más que las secciones del fibrado tangente, pues son aplicaciones que a cada punto de  $M$  le asignan un vector tangente, que es un elemento de la fibra.

Otro ejemplo de secciones en un fibrado lo encontramos en la mecánica cuántica. Las funciones de onda son secciones de lo que se llama un fibrado de línea (*line bundle* en inglés), que es un fibrado vectorial cuyas fibras son isomorfas a  $\mathbb{C}$  [3], puesto que son funciones del espacio-tiempo (la variedad base) en la fibra  $\mathbb{C}$ .

**Definición 2.5** Sea  $E \xrightarrow{\pi} M$  un fibrado vectorial y  $(U, \xi)$  una carta sobre  $M$  con  $\pi^{-1}(U) \cong M \times \mathbb{C}^k$  ó  $\mathbb{R}^k$ . Dado  $p \in U$ , un **referencial local** en  $U$  es una tupla de secciones  $\{e_1(p), \dots, (e_k(p))\}$  que forma una base de  $F_p$ .

En términos del fibrado tangente, un referencial local en  $U$  es una tupla  $(X_1, \dots, X_n)$  de  $n$  campos vectoriales de forma que para todo  $p \in U$ ,  $(X_{1p}, \dots, X_{np})$  es una base ordenada del espacio tangente  $T_pM$ .

Si tenemos dos abiertos  $U_i$  y  $U_j$  con  $p \in U_i \cap U_j$ , podemos considerar un referencial  $\{e_1(p), \dots, e_k(p)\}$  en  $U_i$  y un referencial  $\{\hat{e}_1(p), \dots, \hat{e}_k(p)\}$  de  $U_j$  y preguntarnos cómo podemos relacionar ambos referenciales. Como las secciones forman una base de la fibra por definición, cualquier vector  $V \in \pi^{-1}(p) = F_p$  se expresa como [3]

$$V = V^\alpha e_\alpha = \hat{V}^\alpha \hat{e}_\alpha \quad (2.37)$$

Si introducimos la matriz  $G(p)$  de cambio de base

$$\hat{e}_\beta(p) = e_\alpha(p) G(p)^\alpha_\beta, \quad (2.38)$$

obtenemos que  $\hat{V}^\beta = G^{-1}(p)^\beta_\alpha V^\alpha$ . Por definición, las funciones de transición también nos relacionaban de esta manera vectores de la fibra, por lo que concluimos que las funciones de transición se pueden ver como matrices que describen un cambio de referencial.

### 2.2.2. Conexiones

Supongamos que nuestro espacio-tiempo es una variedad  $M$ . Fijado un referencial local  $(s_1, \dots, s_n)$  podemos escribir (localmente) cualquier sección  $\Phi$  de un fibrado vectorial de la siguiente forma [8]:

$$\Phi = \Phi^j s_j. \quad (2.39)$$

Las  $s_j$  son secciones que cuando evaluemos  $\Phi$  en un punto serán vectores de una base de la fibra y las  $\Phi^j$  son funciones  $C^\infty$  que hacen el papel de las componentes de  $\Phi$ .

Si permitimos que las formas diferenciales tengan llegada a cualquier espacio vectorial en vez de solo a  $\mathbb{R}$ , esta sección  $\Phi$  se puede ver como una función  $C^\infty$  (una 0-forma) que sale de  $M$  y toma valores en el fibrado ya que, de nuevo, a cada punto le asigna un elemento de la fibra [6]. De forma similar al resto de formas vistas hasta ahora, podemos definir una  $k$ -forma sobre un fibrado vectorial  $E \xrightarrow{\pi} M$ . En general, estas formas son del tipo  $\alpha = \alpha^j s_j$  con  $\alpha^j$  una  $k$ -forma con llegada en  $\mathbb{R}$  como las de la sección 1.2 y  $s_j$  los elementos del referencial, de forma que lo que hacen es partir de  $k$  vectores tangentes de  $M$  y devolver un elemento del fibrado  $E$ .

Al espacio vectorial generado por las  $k$ -formas que tienen llegada en el fibrado  $E$  lo llamaremos  $\Omega^k(E)$ .

Si fijamos un punto  $p \in M$  y evaluamos  $\alpha$  en un conjunto de vectores tangentes  $v_1, \dots, v_k \in T_p M$ :

$$(\alpha^j s_j)(v_1, \dots, v_k) = \underbrace{\alpha^j(v_1, \dots, v_k)}_{\in \mathbb{R}} \underbrace{s_j(p)}_{\in F_p} \in F_p \quad (2.40)$$

Ahora vamos a introducir las conexiones, que serán lo que llamaremos derivada covariante [8].

**Definición 2.6** Sea  $E \xrightarrow{\pi} M$  un fibrado. Una conexión es una aplicación  $D : \Omega^k(E) \rightarrow \Omega^{k+1}(E)$  tal que sea lineal y que para toda función  $f : M \rightarrow \mathbb{R}$  de clase  $C^\infty$  y para toda  $k$ -forma  $\alpha$  de  $\Omega^k(E)$  se cumpla

$$D(f\alpha) = df\alpha + fD\alpha. \quad (2.41)$$

Veamos cómo tiene que ser la derivada covariante a partir de las secciones de la forma (2.39):

$$D\Phi = D(\underbrace{\Phi^j}_{\text{funciones}} s_j) = d\Phi^j s_j + \Phi^j Ds_j. \quad (2.42)$$

A priori no sabemos cómo actúa  $D$  sobre el referencial, pero viendo que el primer sumando es una 1-forma, necesitamos que el segundo también lo sea. Por tanto  $Ds_j$  debe poder expresarse combinación

lineal de 1-formas  $A_j^k \in \Lambda_p^1(E)$  de forma que

$$D\Phi = d\Phi^j s_j + \Phi^j A_j^k s_k. \quad (2.43)$$

Notemos que estas formas  $A_j^k = A_{j\mu}^k dx^\mu$  no tienen una expresión cerrada y, por tanto, tenemos cierta libertad para escogerlas. En el fondo, la elección que hagamos de las  $A_j^k$  determina la conexión.

Las  $A_j^k$  son lo que llamamos campos gauge y sus componentes  $A_{j\mu}^k$  son los potenciales gauge. En ocasiones también se les llama conexiones gauge en un ligero abuso de terminología.

Podemos extender la definición de la derivada covariante a una  $k$ -forma sobre el fibrado  $\alpha = \alpha^j s_j$ , donde recordemos que  $\alpha_j$  es una  $k$ -forma ordinaria [8].

$$D(\alpha^j s_j) = d\alpha^j s_j + (-1)^k \alpha^j \wedge (Ds_j) = d\alpha^j s_j + (Ds_j) \wedge \alpha^j \quad (2.44)$$

Si agrupamos los  $A_j^k$  en una matriz de 1-formas  $A$  de forma que  $Ds_j = A_j^k s_k = (As)_j$ , podemos escribir esto en forma matricial

$$D(\alpha s) = (d\alpha)s + \alpha \wedge As = (d\alpha + A \wedge \alpha)s. \quad (2.45)$$

Si queremos librarnos del referencial, podemos escribir  $D\alpha = d\alpha + A \wedge \alpha$ .

La matriz de 1-formas  $A$ , como toda matriz es equivalente a una transformación lineal, no es más que una 1-forma que toma valores en el espacio de los endomorfismos del fibrado  $E$ . Sus entradas son los campos gauge  $A_j^k$ , por lo que también llamaremos indistintamente campo gauge a la matriz  $A$ .

Como nuestro fibrado es vectorial, al evaluar  $A$  en un punto tendremos un endomorfismo de una cierta fibra  $F_p$ . Si recordamos la sección 1.2.4, este espacio de endomorfismos es isomorfo al álgebra de Lie del grupo general lineal.

Como las funciones de transición, que son elementos del grupo de estructura  $G$  del fibrado, controlan el cambio de referenciales, cuando queramos hacer una conexión invariante bajo las transformaciones de  $G$ , podremos restringir el espacio donde toman valores estas formas al álgebra de Lie  $\mathfrak{g}$ .

Veamos ahora qué pasa si hacemos actuar dos veces la derivada covariante sobre un referencial. Para mayor comodidad usamos la notación matricial que acabamos de introducir.

$$D^2(s) = D(D(s)) = D(sA) = (dA)s + (Ds) \wedge A = (dA + A \wedge A)s \equiv Fs \quad (2.46)$$

Que  $A \wedge A$  sea en general distinto de cero es una consecuencia de la no conmutatividad de las formas al permitirles tomar valores en cualquier espacio vectorial. Por ejemplo, si  $A$  y  $B$  son matrices,

$$(A dx + B dy) \wedge (A dx + B dy) = AB dx \wedge dy + B A dy \wedge dx = (AB - BA) dx \wedge dy \neq 0 \text{ en general.} \quad (2.47)$$

$F$  es una matriz de 2-formas  $F^a_{b\mu\nu}$  y se le llama curvatura de la conexión. Notemos que los índices  $a$  y  $b$  son las entradas de la matriz, mientras que  $\mu$  y  $\nu$  son de coordenadas de la variedad y son los índices que le dan la naturaleza de 2-forma.

El espacio de llegada  $F$ , al igual que la matriz  $A$ , es el espacio de endomorfismos del fibrado. Por tanto, si queremos calcular su derivada covariante, tenemos que cambiar un poco la definición, ya que no será lo mismo hacer la derivada covariante de una forma cuyos valores sean matrices. Este tipo de formas las podemos escribir como  $\omega = s\Omega t$  donde  $s$  es un referencial,  $\Omega$  es una  $k$ -forma y  $t$  es el dual del referencial ( $t^i(s_j) = \delta^i_j$ ). El producto  $s\Omega t$  nos da una matriz si vemos  $s$  como un vector columna y  $t$  como un vector fila. Definimos la derivada covariante para estos objetos así [4]:

$$D\omega = d\omega + A \wedge \omega + (-1)^{k+1} \omega \wedge A \quad (2.48)$$

Esta definición viene de que, si aplicamos directamente la derivada covariante,

$$D\omega = Ds\Omega t + s d\Omega t + s\Omega Dt. \quad (2.49)$$

el primer término nos da un  $A$  a la izquierda y el último un  $A$  a la derecha.

En el contexto de la física, la 2-forma  $F$  es el tensor  $F_{\mu\nu}$  que llamamos intensidad de campo o tensor de Faraday en electromagnetismo. En relatividad general, a  $F$  se le llama el tensor de Riemann y los  $A^i_{j\mu}$  son los símbolos de Christoffel [8].

Vamos a dar la forma explícita de las componentes  $F_{\mu\nu}$  de la curvatura:

$$dA = \frac{\partial A^a_{b\mu}}{\partial dx^\sigma} dx^\sigma \wedge dx^\mu = \sum_{\sigma < \mu} \left( \frac{\partial A^a_{b\mu}}{\partial dx^\sigma} - \frac{\partial A^a_{b\sigma}}{\partial dx^\mu} \right) dx^\sigma \wedge dx^\mu \quad (2.50)$$

Para  $A \wedge A$ , como son matrices, nos aprovechamos de (2.47) para obtener

$$\begin{aligned} A \wedge A &= (A_\mu dx^\mu) \wedge (A_\nu dx^\nu) = A_\mu A_\nu dx^\mu \wedge dx^\nu + A_\nu A_\mu dx^\nu \wedge dx^\mu = [A_\mu, A_\nu] dx^\mu \wedge dx^\nu = \\ &= (A^a_{c\mu} A^c_{b\nu} - A^a_{c\nu} A^c_{b\mu}) dx^\mu \wedge dx^\nu. \end{aligned} \quad (2.51)$$

Entonces, juntando los dos términos,

$$F^a_{b\mu\nu} = \frac{\partial A^a_{b\mu}}{\partial dx^\nu} - \frac{\partial A^a_{b\nu}}{\partial dx^\mu} + (A^a_{c\mu} A^c_{b\nu} - A^a_{c\nu} A^c_{b\mu}). \quad (2.52)$$

Vemos que en forma matricial (y con una notación más estándar) esto es equivalente a

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \quad (2.53)$$

Por último, veamos cómo transforma todo esto bajo un cambio de referencial. Tomemos el grupo de estructura  $G$  de nuestro fibrado y dos referenciales  $s = (s_1, \dots, s_n)$ ,  $s' = (s'_1, \dots, s'_n)$  tales que  $s = s'g$ , donde  $g \in G$  es una función de transición. En general, las entradas de  $g$  serán funciones de las coordenadas de la fibra, por lo que su derivada exterior está bien definida. Entonces, podemos dar la relación entre la 1-forma  $A$  del referencial  $s$  y la 1-forma  $A'$  del referencial  $s'$ :

$$sA = Ds = D(s'g) = (Ds')g + s'dg = s'A'g + s'dg \stackrel{s'=sg^{-1}}{=} sg^{-1}(A'g + dg) \Rightarrow \quad (2.54)$$

$$\Rightarrow sA = s(g^{-1}A'g + g^{-1}dg) \Rightarrow A = g^{-1}A'g + g^{-1}dg. \quad (2.55)$$

El último sumando  $g^{-1}dg$  nos indica que los campos gauge  $A$  no son cantidades que transforman de manera covariante (si un referencial mide  $A = 0$ , otro referencial distinto puede medir  $A \neq 0$ ) y por tanto no serán cantidades que tengan sentido físico.

Veamos qué le pasa a la curvatura. Teniendo en cuenta que al hacer la derivada covariante de 1-formas sale un signo menos con el  $Ds$ ,

$$\begin{aligned} D^2(s) &= d^2(s'g) = D((Ds')g + s'dg) = (D^2s')g - Ds' \wedge dg + Ds' \wedge dg + s' \underbrace{d^2g}_{d^2=0} = \\ &= sF = D^2(s) = (D^2s')g = s'F' = sg^{-1}F'g \Rightarrow F = g^{-1}F'g \end{aligned} \quad (2.56)$$

Vemos que  $F$  sí que transforma de manera covariante, por lo que es una candidata a partir de la cual construir una cantidad con sentido físico.

Una última observación es que, como la traza de una matriz es cíclica,

$$tr(ABC) = tr(CAB) = tr(BCA) \Rightarrow tr(F) = tr(g^{-1}F'g) = tr(F'gg^{-1}) = tr(F'). \quad (2.57)$$

La traza de  $F$  es un invariante bajo cambios de referencial y, por tanto, la traza de una 4-forma, pues en dimensión 4 solo podemos integrar 4-formas, sería un muy buen candidato a lagrangiano en nuestra teoría.

## 2.3. Teorías de Yang-Mills

En este trabajo vamos a usar las conexiones para trabajar en las llamadas teorías de Yang-Mills. Vamos a trabajar con el espacio de Minkowski, que denotaremos  $M$ , y con un fibrado vectorial sobre  $M$  con grupo de estructura  $G$ . Como es usual hacer en el espacio de Minkowski, denotaremos las coordenadas que utilizemos como  $(x^0, x^1, x^2, x^3)$ , donde  $x_0$  es el tiempo.

