

## EXISTENCIA DE YANG–MILLS Y DEL SALTO DE MASA

ÓSCAR GARCÍA-PRADA

RESUMEN. En estas notas se da una introducción al problema propuesto por el Instituto Clay en su lista de problemas del milenio de **Existencia de Yang–Mills y del salto de masa**: Probar que para todo grupo de Lie compacto simple  $G$ , la teoría cuántica de Yang–Mills en  $\mathbb{R}^4$  existe y tiene un salto de masa  $\Delta > 0$ .

### INTRODUCCIÓN

En 1954 Chen-Ning Yang y Robert L. Mills introdujeron una teoría para describir la interacción débil (responsable entre otras cosas de ciertas formas de radiactividad) y la interacción fuerte (responsable entre otras cosas de la unión de protones y neutrones para formar un núcleo). Esta teoría ha sido fundamental en el estudio de partículas elementales y física nuclear en los últimos casi sesenta años.

La teoría de Yang–Mills, que emerge por un lado de la geometría diferencial y por otro lado de la física moderna, en particular de la mecánica cuántica, es una generalización de la teoría de Maxwell del electromagnetismo. No obstante, hay una diferencia esencial entre las fuerzas nucleares y la fuerza electromagnética: La fuerza electromagnética se extiende a distancias muy largas, mientras que las fuerzas nucleares son de muy corto alcance. Esto se traduce en que los campos responsables de las interacciones nucleares tienen que tener masa (en contraste con lo que sucede con los fotones responsables de la interacción electromagnética). Se dice en este caso que existe un salto de masa.

El problema es que en la teoría clásica de Yang–Mills las partículas no tienen masa. Sin embargo todos los experimentos, y en particular el fenómeno denominado libertad asintótica, indican que en la teoría cuántica los campos de Yang–Mills que describen las interacciones nucleares tienen masa no nula. El problema propuesto por el Instituto Clay de Matemáticas consiste en demostrar de modo matemáticamente riguroso la existencia de la teoría de Yang–Mills cuántica y la existencia del salto de masa.

El objetivo fundamental de estas notas es dar una introducción a este problema y describir los ingredientes básicos involucrados en el mismo.

---

Deseo dar las gracias a los organizadores de estas Jornadas por haberme invitado a dar estas charlas y por su generoso trabajo. Mi agradecimiento al Instituto Newton por las condiciones ideales durante la preparación de estas notas. Este trabajo está financiado en parte por el proyecto MTM2010–17717 del Ministerio de Ciencia e Innovación.

El plan de las notas es el siguiente: En la Sección 1 se hace un repaso rápido de la interacción electromagnética y las interacciones nucleares, así como del principio de invarianza gauge, que lleva al descubrimiento de la teoría de Yang–Mills. En la Sección 2 se aborda la estructura geométrica de las teorías de Yang–Mills y sus orígenes históricos en la geometría riemanniana y la relatividad general de Einstein. En la Sección 3 se aborda el problema de la cuantización de las teorías de Yang–Mills y se enuncia de modo preciso el problema de existencia de Yang–Mills y salto de masa. Finalmente se explica cuál es la situación del problema y se indican algunas de las posibles estrategias a seguir para abordar el problema.

Con la intención de hacer una descripción lo más simple y concisa posible, se ha elegido un estilo informal y esquemático. Por supuesto, estas notas son incompletas, y sólo pretenden ser una posible introducción al problema de Yang–Mills y salto de masa. El lector interesado en profundizar puede indagar en las referencias que se dan en la descripción oficial del problema por Arthur Jaffe y Edward Witten [3] y en el informe de Michael Douglas sobre el estado del problema [1].

## 1. INTERACCIONES FUNDAMENTALES Y SIMETRÍA GAUGE

**1.1. Electromagnetismo.** La interacción electromagnética la conocemos a través de los fenómenos eléctricos y magnéticos de la vida cotidiana. Para entender estos fenómenos tenemos que postular que la materia tiene un atributo al que llamamos *carga eléctrica*.



Charles A. Coulomb (1736–1806).

La ley de Coulomb establece que dos cargas estacionarias ejercen una fuerza mutua que es inversamente proporcional al cuadrado de la distancia que las separa. Esta fuerza es similar a la fuerza gravitatoria ejercida entre dos masas con las siguientes diferencias. Primero, la fuerza eléctrica es mucho más fuerte, con un

factor de  $10^{35}$ . Segundo, la carga eléctrica puede ser positiva o negativa, de tal modo que dos cargas de distinto signo se atraen y de igual signo se repelen.

Del mismo modo que la masa actúa como fuente de un campo gravitatorio, la carga eléctrica es la fuente de un campo eléctrico. Una carga puntual  $q$  crea un campo eléctrico radial alejándose de ella misma, de una magnitud inversamente proporcional a la distancia al cuadrado de la carga. Este es el denominado *campo de Coulomb*:

$$\text{Campo de Coulomb} = \frac{q}{r^2}.$$

Otra carga  $q'$  en este campo experimenta una fuerza radial igual a  $q'$  veces el campo. La fuerza puede ser atractiva o repulsiva dependiendo de si el signo de  $q'$  es distinto o igual al de  $q$ .

Una manera de visualizar el campo eléctrico es a través de las “líneas de fuerza” tangentes a la dirección del campo en cada punto. El “flujo” se define como el número de líneas que cruzan la unidad de área perpendicular a la dirección del campo.

Si dibujamos una esfera de radio  $r$  alrededor de una carga eléctrica, la superficie de la esfera aumenta con  $r$  como  $r^2$ . Como el campo eléctrico decrece como  $r^{-2}$ , el número de líneas de flujo que atraviesa la esfera es una constante que depende de la carga. Esta propiedad geométrica, conocida como *ley de Gauss*, y que es equivalente a la ley de Coulomb, nos hace ver que no hay que pensar en la interacción eléctrica como algo que se hace más débil con el cuadrado de la distancia sino como algo que se propaga. Por supuesto, según nos alejamos el campo es menos intenso en un punto dado, pero la cantidad total de flujo alrededor de la esfera es la misma con la distancia. Vemos pues claramente que la electricidad es una fuerza de largo alcance.

El *potencial de Coulomb* o *potencial escalar* debido a una carga puntual  $q$  se define como

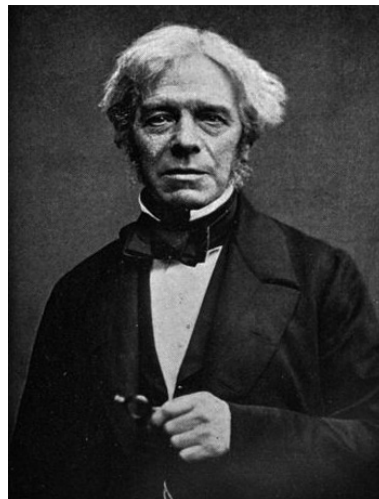
$$\text{Potencial de Coulomb} = \frac{q}{r}.$$

Una colección de cargas define un potencial escalar  $\phi$  que es la suma de los potenciales de Coulomb individuales. El campo eléctrico  $\vec{E}$  se expresa en términos del potencial escalar como

$$\vec{E} = -\vec{\nabla}\phi.$$

Nuestra primera experiencia con el magnetismo tiene que ver con la atracción que un imán ejerce sobre una partícula de hierro. Este fenómeno se describe por medio de un campo magnético que ejerce una fuerza sobre la partícula de hierro.

Hans Christian Oersted hizo el importante descubrimiento de que una corriente eléctrica genera un campo magnético. No hay análogo magnético de la carga. La fuente más simple de un campo magnético no es un “monopolo magnético” sino un “dipolo magnético”, que es equivalente a un lazo de corriente. Esto hace que los fenómenos magnéticos parezcan más complicados que los eléctricos.



Hans Christian Oersted (1777–1851) y Michael Faraday (1791–1857).

Puesto que no hay cargas magnéticas las líneas de fuerza del campo magnético no pueden terminar y tienen que ser líneas cerradas. En términos matemáticos lo que esto significa es que la divergencia del campo magnético  $\vec{B}$  es nula, es decir,  $\vec{\nabla} \cdot \vec{B} = 0$  y, por lo tanto, existe un campo vectorial  $\vec{A}$ , denominado *potencial vectorial* tal que

$$\vec{B} = \vec{\nabla} \times \vec{A}.$$

De acuerdo con los experimentos de Oersted, las cargas en movimiento generan un campo magnético. En otras palabras, un campo eléctrico variable produce un campo magnético. Michael Faraday descubrió el proceso inverso: Un campo magnético variable genera un campo eléctrico. Este proceso se denomina *inducción electromagnética*. Faraday inventó la dinamo basándose en este fenómeno.

La persona que puso todo esto en orden y elaboró la teoría unificada del electromagnetismo fue James Clerk Maxwell. El resultado son las ecuaciones de Maxwell:

$$\begin{aligned}\vec{\nabla} \cdot \vec{E} &= 4\pi\rho \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{E} &= -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \times \vec{B} &= \frac{4\pi}{c} \vec{j} + \frac{1}{c} \frac{\partial \vec{E}}{\partial t},\end{aligned}$$

donde  $\rho$  es la densidad de carga y  $\vec{j}$  es la densidad de corriente, definida como el producto de la densidad de carga por la velocidad, es decir  $\vec{j} = \rho\vec{v}$ .

Estas ecuaciones implican que una perturbación en el campo electromagnético se propaga con velocidad constante  $c$ . Esto le hizo predecir a Maxwell la existencia de



James Clerk Maxwell (1831–1879) y Henrich R. Hertz (1857–1894).

la radiación electromagnética, cosa que Heinrich Hertz verificó 30 años más tarde en el laboratorio, encontrando que  $c$  es igual a la velocidad de la luz. Volveremos a este punto un poco más tarde.

**1.2. Relatividad especial y electromagnetismo.** Los físicos pensaron que puesto que las ondas electromagnéticas se propagan, éstas deben propagarse en un cierto medio (como las ondas en el agua), al que denominaron “éter”. En su famoso experimento, Michelson y Morley detectaron que la velocidad de la luz no depende de la dirección de emisión.

De esto, Einstein concluyó lo que parecía obvio, pero no por ello menos revolucionario: la luz se propaga con una velocidad constante para todos los observadores, y no hay ningún medio salvo el vacío. Esta hipótesis transforma de modo radical nuestra concepción del espacio-tiempo, como veremos en más detalle.

Un principio básico de la física es que una ley física debe ser independiente del observador. Esto quiere decir que la ley debe ser expresada por una ecuación que tiene la misma forma en todos los sistemas de referencia. Los físicos dicen en este caso que la ley física es *covariante* con respecto a las leyes de transformación entre un sistema de referencia y otro.

