DALLA GEOMETRIA ALLA FISICA: LA CURVATURA

JACOPO STOPPA

1. Introduzione

Le due equazioni più importanti della fisica teorica sono l'*equazione di Einstein*

$$\mathrm{Ric}(g) + \left(-\frac{1}{2}s(g) + \Lambda\right) g = T$$

che descrive il campo gravitazionale (teoria della Relatività Generale), e le *equazioni di Yang-Mills*

$$D_A * F(A) = 0$$

che descrivono ogni altro tipo di interazione: in particolare il *campo elettromagnetico* (per il quali le equazioni si riducono a quelle di Maxwell) e i campi che danno origine alle forze nucleari *debole* (radioattività) e *forte* (che tiene uniti neutroni e protoni).

Per quanto nei dettagli molto diverse, queste equazioni hanno in comune un ingrediente fondamentale: *sono entrambe espresse in termini della curvatura di un certo oggetto geometrico.*

Nel caso dell'equazione di Einstein, questo oggetto è la metrica (Lorenziana) g che descrive il nostro spazio-tempo. Le curvature che intervengono sono quelle di Ricci, $\mathrm{Ric}(g)$, e quella scalare $s(g)$: esse descrivono in un senso preciso come il nostro spazio-tempo appare deformato rispetto all'usuale spazio piatto quadridimensionale $\mathbb{R}^4$ (di Minkowski). L'incurvarsi dello spazio-tempo *è* il campo gravitazionale. Queste idee (lo spazio-tempo e la sua curvatura, ecc.) si sono ovviamente fatte strada nell'immaginario collettivo.

Per le equazioni di Maxwell e di Yang-Mills, l'oggetto geometrico in questione è un *campo di gauge* A (terminologia fisica) o *connessione* (matematica). Il campo elettromagnetico, la forza debole (radiazione) ecc. *sono* (collettivamente!) l'increspatura (curvatura) di un unico oggetto geometrico, il campo di gauge (connessione). E' interessante osservare come questa idea, sicuramente altrettanto notevole della curvatura dello spazio-tempo, sia quasi del tutto assente dalla cultura popolare (fantascienza ecc.). Una possibile ragione risiede nel fatto che l'interpretazione intuitiva di A, sebbene possibile, è più delicata di quella di g: in particolare, richiede l'introduzione

Date: 13 giugno 2013.

di ancora altre dimensioni "extra" (ausiliarie) rispetto alle quattro dello spazio-tempo.

Lo scopo di questa lezione è fornire una introduzione, puramente intuitiva, alla *geometria* che interviene nelle equazioni di Yang-Mills, cioè nella descrizione di tutte le interazioni fondamentali osservate in natura, esclusa la forza di gravità. Si tratta di una versione semplificata dell'esposizione nel capitolo 2 delle lezioni Fermiane di M. Atiyah a Pisa (1979).

Strada facendo cerchiamo di spiegare come la teoria di Yang-Mills abbia portato a notevoli progressi matematici, citando per esempio la teoria di Donaldson.

Infine, menzioniamo brevemente uno dei Millennium Problems (Clay Institute 2000), che richiede una dimostrazione matematica dell'esistenza della teoria di Yang-Mills (quantistica).

2. Campi, connessioni, curvatura

2.1. **Particelle con simmetria interna.** Il nostro punto di partenza consiste nel pensare a una particella nello spazio-tempo; per esempio, un elettrone. Dal punto di vista di $\mathbb{R}^4$, questa appare puntiforme proprio come in punti materiali della meccanica Newtoniana. La novità è che vogliamo dotare la particella di una *struttura interna*. Questa struttura riflette le *simmetrie interne* della particella.