Queremos que nuestros campos gauge sean invariantes respecto a  $G$ , así que  $G$  será el grupo que genere las transformaciones gauge. Este grupo de Lie no puede ser cualquiera: tiene que ser compacto y semisimple.

No nos meteremos en qué significa todo esto, pero sí comentaremos sus consecuencias. En el álgebra de Lie de un grupo de este tipo se puede definir una forma bilineal llamada forma de Killing parecida a la traza a partir de la cual podemos construir un lagrangiano invariante bajo transformaciones gauge. Además, en estos grupos, esta forma de Killing es definida negativa y diagonalizable. Esto permite que al pasar de lagrangiano a hamiltoniano se pueda definir una densidad de energía que siempre sea semidefinida positiva y, por tanto, tenga sentido físico [10].

Esto último nos da un detalle importante: la traza de las formas  $A$  y  $F$  (que van a tomar valores en el álgebra de Lie  $\mathfrak{g}$ ) va a ser siempre negativa.

**Definición 2.7** Sea  $G$  un grupo de Lie de dimensión  $n$ ,  $(\mathfrak{g}, [\cdot, \cdot])$  su álgebra de Lie y  $\{T_i\}$  ( $i$  va desde 1 a la dimensión de  $\mathfrak{g}$ ) un conjunto de generadores de  $\mathfrak{g}$ . Se definen las constantes de estructura  $c_{ij}^k \in \mathbb{R}$  como

$$[T_i, T_j] = c_{ij}^k T_k. \quad (2.58)$$

Las constantes de estructura son completamente antisimétricas para los grupos compactos y semisimples y, dependiendo de si se quiere trabajar con generadores hermíticos o antihermíticos, puede aparecer un factor  $i$  extra en el lado derecho de (2.58) [12]. También, dependiendo de la representación del álgebra con la que trabajemos, los generadores serán distintos.

En nuestro caso trabajaremos con generadores antihermíticos. Usando las propiedades de los conmutadores, podemos encontrar una relación entre nuestros generadores y sus análogos hermíticos. Definamos  $J_i = iT_i$ . Entonces

$$[J_i, J_j] = -[T_i, T_j] = -c_{ij}^k T_k = ic_{ij}^k (iT_k) = ic_{ij}^k J_k \quad (2.59)$$

Tomando el adjunto de esto, se ve fácilmente que estos operadores son hermíticos:

$$-ic_{ij}{}^k J_k^\dagger = ([J_i, J_j])^\dagger = -([T_i, T_j])^\dagger = [T_i, T_j] = c_{ij}{}^k T_k = -ic_{ij}{}^k (iT_k) = -ic_{ij}{}^k J_K \Rightarrow J_k^\dagger = J_k \quad (2.60)$$

Ya que tenemos un fibrado, podemos definir una conexión con un campo gauge  $A$  que toma valores en el álgebra de Lie  $\mathfrak{g}$ , tal como acabamos de hacer en la sección 2.2. La derivada covariante es

$$D\alpha = d\alpha + A \wedge \alpha. \quad (2.61)$$

Como tenemos los generadores de  $\mathfrak{g}$ , podemos expresar tanto los campos gauge como la curvatura en términos de ellos [3],

$$(A)_\mu = A_\mu = A_\mu^i T_i \quad F_{\mu\nu} = F_{\mu\nu}^i T_i \quad F_{\mu\nu}^i = \partial_\mu A_\nu^i - \partial_\nu A_\mu^i + c_{jk}{}^i A_\mu^j A_\nu^k. \quad (2.62)$$

Usando el operador estrella de Hodge (1.58) podemos definir lo que se suele llamar el dual de  $F$  [4]:

$$(\star F)_{\mu\nu} \equiv \star F_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\alpha\beta} F^{\alpha\beta}. \quad (2.63)$$

A diferencia del campo gauge, los campos que van a acabar siendo los campos de materia viven en cierto espacio vectorial al que podemos llegar por medio de una representación  $R$  de  $G$ . Normalmente estos campos serán vectores de  $\mathbb{C}^n$ ,  $\psi = (\psi^1, \dots, \psi^m)$ . Entonces los generadores pasan a ser otro tipo de matrices  $T_a (R)^i{}_j$  y los potenciales gauge y la derivada covariante toman la forma [12]

$$A^i{}_{j\mu} = A_\mu^k T_k (R)^i{}_j \quad D_\mu \psi^i = \partial_\mu \psi^i - i A^i{}_{j\mu} \psi^j. \quad (2.64)$$

Para el tratamiento que vamos a hacer de los instantones, no necesitaremos introducir campos de materia en nuestra teoría, pero no está de más saber cómo se definen en una teoría Yang-Mills.

Como veníamos anticipando en la sección anterior, el lagrangiano va a tener la siguiente forma [4]:

$$\mathcal{L} = -\frac{1}{2} \text{Tr}(F_{\mu\nu} F^{\mu\nu}) = -\frac{1}{2} \sum_i F_{\mu\nu}^i F^{i\mu\nu} \quad (2.65)$$

Donde  $F^{\mu\nu} = \eta^{\alpha\mu} F_{\alpha\beta} \eta^{\mu\beta}$ .

En el lagrangiano aparece  $F$  dos veces porque, a la hora de integrar para la acción, vamos a necesitar una 4-forma. En las variedades con métrica se puede definir una 4-forma de volumen  $dv = \sqrt{|\det(\eta_{\mu\nu})|} dx^0 \wedge \dots \wedge dx^3$  de forma que, como para la métrica de Minkowski  $\det(\eta_{\mu\nu}) = 1$ , tenemos que

$$S_{YM} = \int_M -\frac{1}{2} \text{tr}(F_{\mu\nu} F^{\mu\nu}) dv = - \int_M \text{tr}(F \wedge \star F). \quad (2.66)$$

Esta última expresión para la acción sirve aunque estemos en dimensión  $d \neq 4$ , ya que  $\star F$  es una  $(d-2)$ -forma. La última igualdad de (2.66) la podemos ver de la siguiente manera:

$$\begin{aligned} F \wedge \star F &= (F_{\mu\nu} dx^\mu \wedge dx^\nu \wedge (\star F_{\rho\sigma} dx^\rho \wedge dx^\sigma)) = F_{\mu\nu} (\star F_{\rho\sigma}) dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma = \\ &= \frac{1}{2} F_{\mu\nu} (\epsilon_{\rho\sigma\alpha\beta} F^{\alpha\beta}) dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma. \end{aligned} \quad (2.67)$$

Ahora, nos damos cuenta de que los productos exteriores de los  $dx$  tienen que ser alguna permutación de (1234), porque si algún índice es el mismo, el término se anula por antisimetría. Entonces, podemos reordenar estos productos a  $dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$ , añadiendo un factor  $\epsilon_{\mu\nu\rho\sigma}$  para tener en cuenta la antisimetría de las formas. Entonces,

$$F \wedge \star F = \frac{1}{2} \underbrace{\epsilon_{\mu\nu\rho\sigma} \epsilon_{\rho\sigma\alpha\beta}}_{\delta_\alpha^\mu \delta_\beta^\nu} F_{\mu\nu} F^{\alpha\beta} dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3 = \frac{1}{2} F_{\mu\nu} F^{\mu\nu} dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3. \quad (2.68)$$

Tomando la traza de esto, obtenemos la primera expresión de  $S_{YM}$ . La derivada covariante se define exactamente igual que en la sección 2.2. De hecho, si calculamos la derivada covariante de  $F$  podemos llegar a la llamada identidad de Bianchi [4]:

$$\begin{aligned} DF &= d(dA + A \wedge A) + A \wedge (dA + A \wedge A) - (dA + A \wedge A) \wedge A = \\ &= \cancel{d^2 A} + d(A \wedge A) + A \wedge dA + \cancel{A \wedge A \wedge A} - dA \wedge A - \cancel{A \wedge A \wedge A} = \\ &= dA \wedge A - A \wedge dA + A \wedge dA - dA \wedge A = 0 \end{aligned} \quad (2.69)$$

Esta identidad se satisface siempre para cualquier curvatura asociada a un campo gauge.

La acción de Yang-Mills es invariante bajo lo que llamaremos transformaciones gauge. Estas transformaciones definen cómo cambian  $A$  y  $F$  bajo cambios de referencial en el fibrado y ya las encontramos en las ecuaciones (2.55) y (2.56):

$$A(x) \rightarrow A'(x) = g^{-1}(x)A(x)g(x) + g^{-1}(x)dg(x) \quad F(x) \rightarrow F'(x) = g^{-1}(x)F(x)g(x). \quad (2.70)$$

El grupo de estas transformaciones es el grupo de aplicaciones  $\mathbb{R}^n \rightarrow G$  ( $A(x)$  con  $A \in G$ ), pero se suele decir que  $G$  es el grupo gauge. Esta es la forma que tenemos de pasar las formas  $\mathfrak{g}$ -valuadas al espacio-tiempo.

También tenemos que la acción es independiente de la representación de  $\mathfrak{g}$  que utilicemos. Recordemos que las componentes  $F^i_{\mu\nu}$  dependen de la representación y por tanto la traza también, como

vimos en la ecuación (2.62). En este caso hemos estado presuponiendo que estamos en la representación fundamental.

La traza de la acción la podemos expresar de la siguiente manera en función de los generadores  $T_i$  del álgebra de Lie:

$$tr(F_{\mu\nu}F^{\mu\nu}) = tr(F_{\mu\nu}^i T_i F^{j\mu\nu} T_j) = F_{\mu\nu}^i F^{j\mu\nu} tr(T_i T_j), \quad (2.71)$$

por lo que si cambiamos a una representación  $R$ , vamos a tener unos nuevos generadores  $R(T_i)$  y una nueva traza  $tr_R$ , de forma que

$$tr(F_{\mu\nu}F^{\mu\nu}) = tr_R(F_{\mu\nu}^i R(T_i) F^{j\mu\nu} R(T_j)) = F_{\mu\nu}^i F^{j\mu\nu} tr_R(R(T_i)R(T_j)). \quad (2.72)$$

Un grupo de Lie  $G$  actúa sobre su álgebra de Lie  $\mathfrak{g}$  usando lo que se llama la representación adjunta [3], que se define para elementos  $X \in \mathfrak{g}$  como

$$Ad_g(X) = gXg^{-1} \in \mathfrak{g} \quad g \in G. \quad (2.73)$$

Entonces

$$tr_R(R(T_i)R(T_j)) = tr_R(gT_i g^{-1} gT_j g^{-1}) = tr_R(gT_i T_j g^{-1}) = tr_R(T_i T_j), \quad (2.74)$$

donde hemos usado que la traza es cíclica. Se puede ver [12] que la traza se puede diagonalizar de forma que podemos imponer que  $tr_R(T_i T_j) = c_R \delta_{ij}$ , donde  $c_R$  es una constante que depende de la representación. Entonces, con respecto a la traza original que correspondería a la representación fundamental, nuestra acción debe de cambiar como:

$$S_{YM} = \int -\frac{1}{2} tr(F_{\mu\nu}F^{\mu\nu}) = \int -\frac{1}{2c_R} tr_R(F_{\mu\nu}F^{\mu\nu}) \quad (2.75)$$

Vamos a calcular esta  $c_R$  para  $SU(2)$ , que será el caso que manejaremos. Como  $\mathfrak{su}(2)$  tiene dimensión 3, tenemos que, si la representación  $R$  es la de espín  $j$ ,

$$3c_R = tr_R(T_i T_i) \quad (2.76)$$

Ahora bien,  $T_i T_i$  es lo que se llama el elemento Casimir del álgebra de Lie y en nuestro caso coincide con el operador de momento angular total  $J^2$  [3], por lo que su traza, que es la suma de sus autovalores, es  $tr(T_i T_i) = j(j+1)$ . Juntando esto con que la dimensión de la representación de espín  $j$  es  $2j+1$  tenemos:

$$3c_R = j(j+1)(2j+1) \Rightarrow c_R = \frac{j(j+1)(2j+1)}{3} \quad (2.77)$$

Ahora vamos calcular las ecuaciones de movimiento de nuestra teoría de Yang-Mills. Dada la forma complicada que puede tener el lagrangiano totalmente expandido, calcularlas usando las ecuaciones de Euler-Lagrange es muy tedioso (pero no imposible). En su lugar utilizaremos el siguiente método [8]:

Supongamos que tenemos dos campos gauge  $A$  y  $\tilde{A}$  en un mismo referencial. Si les hacemos la misma transformación gauge a  $A$  y  $A'$ , la cantidad  $A - \tilde{A}$  transforma como  $g^{-1}(A' - \tilde{A}')g$ .

Podemos cuantificar la variación de  $A$  como  $A_t = A + t\beta$ , con  $\beta$  una 1-forma con valores matriciales. La curvatura asociada a  $A_t$  es

$$F_t = dA_t + A_t \wedge A_t = dA + A \wedge A + t(d\beta + A \wedge \beta + \beta \wedge A) + t^2\beta \wedge \beta = F + tD\beta + O(t^2). \quad (2.78)$$

Como al final vamos a considerar variaciones pequeñas, nos quedaremos siempre a primer orden en  $t$ . Ignorando el término con  $t^2$  y definiendo el producto escalar

$$\langle \alpha, \beta \rangle = \frac{-1}{2} \int_M \alpha \wedge \star \beta, \quad (2.79)$$

podemos construir la acción correspondiente a  $F_t$ , que no es más que el producto escalar de  $F_t$  consigo misma, y calcular su derivada respecto a  $t$ :

$$\begin{aligned} \frac{d}{dt} \langle F_t, F_t \rangle &= \frac{d}{dt} \left( \int_M (F + tD\beta) \wedge \star (F + tD\beta) \right) = \frac{d}{dt} \left( \int_M F \wedge \star F + t(D\beta \wedge \star F) + t^2 D\beta \wedge \star \beta \right) = \\ &= \frac{d}{dt} (S_{YM} + \langle D\beta, F \rangle t + O(t^2)) = \langle D\beta, F \rangle + O(t). \end{aligned} \quad (2.80)$$

Si evaluamos en  $t = 0$ , el último término desaparece y, usando el operador adjunto de la derivada covariante, obtenemos

$$\frac{d}{dt} (\langle F_t, F_t \rangle)|_{t=0} = \langle D\beta, F \rangle = \langle \beta, D^* F \rangle = 0 \Leftrightarrow D^* F = 0. \quad (2.81)$$

La forma explícita de  $D^*$  es  $D^* = \pm \star D \star$  [8], al igual que la codiferencial exterior. Estas son las ecuaciones de campo de Yang-Mills. Se puede ver [4] que estas ecuaciones también se pueden expresar como

$$D(\star F) = 0 \quad (2.82)$$

Por último vamos a tratar una cuestión que hemos dejado pasar desapercibida hasta ahora: ¿dónde están las unidades?