Consideremos dos observadores que se mueven a una velocidad relativa  $v$ . El sentido común nos dice que el tiempo  $t$  para ambos observadores es el mismo, y que la posiciones de un objeto medida por estos observadores, denotadas por  $x$  y  $x'$ , respectivamente, difieren en una cantidad determinada por la velocidad relativa, es decir

$$\begin{aligned}t' &= t \\x' &= x - vt.\end{aligned}$$



Albert Einstein (1879–1955).

Esta ley de transformación se denomina *transformación galileana*.

La ecuación de Newton  $\vec{F} = m\vec{a}$  es covariante con respecto a la transformación galileana ya que las componentes de los vectores  $\vec{F}$  y  $\vec{a}$  cambian de un sistema de referencia a otro, pero la relación es la misma en todos los sistemas.

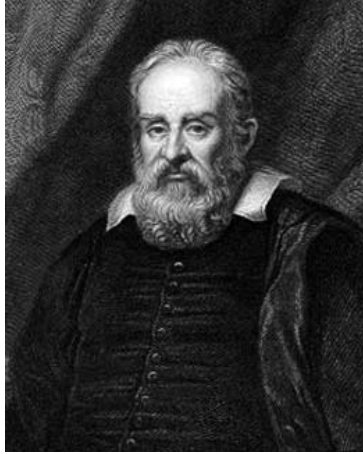
Sin embargo, las ecuaciones de Maxwell no son covariantes con respecto a las transformaciones galileanas, ya que la velocidad de la luz debe ser constante en todos los sistemas de referencia de acuerdo a Einstein.

Esto significa que debemos encontrar la ley de transformación bajo la cual las ecuaciones de Maxwell sean covariantes, además de corregir la ecuación de Newton de modo que la nueva ecuación sea covariante con respecto a la nueva ley de transformación.

La clave para encontrar estas transformaciones está en combinar el tiempo con las tres coordenadas espaciales, formando un espacio-tiempo de dimensión 4 y observar que ni el intervalo temporal ni el intervalo espacial de dos eventos relativamente simultáneos tiene sentido sino que la cantidad que es independiente del observador es el intervalo espacio temporal definido por

$$I = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2.$$

Esta expresión define una noción de “métrica” en el espacio-tiempo de modo análogo a la distancia euclídea. Pues bien, las transformaciones del espacio-tiempo que buscamos deben preservar esta métrica. Este es el modo en el que Einstein formuló el problema. El resultado viene dado por las *transformaciones de Lorentz*:



Galileo Galilei (1564–1642) e Isaac Newton (1643–1727).

$$t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}}$$

$$x' = \frac{x - vt/c^2}{\sqrt{1 - v^2/c^2}}.$$

Cuando  $v/c \rightarrow 0$ , estas transformaciones se reducen a las transformaciones galileanas.

Podríamos a partir de aquí derivar la forma covariante de la ley de Newton, etc., pero nos conformaremos con escribir la forma covariante de la energía. Sabemos que en mecánica clásica la energía es

$$E = \text{Energía cinética} + \text{Energía potencial}$$

$$= \vec{p} \cdot \vec{p}/2m + V(x),$$

donde  $\vec{p} = m\vec{v}$  es el momento.

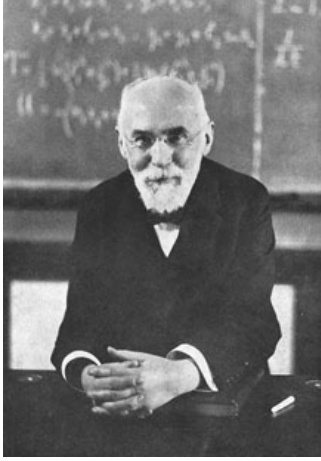
En relatividad

$$(1) \quad E^2 = \vec{p} \cdot \vec{p}c^2 + m^2c^4.$$

En particular, si  $p = 0$ , tenemos la famosa fórmula  $E = mc^2$ .

Podemos ahora también escribir la forma covariante de las ecuaciones de Maxwell. Para ello trabajamos con coordenadas en el espacio-tiempo 4-dimensional, usando la notación

$$\text{4-vector : } x^\mu = (ct, x, y, z) \quad (\mu = 0, 1, 2, 3).$$



Hendrik A. Lorentz (1853–1928) y Hermann Minkowski (1864–1909).

Por simplicidad denotaremos un 4-vector por  $x$  en lugar de  $x^\mu$  si no hay ambigüedad. Su forma covariante  $x_\mu$  (o elemento correspondiente del espacio vectorial dual usando la métrica de Lorentz) es

$$x_\mu = (ct, -x, -y, -z).$$

Así pues el producto Lorentziano de dos 4-vectores  $x$  e  $y$  es  $x \cdot y = x^\mu y_\mu$ , donde usamos la convención de Einstein de sumar sobre índices repetidos. El espacio  $\mathbb{R}^4$  con este producto se denomina, a veces, espacio-tiempo de Minkowski.

Volviendo a las ecuaciones de Maxwell, podemos definir el potencial 4-vector

$$A = (\phi, \vec{A}),$$

y el 4-vector de densidad de corriente

$$j = (c\rho, \vec{j}).$$

Los campos eléctrico y magnético son componentes de un campo tensorial antisimétrico definido por medio del potencial vector como

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu.$$

Explícitamente,

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}.$$

Este campo tensorial tiene un dual  $\tilde{F}^{\mu\nu}$ , obtenido cambiando  $\vec{E} \mapsto \vec{B}$  y  $\vec{B} \mapsto -\vec{E}$ . En su forma covariante las ecuaciones de Maxwell tienen la forma

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= -\frac{4\pi}{c} j^\nu \\ \partial_\mu \tilde{F}^{\mu\nu} &= 0.\end{aligned}$$

Desde el punto de vista de la covarianza, el potencial vector es la variable esencial. No obstante la definición  $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$  no determina unívocamente  $A^\mu$ . La transformación

$$A^\mu \rightarrow A^\mu + \partial^\mu \chi,$$

donde  $\chi$  es una función en el espacio-tiempo, define un potencial vector que determina el mismo tensor  $F^{\mu\nu}$ . Estas transformaciones se llaman *transformaciones gauge* por razones históricas que se explicarán más tarde. Por este motivo a  $A$  se le denomina “potencial gauge”. Todas las propiedades físicas dependen solamente de los campos eléctrico y magnético y por lo tanto deben ser “invariantes del gauge”.

En términos del campo gauge la ecuación  $\partial_\mu \tilde{F}^{\mu\nu} = 0$  se satisface automáticamente y la primera se reduce a

$$\frac{1}{c^2} \frac{\partial^2 A}{\partial t^2} - \nabla^2 A = \frac{4\pi}{c} j.$$

Esto nos dice que la densidad de corriente es la fuente del campo gauge, y el campo se propaga como una onda que viaja a velocidad constante  $c$ . Además la densidad de energía de interacción viene dada por  $j \cdot A$ .

Observemos que cualquier campo que satisfaga una ecuación de onda como ésta describe un fenómeno de largo alcance (algo que se propaga infinitamente). Una manera de describir un fenómeno de corto alcance sería el dado por una ecuación como la anterior pero modificada añadiendo un término extra del modo siguiente:

$$(2) \quad \frac{1}{c^2} \frac{\partial^2 A}{\partial t^2} - \nabla^2 A = -\frac{1}{L^2} A.$$

Esta ecuación daría soluciones de baja energía localizadas en una distancia  $L$  de modo que

$$A \sim \exp(r/L).$$

Esto describe una fuerza de corto alcance con escala de distancia  $L$ . Volveremos más tarde a esta ecuación.

**1.3. Gauge de Weyl.** La idea fundamental del principio de simetría gauge apareció por primera vez en el intento temprano de Hermann Weyl de unificar el electromagnetismo y la teoría de la gravitación de Einstein. La teoría de la relatividad general de Einstein, de la que hablaremos más adelante, reduce la interacción gravitatoria a la geometría del espacio-tiempo, de modo que la fuerza gravitatoria es una consecuencia de la curvatura del mismo. En presencia de curvatura, cuando un vector es transportado de manera paralela a lo largo de una curva cerrada, el vector transportado forma un ángulo con el vector original que es proporcional al

flujo del campo gravitatorio que atraviesa la curva, y da una medida de la curvatura del espacio-tiempo (esto puede experimentarse, por ejemplo, sobre la superficie de una balón de fútbol).



Hermann Weyl (1885–1955).

Weyl había observado la invarianza conforme de las ecuaciones de Maxwell y se propuso explotar este hecho interpretando el campo electromagnético como una distorsión de la longitud relativista producida cuando un vector se mueve en transporte paralelo a lo largo de una curva cerrada. Más precisamente, lo que propuso fue modificar el tensor métrico de Einstein (véase la Sección 2) con un factor de escala

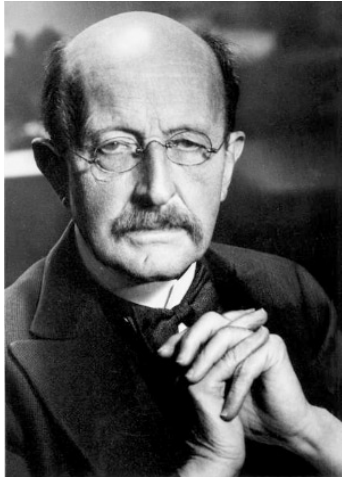
$$\exp \frac{q}{\gamma} \int A_{\mu} dx^{\mu},$$

donde  $A_{\mu}$  es el potencial gauge,  $q$  es la carga y  $\gamma$  es una constante.

La idea de Weyl es de una gran belleza, pero resultó ser incorrecta. Como fue observado por Einstein, desde el punto de vista físico la idea de Weyl es insostenible ya que si la longitud de nuestra regla de medir se acorta cada vez que damos una vuelta en un camino cerrado, entonces la idea de longitud (relativista) no tiene ningún sentido. En contraste a lo que sucede con la dirección del vector, la longitud tiene que tener un único valor.

Resulta, sin embargo que la idea de Weyl era casi correcta, pero en un contexto totalmente diferente: la mecánica cuántica. Esta teoría viene descrita por funciones de onda complejas y pronto resultó claro que un cambio de fase (haciendo  $\gamma = i\hbar$ , donde  $i = \sqrt{-1}$  y  $\hbar$  es la constante de Planck dividida por  $2\pi$ ), más que de escala, era lo correcto para las ecuaciones de Maxwell. O en lenguaje moderno, como explicaremos más adelante, que el grupo gauge era el círculo y no el grupo multiplicativo de los números reales positivos.

Resultó pues que, mientras que los cambios de escala cuadraban bien en la teoría de Einstein reemplazando la métrica por una estructura conforme, no había



Max Planck (1858–1947) y Louis De Broglie (1892–1987).

espacio en la relatividad general para incorporar una fase. Más bien la teoría gauge tenía que ser impuesta como una estructura adicional (como veremos más tarde) y la unificación perseguida por Weyl se desvaneció.

No obstante, el nombre de “gauge” (que significa “escala” o “calibre”) se ha mantenido al referirse al cambio de fase.