Le simmetrie interne della particella formano un *gruppo* G: possono essere composte tra loro (legge di composizione), e una simmetria può sempre essere invertita per tornare al punto di partenza. Un buon esempio di un gruppo di simmetrie (finito) sono le trasformazioni rigide di un quadrato in se. Nel caso del nostro elettrone il gruppo rilevante è quello delle trasformazioni rigide di una circonferenza in se, che preservano il suo verso di percorrenza: è un gruppo infinito di rotazioni. I matematici lo chiamano $U(1)$. Una breve riflessione permette di identificare questo gruppo di nuovo con un oggetto geometrico, precisamente la circonferenza unitaria S^1: $x^2 + y^2 = 1$ (per vederlo, basta pensare all'angolo di rotazione in radianti).

2.2. **Simmetria interna e sfere.** Nulla ci vieta di considerare la stessa equazione con un numero maggiore di variabili, per esempio $x^2+y^2+z^2 = 1$. Geometricamente otteniamo la superficie della sfera unitaria S^2 nello spazio Euclideo. Anche se le equazioni sono simili, c'è una profonda differenza fra la circonferenza S^1 e la sfera S^2: infatti i matematici hanno dimostrato che quest'ultima *non può essere realizzata come il gruppo di trasformazioni continue di un oggetto in se.* Questo teorema è un corollario di quello, più famoso, detto della "palla pelosa": *non è possibile pettinare una sfera coperta di pelliccia in modo uniforme, senza che compaia almeno un punto di scriminitura.*

La cosa torna miracolosamente ad essere vera se aggiungiamo un'altra variabile: $x^2 + y^2 + z^2 + w^2 = 1$. Le soluzioni di questa equazione formano un oggetto geometrico chiamato (non sorprendentemente) sfera tridimensionale S^3. Per visualizzare S^3, possiamo operare l'analogo della proiezione stereografica di S^2: in questo modo S^3 diventa lo spazio Euclideo, con un punto all'infinito. È possibile dimostrare che S^3 *è un gruppo continuo di trasformazioni geometriche* (cioè i suoi punti corrispondono, senza strappi, all'insieme di tutte le trasformazioni di un certo oggetto geometrico in se). Questo è un risultato altamente non banale, e purtroppo non è possibile descrivere in modo semplice la legge di composizione tra i punti della sfera tridimensionale. Quando dotata di questa particolare legge di composizione, S^3 prende un nuovo nome e diviene il gruppo $SU(2)$.

Il fatto sorprendente (scoperto da Yang e Mills negli anni '50 del secolo scorso) è che, proprio come la circonferenza (sfera unidimensionale) S^1 costituisce la struttura interna di un elettrone che interagisce con un campo elettromagnetico (QED), così la sfera tridimensionale S^3 è precisamene il gruppo $SU(2)$ rilevante per l'interazione debole (QFD; radioattività). Allo stesso tempo, scoprirono che le leggi che governano i due fenomeni sono espresse da una unica equazione: l'equazione di Yang-Mills. Tutto quello che cambia è il gruppo di struttura da inserire all'interno dell'equazione: la sfera unidimensionale, rispettivamente quella tridimensionale.

2.3. **Lo spazio degli stati.** Per ora abbiamo solo descritto alcuni esempi di gruppi di struttura interni di particelle. Torniamo alla costruzione geometrica che rappresenta il moto di una particella. In ogni punto dello spazio e del tempo, la particella si trova in uno stato espresso da un elemento $g \in G$ (per esempio, un punto di S^1, o di S^3). L'insieme di tutti gli stati possibili è di solito chiamato P. Per punti distinti $x \neq y$ di $\mathbb{R}^4$, gli stati interni non sono identificati, $G_x \neq G_y$. Pertanto P è una collezione di copie di G disgiunte, una per ogni punto dello spazio-tempo: i matematici chiamano queste copie *fibre*, e P il *fibrato principale* (banale).

