Relatività ristretta e formalismo tensoriale

(per il corso di Fisica Nucleare e Subnucleare 2019/20)

Fiorenzo Bastianelli

1 Introduzione

Agli inizi del 1900 erano conosciute due grandi teorie fisiche:

1) la meccanica di Galileo e Newton,

2) l'elettrodinamica classica, sintetizzata dalle equazioni di Maxwell.

Queste teorie però non sembravano essere compatibili a vicenda. La meccanica newtoniana è consistente per cambi di sistema di riferimento inerziali, definiti dalle *trasformazioni galileiane*. Un sistema di riferimento inerziale K permette di misurare il tempo t e la posizione $\vec{x}$, per cui possiamo indicare tale sistema in modo schematico con $K = \{t, \vec{x}\} = \{t, x^1, x^2, x^3\} = \{t, x, y, z\}$. Similmente possiamo considerare un secondo sistema di riferimento inerziale $K' = \{t', \vec{x}'\}$.

Se il secondo sistema K' si muove con velocità v diretta lungo l'asse x rispetto a K, le relazioni che collegano le coordinate dei due sistemi di riferimento inerziali sono date dalla trasformazione galileiana

$$\begin{aligned} t' &= t \\ x' &= x - vt \\ y' &= y \\ z' &= z \end{aligned} \tag{1}$$

Sono proprio queste le trasformazioni che lasciano *invariate in forma* le equazioni di Newton.

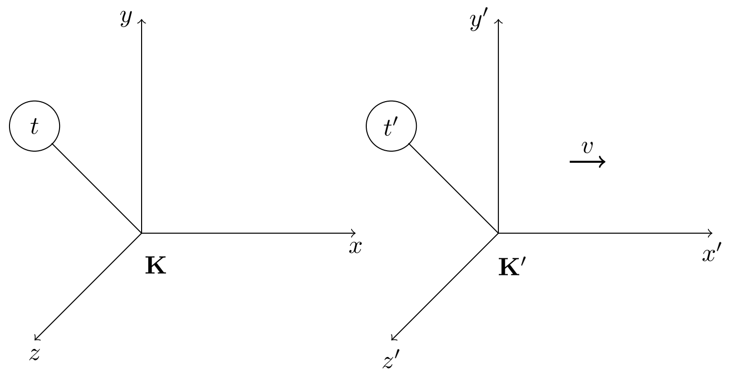


Figure 1: *I sistemi di riferimento inerziali K e K' con i rispettivi orologi per la misura del tempo.*

L'equazione $t' = t$ ci dice che esiste un tempo assoluto, indipendente dallo stato di moto dell'osservatore. È facile verificare che le equazioni del moto libere di Newton assumono la stessa forma (sono invarianti in forma) per queste trasformazioni

$$m\frac{d^2\vec{x}}{dt^2} = 0 \quad \leftrightarrow \quad m\frac{d^2\vec{x'}}{dt'^2} = 0 \; .$$

Questa situazione fu efficacemente descritta da Galileo, con l'immagine di una nave in moto costante rispetto alla terraferma (in "Dialogo sopra i due massimi sistemi del mondo" (1632)).

Viceversa, le equazioni del moto di Newton non sono invarianti per una trasformazione con $x' = x - \frac{1}{2}gt^2$, che produce

$$m\frac{d^2\vec{x}}{dt^2} = 0 \quad \leftrightarrow \quad m\frac{d^2\vec{x'}}{dt'^2} = -m\vec{g}$$

dove $\vec{g} = (g, 0, 0)$. Nel nuovo sistema di riferimento compare una forza apparente, cioè non collegata a forze reali dovute ad interazioni con altre particelle, per cui non c'è una simmetria tra i due sistemi di riferimento: essi non sono equivalenti. Il primo sistema di riferimento è inerziale, mentre il secondo non lo è a causa della presenza di forze apparenti.

Ricordiamo che, in generale, si parla di "simmetria" o "trasformazione di simmetria" quando le equazioni del moto sono invariati in forma sotto opportune trasformazioni delle variabili dinamiche. Il linguaggio matematico appropriato per descrivere le simmetrie è la *teoria dei gruppi*, cruciale nello studio delle moderne teorie fisiche.