**1.4. Mecánica cuántica.** Max Planck en 1900 observó que muchas propiedades de la luz no podían ser explicadas aplicando simplemente las ecuaciones de Maxwell. Era necesario considerar una propiedad adicional: la energía de la luz da saltos o está “cuantizada” en paquetes de valor

$$E = h\nu,$$

donde  $\nu$  es la frecuencia y  $h$  es la constante de Planck:

$$h = 6,63 \times 10^{-27} \text{ ergios-segundo.}$$

Esta idea fue utilizada por Niels Bohr para explicar los niveles de energía del átomo de hidrógeno.

De esto, Planck concluyó que las ondas se comportan en cierto modo como partículas. Louis De Broglie dio un paso más proponiendo que si las ondas se comportan como partículas ¿por qué no podemos considerar que las partículas (electrón, protón o cualquier otra partícula) se comportan como ondas? Es decir, todo son ondas y partículas al mismo tiempo (los dos puntos de vista son compatibles y dan información). Ciertamente esto suena extraño, pero ya nos previene Bohr de las peculiaridades de la mecánica cuántica cuando afirma:

*Cualquiera que no esté impactado con la teoría cuántica es que no la ha entendido.*



Niels Bohr (1885–1962).

Así pues si tenemos una partícula, por ejemplo un electrón, éste puede ser descrito como una onda, donde la energía está relacionada con la frecuencia mediante la fórmula

$$E = h\nu$$

y el momento está relacionado con la longitud de onda mediante la fórmula

$$p = \frac{h}{\lambda}.$$

Pero si las partículas son ondas, la pregunta natural es cuál es la ecuación que gobierna a estas ondas. La persona que dio con esta ecuación fue Erwin Schrödinger. Lo que propuso es lo siguiente: la ecuación involucra una función  $\psi(x)$  que toma valores en los números complejos. El valor de  $|\psi(x)|^2$  da una medida de la probabilidad de encontrar la partícula en el punto  $x$ . La fase de  $\psi(x)$  da lugar a fenómenos de interferencia característicos de las ondas. Para obtener la ecuación que satisface la función de onda  $\psi$ , a partir de la ecuación clásica para la energía y el momento se reemplaza  $E$  por  $i\hbar\frac{\partial}{\partial t}$  y  $\vec{p}$  por  $-i\hbar\vec{\nabla}$ , donde  $\hbar = h/2\pi$ .

Si tomamos la ecuación no relativista

$$E = \vec{p} \cdot \vec{p}/2m + V(x)$$

obtenemos la famosa ecuación de Schrödinger

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\vec{\nabla}^2\psi + V(x)\psi.$$

Si ahora consideramos la ecuación relativista que relaciona la energía y el momento dada por

$$E^2 = \vec{p} \cdot \vec{p}c^2 + m^2c^4,$$

obtenemos la ecuación de Klein–Gordon (de hecho, obtenida originalmente por Schrödinger)

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -\hbar^2 c^2 \vec{\nabla}^2 \psi + m^2 c^4 \psi,$$

y dividiendo por  $-\hbar^2 c^2$ ,

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} &= \vec{\nabla}^2 \psi - \frac{m^2 c^2}{\hbar^2} \psi \\ &= \vec{\nabla}^2 \psi - \frac{1}{L^2} \psi. \end{aligned}$$

donde  $L = \hbar/mc$  se denomina “longitud de onda de Compton”.

Esta es precisamente la ecuación (2) que mencionamos anteriormente para describir un fenómeno cuya propagación decrece rápidamente.

Vemos que exactamente esa ecuación se deriva de la mecánica cuántica y que justamente la razón con la que disminuye la propagación tiene que ver con la masa, de tal modo que si la partícula tiene masa  $m \neq 0$  tenemos un fenómeno de corto alcance de longitud típica  $1/m$ . Las partículas más pesadas dan lugar a un rango más pequeño. Si la partícula no tiene masa entonces el término  $\frac{m^2 c^2}{\hbar^2} \psi$  desaparece de la ecuación, tenemos la ecuación de ondas y todo se propaga indefinidamente en el espacio.

Esto es exactamente el salto de masa en el problema de Yang–Mills. El problema del salto de masa consiste en descubrir por qué los objetos que juegan el papel de la luz en las interacciones nucleares tienen masa.

**1.5. El principio gauge cuántico.** Consideremos ahora la adaptación de la teoría gauge de Weyl a la mecánica cuántica: Una transformación gauge en mecánica cuántica involucra tanto al potencial gauge  $A$  como a la función de onda  $\psi$  que describe a una partícula cargada con carga  $q$ . Consiste en la operación conjunta

$$A \mapsto A + \partial\alpha, \quad \psi \mapsto U\psi,$$

donde  $U$  es el factor de fase dado por:

$$U = \exp\left(\frac{iq}{\hbar c}\alpha\right).$$

Ahora bien, para que la ecuación de Schrödinger sea invariante por transformaciones gauge debemos reemplazar  $\partial$  por

$$D = \partial + \frac{iq}{\hbar c}A.$$

Este operador se denomina “derivada covariante” y como veremos en más detalle en la Sección 2 este objeto constituye un ejemplo de lo que en geometría se conoce como conexión en un fibrado, más concretamente, una  $U(1)$  conexión. Es claro que la ecuación de Schrödinger es ahora invariante gauge porque el término proveniente

de  $A$  se cancela a través de la acción de  $\partial$  sobre el factor de fase  $U$  que multiplica la función de onda.

El principio de simetría gauge puede establecerse del siguiente modo. Para que un sistema pueda acoplarse con el electromagnetismo, éste debe ser invariante por una transformación gauge global. Esto significa que la ecuación de Schrödinger debe ser invariante por un cambio de fase constante, es decir,

$$\psi \mapsto e^{i\alpha}\psi,$$

donde  $\alpha$  es una constante. Por supuesto la ecuación de Schrödinger usual tiene tal invarianza. Ahora bien, para que el sistema sea invariante por una *transformación gauge local*, es decir, permitiendo que  $\alpha$  sea una función, entonces tenemos que reemplazar  $\partial$  por  $D$ . Por decirlo de otro modo, el acoplamiento a un campo gauge eleva una invarianza gauge global a una invarianza gauge local.

**1.6. Fuerzas nucleares y teoría de Yang–Mills.** Hemos visto que la interacción electromagnética se introduce aplicando el principio de simetría gauge al grupo  $U(1)$  de los números complejos de norma unidad. Yang y Mills generalizaron este principio al grupo  $SU(2)$  de matrices complejas  $2 \times 2$  unitarias con determinante unidad, obteniendo una generalización de las ecuaciones de Maxwell que describe todas las interacciones fundamentales entre partículas elementales.



Chen-Ning Yang (1922– ) y Robert L. Mills (1927–1999).

Para explicar lo que Yang y Mills hicieron, pongámonos en el contexto de la física nuclear de los años 1950. En esos años los físicos estaban estudiando neutrones, protones y mesones  $\pi$ . Hay tres tipos de mesones  $\pi$  según su carga sea neutra, positiva o negativa:  $\pi^0, \pi^+, \pi^-$ . La cosa curiosa es que los mesones  $\pi$  podían hacer que un protón se convirtiese en un neutrón o que un neutrón se convirtiese en un protón, o podían ser absorbidos por un protón o por un neutrón:

$$p \longleftrightarrow n + \pi^+,$$

$$n \longleftrightarrow p + \pi^-.$$

La idea clave para explicar estos fenómenos está en la conservación del denominado *isoespín*, un atributo similar al espín. Desde el punto de vista de las fuerzas nucleares, el protón y el neutrón se comportan básicamente del mismo modo. Podemos pensar en ellos como dos estados distintos de una misma partícula a la que llamamos *nucleón*, que tiene isoespín  $1/2$ . Matemáticamente podemos describir un nucleón con una función de ondas con dos componentes:

$$\psi = \begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix} \in \mathbb{C}^2.$$

Podemos ahora representar a los mesones por medio de matrices  $2 \times 2$  del modo siguiente:

$$\begin{aligned} \pi^+ &\sim \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ \pi^- &\sim \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ \pi^0 &\sim \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

La conservación del isoespín es equivalente a la invarianza bajo la acción de  $SU(2)$  (una simetría gauge global con grupo  $SU(2)$ ). La gran idea de Yang-Mills fue hacer lo que Weyl había propuesto en el caso del electromagnetismo pero con matrices. En otras palabras, considerar derivadas covariantes

$$D_\mu = \partial_\mu + A_\mu,$$

donde  $A_\mu$  son matrices  $2 \times 2$  (más precisamente, matrices antihermíticas de traza nula, es decir elementos del álgebra de Lie de  $SU(2)$ ). Como en el caso del electromagnetismo, podemos definir un campo tensorial  $F$  y ecuaciones similares a las de Maxwell (más adelante daremos la forma precisa de  $F$  y de estas ecuaciones).

La gran diferencia es que ahora, debido a que  $SU(2)$  es un grupo no abeliano,  $F$  involucra términos cuadráticos en  $A$ , dando lugar a ecuaciones diferenciales que, en contraste con las ecuaciones de Maxwell, son no lineales, y como consecuencia los campos de fuerza (mesones  $\pi$ ) actúan sobre sí mismos, cosa que no sucede con los rayos de luz.

Sólo hay algo que no funciona en ésta teoría. El problema es que los protones y los neutrones no son partículas elementales. Estos están formados por *quarks*. Esta idea fue promovida por Richard P. Feynman y Murray Gell-Mann. La idea es que hay dos estados para un quark:  $u$  y  $d$  (“up” y “down” en inglés, respectivamente). Un nucleón está constituido por tres quarks y los mesones son la combinación de un quark y un antiquark:

$$\begin{aligned} p &= \{uud\}, & n &= \{udd\} \\ \pi^+ &= \{u\bar{d}\} & \pi^- &= \{d\bar{u}\} \\ \pi^0 &= \text{combinación de } u\bar{u} \text{ y } d\bar{d}. \end{aligned}$$



Richard P. Feynman (1918–1988) y Murray Gell-Mann (1929– ).

Aunque ahora las partículas fundamentales no sean los nucleones, podemos aplicar a los quarks la misma idea y considerar los quarks  $u$  y  $d$  como dos componentes de un mismo objeto básico y aplicar las ideas de Yang y Mills. Este proceso culmina en el desarrollo del denominado *modelo estándar* por Sheldon Glashow, Abdus Salam y Steven Weinberg, que describe las interacciones electromagnética, débil y fuerte.

Un quark viene descrito por una función

$$\psi_{\text{quark}} = \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix}.$$

El grupo de matrices  $SU(2)$  actúa mezclando  $\psi_u$  y  $\psi_d$ . Este mismo grupo mezcla también electrones y neutrinos.