Tal y como está definida, la acción (2.66) es adimensional, en lugar de tener unidades de  $\hbar$ . Esto se soluciona introduciendo lo que se llama una constante de acoplo<sup>1</sup> a la teoría, que se suele denotar  $g_{YM}$ . Haciendo esto, se escribe la acción de la siguiente manera [12]:

$$S_{YM} = \int_M \frac{-1}{2g_{YM}^2} \text{tr}(F_{\mu\nu}F^{\mu\nu}) dv. \quad (2.83)$$

Alternativamente, hay quien prefiere introducir la constante de acoplo en la definición de los campos gauge

$$\tilde{A} = \frac{1}{g_{YM}} A \Rightarrow \tilde{F}_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu - ig_{YM} [\tilde{A}_\mu, \tilde{A}_\nu] \quad (2.84)$$

de forma que la acción queda de la forma (2.66) solo que con la curvatura reescalada  $\tilde{F}_{\mu\nu}$ . Las unidades de  $g_{YM}$  se escogen de manera que  $S_{YM}$  tenga unidades de acción. La hemos ignorado hasta ahora porque, como es un factor que multiplica a la acción, no tiene ningún efecto a la hora de calcular las ecuaciones del movimiento ni afectará a sus soluciones.

---

<sup>1</sup>En este trabajo estamos haciendo teoría clásica de campos, por lo que la constante de acoplo es constante. A la hora de cuantizar la teoría, la constante de acoplo pasa a depender de la escala de energía [12].

## Capítulo 3

# Instantones

Supongamos que tenemos una teoría Yang-Mills cualquiera con un grupo gauge  $G$  y obtenemos una solución de acción euclídea finita. Tras hacer una rotación de Wick ( $t \rightarrow \tau = it$ ) al espacio-tiempo de Minkowski, la acción euclídea queda de la siguiente manera:

$$S_E = -iS_{YM} = \int_M d(it) d^3x \frac{1}{2} \text{tr}(F_{\mu\nu}^i F^{i\mu\nu}) = \int_{\mathbb{R}^4} d^4x \frac{1}{2} \text{tr}(F_{\mu\nu}^i F^{i\mu\nu}) = \int \text{tr}(F \wedge \star F). \quad (3.1)$$

Como el integrando es un cuadrado, es semidefinido positivo, así que si queremos que esta integral sea finita, el integrando debe irse a 0 cuando  $\|x\|$  tienda a infinito. Esto equivale a que

$$F_{\mu\nu}^i(x) = 0 \forall i, \mu\nu \Rightarrow F(x) = DA(x) = 0 \text{ si } \|x\| \rightarrow \infty. \quad (3.2)$$

La solución más general a  $F = DA = 0$  no es  $A = 0$ , sino  $A = \Omega d\Omega^{-1}$  con  $\Omega \in G$  [12] (al menos localmente, no tiene por qué cumplirse  $F = 0$  de forma global). A este tipo de campos se les conoce como gauge puro. Podemos ver que  $F = 0$  haciéndole una transformación gauge al campo  $A$  y viendo que es equivalente al gauge  $A' = 0$ :

$$A = \Omega d\Omega^{-1} \rightarrow A' = g^{-1}(\Omega d\Omega^{-1})g + g^{-1}dg \quad (3.3)$$

Si escogemos  $g \equiv \Omega$ :

$$A' = \Omega^{-1}\Omega(d\Omega^{-1})\Omega + \Omega^{-1}d\Omega = d\Omega^{-1}\Omega + \Omega^{-1}d\Omega = d(\Omega\Omega^{-1}) = d(Id) = 0 \quad (3.4)$$

Claramente  $F' = DA' = 0$  pero, como  $F$  transforma según (2.70),

$$F = \Omega F' \Omega^{-1} = \Omega 0 \Omega^{-1} = 0. \quad (3.5)$$

Entonces podemos afirmar que los campos con acción finita deben irse a un gauge puro en el infinito. Esto nos dice que para el campo  $A(x)$  debe cumplirse que  $A(x) \rightarrow \Omega(x)d\Omega^{-1}(x)$  cuando  $|x| \rightarrow \infty$ .

Fijémonos también en que, como estamos en dimensión 4 y  $F$  es una 2-forma, su dual de Hodge  $\star F$  también es una 2-forma. Esto abre la posibilidad de que exista algún campo gauge tal que tenga acción finita y además su curvatura asociada cumpla  $F = \lambda \star F$ , es decir, que el dual de Hodge sea proporcional a  $F$ .

Para ver cuál es este valor observamos que por 1.62, va a ser una raíz de 1 ó  $-1$ . Ahora mismo, como consecuencia de la rotación de Wick, estamos en  $\mathbb{R}^4$ , donde la métrica es la identidad, así que para el caso de  $F$  tenemos que  $\star \star F = F = \lambda^2 F$ . Entonces las únicas posibilidades son  $\lambda = \pm 1$ .

Estas configuraciones con acción finita y que cumplen  $F = \pm \star F$  son lo que se conocen como los instantones. Notemos que por la identidad de Bianchi (2.69), los instantones satisfacen las ecuaciones de campo de Yang-Mills (2.81) automáticamente.

Esta propiedad de "auto-dualidad" tiene una consecuencia muy interesante: los instantones son extremos absolutos de la acción de Yang-Mills [4].

Para demostrarlo, vamos a separar una curvatura  $F$  cualquiera en sus partes  $F^+$  y  $F^-$ , definidas tales que  $\star F^\pm = \pm F^\pm$ . Esto siempre lo podemos hacer porque, al igual que con las partes real e imaginaria de un número complejo, tenemos que  $F^+ = \frac{1}{2}(F + \star F)$ ,  $F^- = \frac{1}{2}(F - \star F)$ . Usamos esto para reescribir la acción de Yang-Mills

$$S = - \int tr(F^+ + F^-) \wedge \star(F^+ + F^-) = - \int tr(F^+ \wedge \star F^+ + F^+ \wedge \star F^- + F^- \wedge \star F^+ + F^- \wedge \star F^-). \quad (3.6)$$

Como  $\star F^\pm = \pm F^\pm$ , tenemos que

$$F^+ \wedge \star F^- + F^- \wedge \star F^+ = F^- \wedge F^+ - F^+ \wedge F^-. \quad (3.7)$$

Al hacer la traza de esto en componentes, va a aparecer un factor  $tr([F_{\mu\nu}, F_{\rho\sigma}])$ , que es cero porque  $tr(AB) = tr(BA)$  para dos matrices  $A$  y  $B$  cualesquiera, por tanto  $tr([A, B]) = tr(AB - BA) = tr(AB) - tr(BA) = 0$ . Entonces,

$$S = - \int tr(F^+ \wedge F^+) - \int tr(F^- \wedge F^-) \equiv \|F^+\|^2 + \|F^-\|^2. \quad (3.8)$$

La notación de la norma  $\|F^\pm\|$  viene inspirada por el producto escalar (2.79).

Si ahora definimos  $n = +\frac{1}{8\pi^2} \int \text{tr}(F \wedge F)$  (en la siguiente sección veremos de dónde viene esto), tenemos que

$$n = \frac{1}{8\pi^2} \int \text{tr}((F^+ + F^-) \wedge (F^+ + F^-)) = \int \text{tr}(F^+ \wedge F^+) + \int \text{tr}(F^- \wedge F^-) = \|F^+\|^2 - \|F^-\|^2 \quad (3.9)$$

Como para todo par de números  $a$  y  $b$  se cumple que  $a^2 + b^2 \geq |a^2 - b^2|$ , aplicádoselo a la acción vemos que  $S \geq 8\pi^2|n|$ .

De esta desigualdad vemos que los mínimos absolutos de la acción vienen dados por  $S = 8\pi^2|n|$  y se alcanzarán cuando  $F^+$  y/o  $F^-$  sean cero. Notemos que, dependiendo del signo de  $n$ , esto nos da algo de información sobre la curvatura  $F$ :

- Si  $n = 0$  es fácil ver que  $F = 0$ .
- Si  $n > 0$  se cumple

$$S = 8\pi^2 n = \|F^+\|^2 - \|F^-\|^2 = \|F^+\|^2 + \|F^-\|^2 \Rightarrow \|F^-\|^2 = 0 \Rightarrow F^- = 0 \Rightarrow F = \star F \quad (3.10)$$

- Por el mismo argumento, si  $n < 0$ ,

$$S = -8\pi^2 n = \|F^+\|^2 - \|F^-\|^2 = -\|F^+\|^2 - \|F^-\|^2 \Rightarrow \|F^+\|^2 = 0 \Rightarrow F^+ = 0 \Rightarrow F = -\star F. \quad (3.11)$$

### 3.1. Winding number

Acabamos de ver que los campos gauge  $A$  de los instantones son tales que  $A(x) \rightarrow \Omega d\Omega^{-1}(x)$  si  $\|x\| \rightarrow \infty$  con  $\Omega \in G$ .

Como al hacer el límite  $\|x\| \rightarrow \infty$  estamos haciendo el límite en una esfera de  $\mathbb{R}^4$  ( $S^3$ ) de radio cada vez más grande y al final nos sale siempre la misma "constante" ( $\Omega d\Omega^{-1}$ ), realmente la transformación gauge  $A(x)$  (y por tanto  $g(x)$ ) nos está definiendo una aplicación de  $S^3$  a  $G$ , porque podemos reescalar esa esfera de radio tendiendo a infinito a la esfera de radio 1.

Además, al llevar todo el "infinito" de  $\mathbb{R}^4$  al mismo sitio estamos haciendo lo que en topología se conoce como compactificación por un punto [2] y podemos ver que nuestro espacio-tiempo euclídeo  $\mathbb{R}^4$ , añadiendo el infinito, es homeomorfo a la 4-esfera  $S^4$  en estas condiciones.

Si nos acordamos de la sección 1.1,  $g(x)$  es un 3-lazo en  $G$ , que es un espacio topológico al ser un grupo de Lie. Entonces esto nos dice que podemos agrupar todos los campos gauge de acción euclídea

finita (porque en el límite  $A = g d g^{-1}$ ) según las clases de equivalencia del grupo de homotopía  $\pi_3(G)$ . [3]

Todo lo que hemos comentado hasta ahora sirve para cualquier teoría de Yang-Mills. Incluso en otras dimensiones, las soluciones de acción finita vienen dadas por  $\pi_d(G)$  donde  $d$  es la dimensión de la variedad. Naturalmente, si  $n \neq 4$  no se podrá dar que  $F = \star F$  en ningún caso [3].

A partir de ahora, dejaremos de lado la generalidad para estudiar una teoría en concreto que tiene un gran interés, sobre todo a nivel introductorio: tomaremos como grupo gauge  $G = SU(2)$ .

Como vimos anteriormente, tenemos que  $SU(2) \cong S^3$  y por tanto en esta teoría las soluciones de acción finita vienen dadas por  $\pi_3(SU(2)) = \pi_3(S^3) = \mathbb{Z}$ .

El número entero  $n$  asociado a cada clase de equivalencia se llama **winding number** y cuenta el número de "vueltas" que dan los lazos a  $SU(2)$ . Se puede demostrar [4] que el winding number clasifica los fibrados principales cuyo grupo es  $SU(2)$  definidos sobre el euclídeo compactificado  $S^4$ . De hecho, las transformaciones gauge de  $\pi_3(SU(2))$  son las funciones de transición de dichos fibrados. Como  $A$  tiende asintóticamente a gauge puro, el winding number también nos dice que podemos relacionar  $A$  con un fibrado principal de winding number  $n$ .

La forma exacta de un representante de cada clase de  $\pi_3(SU(2))$  es la siguiente [3]:

- La clase del 0, que es el elemento neutro de  $\mathbb{Z}$ , va a ser a la que pertenezca el camino constante  $g_0(x) = \mathbb{1}_2 \in SU(2)$ , que vimos que era el elemento neutro de los grupos de homotopía.
- Para el resto de clases, notamos que toda matriz de  $SU(2)$  se puede expresar de la siguiente manera:

$$U = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + id \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + ic \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + ib \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.12)$$

Vemos que esto es una combinación de la identidad y las matrices de Pauli con los 4 coeficientes que, si nos acordamos de (1.73), las podemos ver como las coordenadas de la 3-esfera. Primero centrémonos en el caso  $n = 1$ . Inspirándonos en que la identidad de la esfera tiene winding number 1 y, por tanto, nuestro representante debe ser homotópico a la identidad, imponemos que las componentes en la "base" de las matrices de Pauli de  $g_1$  no dependan de los ángulos y solo sean funciones reales del radio  $x^2 = (x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2$ .

$$g_1(x) = g_0(x^2) \mathbb{1}_2 + i \sum_{i=1}^3 g_i \sigma_i \equiv g_\mu(x^2) \tau^\mu, \quad (3.13)$$

donde hemos definido  $\tau_0 = \mathbb{1}_2$ ,  $\tau_i = i\sigma_i$  y, siguiendo el procedimiento estándar para subir índices,  $\tau^\mu = \eta^{\mu\nu}\tau_\nu$ . En el euclídeo, como la métrica es la identidad, tenemos  $\tau^\mu = \tau_\mu$ .

Las  $g_i(x^2)$  podemos determinarlas usando que  $g_1$  tiene que ser una matriz de  $SU(2)$ :

$$g^\dagger g = \mathbb{1}_2 \Leftrightarrow g^\mu(x^2)g^\nu(x^2)\tau_\mu\tau_\nu = \mathbb{1}_2 \quad (3.14)$$

Ahora vamos a notar que, como  $A$  es una 1-forma sobre una variedad, sus funciones componente tienen que ser de clase  $C^\infty$ , de forma que las podemos poner como derivada de alguna función  $h$  [4]

$$g^\mu(x^2) = \partial_\mu h(x^2) = 2x_\mu h'(x^2). \quad (3.15)$$

Sustituyendo en (3.14) obtenemos

$$4(h'(x^2))^2 x_\mu x_\nu (\tau^\mu)^\dagger \tau^\nu = 1 \Rightarrow 2(h'(x^2))^2 x_\mu x_\nu \{(\tau^\mu)^\dagger, \tau^\nu\} = \mathbb{1}_2 \quad (3.16)$$

Usando que para las matrices de Pauli  $\{\sigma_i, \sigma_j\} = 2\delta_{ij}\mathbb{1}_2$ ,

$$2(h'(x^2))^2 x_\mu x_\nu \{(\tau^\mu)^\dagger, \tau^\nu\} = \mathbb{1}_2 \Rightarrow 2(h'(x^2))^2 x^2 = 1. \quad (3.17)$$

Integrando obtenemos

$$h'(x^2) = \frac{1}{2(x^2)^{1/2}} \Rightarrow h(x^2) = (x^2)^{1/2} + C. \quad (3.18)$$

Derivando obtenemos la expresión de  $g_1$ :

$$g_1(x) = \frac{x_\mu \tau^\mu}{(x^2)^{1/2}}. \quad (3.19)$$

Se puede ver que  $g_1$  es, de hecho, la aplicación identidad de  $S^3$  en  $SU(2)$  visto como la 3-esfera [3].

- Para un  $n$  cualquiera, se cumple [3] que  $g_n(x) = (x^2)^{-n/2} [x_\mu \tau^\mu]^n = g_1(x)^n$  pertenece a la clase de equivalencia del número  $n$ .