Le equazioni dell'elettrodinamica non sono invarianti per la trasformazione galileiana in (1). Infatti, dopo la formulazione finale dovuta a Maxwell, si capì che le sue equazioni erano invarianti per un altro tipo di trasformazione, la trasformazione di Lorentz

$$\begin{aligned} t' &= \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \\ x' &= \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \\ y' &= y \\ z' &= z \end{aligned} \tag{2}$$

in cui compare la velocità della luce c (da interpretare come costante fondamentale della natura). Usando queste nuove trasformazioni come le corrette leggi di trasformazione che collegano i due sistemi di riferimento inerziali, si riconosce che il tempo è relativo al sistema di riferimento. In particolare, il concetto di *simultaneità* perde il suo carattere assoluto, e ha un significato preciso solo all'interno di uno specifico sistema di riferimento.

Fu Einstein che riuscì a chiarire che le proprietà di simmetria dell'elettrodinamica erano quelle corrette. Sono quindi le trasformazioni di Lorentz le trasformazioni che collegano i due sistemi di riferimento in moto relativo lungo l'asse x. Quelle di Galileo emergono solo come un'approssimazione quando $v \ll c$. Einstein concluse che occorreva modificare la meccanica di newtoniana in una *meccanica relativistica* per renderla compatibile con l'elettrodinamica e con le corrette leggi di trasformazione che collegano i vari sistemi di riferimento inerziali.

I punti chiave su cui si basa la meccanica relativistica sono:

- le leggi fisiche sono identiche in tutti i sistemi di riferimento inerziali
- la "velocità della luce" è la stessa in tutti i sistemi di riferimento inerziali.

Queste due proprietà fondamentali della meccanica relativistica sono state verificate sperimentalmente con grande precisione, e continuano ad essere verificate con sempre maggior precisione. Possono essere usate per riderivare le trasformazioni di Lorentz e le loro proprietà generali. (N.B. "velocità della luce" = velocità limite di propagazione delle interazioni, numericamente uguale a $c \sim 300\,000$ Km/s).

2 Alcune conseguenze delle trasformazioni di Lorentz

La trasformazione di Lorentz in eq. (2) mostra che il tempo t non è più assoluto, ma deve essere considerato un parametro relativo al sistema di riferimento scelto, proprio come le coordinate spaziali $\vec{x}$. In particolare, il concetto di simultaneità perde il suo carattere assoluto, ed ha un significato preciso solo all'interno di uno specifico sistema di riferimento (dimenticare questo fatto conduce spesso alla formulazione di falsi paradossi). Una conseguenza delle trasformazioni di Lorentz è che la velocità della luce è la stessa nei due sistemi di riferimento inerziali K e K' ed è considerata una costante fondamentale della natura.

Riscriviamo in maniera compatta la trasformazione di Lorentz precedente usando le notazioni di β (velocità in unità di velocita della luce) e γ (fattore di dilatazione relativistico) definite da

$$\beta \equiv \frac{v}{c}, \qquad \gamma \equiv \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{3}$$

che possono assumere i valori $0 \leq \beta < 1$ e $1 \leq \gamma < \infty$. Con queste notazioni la trasformazione di Lorentz si scrive come

$$\begin{aligned} t' &= \gamma(t - \frac{\beta}{c}x) \\ x' &= \gamma(x - \beta ct) \\ y' &= y \\ z' &= z \end{aligned} \tag{4}$$

e la sua inversa come

$$\begin{aligned} t &= \gamma(t' + \frac{\beta}{c}x') \\ x &= \gamma(x' + \beta ct') \\ y &= y' \\ z &= z' \end{aligned} \tag{5}$$

che sono derivabili dalla (4) (per ottenerle è sufficiente scambiare le coordinate di K con quelle di K' e cambiare $v \to -v$, come evidente da considerazioni di simmetria).