El siguiente punto importante es que los quarks tienen una simetría extra denominada *color* y que existen tres colores (rojo, azul, amarillo, digamos), es decir,

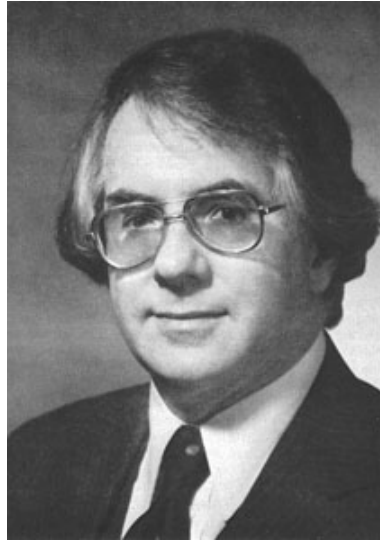
$$\psi_u = \begin{pmatrix} \psi_{u,\text{rojo}} \\ \psi_{u,\text{azul}} \\ \psi_{u,\text{amarillo}} \end{pmatrix} \in \mathbb{C}^3$$

y análogamente para  $\psi_d$ .

La conservación del color es equivalente a la invarianza bajo la acción de  $SU(3)$  (simetría gauge global con grupo  $SU(3)$ ). Las matrices de  $SU(3)$  mezclan las tres componentes de  $\psi_u$  y  $\psi_d$ . Además, como todas las funciones son complejas, podemos multiplicar por una fase en  $U(1)$ .

Vemos entonces que el modelo estándar está basado en el grupo  $SU(3) \times SU(2) \times U(1)$ .

El grupo  $SU(3)$  proporciona ocho tipos de campos denominados *gluones*, que transportan la interacción fuerte y que son responsables de que los constituyentes de un protón o de un neutrón estén juntos formando una sola unidad.



Sheldon Glashow (1932- ).

Los tres campos de  $SU(2)$  y el único campo de  $U(1)$  se combinan para dar los campos  $W^+$ ,  $W^-$  y  $Z$  de la interacción débil y el fotón del electromagnetismo. Es importante observar que el campo gauge del electromagnetismo no se obtiene directamente del factor  $U(1)$  sino combinando éste con  $SU(2)$ . Este es justamente el contenido de la unificación de las interacciones electromagnética y débil de Glashow, Salam y Weinberg en una única fuerza denominada interacción *electrodébil*.

El modelo estándar es un gran logro de la física moderna y es ciertamente la base para comprender las interacciones entre las partículas elementales. No obstante hay grandes problemas que todavía no se comprenden. Desde el punto de vista físico está el problema del salto de masa. ¿Por qué las fuerzas nucleares son de corto alcance? O de otro modo, ¿por qué hay partículas gauge que tienen masa? En el modelo de Yang-Mills los campos gauge no tienen masa, e introducir un término de masa en la ecuación de onda rompe la simetría gauge. Otro problema relacionado para el que todavía tampoco hay explicación es el problema del *confinamiento de los quarks*: ¿Por qué no vemos nunca los quarks desnudos y éstos están siempre agrupados formando neutrones, protones, etc.?

Desde el punto de vista matemático el problema que surge es cómo incorporar la teoría cuántica de campos en la teoría de Yang-Mills de modo matemáticamente riguroso. La mecánica cuántica es perfectamente rigurosa pero es incompleta ya que no puede describir procesos de creación y aniquilación de partículas como



Abdus Salam (1926–1996) y Steven Weinberg (1933– ).

los que hemos descrito anteriormente en la interacción entre mesones y nucleones. Esto sí que se puede hacer con la teoría cuántica de campos. No obstante, al hacerlo aparecen otras dificultades, concretamente al hacer cálculos aparecen cantidades infinitas. Estas dificultades se tratan con un mecanismo denominado *renormalización*. Las cantidades infinitas se deben a la existencia de partículas con momento muy grande. Un modo de corregir esto es aproximar la teoría utilizando retículos, para impedir que los valores del momento sean muy altos. Pero esto no da la verdadera teoría y en cada proceso de aproximación hay que reajustar los parámetros, y luego tomar límites. Este es el proceso de renormalización. Este proceso ha sido descrito en detalle sólo para modelos muy simples, en dimensión baja o para campos escalares. Pero ciertamente no para las teorías de Yang–Mills en dimensión 4.

## 2. GEOMETRÍA DE LAS TEORÍAS GAUGE

**2.1. Hacia una teoría de conexiones.** La idea básica del principio de simetría gauge es que si un sistema físico es invariante bajo la acción de un grupo de Lie  $G$  rígido (es decir, independiente del espacio-tiempo), entonces permanece invariante cuando  $G$  se hace local, es decir se reemplaza por  $G(x)$ , donde  $x = x_\mu$ ,  $\mu = 0, 1, 2, 3$  son las coordenadas del espacio tiempo, a condición de que la derivada usual  $\partial_\mu$  del espacio-tiempo se substituya por la derivada covariante  $D_\mu$ . Esta derivada covariante tiene la forma  $D_\mu = \partial_\mu + A_\mu(x)$ , donde  $A_\mu(x)$  son campos vectoriales que toman valores en el álgebra de Lie de  $G$  y que se transforman de tal modo que  $D_\mu$  se transforma de manera covariante con respecto a la acción del grupo local. Esto significa que la invarianza con respecto al grupo local fuerza la introducción de los campos vectoriales  $A_\mu(x)$  y determina la manera en la que éstos interactúan entre ellos y con la materia. Estos campos  $A_\mu(x)$  resultan ser los campos de radiación de la física de partículas, como hemos visto.

Resulta extraordinario que campos vectoriales tan variados, que juegan papeles tan diversos desde el punto de vista fenomenológico, sean todas manifestaciones de un principio tan simple. Y resulta todavía más extraordinario que el modo en el que estos interactúan con la materia esté prescrito de antemano. No es sorprendente



Neils Henrik Abel (1802–1829) y Marius Sophus Lie (1841–1899).

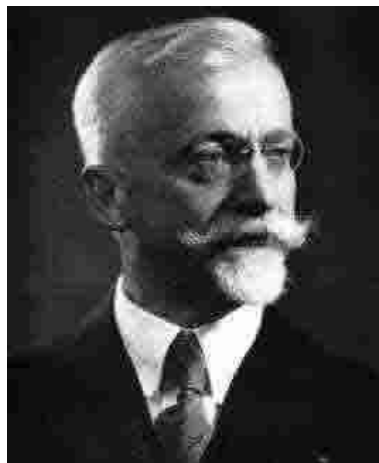
por tanto encontrar que la derivada covariante tenga un significado geométrico profundo. La geometría diferencial moderna está formulada en términos de fibrados y en este contexto los  $G(x)$  corresponden a una sección de un  $G$ -fibrado principal y los campos de radiación  $A_\mu(x)$  corresponden a conexiones en el  $G$ -fibrado.

Históricamente fue la teoría de la gravitación de Einstein la que abrió el camino a la comprensión de la invarianza gauge y su significado geométrico. El propósito original de Einstein, por supuesto, era explicar la equivalencia entre masa inercial y gravitatoria, pero en el proceso de hacer esto revolucionó la teoría de la gravitación mostrando que ésta podía ser atribuida enteramente a la geometría del espacio tiempo. No abundaremos aquí en este aspecto ya que nuestro énfasis es en el estudio de la geometría de las interacciones electromagnética, débil y fuerte. Pero lo que es interesante para nosotros es que la teoría de la relatividad general de Einstein, que se basaba en la geometría riemanniana, fue la inspiración para el desarrollo de la geometría no riemanniana y la geometría de fibrados.

El primer paso, y quizá el más importante, dado en este sentido fue cuando Tullio Levi-Civita, tan sólo un año después de la aparición de la teoría de Einstein, introdujo el concepto de *transporte paralelo*. Lo que Levi-Civita observó es que la covarianza de las derivadas y del tensor de curvatura de Riemann, que se expresan en términos de la conexión de Christoffel

$$\Gamma_{\mu\beta}^{\alpha} = \frac{1}{2}g^{\alpha\sigma}(\partial_{\mu}g_{\beta\sigma} + \partial_{\beta}g_{\sigma\mu} - \partial_{\sigma}g_{\mu\nu}),$$

se debía a las propiedades de transformación de la conexión de Christoffel bajo cambios de coordenadas, y no al hecho de que ésta se derivase de la métrica  $g_{\mu\nu}$  de la variedad riemanniana o espacio-tiempo.



Tullio Levi-Civita (1873–1941) y Élie Cartan (1869–1951).

El paso siguiente estaba claro: introducir la noción de conexión en la variedad como algo independiente de la métrica. Así se definió una conexión como un conjunto de funciones  $\Gamma_{\mu\beta}^{\alpha}$  que se transforman como la conexión de Christoffel. Esta teoría fue desarrollada por el mismo Levi-Civita, Cartan, Weyl y otros.

Detrás del formalismo algebraico está la idea geométrica de transporte paralelo, por el cual un vector  $v(x)$  es transportado a lo largo de una curva usando no el incremento infinitesimal  $dv = (\partial_{\mu}v)dx^{\mu}$ , sino incrementos infinitesimales de la forma

$$\delta v = (\nabla_{\mu}v)dx^{\mu},$$

donde  $\nabla_{\mu}$  es la derivada covariante definida por la conexión ( $\nabla = \partial + \Gamma$ ).

Estos desarrollos fueron continuados por matemáticos, como de Rham, Whitney, Hodge, Chern, Steenrod, Ehresman, etc. culminando en la construcción de la teoría de fibrados y conexiones a principios de los años 1950.

**2.2. Conexiones y curvatura.** A continuación describiremos brevemente algunos de los ingredientes más básicos de la teoría de conexiones en fibrados.

Sea  $G$  un grupo de Lie. Un  $G$ -fibrado *principal*  $P$  sobre una variedad diferenciable  $X$  es una variedad diferenciable con una acción lisa (por la derecha) del grupo  $G$  y espacio de orbitas  $P/G = X$ . Pedimos que la acción admita una estructura local de producto, es decir, que sea localmente equivalente a la acción obvia de  $G$  en  $U \times G$ , donde  $U$  es un abierto de  $X$ . Así pues tenemos una fibración  $\pi : P \rightarrow X$  y decimos que  $P$  tiene grupo de estructura  $G$ .

Tres maneras útiles de definir una *conexión* en un tal fibrado son:

- Como un campo de “subespacios horizontales”  $H \subset TP$  transverso a las fibras de  $\pi$ . Esto es, para  $p \in P$  tenemos una descomposición

$$TP_p = H_p \oplus T(\pi^{-1}(x)),$$



Georges de Rham (1903–1990) y William Hodge (1903–1975).

donde  $\pi(p) = x$ . Se requiere además que el campo de subespacios sea preservado por la acción de  $G$ .