Per il momento P è ancora un oggetto inerte: è solo l'arena dove può propagarsi una particella. Per esempio, è lo scenario dove può propagarsi un elettrone, ma non abbiamo ancora acceso un campo elettromagnetico esterno. Dal punto di vista geometrico questo si esprime *identificando* P con il prodotto Cartesiano $\mathbb{R}^4 \times G$. Per esempio ci troviamo a considerare i prodotti $\mathbb{R}^4 \times S^1$ oppure $\mathbb{R}^4 \times S^3$: l'assenza di campo ci permette di effettuare una identificazione coerente di tutte le fibre con una sola copia di S^1 o S^3. Attenzione: questa identificazione, sebbene possibile, *non è unica*. I fisici chiama una particolare identificazione *gauge* (calibrazione). I matematici usano invece il termine *banalizzazione*.

2.4. **Campi.** L'idea fondamentale è che l'effetto di accendere un campo (per esempio il campo elettromagnetico) è quello di distorcere l'allineamento relativo delle fibre di P. In altre parole, la presenza di un campo significa che non ci è più permesso di identificare in modo coerente le diverse fibre G_x e G_y in punti dello spazio-tempo distinti, $x \neq y$.

Per esempio, il campo elettromagnetico *è* il modo in cui le diverse circonferenze S^1, una per ogni punto dello spazio-tempo, sono state tra loro scombinate. Lo stesso vale per il campo che soggiace alla forza debole: esso *è* il modo in cui la nostra collezione di sfere S^3, una per ogni punto di $\mathbb{R}^4$, risulta scombinata rispetto alla collezione regolare.

2.5. **Curvatura.** Come misurare questa distorsione? Pensiamo di fissare punti distinti $x \neq y$ nello spazio-tempo. Una particella si propaga da x a y, *portando con se il proprio spazio di simmetrie interne.* In altre parole, anche se le fibre sono tra loro del tutto scombinate, ci è ancora possibile identificarle due a due non in tutto lo spazio-tempo, ma lungo la traiettoria di una particella. I matematici chiamano questa operazione *trasporto parallelo.* L'idea ora è di considerare due diverse traiettorie tra x e y. Lo stato iniziale della particella è un elemento di G_x, e lo possiamo trasportare parallelamente a un elemento di G_y lungo le due traiettorie distinte. *Il risultato finale, in generale, non è lo stesso.* Questo avviene perchè le varie copie di G sono tra di loro scombinate (dalla presenza del campo). I due diversi trasporti paralleli sono però legati tra loro da una simmetria interna nel punto y, cioè un elemento g di G_y.

Per esempio, nel caso del campo elettromagnetico, lo stato interno di un elettrone risulterà variato per una rotazione, detta "shift di fase".

I matematici chiamano l'elemento g la *curvatura totale* o *distorsione* di P sopra una superficie racchiusa tra le due traiettorie distinte. Questa distorsione *è* (in un certo senso, vedi sotto) il campo.

2.6. **Versione infinitesimale.** Come le leggi del moto nella meccanica Newtoniana, anche le leggi per i campi descritte discorsivamente finora sono in realtà espresse in forma infinitesimale. Ad esempio, si considera il trasporto parallelo infinitesimale a partire da un punto x lungo una direzione fissata (cioè lungo traiettorie sempre più corte che partono da x in direzione fissata). Nel limite otterremo uno *spostamento infinitesimale A all'interno della fibra G_x.* I matematici chiamano A la *connessione.* La sua *curvatura infinitesimale* F dipende da due direzioni nel punto x e di nuovo ha valori negli spostamenti infinitesimali di G_x in se. Le trasformazioni infinitesimali di un gruppo come S^1 o S^3 in se stesso vengono chiamate dai matematici *algebre di Lie.* Ad esempio si può dimostrare che l'algebra di Lie di S^1 è costituita dalla retta dei numeri reali (o meglio, dei numeri immaginari puri $i\mathbb{R}$). L'algebra di Lie della sfera S^3 è più complicata e si chiama algebra delle matrici complesse 2×2 anti-hermitiane a traccia nulla (brevemente: $\mathfrak{su}(2)$).