Alcune conseguenze immediate delle trasformazioni di Lorentz sono: *i*) somma delle velocità, *ii*) contrazione delle lunghezze, *iii*) dilatazione dei tempi.

i) Somma delle velocità

Consideriamo una particella con velocità V_x diretta lungo x nel sistema di riferimento K, come in figura 2. Dalla definizione di velocità nei sistemi K e K' abbiamo che

$$V_x = \frac{dx}{dt}\ , \qquad V_x = \frac{dx'}{dt'}\ . \tag{6}$$

Dalla trasformazione di Lorentz (4) segue che

$$dx' = \gamma(dx - vdt)\ , \qquad dt' = \gamma(dt - \frac{v}{c^2}dx)$$

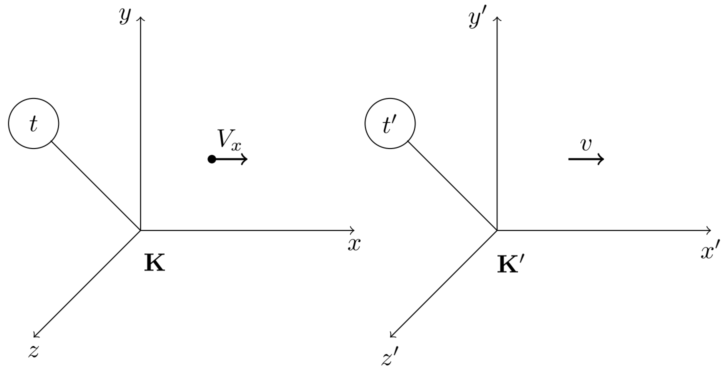


Figure 2: *Particella con velocità V_x nel sistema K.*

da cui

$$V'_x = \frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma(dt - \frac{v}{c^2}dx)} = \frac{V_x - v}{1 - \frac{vV_x}{c^2}} \ . \tag{7}$$

Quindi

$$\boxed{V'_x = \frac{V_x - v}{1 - \frac{vV_x}{c^2}}} \ . \tag{8}$$

N.B. Se $V_x = c$ allora si trova che $V'_x = c$.

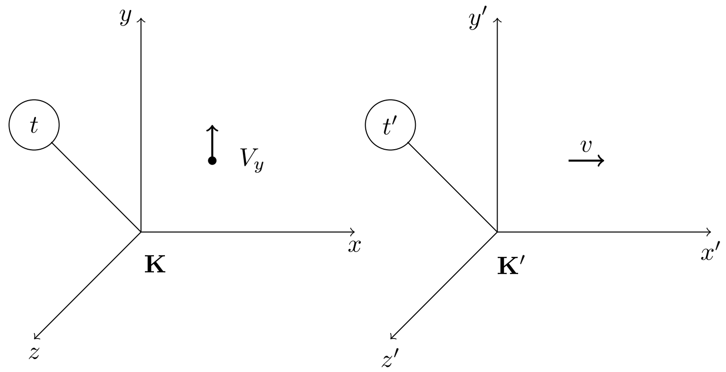


Figure 3: *Particella con velocità V_y nel sistema K.*

Si può procedere in modo simile anche nel caso di una particella che abbia una componente della velocità diretta lungo l'asse y, come in figura 3. Infatti per definizione $V_y = \frac{dy}{dt}$ e $V'_y = \frac{dy'}{dt'}$. Dalla trasformazione di Lorentz (4) segue che

$$dy' = dy \ , \qquad dt' = \gamma(dt - \frac{v}{c^2}dx)$$

da cui

$$V'_y = \frac{dy'}{dt'} = \frac{dy}{\gamma(dt - \frac{v}{c^2}dx)} = \frac{V_y}{\gamma(1 - \frac{vV_x}{c^2})} = \frac{V_y}{1 - \frac{vV_x}{c^2}}\sqrt{1 - \frac{v^2}{c^2}} \tag{9}$$

e quindi

$$\boxed{V_y' = \frac{V_y}{1 - \frac{vV_x}{c^2}}\sqrt{1 - \frac{v^2}{c^2}}} \; . \tag{10}$$