- Como una 1-forma  $A$  en  $P$  con valores en el álgebra de Lie  $\mathfrak{g}$  de  $G$ , es decir, una sección de  $T^*P \otimes \mathfrak{g}$  sobre  $P$ . De nuevo se requiere que esta 1-forma sea invariante por la acción del grupo  $G$  que actúa combinando las acciones de  $G$  en  $P$  y la acción adjunta de  $G$  en  $\mathfrak{g}$ . También  $A$  debe restringirse a la forma canónica invariante por la derecha en las fibras.



Charles Ehresman (1905–1979) y Hassler Whitney (1907–1989).

- Para toda representación lineal  $\rho$  de  $G$  en  $V = \mathbb{C}^n$ , o  $V = \mathbb{R}^n$ , se obtiene un fibrado vectorial  $E$  sobre  $X$  definido como  $E := P \times_{\rho} V$ . Localmente  $E$  tiene la forma  $U \times V$ , donde  $U$  es un abierto de  $X$ . Las fibras de  $E$  son pues copias del espacio vectorial  $V$ . Recíprocamente, dado un fibrado vectorial  $E$  se puede obtener un fibrado principal. Por ejemplo, si  $E$  es un fibrado vectorial complejo de rango  $n$  ( $n = \dim V$ ), podemos obtener un fibrado principal  $P$  con grupo de estructura  $GL(n, \mathbb{C})$  considerando el conjunto de todas las “referencias” en  $E$ . Un punto en la fibra de  $P$  sobre  $x \in X$  es un conjunto de vectores formando una base de  $E_x$ . Si uno tiene una estructura algebraica adicional en  $E$  esto da lugar a un fibrado principal con un grupo de estructura más pequeño. Por ejemplo si  $E$  es un fibrado vectorial complejo con una métrica hermítica (un producto hermítico en cada fibra variando de manera lisa), entonces obtenemos un  $U(n)$  fibrado principal de referencias ortonormales en  $E$ . Si además fijamos una forma de volumen en cada fibra, entonces el grupo de estructura es  $SU(n)$ . Para los grupos clásicos (automorfismos de un espacio vectorial que conservan alguna estructura algebraica lineal), los conceptos de fibrado principal y fibrado vectorial son enteramente equivalentes.

Ahora bien, dado un fibrado vectorial  $E$ , una conexión en el fibrado de referencias puede definirse a través de una *derivada covariante* en  $E$ . Esto es, una aplicación lineal

$$\nabla : \Omega_X^0(E) \longrightarrow \Omega_X^1(E).$$

Aquí  $\Omega_X^p(E)$  denota el espacio de secciones lisas de  $\Lambda^p T^*X \otimes E$  — $p$ -formas con valores en  $E$ . La aplicación  $\nabla$  debe satisfacer la regla de Leibnitz

$$\nabla(f \cdot s) = f \nabla s + df \cdot s$$



Norman Steenrod (1910–1971) y Shiing-Shen Chern (1911–2004).

para cualquier función  $f$  (real o compleja, según sea apropiado). Si  $E$  tiene estructura algebraica adicional, como por ejemplo una métrica  $\langle \cdot, \cdot \rangle$ , una conexión compatible con la métrica debe satisfacer que  $d\langle s, t \rangle = \langle \nabla s, t \rangle + \langle s, \nabla t \rangle$  para todas las secciones  $s, t$ .

Se puede verificar que estas tres definiciones son equivalentes, al menos para los grupos de Lie clásicos.

Denotaremos por  $\text{ad}(E)$  (o  $\text{ad}(P)$  según convenga) al fibrado de álgebras de Lie asociado a la representación adjunta. Así pues  $\text{ad}(E)$  es un subfibrado (real) de  $\text{End } E = E \otimes E^*$ . Si el grupo de estructura es  $U(n)$ , entonces  $\text{ad}(E)$  consiste en el fibrado de endomorfismos antisimétricos del fibrado  $E$ . Si el grupo es  $SU(n)$  pediremos además que la traza sea nula.

Puesto que nuestro principal interés está en las teorías gauge con grupo  $U(n)$  o  $SU(n)$ , adoptaremos fundamentalmente el tercer punto de vista aplicado a los fibrados vectoriales complejos equipados de una métrica hermítica. No obstante para recordar su significado más geométrico, denotaremos la conexión por  $A$  y la derivada covariante por  $\nabla_A$ .

Una propiedad importante es que la diferencia de dos conexiones es un tensor. Supongamos que  $A$  es una conexión en  $E$  y que  $a$  es un elemento de  $\Omega_X^1(\text{ad } E)$ . Entonces el operador  $\nabla_A + a$  es de nuevo una derivada covariante. Recíprocamente la diferencia de dos conexiones en  $E$  es un elemento de  $\Omega_X^1(\text{ad } E)$ . Así pues, el espacio  $\mathcal{A}$  de todas las conexiones en  $E$  es un espacio afín de dimensión infinita modelado sobre  $\Omega_X^1(\text{ad } E)$ .

Para tener una idea más concreta y poder realizar cálculos es útil estudiar las conexiones localmente, es decir usando trivializaciones del fibrado. Para ello suponemos que  $U$  es un abierto de  $X$ . Entonces tenemos la trivialización  $f : E|_U \rightarrow U \times \mathbb{C}^n$ . Sobre  $U$  podemos escribir

$$\nabla_A = d + A$$

donde  $A$  (abusando la notación) es una 1-forma a valores en  $\mathfrak{g}$  (una matriz de 1-formas sobre  $U$ ). Todavía más explícitamente, si elegimos coordenadas  $x_\mu$  en  $U$ , podemos escribir

$$\nabla_A = \sum \nabla_\mu dx_\mu,$$

donde la derivada covariante en la dirección  $x_\mu$  es  $\nabla_\mu$  dada por

$$\nabla_\mu = \partial_\mu + A_\mu,$$

donde las  $A_\mu$  son funciones matriciales.

Por supuesto si nuestra variedad diferenciable  $X$  es  $\mathbb{R}^4$  entonces la descripción que acabamos de dar es global.

Un ingrediente fundamental en la teoría para comprender como se relaciona la descripción de una conexión al usar distintas trivializaciones es el grupo gauge. El grupo gauge  $\mathcal{G}$  de  $E$  está formado por los automorfismos

$$g : E \longrightarrow E$$

que respetan la estructura de las fibras e inducen la identidad en  $X$ . Si el fibrado  $E$  es trivial tenemos que  $\mathcal{G}$  coincide con el conjunto de aplicaciones lisas de  $X$  a  $G$ .

El grupo gauge  $\mathcal{G}$  actúa sobre el espacio  $\mathcal{A}$  de conexiones mediante la regla

$$\nabla_{g(A)} = g\nabla_A(g^{-1}s).$$

De donde obtenemos que

$$g\nabla_A g^{-1} = \nabla_A - (\nabla_A g)g^{-1}.$$

Aquí la derivada covariante de  $g$  se calcula considerando  $g$  como una sección del fibrado  $\text{End } E$ . Así pues

$$g(A) = A - (\nabla_A g)g^{-1}.$$

Observemos que si  $G = U(1)$  y  $g = e^{i\chi}$ , entonces  $g(A) = A - id\chi$  donde  $\chi$  es una función real.

La derivada covariante  $\nabla_A$  puede extenderse a operadores que actúan sobre formas a valores en el fibrado. Se obtienen así las derivadas exteriores

$$d_A : \Omega_X^p(E) \longrightarrow \Omega_X^{p+1}(E),$$

definidas por las siguientes propiedades:

$$\begin{aligned} d_A &= \nabla_A \text{ en } \Omega_X^0(E), \\ d_A(\alpha \wedge \varphi) &= (d\alpha) \wedge \varphi + (-1)^p \alpha \wedge d_A \varphi, \text{ para } \alpha \in \Omega_X^p, \varphi \in \Omega_X^q(E). \end{aligned}$$

En contraste con lo que sucede con la derivada exterior usual  $d$ , no es cierto en general que  $d_A d_A$  sea cero. Por el contrario, la regla de Leibnitz nos dice que esta composición es un operador algebraico (es decir que conmuta con la multiplicación por funciones lisas) que usaremos para definir la curvatura  $F_A$  de una conexión. Así:

$$d_A d_A s = F_A s,$$

donde  $F_A \in \Omega^2(\text{ad}(E))$ . En términos de una trivialización local, la curvatura viene dada en función de la matriz de la conexión  $A$  por una matriz de dos formas:

$$F_A = dA + A \wedge A.$$

En  $A \wedge A$  estamos combinando el producto exterior con la multiplicación de matrices. Desde el punto de vista de fibrados principales es mejor escribir este término como  $1/2[A, A]$ .

Para entender la definición con más claridad, podemos elegir coordenadas en  $X$  además de una trivialización. La matriz de curvatura  $F_A = \sum F_{\mu\nu} dx_\mu dx_\nu$  tiene componentes  $F_{\mu\nu}$  que son los conmutadores de las derivadas covariantes en las diversas direcciones coordenadas:

$$\begin{aligned} F_{\mu\nu} &= [\nabla_\mu, \nabla_\nu] = \left[ \frac{\partial}{\partial x_\mu} + A_\mu, \frac{\partial}{\partial x_\nu} + A_\nu \right] \\ &= \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} + [A_\mu, A_\nu]. \end{aligned}$$

Un punto importante es que bajo una transformación gauge  $g \in \mathcal{G}$  la curvatura se transforma como un tensor, es decir

$$F_{g(A)} = gF_Ag^{-1}.$$

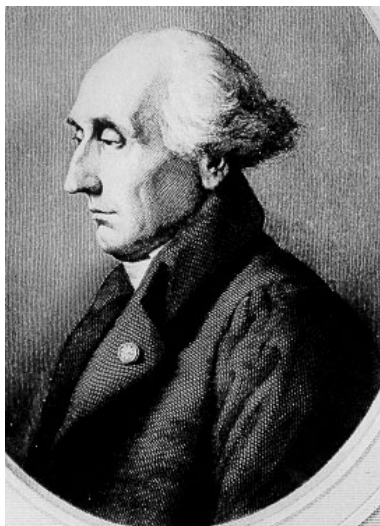
Es un ejercicio obtener la *identidad de Bianchi* para una conexión  $A$ :

$$(3) \quad d_A F_A = 0.$$

**2.3. Ecuaciones de Yang–Mills.** Para obtener las ecuaciones de Yang–Mills se aplica el *principio de mínima acción*.

Recordemos que en mecánica clásica se define el lagrangiano  $L$  como

$$L = \text{Energía cinética} - \text{Energía potencial}.$$



Joseph-Louis Lagrange (1732–1813).

A partir del lagrangiano se define la *acción*  $S(\gamma)$  a lo largo de un camino  $\gamma$  como

$$S(\gamma) = \int_{\gamma} L dt.$$

Según variamos el camino  $\gamma$  cambia la acción. El camino correcto es el que minimiza la acción. Este viene dado por las ecuaciones de Euler–Lagrange, equivalentes a las ecuaciones de Newton.