Vamos a ver que estas clases de equivalencia están muy relacionadas con la acción de los instantones. Sin pérdida de generalidad consideramos un campo gauge  $A$  asociado a un instantón, de forma que su curvatura cumple  $F = \star F$  y tendrá cierto winding number  $n > 0$ <sup>1</sup>. En primer lugar, notamos

<sup>1</sup>Normalmente, se llama instantones a las soluciones con  $n > 0$  y anti-instantones a las soluciones con  $n < 0$ . Esto se debe a que a las condiciones  $F = \star F$  y  $F = -\star F$  se les suele llamar "self-dual condition" y "anti self-dual condition", respectivamente.

que

$$d(\text{tr}(F \wedge F)) = d(\text{tr}(F \wedge \star F)) = 0 \quad (3.20)$$

al ser  $\text{tr}(F \wedge F)$  una 4-forma en una variedad de dimensión 4. Por el lema de Poincaré (teorema 1.4), localmente podemos expresar  $\text{tr}(F \wedge F) = dK$  con  $K$  una 3-forma definida localmente. A  $K$  se le conoce con el nombre de forma de Chern-Simons.

Se cumple que

$$K = \text{tr}(A \wedge dA + \frac{2}{3}A \wedge A \wedge A). \quad (3.21)$$

Es conveniente expresar  $K$  en términos de  $F$  en vez de  $dA$ . Como  $F = dA + A \wedge A \Rightarrow dA = F - A \wedge A$ ,

$$K = \text{tr}(A \wedge (F - A \wedge A) + \frac{2}{3}A \wedge A \wedge A) = \text{tr}(A \wedge F - \frac{1}{3}A \wedge A \wedge A). \quad (3.22)$$

Para demostrar que  $dK = \text{tr}(F \wedge F)$ , simplemente calculamos<sup>2</sup> la derivada exterior de  $K$ :

$$dK = \text{tr} \left( dA \wedge dA + \cancel{A \wedge d^2 A} + \frac{2}{3}(dA \wedge A \wedge A - A \wedge dA \wedge A + A \wedge A \wedge dA) \right). \quad (3.23)$$

Usando de nuevo que  $dA = F - A \wedge A$  en todos los términos que aparece y desarrollando,

$$dK = \text{tr}(F \wedge F - A \wedge A \wedge F - F \wedge A \wedge A + \wedge^4 A + \frac{2}{3}(F \wedge A \wedge A - A \wedge F \wedge A + A \wedge A \wedge F - \wedge^4 A)). \quad (3.24)$$

Usando que la traza es cíclica en las funciones componentes y la antisimetría del producto exterior, podemos ver que

$$\text{tr}(A \wedge F \wedge A) = -\text{tr}(A \wedge A \wedge F) = -\text{tr}(F \wedge A \wedge A) \Rightarrow \text{tr}(A \wedge F \wedge A) = -\frac{1}{2}\text{tr}(A \wedge A \wedge F + F \wedge A \wedge A). \quad (3.25)$$

Usando estas mismas propiedades podemos obtener también que, si hacemos una permutación cíclica a  $\text{tr}(\wedge^4 A)$ ,

$$\text{tr}(\wedge^4 A) = \text{tr}(A_\mu A_\nu A_\rho A_\sigma dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma) = -\text{tr}(A_\sigma A_\mu A_\nu A_\rho dx^\sigma \wedge dx^\mu \wedge dx^\nu \wedge dx^\rho), \quad (3.26)$$

ya que para pasar de (1234) a (4123) necesitamos tres transposiciones. Entonces la traza de  $\wedge^4 A$  es su propio opuesto porque los índices sobre los que estamos sumando son los mismos y concluimos que  $\text{tr}(\wedge^4 A) = 0$ . Entonces, sustituyendo (3.25) en (3.24),

$$dK = \text{tr}(F \wedge F - A \wedge A \wedge F - F \wedge A \wedge A + F \wedge A \wedge A + A \wedge A \wedge F) = \text{tr}(F \wedge F). \quad (3.27)$$

---

<sup>2</sup>Para no saturar las cuentas, denotaremos  $\wedge^n A = A \wedge A \wedge \dots \wedge A$  ( $n$  veces).

Vamos a usar esto para calcular una forma alternativa de la acción del instantón [3]. Recordemos que en las condiciones en las que estamos podemos ver el espacio-tiempo como la 4-esfera  $S^4$  y que la forma de Chern-Simons  $K$  está definida solo localmente. Entonces, para trabajar con  $K$  necesitamos dar una carta de  $S^4$  y definir  $K$  con las coordenadas de dicha carta. Vamos a usar el hemisferio norte  $U_N$  y el hemisferio sur  $U_S$ , que los podemos definir a partir del  $\mathbb{R}^4$  compactificado. En este caso elegimos que el infinito sea el polo sur y el 0 sea el polo norte, de forma que

$$U_N = \{x \in \mathbb{R}^4 \mid \|x\| \leq L + \epsilon \quad (\epsilon > 0, L > 0)\} \quad (3.28)$$

$$U_S = \{x \in \mathbb{R}^4 \mid \|x\| \geq L - \epsilon \quad (\epsilon > 0, L > 0)\} \quad (3.29)$$

Los bordes de  $U_N$  y  $U_S$  son 3-esferas  $S^3$ , al igual que el borde de una semiesfera es un círculo.

En el hemisferio sur de  $S^4$ , que es el que contiene el infinito, podemos fijar sin pérdida de generalidad  $A = 0$ , ya que los gauge puros los podemos convertir en el gauge nulo mediante la transformación adecuada. Como  $U_N \cap U_S$  no es vacía, para que nuestro  $A$  sea consistente debe cumplirse que  $A$  sea un gauge puro  $gdg^{-1}$  en el hemisferio norte [3] (es decir, que los podamos conectar con una función de transición del fibrado de  $S^4$ ).

Ahora aplicamos el teorema de Stokes para calcular la integral de  $K$  en  $\partial U_N = S^3$

$$\int_{S^3} K = \int_{U_N} dK = \int_{S^4} tr(F \wedge F), \quad (3.30)$$

donde la integral de  $tr(F \wedge F)$  la hemos extendido a toda la esfera porque la contribución del hemisferio sur es nula.

Para un gauge puro, como  $F = 0$ , usando la expresión (3.22) obtenemos que  $K = -\frac{1}{3}tr(\wedge^3 A)$ . Por tanto

$$\int_{S^4} tr(F \wedge F) = -\frac{1}{3} \int_{S^3} tr(\wedge^3 A) = -\frac{1}{3} \int_{S^3} d^3x tr(\epsilon^{ijk} (g\partial_i g^{-1})(g\partial_j g^{-1})(g\partial_k g^{-1})) \quad (3.31)$$

Una vez obtenido esto podemos evaluar la integral para los potenciales gauge  $g_n(x)$  de las distintas clases de equivalencia de  $\pi_3(S^3)$ :

- Claramente el caso  $n=0$  en el que  $g_0(x) = \mathbb{1}_2$  es constante nos da  $\int_{S^4} tr(F \wedge F) = 0$ .
- El caso  $n = 1$  nos da el siguiente campo gauge:

$$A = g_1 dg_1^{-1} = \frac{1}{R}((x^0 Id + ix^j \sigma_j) d \left( \frac{1}{R}(x^0 Id - ix^k \sigma_k) \right)) \quad (3.32)$$

Recordemos que como el álgebra de Lie  $\mathfrak{su}(2)$  es un espacio tangente que, en este caso, lo podemos ver como isomorfo a un espacio tangente de  $S^3 = SU(2)$ , la forma de  $A$  podría depender del punto respecto al cual tomemos el espacio tangente. Esto no debería poder ser así ya que  $g_1$  es la identidad de  $S^3$  y  $S^3$  es un espacio homogéneo, lo que quiere decir que todos los espacios tangentes son isomorfos entre sí.

Por tanto, la forma  $tr(\wedge^3 A)$  no debería depender del punto de la 3-esfera en el que "anclamos" el sistema de coordenadas para pasar al espacio tangente [3]. Entonces, podemos escoger el punto  $(1, 0, 0, 0)$  y, al evaluar ahí, nos queda que  $A = -i\sigma_k dx^k$  ( $k = 1, 2, 3$ ). Por tanto, usando las propiedades de los símbolos de Levi-Civita y que, si  $i \neq j$ ,  $\sigma_i \sigma_j = i\epsilon^{ijk} \sigma_k$ ,

$$tr(\wedge^3 A) = -i^3 tr(\sigma_i \sigma_j \sigma_k) dx^i \wedge dx^j \wedge dx^k = i tr(\epsilon^{ijm} \sigma_m \sigma_k) dx^i \wedge dx^j \wedge dx^k. \quad (3.33)$$

Como las matrices de Pauli tienen todas traza cero, la traza de  $\sigma_m \sigma_k$  será cero a no ser que  $m = k$ , en cuyo caso  $\sigma_k^2 = \mathbb{1}_2$  y  $tr(\mathbb{1}_2) = 2$ . Utilizando esto llegamos a que

$$tr(\wedge^3 A) = -2\epsilon_{ijk} dx^i \wedge dx^j \wedge dx^k = -2\epsilon_{ijk} \epsilon_{ijk} dx^1 \wedge dx^2 \wedge dx^3 = -12 dx^1 \wedge dx^2 \wedge dx^3, \quad (3.34)$$

donde al final usamos que  $\sum \epsilon_{ijk} \epsilon_{ijk} = 3! = 6$ .

En general, el volumen de la  $d$ -esfera es

$$V(S^d) = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma\left(\frac{d+1}{2}\right)} \Rightarrow V(S^3) = 2\pi^2. \quad (3.35)$$

Entonces, usando que el volumen de la 3-esfera es  $2\pi^2$ ,

$$\int_{S^3} tr(\wedge^3 A) = -12 \int_{S^3} d^3x = -12(2\pi^2) = -24\pi^2. \quad (3.36)$$

Equivalentemente,

$$n = \frac{1}{8\pi^2} \int_{S^4} F \wedge F = \frac{1}{24\pi^2} \int_{S^3} tr(\wedge^3 A) = 1 \quad (3.37)$$

Antes de continuar, debemos hacer una observación bastante importante. En este trabajo hemos definido el gauge puro como  $gdg^{-1}$ . Sin embargo, en la bibliografía se usa la definición  $g^{-1}dg$ . Esto introduce un signo menos a la hora de evaluar (3.32), lo que acaba teniendo como consecuencia que el winding number sea el opuesto del que acabamos de calcular. Estas dos definiciones no se contradicen, ya que lo que estamos haciendo es escoger el "sentido de giro" positivo en la 3-esfera como el que los libros definen como negativo.

- El caso general para un winding number  $n$  se puede ver de la siguiente manera: la transformación gauge  $g_n(x)$  se puede escribir como un producto de transformaciones gauge de winding number más bajo, concretamente  $g_n(x) = g_1(x)^n$ .

Consideremos el caso  $n = 2$ , en el que tenemos  $g_2 = g_1 g_1$ . En primer lugar, dividimos  $S^3$  en sus hemisferios de la misma forma que hicimos en (3.28):

$$U_N^{(3)} = \{x \in \mathbb{R}^3 \mid \|x\| \leq L + \epsilon\} \quad (3.38)$$

$$U_S^{(3)} = \{x \in \mathbb{R}^3 \mid \|x\| \geq L - \epsilon\} \quad (3.39)$$

El gauge  $g_1(x)$  es homotópico [3] a una aplicación  $g_{1N}(x)$  de winding number 1 y tal que  $g_{1N}(x) = \mathbb{1}_2 \in SU(2)$  si  $x$  está en el hemisferio sur  $U_S^{(3)}$ . De la misma manera, también es homotópico a una aplicación  $g_{1S}(x)$  de winding number 1 y tal que  $g_{1S}(x) = \mathbb{1}_2$  en el hemisferio norte.

Entonces, podemos decir que  $g_2 = g_1 g_1$  es homotópica a

$$g'_2(x) = \begin{cases} g_{1N}(x) & \text{si } x \in U_N^{(3)} \\ g_{1S}(x) & \text{si } x \in U_S^{(3)} \end{cases} \quad (3.40)$$

El gauge puro asociado es  $A_2(x) = g'_2(x) d(g'_2)^{-1}(x)$ , de forma que [3]

$$\begin{aligned} n &= \frac{1}{24\pi^2} \int_{S^3} d^3x \operatorname{tr}(\wedge^3 A_2(x)) = \\ &= \frac{1}{24\pi^2} \left( \int_{U_N^{(3)}} \operatorname{tr}(\wedge^3 (g_{1N} dg_{1N}^{-1})) + \int_{U_S^{(3)}} \operatorname{tr}(\wedge^3 (g_{1S} dg_{1S}^{-1})) \right) = 1 + 1 = 2 \end{aligned} \quad (3.41)$$

Se puede repetir este procedimiento para  $A_n(x) = g_n(x) dg_n^{-1}(x)$  y obtenemos que [3]

$$|n| = \frac{1}{24\pi^2} \int_{S^3} \operatorname{tr}(A \wedge A \wedge A) = \frac{1}{8\pi^2} \int_{S^4} \operatorname{tr}(F \wedge F) = \frac{1}{8\pi^2} S_{YM} \quad (3.42)$$

Aquí ponemos un valor absoluto porque, a la hora de pasar a la acción, los anti-instantones contribuyen con un signo menos ya que  $F = -\star F$ .

Este resultado es extremadamente importante ya que podemos calcular la acción de cualquier instantón (o de cualquier campo que sea gauge puro) a partir de su winding number y viceversa.

### 3.2. Vacíos en Yang-Mills y el vacío $\theta$

Los vacíos en una teoría de campos los podemos definir como los estados de menor energía que podemos construir, es decir, los estados fundamentales del hamiltoniano.

Viendo nuestra acción (2.66) tiene bastante sentido pensar que los vacíos serán los estados que cumplan que la curvatura  $F$  es idénticamente 0. Como ya mencionamos antes, esto no implica  $A = 0$ , sino  $A = g dg^{-1}$  con  $g \in SU(2)$ . Estamos prácticamente en la misma situación que antes, solo que ahora en vez de cumplirse asintóticamente, nuestros campos son directamente gauge puro.

Por tanto, los vacíos de la teoría vienen dados por infinitas clases de equivalencia de  $\pi_3(SU(2)) = \mathbb{Z}$  desconectadas entre sí y etiquetadas por el winding number (3.42). El hecho de que estén desconectadas es bastante relevante porque, al no ser homotópicas, no podemos pasar de manera regular de un vacío a otro mediante una transformación gauge infinitesimal [11]. Sin embargo, estas configuraciones tienen acción euclídea finita, así que según lo que vimos en (2.27) hay una probabilidad no nula de que exista un efecto túnel entre vacíos de diferente winding number.

Antes de ver cómo es esta probabilidad hay que hablar un poco de la parte cuántica de la teoría. Hasta ahora, hemos estado tratando estos vacíos como si fueran clásicos, pero la naturaleza es cuántica y, para calcular propagadores, necesitamos introducir kets que representen un estado.