ii) Contrazione delle lunghezze

Supponiamo che nel sistema K' ci sia un oggetto di lunghezza $L_0 = x_2' - x_1'$ in quiete, disposto lungo l'asse x come in figura 4. La sua lunghezza vista nel sistema di riferimento a riposo con l'oggetto stesso è L_0. Nel sistema K l'oggetto è visto muoversi con velocità v diretta lungo l'asse x, per cui occorrerà misurare simultaneamente la posizione dei suoi estremi, diciamo ad un tempo fissato t, per calcolare la lunghezza L

$$L = x_2(t) - x_1(t). \tag{11}$$

Si ricordi infatti che la simultaneità è un concetto relativo al sistema di riferimento. Ora possiamo chiederci: come sono collegate L ed L_0? Dalle trasformazioni di Lorentz si ottiene

$$L_0 = x_2' - x_1' = \frac{x_2(t) - vt}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{x_1(t) - vt}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{x_2(t) - x_1(t)}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{L}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma L \; . \tag{12}$$

Abbiamo usato le trasformazioni di Lorentz in (4), dove è facile imporre la richiesta che gli estremi dell'oggetto siamo misurati simultaneamente al tempo t nel sistema K, dove l'oggetto è visto in movimento. Dunque la lunghezza vista nel sistema di riferimento in cui l'oggetto è in moto con velocità v risulta contratta

$$\boxed{L = \gamma^{-1} L_0} \; . \tag{13}$$

In generale $v \leq c$, per cui $\gamma \geq 1$ e quindi $L \leq L_0$.

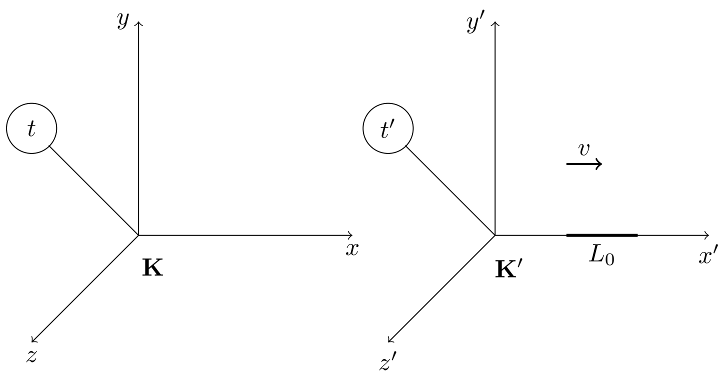


Figure 4: *Un oggetto di lunghezza L_0 posizionato lungo l'asse x' del sistema K'.*

iii) Dilatazione dei tempi

Consideriamo due eventi nel sistema K' che accadono nello stesso punto spaziale di K', ad esempio $E_1 = (t'_1, 0, 0, 0)$ ed $E_2 = (t'_2, 0, 0, 0)$. Questi due eventi sono separti da un intervallo temporale $T_0 = t'_2 - t'_1$. Ora l'intervallo di tempo misurato nel sistema K è dato da

$$T = t_2 - t_1 = \gamma t'_2 - \gamma t'_1 = \gamma(t'_2 - t'_1) = \gamma T_0 \tag{14}$$

dove abbiamo usato le trasformazioni di Lorentz inverse riportate nella (5), dove è facile imporre la condizione che l'intervallo temporale è riferito a due eventi che accadono nello stesso punto spaziale del sistema K' (ad esempio una particella ferma in K' che decade in un certo intervallo temporale). Dunque

$$\boxed{T = \gamma T_0} \tag{15}$$

e quindi $T \geq T_0$, per cui si parla di dilatazione dei tempi. Il tempo che scorre nel sistema di riferimento in quiete con l'oggetto in questione è detto tempo proprio. È il tempo più breve possibile per il fenomeno in questione, poiché in tutti gli altri sistemi di riferimento questo tempo risulta necessariamente dilatato.