Para definir la acción de una conexión se necesita que  $X$  esté equipado de una métrica riemanniana (o lorentziana, como en el caso del espacio-tiempo). Se define entonces

$$S(A) = \int_X |F_A|^2 d\text{vol},$$

donde  $d \text{ vol}$  es la forma de volumen riemanniana y la norma  $|F_A|^2$  se calcula usando conjuntamente una métrica en  $\mathfrak{g}$  (la forma de Killing, por ejemplo) y la métrica de  $X$ .

Para ser más concretos consideraremos el caso en el que la dimensión de  $X$  es 4 —después de todo, nuestro mayor interés está en el caso  $X = \mathbb{R}^4$ —.

En esta situación el operador de Hodge  $*$  envía 2-formas en 2-formas de modo que el producto interno  $L^2$  de dos formas  $\alpha$  y  $\beta$  viene dado por

$$\langle \alpha, \beta \rangle = \int_X \alpha \wedge * \beta.$$

Entonces

$$S(A) = - \int_X \text{Tr}(F_A \wedge * F_A).$$

La ecuación de Euler–Lagrange de  $S(A)$  da las ecuación de Yang–Mills:

$$(4) \quad d_A * F_A = 0.$$

En  $\mathbb{R}^4$  equipado con la métrica euclídea o lorentziana la identidad de Bianchi  $d_A F_A = 0$  y la ecuación de Yang–Mills  $d_A * F_A = 0$  se pueden escribir en términos de derivadas covariantes  $\nabla_\mu$ , respectivamente, como

$$[\nabla_\mu, [\nabla_\nu, \nabla_\sigma]] + [\nabla_\nu, [\nabla_\sigma, \nabla_\mu]] + [\nabla_\sigma, [\nabla_\mu, \nabla_\nu]] = 0$$

y

$$[\nabla_\mu, [\nabla_\mu, \nabla_\nu]] = 0,$$

donde en la segunda ecuación se suma sobre los índices repetidos con el consiguiente signo según la métrica sea euclídea o lorentziana (la velocidad de la luz se ha normalizado a 1).

Observemos que puesto que en dimensión 4 el operador  $*$  actuando sobre 2-formas es invariante conforme, en el sentido de que dos métricas  $g_{\mu\nu}$  y  $\rho(x)g_{\mu\nu}$  definen el mismo operador  $*$ , entonces es claro que las ecuaciones de Yang–Mills dependen sólo de la estructura conforme. Esta importante propiedad de la teoría de Maxwell se conserva pues en el caso no abeliano.

Como ya hemos mencionado varias veces, la teoría de Maxwell es una teoría de Yang–Mills para el grupo  $U(1)$ . La curvatura en este caso es

$$F_A = dA$$

y las ecuaciones de Maxwell son equivalentes a  $dF_A = 0$  (identidad de Bianchi, que en este caso es consecuencia de que  $d^2 = 0$ ) y  $d * F_A = 0$ .

Una consecuencia inmediata de (3) y (4) es que (4) se cumple si  $F_A$  satisface una de las dos ecuaciones

$$*F_A = F_A \quad (\text{autodualidad})$$

$$*F_A = -F_A \quad (\text{antiautodualidad}).$$

Estas son ecuaciones de primer orden no lineales para la conexión  $A$  que implican las ecuaciones de segundo orden de Yang–Mills.

En  $\mathbb{R}^4$  con la métrica euclídea la condición de antiautodualidad corresponde al siguiente sistema de ecuaciones para las matrices de la conexión  $A_\mu$  ( $\mu = 1, 2, 3, 4$ ):

$$\begin{aligned} F_{12} + F_{34} &= 0, \\ F_{14} + F_{23} &= 0, \\ F_{13} + F_{42} &= 0, \end{aligned}$$

donde

$$F_{\mu\nu} = [\nabla_\mu, \nabla_\nu] = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} + [A_\mu, A_\nu].$$

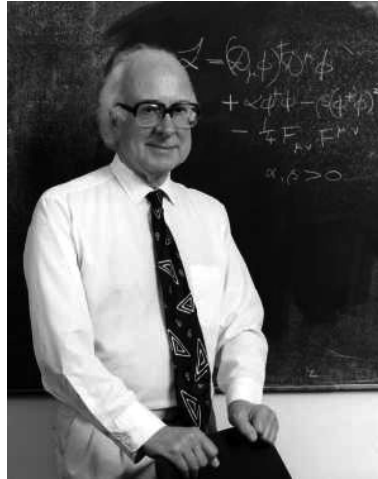
### 3. TEORÍA DE YANG-MILLS CUÁNTICA

**3.1. Teoría de Yang-Mills cuántica y modelo estándar.** En los años 1950 cuando la teoría de Yang-Mills fue descubierta, se sabía que la versión cuántica de la teoría de Maxwell —conocida como Electrodinámica Cuántica (abreviado QED en inglés)— da una descripción extremadamente ajustada de las fuerzas eléctricas y magnéticas.

Así que resultó natural preguntarse si la teoría de Yang-Mills podía describir las otras fuerzas de la naturaleza, concretamente la interacción débil (responsable entre otras cosas de ciertas formas de radiactividad) y la interacción fuerte (responsable entre otras cosas de la unión entre protones y neutrones para formar un núcleo). No obstante, como ya hemos mencionado anteriormente, el hecho de que los campos de Yang-Mills no tuvieran masa fue un serio obstáculo para aplicar la teoría de Yang-Mills a estas fuerzas ya que la interacción fuerte y débil son interacciones de corto alcance y muchas de las partículas tienen masa. Así pues no parecía que este tipo de fenómenos tuvieran que ver con campos de largo alcance que describen partículas sin masa.

En los años 1960 y años 1970, se superaron estos obstáculos en la interpretación física de la teoría gauge no abeliana. En el caso de la interacción débil, esto se logró con la teoría electrodébil de Glashow-Salam-Weinberg con grupo gauge  $SU(2) \times U(1)$ . Introduciendo un “campo de Higgs” adicional mediante el mecanismo de “ruptura espontánea de simetría” se consigue dar masa a las partículas gauge. El campo de Higgs es constante en el vacío y reduce el grupo de estructura de  $SU(2) \times U(1)$  a un subgrupo  $U(1)$  embebido diagonalmente en  $SU(2) \times U(1)$ . Debido a esta reducción, los campos de largo alcance son sólo los del electromagnetismo, de acuerdo con lo que se ve en la naturaleza.

La solución del problema de la ausencia de masa en los campos de Yang-Mills para la interacción fuerte es de una naturaleza completamente distinta. La solución no se obtuvo añadiendo campos adicionales a la teoría de Yang-Mills, sino con el descubrimiento de una propiedad muy importante de la misma teoría cuántica de Yang-Mills, versión cuántica de la teoría clásica descrita en la Sección 2. Esta propiedad se denomina *libertad asintótica*. De modo aproximado, esta propiedad dice que a cortas distancias el campo tiene un comportamiento cuántico muy similar al comportamiento clásico, sin embargo, a distancias largas, la teoría clásica no se corresponde con el comportamiento cuántico del campo.



Peter Higgs (1929– ).

La libertad asintótica, junto con otros experimentos y descubrimientos teóricos hechos en los años 1960 y años 1970, hicieron posible describir la interacción fuerte con una teoría gauge no abeliana con grupo  $SU(3)$ , como ya mencionamos en la Sección 1. Los campos adicionales describen, a nivel clásico, quarks —objetos con espín  $1/2$  de algún modo análogos al electrón, pero que se transforman según la representación fundamental de  $SU(3)$ —. La teoría gauge no abeliana de la interacción fuerte se llama *Cromodinámica Cuántica* (abreviado QCD, en inglés).

El uso de la QCD para describir la interacción fuerte fue motivado por una serie de descubrimientos experimentales y teóricos hechos en los años 1960 y los años 1970, que involucran simetrías y comportamiento a altas energías de la interacción fuerte. Pero la teoría gauge no abeliana clásica es muy diferente de las observaciones experimentales de las interacciones fuertes. Para que la QCD describa las interacciones fuertes con éxito, a nivel cuántico tiene que tener al menos las siguientes dos propiedades, ambas totalmente diferentes del comportamiento de la teoría clásica:

- Tiene que tener un “salto de masa”, es decir, debe existir una constante  $\Delta > 0$  tal que toda excitación del vacío tenga energía al menos  $\Delta$ .
- Tiene que tener “confinamiento de los quarks”, es decir, aunque la teoría se describa en términos de campos elementales, tales como los campos del quark, que se transforman de manera no trivial bajo la acción de  $SU(3)$ , los estados de las partículas físicas —como el protón, neutrón y pión— deben ser  $SU(3)$ -invariantes.

El primer punto es necesario para explicar por qué la interacción fuerte es fuerte pero de rango corto; la segunda es necesaria para explicar por qué no vemos nunca un quark individual. (Otra propiedad también requerida, que hemos obviado por simplificar es el que haya “ruptura de la simetría quiral”).

Los experimentos y simulaciones por ordenador que se han realizado desde finales de los años 1970 indican fuertemente que la QCD debe tener estas propiedades. Estas propiedades pueden verse hasta cierto punto en cálculos teóricos realizados en varios modelos muy simplificados (como las teorías gauge en retículos). Pero no existe en estos momentos una explicación teórica, y mucho menos matemáticamente completa, que demuestre alguna de estas propiedades.

Para comprender los aspectos más básicos del proceso de cuantización de la teoría de Yang-Mills, y formular el problema del salto de masa con más precisión, haremos un repaso breve al proceso de cuantización de un sistema clásico, comenzando con la mecánica cuántica.

**3.2. Mecánica cuántica.** Recordemos la formulación de Hamilton de la mecánica clásica. Ésta está basada en lo que se denomina el “Hamiltoniano”

$$\text{Hamiltoniano} = \text{Energía cinética} + \text{Energía potencial.}$$

Las leyes de Newton son equivalentes a las *ecuaciones canónicas* de Hamilton:

$$\dot{x} = \frac{\partial}{\partial p} H(p, x),$$

$$\dot{p} = -\frac{\partial}{\partial x} H(p, x).$$

Los puntos de vista lagrangiano y hamiltoniano son equivalentes y ambos constituyen la manera más simple de describir un sistema clásico.

Fue Dirac quien primero formuló la mecánica cuántica como una teoría consistente, y demostró que los puntos de vista de Heisenberg y Schrödinger eran equivalentes. La teoría puede resumirse del siguiente modo:



William R. Hamilton (1805–1865) y Erwin Schrödinger (1887–1961).