Estos estados tienen dos fuentes de degeneración: los posibles vacíos distintos que pueda haber dentro de la misma clase de homotopía y las infinitas clases que existen. Eligiendo el campo gauge  $A$  de manera adecuada (se puede consultar [11] para los detalles), podemos considerar que solo tendremos un vacío  $|n\rangle$  por cada winding number. La posibilidad de efecto túnel entre los vacíos hace que éstos puedan cambiar con el tiempo: si un estado cumple  $|\psi, t = 0\rangle = |n\rangle$ , entonces

$$|\psi, t > 0\rangle = \sum_{m \in \mathbb{Z}} c_m |m\rangle, \quad (3.43)$$

en lugar de evolucionar adquiriendo una fase ( $|\psi, t > 0\rangle = e^{i\varphi(t)} |n\rangle$ ). Para construir un vacío que no cambie, definimos un operador  $T$  que lleve un vacío asociado a un winding number  $n$  al vacío de winding number  $n + 1$ :

$$T |n\rangle = |n + 1\rangle \quad (3.44)$$

Esto no es más que un operador de traslación entre los vacíos. Como el lagrangiano es invariante gauge y la acción de  $T$  es básicamente aplicar una transformación gauge (3.19) o similar,  $T$  conmuta con el hamiltoniano. Por el teorema de Bloch, podemos deducir que los autoestados de  $T$  serán también autoestados de energía y, al aplicarles estas traslaciones, saldrá un factor de fase con una cierta periodicidad, de manera parecida a cómo en las redes cristalinas la periodicidad del potencial introduce estas fases.

En particular, bajo transformaciones gauge de winding number  $k$ , podemos decir los autoestados de  $T$  sufren un cambio de fase  $e^{ik\theta}$  [12].

Esto nos lleva a definir lo que vamos a llamar el vacío  $\theta$  como

$$|\theta\rangle = \sum_{n \in \mathbb{Z}} e^{in\theta} |n\rangle. \quad (3.45)$$

$|\theta\rangle$  es un autoestado de  $T$  con autovalor  $e^{i\theta}$  y, por tanto, también es un autoestado del hamiltoniano. Notemos que como el valor de  $\theta$  nos lo acabamos de inventar, estos vacíos  $\theta$  no tienen por qué ser únicos [11]. Vamos a calcular el propagador entre dos estados  $|\theta\rangle$  y  $|\theta'\rangle$ :

$$\langle \theta' | e^{-iHt} | \theta \rangle = \sum_{m, n \in \mathbb{Z}} e^{im\theta'} e^{-in\theta} \langle m | e^{-iHt} | n \rangle = \sum_{m, n \in \mathbb{Z}} e^{im(\theta' - \theta)} e^{-i(n-m)\theta} \langle m | e^{-iHt} | n \rangle. \quad (3.46)$$

Como el hamiltoniano conmuta con  $T$ , el elemento de matriz  $\langle m | e^{-iHt} | n \rangle$  solo debe depender de la diferencia entre  $n$  y  $m$ , que son las cosas que sí se ven afectadas por una transformación gauge. Por tanto podemos "desplazar" los índices [11] y trabajar con la diferencia  $k = n - m$  de forma que

$$\langle \theta' | e^{-iHt} | \theta \rangle = \sum_{m, k} e^{im(\theta' - \theta)} e^{ik\theta} \langle k | e^{-iHt} | 0 \rangle = 2\pi\delta(\theta' - \theta) \sum_k e^{ik\theta} \langle k | e^{-iHt} | 0 \rangle. \quad (3.47)$$

Esto nos dice que la probabilidad de que uno de los estados  $\theta$  evolucione a otro  $\theta'$  es cero, lo que implica que en nuestro mundo solo puede existir un valor de este ángulo  $\theta$  que nunca cambiará.

Para acabar con el tema de los vacíos, vamos a dar una corta descripción de cómo los instantones aparecen al considerar el efecto túnel entre vacíos.

Resulta que estos estados  $|\theta\rangle$  tienen una contribución no nula a la path integral [12] y, por tanto, debemos añadir esta contribución en la acción de Yang-Mills. Como esto solo va a depender, de nuevo, de las diferencias entre el winding number, a la hora de hacer la integral usaremos estados con  $n = 0$  y  $n = k$ . En el espacio de configuraciones, las clases de homotopía las podemos ver como "islas" que estarán desconectadas y separadas una cierta distancia entre sí. Si queremos calcular la integral de caminos entre estos dos vacíos debemos considerar los caminos que empiezan en la "isla" de  $n = 0$  y acaban en la "isla" de  $n = k$ . Esta región de integración la podemos ver como dos esferas  $S^3$  situadas en los instantes  $t = \pm\infty$  [12]. Comparando cada término de la suma de (3.47) con (2.18), tenemos que

$$\langle k | e^{-iHt} | 0 \rangle = \int \mathcal{D}[A] e^{iS_{YM}} \quad (3.48)$$

Para el estado  $n = 0$  la acción es nula y, para el  $n = k$ , la acción viene dada por el winding number (3.42). Entonces, para recuperar el factor  $e^{ik\theta}$  necesitamos añadir a la acción el siguiente término [11]

$$\mathcal{L}_\theta = \frac{\theta}{8\pi^2} \int \text{tr}(F \wedge \star F) = \frac{\theta}{16\pi^2} \int \text{tr}(F_{\mu\nu} F^{\mu\nu}) \quad (3.49)$$

Para tener en cuenta el signo del winding number, ponemos un  $\star F$  que lo introducirá al usar la condición  $F = \pm \star F$ . Vemos que  $e^{ik\theta} \langle k | e^{-iHt} | 0 \rangle = \int \mathcal{D}[A] e^{i(S_{YM} + S_\theta)}$ . A este término extra se le llama el término  $\theta$  del lagrangiano de Yang-Mills. Si nos acordamos de la forma de Chern-Simons, este término va a ser una derivada total y a primera vista, parecería que no tiene efectos observables, y esto es cierto clásicamente y a nivel de diagramas de Feynman [11]. Sin embargo, el término  $\theta$  tiene efectos físicos visibles en la teoría de monopolos o de aislantes topológicos [12].

En general, si queremos hacer un efecto túnel entre los estados  $|n_1\rangle$  y  $|n_2\rangle$  podemos hacer una cuenta parecida:

$$\langle n_1 | e^{-iHt} | n_2 \rangle = \int_{n_1}^{n_2} \mathcal{D}[A] e^{iS[A]} \quad (3.50)$$

A este tipo de integrales se les puede aplicar el método de la fase estacionaria de análisis asintótico para dar una aproximación. Aquí, la fase de la integral viene dada por la acción. La idea principal es hacer un desarrollo de Taylor alrededor de los extremos  $A_{ex}$  de la acción de forma que podamos aproximar  $S \approx \frac{1}{2}k(A - A_{ex})^2$ , por lo que las configuraciones que se alejen mucho de los extremos oscilan muy rápido y no contribuyen a la integral.

Aquí es donde entran los instantones. En la sección anterior vimos que los instantones no solo son un extremo de la acción, sino que son un extremo absoluto. En este caso particular, como únicamente estamos considerando los caminos entre los vacíos de winding number  $n_1$  y  $n_2$ , solo los instantones de winding number  $n_2 - n_1$  contribuirán a la integral [12], pero lo harán de forma dominante con respecto a cualquier otra solución de las ecuaciones de movimiento. Entonces, en una primera aproximación, la probabilidad de efecto túnel entre dos vacíos se puede calcular usando la acción de los instantones.

### 3.3. El instantón $n = 1$

Una vez vista la relevancia física que tienen los instantones, veamos cómo se construyen explícitamente. En primer lugar veremos el caso no trivial más sencillo, que resulta ser el de un instantón de winding number 1.

Para simplificar los cálculos pediremos que el campo gauge del instantón  $A$  tenga simetría esférica, igual que hicimos con el gauge puro  $g_1(x)$  en (3.19). Recordemos también que para los instantones vamos a trabajar en el espacio-tiempo euclídeo.

Como  $A(x) \rightarrow gdg^{-1}(x)$  cuando  $\|x\| \rightarrow \infty$ , nos inspiramos en esto y la simetría esférica para

hacer el siguiente *ansatz* [4]:

$$A_\mu(x) = f(x^2) g(x) \partial_\mu g^{-1}(x), \quad g \in SU(2). \quad (3.51)$$

A la función  $f$  le pediremos que  $f(x^2) \rightarrow 1$  cuando  $\|x\|$  tienda a infinito para que  $A$  tienda a gauge puro. Adicionalmente, tomaremos que dicho gauge puro sea  $g(x) = g_1(x)$  porque vimos que sus "componentes" al expresarlo en términos de las matrices de Pauli son independientes de los ángulos [4]. Usando (3.19) nuestro campo  $A$  toma la forma

$$A_\mu(x) = f(x^2) \frac{x_\nu \tau^\nu}{(x^2)^{1/2}} \left( \frac{(\tau^\mu)^\dagger}{(x^2)^{1/2}} + \frac{x_\beta x_\mu}{(x^2)^{3/2}} (\tau^\beta)^\dagger \right). \quad (3.52)$$

Recordemos que, como las  $\tau$  son matrices de  $\mathfrak{su}(2)$ ,  $\tau = -\tau^\dagger$ . Además,  $\tau_\mu^2 = -\mathbb{1}_2$ . Introduciendo el factor  $x_\nu \tau^\nu$  en el paréntesis llegamos a

$$A_\mu(x) = \frac{f(x^2)}{(x^2)^{1/2}} \left( \frac{x_\nu \tau^\nu}{(x^2)^{1/2}} (\tau^\mu)^\dagger + \frac{x_\mu}{(x^2)^{1/2}} \mathbb{1}_2 \right). \quad (3.53)$$

Lo único que nos falta para tener la forma cerrada del instantón es saber cómo es  $f$ . Para ello vamos a utilizar la condición que nos queda, que es la de  $F = \star F$ .

Notemos que como la identidad conmuta con todo y su derivada es cero, el término proporcional a la identidad no va a contribuir a la hora de calcular la curvatura, de forma que podemos ignorarlo [4] y calcular la curvatura para

$$A_\mu(x) = \frac{f(x^2)}{x^2} x_\nu \tau^\nu (\tau^\mu)^\dagger = a(x^2) b (b'_\mu)^\dagger \quad (3.54)$$

Donde hemos definido  $a(x^2) = f(x^2)/x^2$  y  $b = x_\nu \tau^\nu$  (de forma que derivando,  $b'_\mu = \tau^\mu$  y  $b'' = 0$ ). Usando la expresión por componentes (2.53) calculamos  $F_{\mu\nu}$ :

$$F_{\mu\nu} = a'_\mu b (b'_\nu)^\dagger - a'_\nu b (b'_\mu)^\dagger + a((b'_\mu (b'_\nu)^\dagger) - b'_\nu (b'_\mu)^\dagger) + a^2 [b (b'_\mu)^\dagger, b (b'_\nu)^\dagger]. \quad (3.55)$$

Ahora, notando que  $bb^\dagger = x^2 Id$  y que al derivar  $(bb^\dagger)'_\mu = b'_\mu b^\dagger + b (b'_\mu)^\dagger = 2x_\mu$ , podemos desarrollar el conmutador

$$\begin{aligned} [b (b'_\mu)^\dagger, b (b'_\nu)^\dagger] &= \underbrace{b (b'_\mu)^\dagger}_{2x_\mu - b'_\mu b^\dagger} b (b'_\nu)^\dagger - \underbrace{b (b'_\nu)^\dagger}_{2x_\nu - b'_\nu b^\dagger} b (b'_\mu)^\dagger = \\ &= 2x_\mu b (b'_\nu)^\dagger - b'_\mu \underbrace{b^\dagger b}_{x^2 Id} (b'_\nu)^\dagger + 2x_\nu b (b'_\mu)^\dagger - b'_\nu \underbrace{b^\dagger b}_{x^2 Id} (b'_\mu)^\dagger = \\ &= 2x_\mu b (b'_\nu)^\dagger + 2x_\nu b (b'_\mu)^\dagger - x^2 Id (b'_\mu (b'_\nu)^\dagger) - b'_\nu (b'_\mu)^\dagger \end{aligned} \quad (3.56)$$

Juntándolo todo obtenemos

$$F_{\mu\nu} = (a'_\mu + 2x_\mu a^2)b(b'_\nu)^\dagger - (a'_\nu + 2x_\nu a^2)b(b'_\mu)^\dagger + (a - x^2 a^2)(b'_\mu(b'_\nu)^\dagger - b'_\nu(b'_\mu)^\dagger) \quad (3.57)$$

Con esta forma de  $F_{\mu\nu}$  se puede ver que, por simetría, el término

$$b'_\mu(b'_\nu)^\dagger - b'_\nu(b'_\mu)^\dagger = \tau^\mu(\tau^\nu)^\dagger - \tau^\nu(\tau^\mu)^\dagger \quad (3.58)$$

no cambia al hacer el dual de Hodge [4]. Sin embargo, el resto de términos sí que cambian. Esto nos dice que, para que  $F = \star F$ , los dos primeros sumandos tienen que anularse, lo cual nos pone la siguiente restricción a la función  $a$ :

$$a'_\mu + 2x_\mu a^2 = 0 \quad (3.59)$$

La solución de esta ecuación diferencial no lineal es

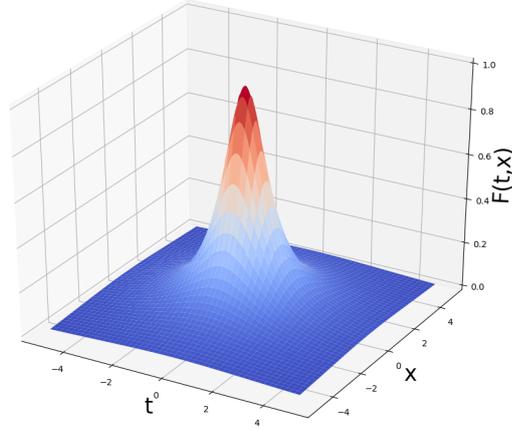
$$a(x^2) = \frac{1}{x^2 + \lambda^2} \Rightarrow f(x^2) = \frac{x^2}{x^2 + \lambda^2}, \quad (3.60)$$

con  $\lambda^2$  una constante de integración. Notemos que, como queríamos,  $f(x^2)$  tiende a 1 asintóticamente.

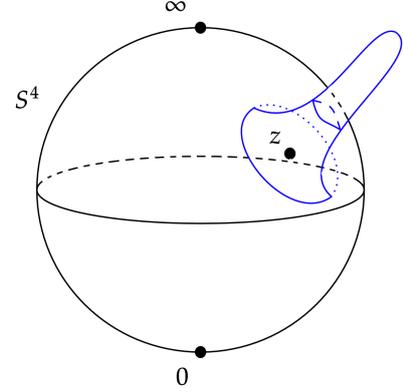
Por tanto la forma cerrada de  $A_\mu$  (ignorando el término de la identidad) es

$$A_\mu(x) = \frac{1}{x^2 + \lambda^2} x_\nu \tau^\nu (\tau^\mu)^\dagger \quad (3.61)$$

Este campo gauge en concreto se conoce como el instantón BPST en referencia a sus descubridores: Belavin, Polyakov, Schwarz y Tyupkin. También podemos ver el por qué del nombre "instantón": si miramos la gráfica de su curvatura, tiene un máximo en el origen y decae muy rápido fuera de ese punto.