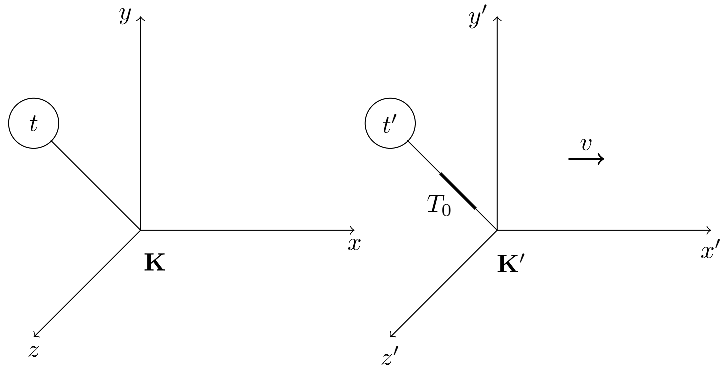


Figure 5: *Esemplificazione della dilatazione dei tempi.*

Per esemplificare le considerazioni precedenti, consideriamo una particella relativistica che viene prodotta dai raggi cosmici nell' alta atmosfera e poi che decade. In particolare, consideriamo un muone, una particella il cui tempo di decadimento è dell'ordine di $\tau \sim 10^{-6}$ s, definito come il tempo in cui in media la particella decade nel suo sistema di riferimento (dove la particella è ferma). Supponiamo che sia stato prodotto nell'alta atmosfera con una velocità molto prossima a quella della luce, in modo tale che il suo γ valga $\gamma = 10^3$. Ci si può chiedere se questo muone abbia o meno la possibilità di raggiungere la superficie della terra. Possiamo subito renderci conto che tale muone viaggia essenzialmente alla velocità della luce, infatti possiamo calcolare

$$\beta = \sqrt{1 - \frac{1}{\gamma^2}} = \sqrt{1 - 10^{-6}} \sim 1 \, .$$

Un osservatore solidale con la terra deve tener conto della dilatazione dei tempi, per cui il tempo di decadimento del muone risulta dilatato, e quindi calcola un tragitto

$$L = v\tau\gamma \sim c\tau\gamma \sim 300 \, \text{Km} \tag{16}$$

sufficiente per raggiungere la terra (possiamo stimare lo spessore dell'atmosfera in circa 10-20 Km, altezza a cui arriva la troposfera). Un osservatore solidale con il muove viceversa non sente l'effetto di dilatazione del tempo, ma vede la terra che si avvicina a lui con una velocità prossima a quella della luce, ed osserva lo spessore dell'atmosfera contratto di un fattore γ^{-1}, per cui condivide la conclusione che il muone raggiungerà la superficie della terra.

3 Spaziotempo di Minkowski

Verifichiamo con un esercizio che le trasformazioni di Lorentz garantiscono che la velocità della luce è identica nei sistemi di riferimento inerziali.

Esercizio 1: *Un lampo di luce emesso al tempo $t = 0$ nel punto $x = y = z = 0$ del sistema K descrive un fronte d'onda sferico di coordinate (t, x, y, z) identificato dalla relazione $-c^2t^2 + x^2 + y^2 + z^2 = 0$. Usando le trasformazioni di Lorentz, verificare che nel sistema K' il fronte di onda sferico assume una forma identica data da $-c^2t'^2 + x'^2 + y'^2 + z'^2 = 0$, per cui la luce si propaga con la stessa velocità c.*

L'esercizio ci fa apprezzare come "*ct*" sia essenzialmente una nuova coordinata del sistema di riferimento con le dimensioni di una lunghezza: scriviamo

$$(ct, x, y, z,) = (x^0, x^1, x^2, x^3) = (x^0, \vec{x}) = x^\mu. \tag{17}$$

Le quattro coordinate x^μ costituiscono le coordinate dello spazio-tempo relativistico, detto *spazio di Minkowski*, e x^μ è chiamato quadrivettore posizione. Un quadrivettore posizione identifica un "evento" dello spazio-tempo: un evento è descritto dalle coordinate spaziotemporali di qualcosa che accade nel punto spaziale di coordinate $\vec{x}$ ad un istante di tempo $t \equiv x^0/c$.

L'esercizio precedente ci fa anche apprezzare come la quantità

$$s^2 \equiv -c^2t^2 + x^2 + y^2 + z^2 \tag{18}$$

sia invariante per trasformazioni di Lorentz, l'esercizio chiedeva di mostrarlo per $s^2 = 0$, ma la dimostrazione è identica anche per $s^2 \neq 0$. (NB: occorre non confondere l'indice che indica la componete della coordinata con una eventuale potenza: tenendo a mente questa ambiguità notazionale non c'è pericolo di confondersi).