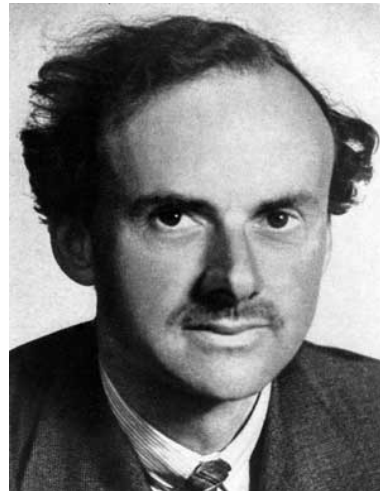
- El estado de un sistema se corresponde con un vector de un *espacio de Hilbert* abstracto. Los vectores  $\psi$  y  $\lambda\psi$ , donde  $\lambda$  es un número complejo, describen el mismo estado (es decir, los estados del sistema vienen descritos por el espacio proyectivo asociado al espacio de Hilbert).
- Los observables (como el momento, por ejemplo) están en correspondencia con operadores que actúan sobre el espacio de Hilbert. Si medimos el observable en uno de sus estados propios obtenemos el correspondiente valor propio. Si medimos el observable en un estado que no es propio obtenemos una distribución estadística de los autovalores. Para asegurarnos de que los autovalores son reales, los operadores deben ser autoadjuntos.
- Una teoría clásica puede ser “cuantizada” convirtiendo el hamiltoniano  $H(p, x)$  en un operador, usando las reglas de conmutación  $[p, x] = -i\hbar$ , donde aquí  $x$  y  $p$  son los operadores correspondientes a la posición y al momento. Este procedimiento se conoce como *cuantización canónica*.
- El hamiltoniano cuántico es un generador de la evolución temporal. Esto se expresa con la ecuación de Schrödinger:

$$H\psi = i\hbar \frac{\partial\psi}{\partial t}.$$

Un método alternativo de cuantización que, veremos es particularmente apropiado para cuantizar las teorías gauge, es el método de la *integral de caminos*, desarrollado por Feynman, basándose en una idea de Dirac.

Este método está basado en dos postulados:

- La probabilidad  $P(b, a)$  de que una partícula se mueva de un punto  $a$  a otro punto  $b$  es el cuadrado del valor absoluto de una función compleja  $K(b, a)$ ,



Werner Heisenberg (1907–1972) y Paul Dirac (1902–1984).

denominada *amplitud de transición*:

$$P(b, a) = |K(b, a)|^2.$$

- La amplitud de transición  $K(b, a)$  viene dada por la suma sobre todos los caminos posibles entre  $a$  y  $b$

$$K(b, a) = \sum_{\text{caminos } \gamma} \exp\left(\frac{i}{\hbar} S(\gamma)\right),$$

donde  $S(\gamma) = \int_{\gamma} L dt$  es la acción a lo largo de un camino  $\gamma$  entre  $a$  y  $b$ .

Puesto que los caminos forman un conjunto continuo, esta suma es realmente una integral, denominada la *integral de caminos de Feynman*.

El límite  $\hbar \rightarrow 0$  nos da la mecánica clásica. Para ver esto, observemos que en este límite cualquier pequeña variación de la acción es aumentada infinitamente. El ángulo de fase  $\exp(iS/\hbar)$  fluctúa enormemente dando un gran número de giros de  $2\pi$ , con lo que su valor se hace esencialmente aleatorio. Así pues, contribuciones de distintos caminos tienden a cancelarse mutuamente, dejando solamente las contribuciones que minimizan la acción clásica, es decir  $\delta S = 0$ , de donde obtenemos las ecuaciones clásicas, como ya vimos anteriormente.

Para un valor finito de  $\hbar$  todos los caminos contribuyen, dando correcciones cuánticas a la física clásica.

Un posible camino constituye lo que podríamos llamar una “realidad virtual”. De acuerdo con Feynman uno puede construir un sistema cuántico eligiendo un sistema permisible de realidades virtuales con acciones clásicas fijadas. El principio de mínima acción seleccionará el límite clásico de la teoría.

**3.3. Teoría cuántica de campos.** En una teoría clásica de campos, como la teoría de Yang–Mills, descrita en la Sección 2, los campos son funciones  $\varphi(x, t)$  del espacio-tiempo (vectores, tensores, conexiones, etc.), como, por ejemplo, los campos gauge  $A_{\mu}(x, t)$ .

En teoría cuántica de campos:

- Los campos  $\varphi(x, t)$  se convierten en operadores que dependen del espacio-tiempo.
- La posición  $x$  y el tiempo  $t$  son números que determinan un punto en el espacio-tiempo —no son operadores—.
- El momento y la energía son operadores  $H$  y  $P$ .

Estos campos cuánticos deben satisfacer una serie de axiomas. Por ejemplo, Gårding y Wightman (ver [7]) dan axiomas matemáticos precisos para una teoría cuántica en  $\mathbb{R}^4$  con signatura lorentziana.

Básicamente se pide que el espacio de Hilbert  $\mathcal{H}$  sobre el que actúa el campo cuántico sea un espacio en el que se representa el grupo de Poincaré (definido como el producto semidirecto del grupo de transformaciones de Lorentz por las translaciones en el espacio-tiempo). El Hamiltoniano  $H$  y el momento  $P$  son los operadores autoadjuntos correspondientes a los elementos del álgebra de Lie del

grupo de Poincaré que generan translaciones en el tiempo y en el espacio, respectivamente.

Un *vector del vacío* es un elemento de  $\mathcal{H}$  que es invariante bajo la representación del grupo de Poincaré. Se asume además que la representación tiene energía positiva, es decir  $H \geq 0$ , y un vector del vacío  $\Omega$  que es único salvo multiplicación por una fase. Las transformaciones gauge de los campos cuánticos también actúan como transformaciones lineales de  $\mathcal{H}$  y se transforman covariantemente bajo la acción del grupo de Poincaré. Los campos cuánticos en regiones del espacio-tiempo que no puedan conectarse por una señal luminosa deben ser independientes, lo que se traduce en la formulación de Gårding y Wightman en que los operadores correspondientes conmutan (o anticonmutan para campos fermiónicos).



Arthur Wightman (1922– ).

Uno de los mayores logros de la teoría cuántica de campos axiomática es el descubrimiento de cómo pasar de una teoría cuántica de campos en el espacio-tiempo euclídeo invariante bajo el grupo euclídeo a una teoría cuántica de campos invariante bajo el grupo de Lorentz en el espacio-tiempo de Minkowski, y viceversa. Usando la condición de energía positiva se puede hacer una continuación analítica de las amplitudes de las teorías de campos en el espacio de Minkowski a las teorías de campo en el espacio euclídeo (rotación de Wick). Esto nos permite trabajar con la signatura que más nos convenga, en muchos casos la euclídea.

**3.4. Problema de existencia de Yang–Mills y salto de masa.** Para establecer la existencia de una teoría gauge con grupo de Lie  $G$ , uno debería definir una teoría cuántica de campos en el sentido anteriormente descrito, con los campos cuánticos en correspondencia con los polinomios locales en la curvatura  $F$  y sus derivadas covariantes, tales como  $\text{Tr } F_{\mu\nu} F_{\sigma\tau}(x)$ . Las funciones de correlación de

los campos cuánticos deben coincidir a distancias cortas con las predicciones de la libertad asintótica y la teoría de renormalización perturbativa.

Puesto que el vector del vacío  $\Omega$  es invariante por el grupo de Poincaré, es un vector propio con energía nula, es decir  $H\Omega = 0$ . El axioma de energía positiva afirma que en cualquier teoría cuántica de campos, el espectro de  $H$  está contenido en el intervalo  $[0, \infty)$ . Una teoría cuántica tiene un *salto de masa*  $\Delta$  si  $H$  no tiene valores propios en el intervalo  $(0, \Delta)$  para algún  $\Delta > 0$ . El supremo de un tal  $\Delta$  es la masa  $m$ , y se requiere que  $m < \infty$ .

EXISTENCIA DE YANG–MILLS Y SALTO DE MASA. *Probar que para todo grupo de Lie compacto simple  $G$ , la teoría cuántica de Yang–Mills en  $\mathbb{R}^4$  existe y tiene un salto de masa  $\Delta > 0$ . La existencia incluye el establecer propiedades axiomáticas al menos tan fuertes como las citadas en [5, 7].*

**3.5. Primeros pasos y estrategias.** Como hemos visto, en la mecánica cuántica de partículas la posición  $x$  y el momento  $p$  se convierten en operadores que no conmutan y un estado cuántico puede ser representado por una función de onda  $\psi(x)$ , lo que nos lleva al dominio del análisis funcional. ¿Qué sucede sin embargo si cuantizamos un campo? Por ejemplo si cuantizamos el campo electromagnético, las componentes  $\vec{E}(x)$  y  $\vec{B}(x)$  de los campos eléctrico y magnético se convierten en operadores que no conmutan y tenemos entonces que considerar un álgebra no conmutativa de dimensión infinita. Este álgebra se puede representar en un espacio de Hilbert consistente en funciones de onda  $\psi(\vec{B})$ , con lo que subimos un grado en la dificultad del problema: el estado cuántico es una función sobre un espacio de funciones (en este caso el espacio de funciones es el espacio de todos los posibles  $\vec{B}$ ). Tenemos pues que hacer análisis funcional en un espacio con infinitas variables, lo que representa un nivel nuevo de dificultad.

El procedimiento más apropiado para cuantizar la teoría de Yang–Mills clásica resulta ser el método de la integral de caminos de Feynman. Recordemos que la acción de Yang–Mills con grupo de estructura  $G = U(n)$  o  $G = SU(n)$  viene dada por

$$S(A) = \frac{1}{4g^2} \int_X \text{Tr}(F_A \wedge *F_A)$$

(aquí consideramos la acción definida en la Sección 2 multiplicada por un factor que involucra la constante de acoplamiento  $g$ ). Para un grupo simple  $G$  arbitrario se usa la forma de Killing en el álgebra de Lie de  $G$ . Entonces, se considera la integral de caminos de Feynman sobre el espacio  $\mathcal{A}$  de conexiones dada formalmente por la expresión

$$Z = \frac{1}{\text{vol}(\mathcal{G})} \int_{\mathcal{A}} DA \exp(-S(A)).$$

Como  $\mathcal{A}$  es un espacio afín, formalmente tiene una medida  $DA$  invariante por traslaciones (única salvo factor constante que se cancelará cuando definamos las funciones de correlación). La integral de Feynman para teorías gauge se formula normalmente sobre el espacio  $\mathcal{A}/\mathcal{G}$ , donde  $\mathcal{G}$  es el grupo de transformaciones gauge. En lugar de hacer esto, aquí hemos dividido por el volumen de  $\mathcal{G}$  para definir  $Z$ .

$Z$  se denomina “función de partición”, nombre que viene de la física estadística (donde la integral es la suma de las amplitudes de probabilidad de todos los estados microscópicos del sistema).