(a) Gráfica de  $\frac{\lambda^2}{t^2+x^2+\lambda^2}$  para  $\lambda = 1$ .



(b) Representación de la curvatura de (3.61) centrada en un punto  $z \in S^4$ . Imagen cedida bajo licencia Creative Commons por Tazerenix.

Para asegurarnos de que hemos hecho las cosas bien, podemos comprobar con la ecuación (3.37) que nuestro  $A$  tiene winding number 1 [4]. Para ahorrarnos paréntesis vamos a trabajar con las matrices de Pauli con los índices abajo. Recordemos que como estamos en el euclídeo,  $\tau^\mu = \tau_\mu$ .

La curvatura (3.57) con la condición (3.59) toma la forma

$$F_{\mu\nu} = (a - x^2 a^2)(\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger) \quad (3.62)$$

Entonces,

$$\begin{aligned} n &= \frac{1}{8\pi^2} \int tr(F \wedge F) = \frac{1}{8\pi^2} \frac{1}{4} \int tr(F_{\mu\nu} F_{\rho\sigma}) \epsilon^{\mu\nu\rho\sigma} d^4x = \\ &= \frac{\epsilon^{\mu\nu\rho\sigma}}{32\pi^2} tr((\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger)(\tau_\rho \tau_\sigma^\dagger - \tau_\sigma \tau_\rho^\dagger)) \int (a - x^2 a^2)^2 d^4x \end{aligned} \quad (3.63)$$

Tenemos un factor  $\frac{1}{2}$  adicional porque podemos escribir  $F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$  por antisimetría.

Notemos que, al estar contrayendo con los  $\epsilon^{\mu\nu\rho\sigma}$ , también por antisimetría tenemos que

$$\epsilon^{\mu\nu\rho\sigma} (\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger) = \epsilon^{\mu\nu\rho\sigma} \tau_\mu \tau_\nu^\dagger - \epsilon^{\mu\nu\rho\sigma} \tau_\nu \tau_\mu^\dagger = \epsilon^{\mu\nu\rho\sigma} \tau_\mu \tau_\nu^\dagger + \epsilon^{\nu\mu\rho\sigma} \tau_\mu \tau_\nu^\dagger = 2\epsilon^{\mu\nu\rho\sigma} \tau_\mu \tau_\nu^\dagger. \quad (3.64)$$

Podemos simplificar un poco el término que precede a la integral:

$$-\frac{\epsilon^{\mu\nu\rho\sigma}}{32\pi^2} tr((\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger)(\tau_\rho \tau_\sigma^\dagger - \tau_\sigma \tau_\rho^\dagger)) = -\frac{1}{8\pi^2} \epsilon^{\mu\nu\rho\sigma} tr(\tau_\mu \tau_\nu^\dagger \tau_\rho \tau_\sigma^\dagger) \quad (3.65)$$

Desarrollando los casos de índice cero del  $\epsilon^{\mu\nu\rho\sigma}$  nos podemos deshacer de algunas matrices identidad (recordemos que  $\tau^0 = \mathbb{1}_2$ ).

$$\begin{aligned}
\epsilon^{\mu\nu\rho\sigma} \text{tr}(\tau_\mu \tau_\nu^\dagger \tau_\rho \tau_\sigma^\dagger) &= \underbrace{\epsilon^{0\nu\rho\sigma}}_{\epsilon^{ijk}} \text{tr}(\mathcal{A} \tau_\nu^\dagger \tau_\rho \tau_\sigma^\dagger) + \underbrace{\epsilon^{\mu 0\rho\sigma}}_{-\epsilon^{ijk}} \text{tr}(\tau_\mu \mathcal{A} \tau_\rho^\dagger \tau_\sigma^\dagger) + \\
&+ \underbrace{\epsilon^{\mu\nu 0\sigma}}_{\epsilon^{ijk}} \text{tr}(\tau_\mu \tau_\nu^\dagger \mathcal{A} \tau_\sigma^\dagger) + \underbrace{\epsilon^{\mu\nu\rho 0}}_{-\epsilon^{ijk}} \text{tr}(\tau_\mu \tau_\nu^\dagger \tau_\rho \mathcal{A}^\dagger) = \\
&= \epsilon^{ijk} \text{tr}(\tau_i^\dagger \tau_j \tau_k^\dagger - \tau_i \tau_j \tau_k^\dagger + \tau_i \tau_j^\dagger \tau_k^\dagger - \tau_i \tau_j^\dagger \tau_k) = 4\epsilon^{ijk} \text{tr}(\tau_i \tau_j \tau_k), \quad (3.66)
\end{aligned}$$

donde al final hemos usado que  $\tau_i^\dagger = -\tau_i$ . Notemos que los índices  $i, j, k$  van todos desde 1 hasta 3. Utilizando  $\tau_i = i\sigma_i$  y  $\sigma_i \sigma_j = i\epsilon_{ijl} \sigma_l$  (si  $i \neq j$ , que aquí es el caso),

$$4\epsilon^{ijk} \text{tr}(\tau_i \tau_j \tau_k) = -4i\epsilon^{ijk} \text{tr}(\sigma_i \sigma_j \sigma_k) = 4\epsilon^{ijk} \underbrace{\epsilon_{ijl}}_{2\delta_l^k} \text{tr}(\sigma^l \sigma_k) = 8\text{tr}(\sigma^l \sigma_l) = 24\text{tr}(\mathbb{1}_2) = 48. \quad (3.67)$$

Sustituyendo esto en (3.63), tenemos que

$$n = \frac{6}{\pi^2} \int (a - x^2 a^2)^2 d^4 x \quad (3.68)$$

Solo nos faltaría hacer la integral. Para ello, nos fijamos en que  $a - x^2 a^2 = a(1 - \frac{x^2}{x^2 + \lambda^2}) = a^2 \lambda^2$ , de forma que

$$\int_{\mathbb{R}^4} (a - x^2 a^2)^2 d^4 x = \int_{\mathbb{R}^4} \frac{\lambda^4}{(x^2 + \lambda^2)^2} d^4 x \stackrel{x=\lambda y}{=} \int_{\mathbb{R}^4} \frac{1}{(y^2 + 1)^2} d^4 y \quad (3.69)$$

Esta integral la podemos hacer en coordenadas hipersféricas  $(r, \theta_1, \theta_2, \theta_3)$ . En este sistema de coordenadas,  $d^4 y = d\Omega_3 r^3 dr$ , con  $\Omega_3$  el ángulo sólido de  $\mathbb{R}^4$ . Entonces, como  $y^2 = r^2$ ,

$$\int_{\mathbb{R}^4} \frac{1}{(y^2 + 1)^2} d^4 y = \int d\Omega_3 \int_0^\infty dr \frac{r^3}{(1+r^2)^4} = \frac{1}{2} \int d\Omega_3 \int_0^\infty dr^2 \frac{r^2}{(1+r^2)^4}. \quad (3.70)$$

La integral del ángulo sólido es el volumen de la 3-esfera, que sabemos que es  $2\pi^2$ . La integral de la parte radial se puede ver, con cualquier programa de cálculo, que da  $1/6$ , de forma que

$$n = \frac{6}{\pi^2} 2\pi^2 \frac{1}{2} \frac{1}{6} = 1 \quad (3.71)$$

Por tanto, podemos convencernos de que el cálculo del instantón BPST es correcto.

### 3.4. Cuaterniones

A los instantones con winding number  $|n| > 1$  se les llama multi-instantones. Para calcularlos, es útil introducir el conjunto de los cuaterniones (también llamados cuaternios) porque, como veremos a continuación, podemos usarlos para generar  $SU(2)$  y  $\mathfrak{su}(2)$  [4]. Los cuaterniones generalizan los números complejos y se definen como:

$$\mathbb{H} = \{x = x_0 + x_1i + x_2j + x_3k \mid x_0, x_1, x_2, x_3 \in \mathbb{R}\} \quad (3.72)$$

$i, j$  y  $k$  son elementos dados por las siguientes ecuaciones:

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k \quad jk = -kj = i \quad ki = -ik = j, \quad (3.73)$$

de las cuales podemos inferir que  $ijk = 1$  y que la multiplicación de cuaterniones no es conmutativa pero sí asociativa. Los cuaterniones (con la suma y la multiplicación ordinarias) se pueden ver como un espacio vectorial isomorfo a  $\mathbb{R}^4$  o a  $\mathbb{C}^2$  [4], por lo que tiene dimensión 4. Al igual que con los números complejos, se pueden definir conceptos como el módulo o el conjugado de un cuaternión,

$$|x| = \sqrt{x_0^2 + x_1^2 + x_2^2 + x_3^2}, \quad \bar{x} = x_0 - x_1i - x_2j - x_3k. \quad (3.74)$$

A  $x_0$  se le llama parte real de  $x$  y a  $x_1i + x_2j + x_3k$  se le llama parte imaginaria de  $x$ .<sup>3</sup>

Nos interesan los cuaterniones de módulo 1. Es decir, los que cumplen  $x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$  y, por tanto, los podemos ver como elementos de  $S^3$ . Resulta que estos cuaterniones son un grupo con la multiplicación y son isomorfos a  $SU(2)$ . Si queremos ver esto explícitamente [4], podemos usar la estructura de  $\mathbb{H}$  como  $\mathbb{C}^2$  y expresar dos cuaterniones  $x$  e  $y$  como

$$x = z_1 + z_2j, \quad y = w_1 + w_2j \quad z_1, z_2, w_1, w_2 \in \mathbb{C}. \quad (3.75)$$

Al expresar  $x$  e  $y$  de esta forma, puede no estar del todo claro dónde aparece el número imaginario  $k$ . Observemos que

$$w_2j = (\alpha_2 + i\beta_2)j = \alpha_2j + \beta_2(ij) = \alpha_2j + \beta_2k. \quad (3.76)$$

Si calculamos el producto de  $x$  e  $y$ ,

$$xy = z_1w_1 + z_1w_2j + z_2jw_1 + z_2jw_2j. \quad (3.77)$$

---

<sup>3</sup>También se les llama parte escalar y parte vectorial, respectivamente.

Desarrollando las partes real e imaginaria de cada número complejo y recordando que  $ij = -ji$ , tenemos que, por ejemplo,  $z_2 j w_1 = z_2 \overline{w_1} j$ , porque al mover la  $j$  cambia el signo de la parte imaginaria de  $w_1$ . Con un poco de cálculo llegamos a

$$xy = (z_1 w_1 - z_2 \overline{w_2}) + (z_1 w_2 + z_2 \overline{w_1})j = \begin{pmatrix} w_1 & w_2 \\ -\overline{w_2} & \overline{w_1} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = Yx. \quad (3.78)$$

La matriz  $Y$  es una representación de  $y$ . También es una matriz de la forma (1.71). Si imponemos que  $y$  tenga módulo 1, tenemos que  $Y^\dagger Y = \mathbb{1}_2$ ,  $\det(Y) = 1$  de forma que  $Y \in SU(2)$ , por lo que de verdad tenemos que los cuaterniones de módulo 1 son homeomorfos a  $SU(2)$ . En particular, como  $SU(2)/\mathbb{Z}_2 \cong SO(3)$ , los cuaterniones también generan el grupo de las rotaciones en el espacio (por eso en algunos libros de texto llaman  $i$ ,  $j$  y  $k$  a los vectores unitarios de la base canónica de  $\mathbb{R}^3$ ).

Si nos fijamos en los números imaginarios  $i$ ,  $j$  y  $k$  (que corresponden a  $\text{Im}(w_1) = 1$ ,  $\text{Re}(w_2) = 1$  y  $\text{Im}(w_2) = 1$  respectivamente en la ecuación (3.78), con todos los demás números nulos), resulta que podemos representarlos mediante las matrices de Pauli [4].

$$i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i\sigma_3 \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\sigma_2 \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = i\sigma_1. \quad (3.79)$$

Como las matrices de Pauli generan  $\mathfrak{su}(2)$ , los cuaterniones puramente imaginarios y de módulo 1 también generan  $\mathfrak{su}(2)$ .

Una vez vista esta introducción, volvamos a nuestra teoría de Yang-Mills en  $SU(2)$ . Un campo gauge toma valores en  $\mathfrak{su}(2)$ , pero acabamos de ver que es lo mismo que decir que toma valores en los cuaterniones puramente imaginarios, así que deberíamos poder encontrar una expresión de  $A = A_\mu(x)dx^\mu$  en términos de cuaterniones.

Para ello, definimos la siguiente diferencial valorado en  $\mathbb{H}$  [4]:

$$dx = dx^0 + dx^1 i + dx^2 j + dx^3 k. \in \mathbb{H} \quad (3.80)$$

Aquí los  $dx^0$ ,  $dx^1$ ,  $dx^2$  y  $dx^3$  son los mismos que llevamos usando hasta ahora. La diferencia es que ahora vamos a ver las coordenadas  $x = (x^0, x^1, x^2, x^3)$  como un cuaternión  $x = x^0 + x^1 i + x^2 j + x^3 k$ .

Este diferencial  $dx$  es una 1-forma cuyo valor es un cuaternión arbitrario, por lo que si queremos algo puramente imaginario debemos imponer que  $A = \text{Im}(\tilde{A}(x))dx = A_\mu dx^\mu$ , donde  $\tilde{A}(x) = \tilde{A}_0(x) + \tilde{A}_1(x)i + \tilde{A}_2(x)j + \tilde{A}_3(x)k$ .