Se facciamo una scelta di gauge ben definita (cioè scegliamo una banalizzazione) allora possiamo rappresentare la deformazione infinitesimale A e la sua curvatura infinitesimale $F(A)$ esplicitamente: dato un punto $x \in \mathbb{R}^4$ e una direzione di spostamento (spaziale o temporale) $\mu = 0, 1, 2, 3$, la deformazione $A_\mu(x)$ è un elemento dell'algebra di Lie (ad esempio, semplicemente un numero immaginario puro per il campo elettromagnetico). Similmente la curvatura $F_{\mu\nu}(x)$ dipende dal punto e da due direzioni spazio-temporale (due traiettorie infinitesimali). La distorsione infinitesimale A viene chiamato *campo di gauge* ed è a tutti gli effetti il campo fisico. La sua curvatura $F_{\mu\nu}$ è una quantità derivata: la relazione esplicita è

$$F_{\mu\nu} = [d_\mu + A_\mu, d_\nu + A_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Per il campo elettromagnetico,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

La curvatura $F(A)$ può essere pensata come la distorsione prodotta dal campo A, o può essere identificata con il campo se misuriamo i campi tramite i loro effetti locali.

Questa identificazione dei campi con la distorsione di un oggetto geometrico è naturalmente alla base della teoria di Einstein della gravità (teoria della Relatività Generale). Il punto fondamentale è che qui ciò che viene distorto non è lo spazio-tempo in cui viviamo, ma la geometria di uno spazio degli stati interni sovrapposto allo spazio tempo (che può tranquillamente restare piatto). Questa è una differenza importante e storicamente lo studio della geometria dei fibrati principali e delle connessioni è iniziato dopo quella dello spazio-tempo.

2.7. **Principio di minima azione.** Proprio come le equazioni della meccanica Newtoniana, le equazioni di Yang-Mills sono derivate da un principio di minima azione: a livello classico (non quantistico) i campi (elettromagnetico ecc.) si devono propagare in modo da minimizzare, in un senso opportuno, la curvatura che inducono. In termini matematici, si dice che devono essere *punti critici per il funzionale di Yang-Mills*

$$\mathcal{L} = -\frac{1}{2} \int_{\mathbb{R}^4} \text{Traccia}(F_{\mu\nu} F^{\mu\nu}) dx_0 dx_1 dx_2 dx_3.$$

Questo vincolo è espresso proprio dall'equazione di Yang-Mills riportata nell'introduzione. Un altro modo per scriverla è porre $D_\mu = \partial_\mu + A_\mu$. Si può mostrare che allora l'equazione diviene

$$[D_\mu, [D_\mu, D_\nu]] = 0.$$

2.8. **Dualità.** Lo spazio-tempo $\mathbb{R}^4$ gode di una proprietà notevole, cioè l'esistenza di quello che viene chiamato *operatore star di Hodge*, denotato con $*$. Si tratta di una particolare simmetria che agisce su ogni oggetto geometrico definito sullo spazio-tempo e soddisfa $*^2 = 1$. Per esempio $*$ agisce sulla curvatura $F_{\mu\nu}$, permutando gli indici: F_{01} viene trasformato in F_{23}, e così via (con un segno opportuno). Facendo uso di $*$ si scopre che il funzionale di Yang-Mills può riscriversi come

$$\mathcal{L} = -\int_{\mathbb{R}^4} \text{Traccia}(F \wedge *F).$$

Questo fatto ha due conseguenza fondamentali:

(1) Ogni teoria di Yang-Mills può essere formulata anche su uno spazio-tempo curvo. Tutto quello che ci serve è un operatore che agisca come $*$. I fisici dicono che le teorie di Yang-Mills sono *covarianti* o che possono essere accoppiate alla gravità. In effetti, la dipendenza dalla curvatura dello spazio-tempo è solo attraverso $*$. In particolare, due spazi-tempo curvi g e λg dove $\lambda > 0$ vedono precisamente la stessa teoria di Yang-Mills.