A continuación se eligen puntos  $x_i \in X$  y “operadores locales”  $\mathcal{O}_i(x_i)$  que sean polinomios en la curvatura y sus derivadas covariantes, invariantes por transformaciones gauge, en los puntos  $x_i$ . Definimos

$$Z_{\mathcal{O}} = \frac{1}{\text{vol}(\mathcal{G})} \int_{\mathcal{A}} DA \exp(-S(A)) \prod_i \mathcal{O}_i(x_i).$$

Finalmente definimos los “valores esperados” o “funciones de correlación”

$$\langle \prod_i \mathcal{O}_i(x_i) \rangle = Z_{\mathcal{O}}/Z.$$

Para demostrar la existencia de la teoría de Yang–Mills cuántica, se debe dar sentido riguroso a la función de partición y a las funciones de correlación (dando sentido a las integrales de caminos definidas heurísticamente) y mostrar que satisfacen ciertos axiomas relacionados con el hecho de que los  $\mathcal{O}_i(x_i)$  pueden ser interpretados como operadores que actúan en un espacio de Hilbert.

Una manera de abordar el problema (utilizada en simulaciones numéricas) y dar sentido a la integral de caminos consiste en aproximar el espacio  $\mathcal{A}$  de conexiones por un espacio de dimensión  $k$  y luego tomar el límite  $k \rightarrow \infty$ . Esto se puede hacer, por ejemplo, considerando la teoría de Yang–Mills en un “retículo”, en otras palabras, con un grafo  $\Gamma$  con vértices, lados y un conjunto de caras o “plaquetas”, cada una de las cuales es un lazo (camino cerrado) en  $\Gamma$ . Por ejemplo, podemos tomar como vértices los puntos enteros  $\mathbb{Z}^4 \subset \mathbb{R}^4$ ; los lados son las líneas rectas que unen dos puntos con distancia unidad, y las plaquetas los lazos de longitud cuatro. Una  $G$ -conexión sobre  $\Gamma$  es entonces una aplicación de los lados a  $G$ , y la curvatura de la conexión en una plaqueta concreta es menos la traza de la holonomía alrededor del lazo menos la matriz identidad. La acción de Yang–Mills puede definirse como la suma del cuadrado de las curvaturas.

Con esta descripción explícita del espacio de configuraciones se puede definir la teoría cuántica reticular de Yang–Mills en un subgrafo finito  $\gamma$  de  $\Gamma$  mediante la integral de Feynman

$$Z[\gamma, g^2, G] = \int \prod \exp(-S) dU_i,$$

donde la integral es sobre todas las holonomías en  $\gamma$ , en otras palabras, todas las aplicaciones de los lados en  $G$ ; la medida es el producto de la medida de Haar para la holonomía en cada lado y  $S$  es la acción de Yang–Mills. Las otras cantidades físicas de interés son valores esperados obtenidos utilizando esta medida.

Tomando como punto de partida estas integrales en dimensión finita, el asunto de la existencia de una teoría cuántica de Yang–Mills se traduce esencialmente en si existe una manera razonable de definir el límite de  $Z[\gamma, g^2, G]$  sobre subgrafos  $\gamma$  de  $\Gamma$  cada vez más grandes para definir la integral de Feynman en  $\Gamma$ . Tomar este límite involucrará claramente el método de renormalización, y de hecho el estudio

de este ejemplo por Kenneth Wilson fue su motivación original para estudiar el denominado *grupo de renormalización*.

Se sabe bastante de las propiedades esperadas de este límite utilizando una gran variedad de argumentos físicos, como la libertad asintótica, etc. Una de las propiedades más importantes es que en el límite  $g^2 \rightarrow 0$  y  $\gamma$  grande, se cree que la elección específica de  $\Gamma$  que aproxima a  $\mathbb{R}^4$  no es importante, y los correspondientes valores esperados convergerán a las funciones de correlación en la teoría cuántica de campos continua. Además, éstas deberían satisfacer propiedades formales, como la invarianza por el grupo de simetrías de la métrica plana en  $\mathbb{R}^4$  y otros axiomas formalizados por Osterwalder-Schrader en [5]. Los axiomas adicionales proporcionan las condiciones suficientes para la construcción de un espacio de Hilbert y la interpretación en términos de operadores actuando en este espacio de Hilbert.

La parte correspondiente a la “existencia” en el problema, consiste en establecer estos axiomas, mientras que la parte correspondiente al salto de masa involucra el decaimiento de las funciones de correlación con la distancia. Por otro lado, no es esencial seguir el método de las teorías gauge en retículos. Existen otros modos de aproximar la integral funcional de Feynman. Lo que es importante es poder definir el límite a la teoría continua de campos. También sería en principio posible abordar el problema sin tomar límites.

Hasta donde el autor de estas notas sabe, no ha habido avances importantes en este problema. Si bien es cierto que se han realizado avances en teorías de campos en dimensión más baja, no parece que haya habido progreso en la construcción matemática rigurosa de la teoría de Yang-Mills cuántica. Los trabajos más relevantes en este sentido siguen siendo los trabajos de Balaban y Magnen, Rivasseau y Sénéor citados en la descripción oficial del problema por Jaffe y Witten [3].

Hay dos clases de teorías cuánticas que se cree son muy similares a la teoría de Yang-Mills en dimensión cuatro. La primera es el modelo sigma no lineal en dimensión dos. En esta teoría los campos son aplicaciones de un espacio-tiempo de dimensión dos en un espacio simétrico riemanniano con curvatura positiva. Aunque al parecer no hay una construcción matemática rigurosa de esta teoría en el sentido de la descripción del problema de Yang-Mills, se ha visto que sí que existe salto de masa en una cierta versión del modelo sigma para el grupo  $O(n)$  [4].

La otra gran clase de modelos muy similares a la teoría de Yang-Mills son las teorías supersimétricas de Yang-Mills en dimensión cuatro. Estas son modificaciones de la teoría de Yang-Mills que además de conexiones involucran otros campos, en particular campos “fermiónicos” (las conexiones son campos “bosónicos” y la supersimetría intercambia bosones y fermiones). La condición básica es que el operador Hamiltoniano que actúa sobre el espacio de Hilbert de la teoría, y que genera la translación temporal, tenga una raíz cuadrada que se denomina “supercarga”.

La supersimetría introduce grandes simplificaciones en la teoría. Desde el punto de vista físico la más importante es que el problema de la renormalización es más abordable. Aunque introduciendo la supersimetría se cambia el problema original, todavía existe una gran relación con el problema. Por ejemplo, se puede comenzar con la teoría supersimétrica y añadir a la acción términos que rompan la

supersimetría, y que sean importantes sólo a largas distancias, para obtener una teoría con las mejores propiedades de renormalización de la teoría supersimétrica a distancias cortas, pero que reduce a la teoría convencional de Yang–Mills a largas distancias. Así pues una solución al problema en una clase suficientemente general de teorías supersimétricas implicaría de hecho la solución al problema original.

Aunque en el momento actual no se está más cerca de este nuevo objetivo que del objetivo original, lo cierto es que ha habido un gran progreso en el estudio de las teorías supersimétricas en los últimos años (véanse referencias en [1]). La primera de estas teorías es la célebre teoría de Seiberg–Witten de 1994. Aunque los matemáticos tienen más familiaridad con las aplicaciones de este trabajo a la topología de las variedades de dimensión cuatro (véase [2] y las referencias que se dan allí para este desarrollo), la motivación física original y uno de sus principales logros fue precisamente el abordar el problema del salto de masa. De hecho, la razón básica de la simplicidad de los invariantes de Seiberg–Witten comparados con los invariantes de Donaldson es la propiedad del salto de masa, que reduce el cálculo de los invariantes a un problema de dimensión más baja.

Otra teoría supersimétrica que parece relevante en nuestro contexto es la teoría supersimétrica de Yang–Mills con  $N = 4$  ( $N$  hace referencia al número de cargas supersimétricas; la teoría de Seiberg–Witten tiene  $N = 2$ ). El punto de partida de esta teoría es la “correspondencia AdS/CFT” de Maldacena, de acuerdo a la cual la teoría supersimétrica de Yang–Mills con  $N = 4$  se puede reformular como una teoría de cuerdas en el espacio de anti de Sitter (AdS). Lo interesante es que, al menos para  $SU(n)$ , cuando  $n$  se hace muy grande esta teoría es resoluble, como el modelo sigma mencionado anteriormente. La teoría resulta todavía muy misteriosa, pero podría dar lugar a una relación muy precisa entre las teorías en dimensión dos (mucho mejor entendidas) y el problema original de Yang–Mills en dimensión cuatro, que podría proporcionar una manera eficaz de demostrar el salto de masa (y otras propiedades como el confinamiento de los quarks, y la ruptura de la simetría quirial). Esta idea de relacionar la teoría gauge con una teoría de cuerdas es una vieja idea que se remonta a una sugerencia de 't Hooft en 1973.

En cualquier caso, parece claro que el problema de Yang–Mills propuesto por el Instituto Clay de Matemáticas resulta muy duro por ahora y parece aconsejable abordar por el momento problemas más fáciles (para irse entrenando) que minimicen en lo posible las dificultades técnicas y optimicen el interés geométrico. El éxito en el estudio de teorías cuánticas de campos más simples y de interés fuera del contexto usual de la teoría cuántica de campos puede ser la clave para atraer mayor interés matemático, y aportar nuevas energías y métodos en este campo.

#### REFERENCIAS

- [1] M.R. DOUGLAS, Report on the status of the Yang–Mills millenium prize problem: [http://www.claymath.org/millennium/Yang-Mills\\_Theory/](http://www.claymath.org/millennium/Yang-Mills_Theory/).
- [2] O. GARCÍA-PRADA, Seiberg–Witten invariants and vortex equations, *Quantum Symmetries*, Les Houches 1995, Elsevier, 1998, 885–932.
- [3] A. JAFFE Y E. WITTEN, Quantum Yang–Mills theory, *The Millennium Prize Problems*, Clay Mathematical Institute, AMS, 2006, 129–152.

- [4] C. KOPPER, Mass generation in the large  $N$  nonlinear sigma model, *Comm. Math. Phys.* **202** (1999), 89–126.
- [5] K. OSTERWALDER Y R. SCHRADER, Axioms for Euclidean Green's function, *Comm. Math. Phys.* **31** (1973), 83–112 y *Comm. Math. Phys.* **42** (1975), 281–305.
- [6] L. SADUN, Lecture at University of Texas (vídeo):  
<http://www.claymath.org/millennium/Yang-Mills-Theory/>.
- [7] R. STREATER Y A. WIGHTMAN, PCT, Spin and Statistics and all That, W.A. Benjamin, New York, 1964.
- [8] E. WITTEN, Physical law and the quest for mathematical understanding, *Bulletin of the AMS* **40** (2001), 1–14.

INSTITUTO DE CIENCIAS MATEMÁTICAS, CSIC-UAM-UC3M-UCM, c/ NICOLÁS CABRERA,  
13–15, 28049 CANTOBLANCO, MADRID

*Correo electrónico:* [oscar.garcia-prada@icmat.es](mailto:oscar.garcia-prada@icmat.es)

*URL:* <http://www.mat.csic.es/webpages/garcia-prada>