Para aclarar un poco qué significa todo esto, vamos a encontrar la relación entre las componentes  $A_\mu$  que manejábamos hasta ahora y las componentes cuaterniónicas  $\tilde{A}_\mu$ :

$$\begin{aligned}\tilde{A}(x)dx &= (\tilde{A}_0 + \tilde{A}_1i + \tilde{A}_2j + \tilde{A}_3k)(dx^0 + dx^1i + dx^2j + dx^3k) = \\ &= (\tilde{A}_0 + \tilde{A}_1i + \tilde{A}_2j + \tilde{A}_3k)dx^0 + (-\tilde{A}_2 + \tilde{A}_0i + \tilde{A}_3j - \tilde{A}_2k)dx^1 + \\ &+ (-\tilde{A}_2 - \tilde{A}_3i + \tilde{A}_0j + \tilde{A}_1k)dx^2 + (-\tilde{A}_3 + \tilde{A}_2i - \tilde{A}_1j + \tilde{A}_0k)dx^3.\end{aligned}\quad (3.81)$$

Para este cálculo solo nos ha hecho falta usar las reglas (3.73). Entonces, tomando la parte imaginaria podemos identificar

$$A_0(x) = \tilde{A}_1i + \tilde{A}_2j + \tilde{A}_3k \quad (3.82)$$

$$A_1(x) = \tilde{A}_0i + \tilde{A}_3j - \tilde{A}_2k \quad (3.83)$$

$$A_2(x) = -\tilde{A}_3i + \tilde{A}_0j + \tilde{A}_1k \quad (3.84)$$

$$A_3(x) = \tilde{A}_2i - \tilde{A}_1j + \tilde{A}_0k. \quad (3.85)$$

Solo nos faltaría saber cómo se calcula la curvatura. Podríamos calcular la derivada covariante de  $Im(\tilde{A}(x)dx)$  directamente, pero como tomar la parte imaginaria lo podemos ver como hacer una "proyección" de los cuaterniones de módulo 1 a  $\mathfrak{su}(2)$  y nuestro  $\tilde{A}$  toma valores en  $\mathfrak{su}(2)$ , deducimos [4] que la derivada covariante debe conmutar con la operación "tomar la parte imaginaria". Por tanto

$$F = Im(D\tilde{A}(x)dx) = Im(d\tilde{A}(x) \wedge dx + (\tilde{A}(x)dx) \wedge (\tilde{A}(x)dx)). \quad (3.86)$$

Ya que hemos desarrollado todo esto, vamos a poner la solución de  $n = 1$  en términos de cuaterniones. Esto es bastante fácil ya que los números imaginarios no son más que las matrices de Pauli multiplicadas por  $i$ , que son las que llamamos  $\tau^\mu$ . Comparando con (3.61) para cada valor de  $\mu$  obtenemos:

$$\tilde{A}_0(x) = \frac{x_1i + x_2j + x_3k}{|x|^2 + \lambda^2} \quad (3.87)$$

$$\tilde{A}_1(x) = \frac{-x_0i - x_3j + x_2k}{|x|^2 + \lambda^2} \quad (3.88)$$

$$\tilde{A}_2(x) = \frac{x_3i - x_0j - x_1k}{|x|^2 + \lambda^2} \quad (3.89)$$

$$\tilde{A}_3(x) = \frac{-x_2i + x_1j - x_0k}{|x|^2 + \lambda^2} \quad (3.90)$$

Si ahora definimos el conjugado de  $dx$ ,  $d\bar{x} = dx^0 - dx^1i - dx^2j - dx^3k$ , podemos escribir esto de forma más compacta [4],

$$A(x) = Im\left(\frac{xd\bar{x}}{|x|^2 + \lambda^2}\right). \quad (3.91)$$

Para calcular la curvatura notamos que  $|x|^2 = x\bar{x}$  como en los números complejos.

$$F = \text{Im} \left( d \left( \frac{x}{x\bar{x} + \lambda^2} \right) \wedge d\bar{x} + \frac{(xd\bar{x}) \wedge (xd\bar{x})}{(x\bar{x} + \lambda^2)^2} \right) \quad (3.92)$$

Calculemos el primer sumando:

$$\begin{aligned} d \left( \frac{x}{x\bar{x} + \lambda^2} \right) \wedge d\bar{x} &= \frac{dx \wedge d\bar{x}}{x\bar{x} + \lambda^2} - \frac{x}{(x\bar{x} + \lambda^2)^2} d(x\bar{x}) \wedge d\bar{x} = \\ &= \frac{dx \wedge d\bar{x}}{x\bar{x} + \lambda^2} - \frac{|x|^2 dx \wedge d\bar{x}}{(x\bar{x} + \lambda^2)^2} - \frac{(xd\bar{x}) \wedge (xd\bar{x})}{(x\bar{x} + \lambda^2)^2} \end{aligned} \quad (3.93)$$

Hay una cancelación entre el segundo sumando de (3.93) y el segundo sumando de (3.92).

Notamos ahora que la forma  $dx \wedge d\bar{x}$  es puramente imaginaria:

$$\begin{aligned} dx \wedge d\bar{x} &= (dx^0 + dx^1 i + dx^2 j + dx^3 k) \wedge (dx^0 + dx^1 i + dx^2 j + dx^3 k) = \\ &= dx^0 \wedge dx^1 i + dx^0 \wedge dx^2 j + dx^0 \wedge dx^3 k + dx^1 \wedge dx^0 i + dx^2 \wedge dx^0 j + dx^3 \wedge dx^0 k + \\ &+ dx^1 \wedge dx^2 i j + dx^2 \wedge dx^1 j i + dx^1 \wedge dx^3 i k + dx^3 \wedge dx^1 k i + dx^2 \wedge dx^3 j k + dx^3 \wedge dx^2 k j. \end{aligned} \quad (3.94)$$

Los términos de la segunda línea se cancelan todos entre sí porque  $dx^i \wedge dx^j = -dx^j \wedge dx^i$ . Sin embargo, en la tercera línea, se suman porque el signo menos se compensa con las reglas (3.73). Por tanto,

$$dx \wedge d\bar{x} = 2dx^2 \wedge dx^3 i + 2dx^3 \wedge dx^1 j + 2dx^1 \wedge dx^2 k. \quad (3.95)$$

Entonces, vemos que  $F$  es puramente imaginaria y viene dada por una expresión bastante sencilla:

$$F = \left( 1 - \frac{|x|^2}{|x|^2 + \lambda^2} \right) \frac{dx \wedge d\bar{x}}{|x|^2 + \lambda^2} = \frac{\lambda^2}{(|x|^2 + \lambda^2)^2} dx \wedge d\bar{x} \quad (3.96)$$

A partir de esto se pueden crear otros instantones del mismo winding number mediante una transformación lineal  $x \rightarrow a(x + b)$  con  $a \in \mathbb{R}$  y  $b \in \mathbb{H}$ . De hecho, se puede probar [4] que estas son las únicas soluciones posibles.

La ventaja de estar trabajando con cuaterniones es que, una vez tenemos esta expresión para  $F$ , podemos construir el instantón de winding number -1 intercambiando  $x$  y  $\bar{x}$ . En términos del gauge, esto es equivalente a cambiar el papel de  $g$  y  $g^{-1}$  [4]. Además, como las matrices de Pauli tienen traza cero, para calcular, por ejemplo, la traza de  $F \wedge F$ , solo tenemos que tomar la parte real de su forma cuaterniónica.

Una última observación sobre esto es que hay que tener cuidado con esta expresión de  $A$  cuando trabajemos con el espacio compactificado  $S^4$ . Como  $S^4$  no se puede cubrir con una sola carta y las formas dependen explícitamente de las cartas, en algún momento vamos a tener que hacer un cambio. Para verlo, consideramos los abiertos  $0 \leq |x| \leq \epsilon$ , que en la esfera  $S^4$  serían entornos de los polos. Cuando pasamos a la zona de  $|x| \rightarrow \infty$ , el campo (3.91) tiende a gauge puro (en concreto a  $\bar{x}^{-1} d\bar{x}$ ), pero cuando  $|x| \rightarrow 0$ , este gauge puro diverge. Esto se arregla haciendo primero la transformación gauge (2.70) con  $g = \bar{x}^{-1}$  (a la que tiende  $A$ ) y luego, en la carta problemática, hacer el cambio de variable  $y = 1/x$ , lo que elimina la divergencia. La necesidad de este paso previo de hacer la transformación gauge pone de manifiesto que los fibrados principales que están asociados a los instantones no son triviales [4].

### 3.5. La construcción ADHM

Ahora, por fin, pasemos a calcular multi-instantones. En torno a los años 70, Atiyah, Drinfeld, Hitchin y Manin desarrollaron lo que se llama la construcción ADHM para calcular multi-instantones, tanto en  $SU(2)$  como en otros grupos gauge más complicados [11]. El desarrollo que haremos ahora con los cuaterniones es una forma particular de interpretar esta construcción.

En primer lugar, introduciremos el espacio  $\mathbb{H}^n$  de los vectores con  $n$  componentes que sean cuaterniones. Debería ser fácil ver que  $\mathbb{H}^n$  se puede tratar como un espacio real de dimensión  $4n$ .

Haremos otro *ansatz* para la forma del instantón con winding number  $n = k$  a partir de lo que acabamos de hacer para el caso  $n = 1$ ,

$$A = \text{Im} \left( \frac{\bar{v}^T(x) dv}{1 + |v|^2} \right), \quad (3.97)$$

donde  $v(x)$  es un vector columna de  $k$  cuaterniones.  $\bar{v}^T$  sería el vector fila  $(\bar{v}_1 \dots \bar{v}_k)^T$  con las componentes conjugadas y, como nos podríamos imaginar,  $|v|^2 = \sum_{i=1}^k \bar{v}_i v_i = \bar{v}^T v$ . Generalizando la transformación  $a(x+b)$  del caso  $n = 1$ , vamos a tomar el vector  $v$  de la siguiente forma:

$$v(x) = \bar{w}^T(x), \quad w = \vec{\lambda}(B - x \mathbb{1}_k) \quad (3.98)$$

$B$  es una matriz simétrica  $k \times k$  de cuaterniones y  $\vec{\lambda}$  es un vector fila de  $k$  cuaterniones. Estas dos cantidades no pueden ser arbitrarias. Para que  $A$  tal como lo hemos definido en (3.97) sea realmente un multi-instantón se deben cumplir las siguientes condiciones [4]

1.  $B^\dagger B + \vec{\lambda}^\dagger \vec{\lambda}$  es una matriz real

$$2. \text{ Si } \left. \begin{array}{l} (B - xI)z = 0 \\ \lambda \cdot z = 0 \end{array} \right\} \text{Entonces } z = 0 \quad (z \in \mathbb{H}^k)$$

La condición 1 nos asegura que al tomar el dual de Hodge de la curvatura  $F$  tengamos  $F = \star F$  [4].

La condición 2 nos dice que la matriz  $(k + 1) \times k$

$$\Lambda = \begin{pmatrix} B - xI \\ \vec{\lambda} \end{pmatrix} \quad (3.99)$$

tiene rango máximo. Se puede ver que estas condiciones son equivalentes a que la matriz  $R = \bar{\Lambda}^T \Lambda$  de dimensión  $k \times k$  sea real e invertible [11].

Calculemos la curvatura de (3.97). Como para el caso anterior, primero calculamos la derivada exterior

$$d \left( \frac{\bar{v}^T dv}{1 + |v|^2} \right) = \frac{d\bar{v}^T \wedge dv (1 + |v|^2)}{(1 + |v|^2)^2} - \frac{\bar{v}^T dv}{(1 + |v|^2)^2} d(\bar{v}^T v). \quad (3.100)$$

El segundo término por separado es

$$\frac{\bar{v}^T dv}{(1 + |v|^2)^2} d(\bar{v}^T v) = \frac{(\bar{v}^T dv) \wedge (d\bar{v}^T v)}{(1 + |v|^2)^2} + \frac{(\bar{v}^T dv) \wedge (\bar{v}^T \wedge dv)}{(1 + |v|^2)^2} = \frac{|v|^2 d\bar{v}^T \wedge dv}{(1 + |v|^2)^2} + \frac{(\bar{v}^T dv) \wedge (\bar{v}^T \wedge dv)}{(1 + |v|^2)^2} \quad (3.101)$$

El segundo término de (3.101) es exactamente  $A \wedge A$ , que se cancelará al hacer  $dA + A \wedge A$ . El cambio en el primero se debe a las condiciones que estamos imponiendo a  $v$  y a las matrices  $\vec{\lambda}$  y  $B$  que lo definen [11]. Juntándolo todo, obtenemos un resultado parecido a (3.96),

$$dA + A \wedge A = \left( 1 - \frac{|v|^2}{1 + |v|^2} \right) \frac{d\bar{v}^T \wedge dv}{1 + |v|^2} \Rightarrow F = \frac{d\bar{v}^T \wedge dv}{1 + |v|^2} \quad (3.102)$$

Existe una forma alternativa del *ansatz* que hemos hecho para  $A$  utilizando un vector columna  $N$  de  $k + 1$  cuaterniones tal que  $\bar{N}^T \Lambda = 0$  y  $\bar{N}^T N = \mathbb{1}_{k+1}$ . En este caso,  $A_\mu = \bar{N}^T \partial_\mu N = -(\partial_\mu \bar{N}^T) N$  y las componentes de la curvatura son [11]

$$F_{\mu\nu} = N^\dagger C R^{-1} (\tau_p \hat{e}_q - \tau_q \hat{e}_p) C^\dagger N, \quad (3.103)$$

donde  $C$  es la matriz que tiene unos por encima de la diagonal principal y el resto de sus entradas son cero y  $\hat{e}_1 = i, \hat{e}_2 = j, \hat{e}_3 = k, \hat{e}_0 = 1$ .

Para terminar, hagamos un ejemplo de multi-instantón utilizando lo que acabamos de ver. Escojamos un winding number  $n = k$  y una matriz  $B$  diagonal:

$$B = \text{diag}(b_1, b_2, \dots, b_k) \quad b_i \in H \forall i \quad (3.104)$$

El vector  $\vec{\lambda}$  lo vamos a escoger con componentes reales y positivas. Estas dos definiciones hacen que la condición 1 se cumpla trivialmente. Para la segunda condición construimos la matriz  $\Lambda$ :

$$\Lambda = \begin{pmatrix} b_1 - x & & & \\ & b_2 - x & & \\ & & \ddots & \\ & & & b_k - x \\ \lambda_1 & \lambda_2 & \dots & \lambda_k \end{pmatrix} \quad (3.105)$$

Si queremos que para cualquier  $z \in \mathbb{H}^k$  la única solución de  $\Lambda z = 0$  sea  $z = 0$ , basta con que todos los  $b_j$  ( $j = 1, \dots, k$ ) sean diferentes, ya que como los  $\lambda_i$  los tomamos reales y positivos, la ecuación que genera la última fila de  $A$  cumple que  $z = 0$  directamente.

Utilizando las definiciones, podemos llegar a la forma de  $\Lambda$  para estos valores de  $\vec{\lambda}$  y  $B$  [4],

$$A = \text{Im} \left( \frac{(-\lambda_1^2(b_1 - x) - \lambda_2^2(b_2 - x) - \dots - \lambda_k^2(b_k - x))d\bar{x}}{1 + \lambda_1^2|b_1 - x|^2 + \dots + \lambda_k^2|b_k - x|^2} \right). \quad (3.106)$$

Este  $A$  tiene la forma de  $k$  instantones del tipo BPST, ecuación (3.91), si les aplicamos una transformación  $x \rightarrow \lambda_i(b_i - x)$ .

Este tipo de soluciones no son las únicas que se pueden obtener. El multi-instantón (3.97) es invariante bajo las siguientes transformaciones:

- (a)  $\vec{\lambda} \rightarrow \vec{\lambda}T$ ,  $B \rightarrow T^{-1}BT$ ,  $T \in O(k)$ . Como  $\lambda$  y  $B$  son vectores y matrices de constantes, la matriz  $T$  también debe ser una matriz de constantes y por tanto  $dT = dT^{-1} = 0$ . Calculamos primero cómo cambia  $w$ :

$$w \rightarrow \vec{\lambda}T(T^{-1}BT - xId) = \vec{\lambda}(BT - TxId) = \vec{\lambda}(B - xId)T = wT. \quad (3.107)$$

Entonces,  $\vec{w}^T = v \rightarrow T^{-1}\vec{w}^T = T^{-1}v$ . Haciendo la derivada exterior obtenemos

$$\begin{aligned} dv \rightarrow d(T^{-1}v) &= T^{-1}dv \Rightarrow A \rightarrow \text{Im} \left( \frac{\vec{v}^T T(T^{-1}dv)}{1 + |v|^2} \right) = \\ &= \text{Im} \left( \frac{\vec{v}^T dv}{1 + |v|^2} \right) = A \end{aligned} \quad (3.108)$$

- (b)  $\vec{\lambda} \rightarrow u\vec{\lambda}$ , con  $u \in \mathbb{H}$ ,  $|u| = 1$ . Esta transformación no cambia nada del módulo de  $v$  y es fácil ver que en el numerador  $\vec{v}^T dv$  tampoco introduce cambios.