La grandezza s^2 è uno *scalare*: significa che è un invariante per trasformazioni di Lorentz e dunque può essere calcolato a scelta con le coordinate x^μ o con le coordinate x'^μ (proprio come il modulo quadrato di un vettore usuale che può essere calcolato con il teorema di Pitagora, usando le componenti del vettore lungo gli assi cartesiani del sistema di riferimento scelto oppure usando le componenti del vettore lungo assi cartesiani ruotati rispetto a quelli precedenti). Dunque la grandezza s^2 è interpretabile come la distanza invariante al quadrato dell'evento di coordinate x^μ dall'origine del sistema di riferimento spaziotemporale con coordinate $x_0^\mu = (0, 0, 0, 0)$. In generale, dati due eventi di coordinate x^μ e y^μ, il quadrato della distanza invariante è data da

$$s^2 = -(x^0 - y^0)^2 + (x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2 \ .$$

Infatti, la scelta dell'orgine del sistema di riferimento è arbitraria, e non riveste nessun significato particolare (lo spazio tempo è uno spazio affine, piuttosto che uno spazio vettoriale).

Nel caso di punti dello spazio-tempo collegati dalla propagazione della luce si ha che $s^2 = 0$, come si vede dall'esercizio 1. In generale si può avere la seguente classificazione

$$\begin{array}{ll} s^2 < 0 & \text{(distanze di tipo tempo)} \\ s^2 = 0 & \text{(distanze di tipo luce)} \\ s^2 > 0 & \text{(distanze di tipo spazio).} \end{array}$$

Questa classificazione è utile perché non dipende dalla scelta del sistema di riferimento inerziale: è una classificazione invariante (vedi fig. 6).

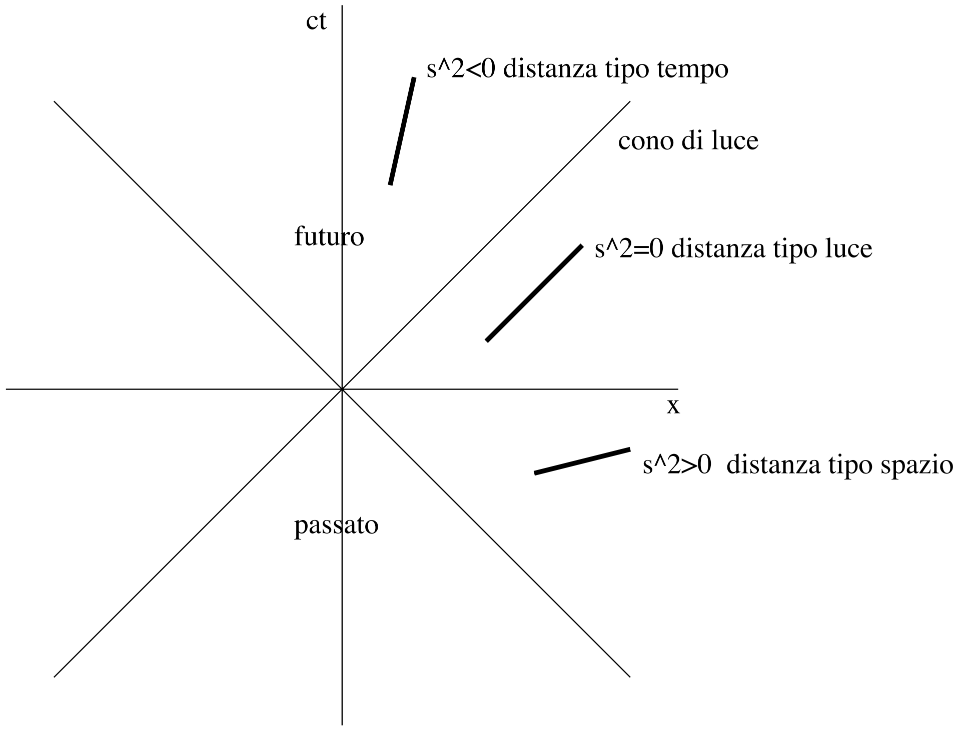
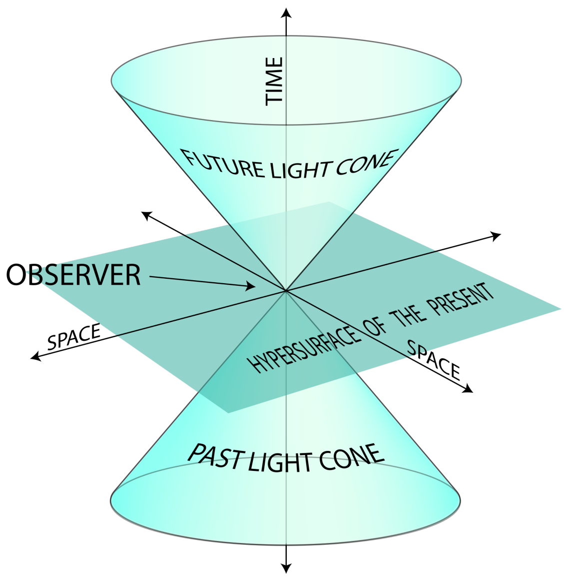