(2) Se la curvatura F è *auto-duale*
$$F = *F$$
oppure *anti-autoduale*
$$F = - * F$$
allora F è automaticamente soluzione delle equazioni di Yang-Mills. Abbiamo così equazioni del primo ordine (nonlineari) nel campo di gauge che implicano le equazioni di Yang-Mills.

2.9. **Teoria di Donaldson.** Negli anni '80 del secolo scorso S. K. Donaldson ha scoperto come la teoria di Yang-Mills possa essere utilizzata per dimostrare alcuni spettacolari risultati matematici. In particolare egli ha dimostrato il fatto fondamentale che in dimensione 4 (quella dello spazio-tempo $\mathbb{R}^4$) non è sempre possibile stirare le pieghe di un oggetto geometrico localmente regolare (cioè che appare localmente come $\mathbb{R}^4$, a meno di trasformazioni continue, senza strappi), per farlo apparire liscio globalmente. Questo è in netto contrasto con quanto avviene in dimensione 2 e 3 (per esempio, è facile convincersi che tutte le superfici di solidi senza buchi nello spazio Euclideo, come un cubo o un tetraedro ecc., possono essere stirate per diventare una sfera S^2).

L'ingrediente fondamentale del teorema di Donaldson è lo studio delle soluzioni dell'equazione di Yang-Mills con gruppo $SU(2)$ sopra uno spazio-tempo regolare M che (per ipotesi) può essere stirato. Lo spazio di tutte lo soluzioni è di nuovo un oggetto geometrico $\mathfrak{M}$, stavolta con dimensione

5 (una in più). Uno studio delicato permette di dimostrare che $\mathfrak{M}$ ha un bordo naturale, costituito proprio da M, e da un numero finito di oggetti geometrici chiamati piani proiettivi complessi (simili a delle sfere di dimensione 4). Pertanto $\mathfrak{M}$ appare come una membrana tesa fra il nostro M e la collezione di piani. Infine questo fornisce un vincolo non banale sullo spazio M di partenza, nella sola ipotesi che fosse globalmente regolare: si possono fare esempi concreti di spazi localmente regolari che non soddisfano questo vincolo.

3. Un problema da un milione di dollari

Quando la geometria interna è quella di una circonferenza S^1, le equazioni di Yang-Mills si riducono alle classiche equazioni di Maxwell che descrivono il propagarsi del campo elettro-magnetico. Possiamo osservare le soluzioni delle equazioni di Maxwell in ogni istante (con l'occhio umano o con strumenti a bassa tecnologia, come una radio portatile): sono le onde elettromagnetiche (luce, onde radio ecc.)

La situazione è radicalmente diversa per geometrie interne più complicate: sebbene previste teoricamente, nessuno di noi ha mai osservato le soluzioni *classiche* delle equazioni di Yang-Mills per $SU(2)$ o per gruppi più complicati. L'unico modo per vedere queste equazioni di Yang-Mills all'azione è spostarsi nel mondo della teoria quantistica dei campi, con i vari fenomeni di scattering, decadimento ecc. osservati dagli acceleratori; in altre parole dobbiamo servirci di strumenti complessi.

Questo fenomeno (l'impossibilità di osservare "onde di campo nucleare debole o forte") viene chiamato dai fisici *gap di massa*. Per il gruppo $SU(2)$, esiste una spiegazione matematica abbastanza (anche se non del tutto) rigorosa per giustificare il gap di massa. Nel caso di gruppi più complicati, in particolare il gruppo $SU(3)$ dell'interazione forte, non è ancora stata scoperta alcuna giustificazione matematica, pienamente rigorosa o no, che spieghi questo fenomeno. Nell'anno 2000 il Clay Mathematics Institute ha messo in palio un milione di dollari per la risoluzione di questo problema. Enunciato ufficiale: *dimostrare che per ogni gruppo di gauge compatto e semplice G, esiste una teoria di Yang-Mills quantistica con gruppo G, e gap di massa strettamente positivo.*

Dipartimento di Matematica "F. Casorati", Università di Pavia
E-mail address: jacopo.stoppa@unipv.it