Teniendo en cuenta todos los objetos que intervienen en (3.97) y estas dos transformaciones, se puede calcular el número de parámetros libres y, por tanto, la dimensión del espacio de multi-instantones de winding number  $k$ :

- $\vec{\lambda}$  tiene  $4k$  parámetros ( $\mathbb{H}^k$  es un espacio real de dimensión  $4k$ )
- $B$ , al ser simétrica, tiene  $4k(k+1)/2$  parámetros.
- $T$ , al ser una matriz real de  $O(k)$ , tiene  $k(k-1)/2$  parámetros.
- El cuaternión  $u$  de la transformación (b) tiene 3 parámetros libres (el cuarto lo fija  $|u|^2 = 1$ ).

Ahora bien, a esto hay que restarle todas las restricciones que imponen las condiciones 1 y 2. Contando las ecuaciones reales [4], obtenemos  $3k(k-1)/2$ . Entonces, el número total de parámetros libres es:

$$4k + \frac{4k(k+1)}{2} - \frac{k(k-1)}{2} - 3 - \frac{3k(k-1)}{2} = 8k - 3 \quad (3.109)$$

Usando unos resultados llamados teoremas del índice, es posible demostrar [11] que estos  $8k - 3$  parámetros generan todos los posibles multi-instantones de winding number  $k$ .

### 3.6. Fibrados en el espacio de twistores

Lo que hemos visto de la construcción ADHM es su formulación más práctica en términos de cuaterniones. Detrás de todas las condiciones que hemos ido imponiendo en la sección anterior, hay aplicados numerosos resultados de la teoría de fibrados que está en la base de las teorías gauge.

Muchos de los detalles matemáticos son demasiado complicados para lo que se quiere hacer en este trabajo, así que solo daremos unas pinceladas.

En primer lugar vamos a definir los llamados espacios proyectivos. Para ello consideramos un punto  $0 \neq z \in \mathbb{C}^n$  y definimos una relación que identifique  $z$  con  $\lambda z$  para todo  $\lambda \in \mathbb{C} \setminus \{0\}$ . Al tomar el cociente de  $\mathbb{C}^n$  bajo esta relación obtenemos lo que se llama el espacio proyectivo  $\mathbb{C}\mathbb{P}^{n-1}$ , que es el conjunto de todas las "rectas" de  $\mathbb{C}^n$  que pasan por el origen. [4]

Estos espacios también se pueden ver como cocientes de grupos de simetrías de la siguiente forma: el espacio total  $\mathbb{C}^n$  es invariante bajo las transformaciones del grupo  $U(n)$ . Sin embargo, si queremos transformar todo  $\mathbb{C}^n$  dejando completamente fija una recta, que tiene grupo de simetría  $U(1)$ , perdemos un número considerable de simetrías. El espacio proyectivo se puede ver, entonces, como un

cociente del grupo de simetrías de  $\mathbb{C}^n$  sin la recta y el grupo de simetrías de  $\mathbb{C}^n$  con la recta:

$$\mathbb{C}\mathbb{P}^{n-1} = U(n)/(U(1) \times U(n-1)). \quad (3.110)$$

Estos procedimientos de la relación de equivalencia y de los grupos de simetría se puede hacer en  $\mathbb{R}^n$  o incluso con los cuaterniones, considerando el grupo adecuado.<sup>4</sup>

A nosotros nos va a interesar el espacio proyectivo de  $\mathbb{C}^4$ , que se suele llamar espacio de twistores y se denota  $\mathbb{C}\mathbb{P}^3 = \mathbb{T}\mathbb{P}$ .  $\mathbb{T}\mathbb{P}$  es una variedad compleja de tres dimensiones (es decir, 6 dimensiones reales).

Construimos ahora el espacio proyectivo al espacio cuaterniónico  $\mathbb{H}^2$ . Las rectas que pasan por el origen se pueden ver, a su vez, como rectas que intersecan a la esfera en  $\mathbb{H}^2$ , (que se puede ver como una 7-esfera real  $S^7$ ), en un punto  $(aq^1, aq^2)$ , donde  $q^1, q^2$  son cuaterniones y  $a$  es un cuaternión de módulo 1, es decir, de  $SU(2) \cong S^3$ .

Entonces, podemos decir [4] que  $\mathbb{H}\mathbb{P}^1$  es isomorfo a  $S^7/S^3 \cong S^4$ , que casualmente es la compactificación del espacio euclídeo  $\mathbb{R}^4$ .

Con estos dos espacios se puede construir un fibrado entre el espacio de twistores y  $S^4$  [4].

$$\begin{array}{ccc} \mathbb{C}\mathbb{P}^3 & = & \mathbb{T}\mathbb{P} \\ \downarrow \pi & & \downarrow \pi \\ \mathbb{H}\mathbb{P}^1 & = & S^4 \end{array}$$

Figura 3.2: Fibrado del espacio de twistores.

Entonces, podemos conectar el espacio de twistores con  $\mathbb{R}^4$  de la siguiente manera: compactificamos  $\mathbb{R}^4$  en un punto, obteniendo una 4-esfera  $S^4$ . Usando el fibrado entre  $\mathbb{T}\mathbb{P}$  y  $S^4$ , tenemos que para cada punto de  $\mathbb{R}^4$  hay una correspondencia con una fibra (que en este caso resulta ser una esfera  $S^2$ ) dentro de  $\mathbb{T}\mathbb{P}$ <sup>5</sup>.

El único detalle matemático que nos por falta ver es que en  $\mathbb{C}^4$  se pueden definir dos tipos especiales de planos. Tomando dos vectores  $X$  e  $Y$  linealmente independientes y tangentes a un plano  $P \subset \mathbb{C}^4$

<sup>4</sup>También, en vez de líneas, se pueden fijar objetos de más dimensiones para definir lo que se llama el grassmanniano.

<sup>5</sup>Como curiosidad, este es un caso particular de un fibrado de Hopf. Un fibrado de Hopf  $E \xrightarrow{\pi} M$  es tal que  $M = S^n$ ,  $F = S^p$  y  $E = S^{n+p}$ .

se puede definir la siguiente matriz antisimétrica  $4 \times 4$ :

$$\omega = Y \otimes X - X \otimes Y \quad (3.111)$$

$\omega$  es una 2-forma con valores matriciales, igual que la curvatura  $F$ . Entonces, podemos definir dos tipos de planos:

- Si  $\omega$  cumple  $\star\omega = \omega$ , se dice que  $P$  es un  $\alpha$ -plano.
- Si  $\star\omega = -\omega$ , se dice que  $P$  es un  $\beta$ -plano.

Estos dos tipos de planos se pueden construir a partir de los puntos  $(z_1, z_2)$ , definidos como

$$z_1 = \lambda_1(e_1 + ie_3) \quad z_2 = \lambda_2(e_2 + ie_4), \quad (3.112)$$

donde  $\lambda_1, \lambda_2 \in \mathbb{C}$  y  $\{e_1, e_2, e_3, e_4\}$  es una base de  $\mathbb{C}^4$ . El plano que forman estos puntos es un  $\beta$ -plano. Usando matrices del grupo ortogonal  $O(4, \mathbb{C})$  podemos transformar este plano en cualquier  $\alpha$  ó  $\beta$ -plano [4].

El conjunto de  $\beta$ -planos resulta ser  $\mathbb{CP}^3$  y el conjunto de  $\alpha$ -planos es parecido a  $\mathbb{CP}^3$  en el sentido de que los puntos y las líneas cambian sus papeles a la hora de definir el espacio proyectivo (a esto se le llama dual en el sentido de la geometría proyectiva) [4].

La relevancia de estos dos tipos de planos es que una curvatura  $F$  cumple  $F = \pm \star F$  si y solo si se hace idénticamente cero al restringirla a los  $\alpha$  (-) ó  $\beta$  (+) planos.

Este hecho se puede ver por componentes usando las formas  $\omega$  definidas en (3.111): por ejemplo, sea  $F$  tal que  $F = \star F$  y  $\omega$  tal que  $\omega = -\star\omega$  (es decir, define un  $\beta$ -plano). Entonces

$$F_{\mu\nu}\omega_{\mu\nu} = \left(\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}\right)\left(-\frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\omega_{\alpha\beta}\right) = -\frac{1}{4}\underbrace{\epsilon^{\mu\nu\rho\sigma}\epsilon^{\mu\nu\alpha\beta}}_{\delta^{\rho\alpha}\delta^{\sigma\beta}}F_{\rho\sigma}\omega_{\alpha\beta} = -F_{\rho\sigma}\omega_{\rho\sigma} \quad (3.113)$$

Concluimos que  $F_{\mu\nu}\omega_{\mu\nu} = 0$ .

Después de esto, vamos a dar una idea de cómo son exactamente estos fibrados en el espacio de twistores [13].

El punto de partida es una aplicación entre espacios vectoriales complejos,  $A(z) : W \rightarrow V$ , con  $W$  de dimensión  $k$  y  $V$  de dimensión  $2k + 2$ . El número  $k$  va a ser nuestro winding number.  $z$  es un vector de  $\mathbb{C}^4$  y  $A$  se puede expresar como

$$A = \sum_{i=1}^4 A_i z_i \quad (3.114)$$

Vamos a suponer que  $V$  tiene definido un producto escalar antisimétrico y no degenerado  $\langle v_1, v_2 \rangle = -\langle v_2, v_1 \rangle$ , de manera que podemos definir el aniquilador

$$U^0 = \{v \in V \mid \langle u, v \rangle = 0 \text{ para todo } u \in U\} \quad (3.115)$$

Le imponemos la siguiente condición a  $A(z)$ :

$$\text{Si } z \neq 0, U_z = A(z)(W) \subseteq V \text{ tiene dimensión } k \text{ y } U_z \subset (U_z)^0 \quad (3.116)$$

Bajo esta condición se puede ver [13] que el cociente  $E_z = (U_z)^0/U_z$  tiene dimensión 2 y  $E_z = E_{\lambda z}$  para todo  $\lambda \in \mathbb{C}$  (es decir, son un espacio proyectivo  $\mathbb{CP}^2$ ), de forma que podemos definir un fibrado  $E$  sobre el espacio de twistores  $\mathbb{TP} = \mathbb{CP}^3$  con los  $E_z$  como fibras [13].

Se pueden dar unas condiciones sobre el fibrado  $E$  [13] para que corresponda a un instantón de winding number  $k = \dim W$  de nuestra teoría de Yang-Mills  $SU(2)$ .

Para construir el campo gauge asociado, se tomamos un punto  $x \in S^4$  representado por una recta de  $\mathbb{CP}^3$  que junte los puntos

$$z = (z_1, z_2, z_3, z_4) \quad \text{y} \quad \sigma z = (-\bar{z}_2, \bar{z}_1, -\bar{z}_4, \bar{z}_3). \quad (3.117)$$

Si definimos un subespacio  $F_x \subset V$  de dos dimensiones que sea ortogonal a  $U_z$  y a  $U_{\sigma z}$  a la vez, obtenemos un fibrado  $F$  sobre  $S^4$  cuyas fibras son estos espacios  $F_x$ . A partir de este fibrado, se define una derivada covariante y de ahí se obtiene el campo gauge  $A$ . Además, existe una correspondencia entre la conexión de este fibrado  $F$  y el fibrado  $E$  construido sobre el espacio de twistores [13].

En [13] se muestra también que, al igual que en el caso de los cuaterniones, el número de parámetros libres que tiene el multi-instantón es  $8k - 3$ .

# Conclusiones

En este trabajo hemos desarrollado una introducción a la topología y geometría necesarias para entender con detalle las teorías gauge, que son una continuación de lo que se enseña en las asignaturas de Topología y Análisis de los grados de Matemáticas.

Juntando esto con lo aprendido en la asignaturas de Mecánica y Ondas, Mecánica Cuántica y Teoría Clásica de Campos, conseguimos desarrollar la acción y las ecuaciones del movimiento de una teoría gauge, para luego profundizar en los instantones de una teoría  $SU(2)$ . Calculamos explícitamente algunas de estas soluciones, tanto directamente como con la construcción ADHM, y pudimos explicar parte de su importancia en la práctica, como ser la contribución dominante en el cálculo de probabilidades de efecto túnel entre vacíos.

Como posibles expansiones de este trabajo, se podría profundizar más en la teoría de fibrados para ver la construcción de instantones con todo el detalle posible o ver cómo se adapta todo el contenido a una teoría de campos con fermiones. También se pueden ver el resto de consecuencias físicas que tienen los instantones en teoría de campos, algunas de las cuales ya mencionamos en la introducción del trabajo y que son relevantes dentro y más allá del Modelo Estándar.

# Bibliografía

- [1] G. 't Hooft, *The conceptual basis of quantum field theory*, in *Philosophy of Physics*, J. Butterfield and J. Earman, eds., Handbook of the Philosophy of Science, (Amsterdam), pp. 661–729, North-Holland (2007), DOI.
- [2] J. Munkres, *Topology*, Featured Titles for Topology, Prentice Hall, Incorporated (2000).
- [3] M. Nakahara, *Geometry, Topology and Physics*, CRC Press (2018).
- [4] C. Nash and S. Sen, *Topology and Geometry for Physicists*, Dover Books on Mathematics, Dover Publications (2013).
- [5] S. Fushida-Hardy, “A non-visual proof that higher homotopy groups are abelian.” Universidad de Stanford, <https://stanford.edu/sfh/homotopy.pdf>.
- [6] J. Lee, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, Springer New York (2013).
- [7] M. Spivak, *Calculus On Manifolds: A Modern Approach To Classical Theorems Of Advanced Calculus*, Avalon Publishing (1971).
- [8] A. Viña Escalar, *Geometría diferencial*, Universidad de Oviedo, Servicio de Publicaciones (2000).
- [9] D. Bleeker, *Gauge Theory and Variational Principles*, Dover Books on Mathematics, Dover Publications (2013).
- [10] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Graduate Texts in Mathematics, Springer International Publishing (2015).

- [11] E. Weinberg, *Classical Solutions in Quantum Field Theory: Solitons and Instantons in High Energy Physics*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (2012).
- [12] D. Tong, *Gauge theory*, Universidad de Cambridge, <https://www.damtp.cam.ac.uk/user/tong/gaugetheory.html>, 2018.
- [13] M. Atiyah, N. Hitchin, V. Drinfeld and Y. Manin, *Construction of instantons*, *Physics Letters A* **65** (1978) 185.