Figure 6: *Spazio di Minkowski: sono mostrate le coordinate* (ct, x) *mentre le coordinate* y *e* z *non sono riportate in figura. È indicato il cono di luce rispetto all'origine* $(0, 0)$*: la parte superiore interna al cono di luce descrive il futuro assoluto del punto* $(0, 0)$*, mentre la parte inferiore ne descrive il suo passato assoluto. Inoltre sono riportati dei segmenti le cui lunghezze Minkowskiane sono di tipo tempo, luce e spazio.*

Ecco qui sotto un'altra rappresentazione grafica dello spazio tempo

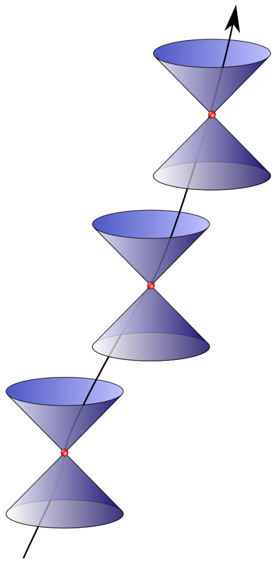


Riprendiamo il concetto di tempo proprio per vedere come può essere riformulato scrivendolo come invariate relativistico. Per ogni oggetto, il tempo proprio è quello che scorre nel sistema di riferimento solidale all'oggetto stesso. Indichiamo ora con τ il tempo proprio (è la notazione standard). Possiamo riscrivere il tempo proprio infinitesimo $d\tau$ per un'oggetto solidale con il sistema di riferimento K' (possiamo pensare ad una particella ferma posta nell'origine spaziale del sistema K') in modo invariante nel seguente modo

$$d\tau \equiv dt' = \gamma^{-1} dt = dt\,\sqrt{1-\frac{v^2}{c^2}} = \sqrt{dt^2 - \frac{dx^2+dy^2+dz^2}{c^2}} = \sqrt{-\frac{ds^2}{c^2}} \tag{19}$$

dove nell'ultimo passaggio abbiamo usato la definizione già introdotta di lunghezza minkowskiana quadrata che infatti è negativa per distanze di tipo tempo. Questa lunghezza è un invariante di Lorentz, e quindi facilmente calcolabile in qualunque sistema di riferimento. Dalla relazione (19) risulta che il tempo proprio è un invariante relativistico.

Una particella nel suo moto descrive una linea nello spazio tempo, detta *linea di mondo* o linea d'universo. Questa linea in punto deve naturalmente indirizzarsi all'interno del cono di luce del punto in considerazione, poichè la velocità della luce è insuperabile.



Il tempo proprio essenzialmente misura (in maniera invariante) la lunghezza della linea di mondo della particella ($d\tau = \frac{1}{c}\sqrt{-ds^2}$). Più avanti sarà utilizzato per costruire un' azione per una particella relativistica massiva, cosi da dedurre le equazioni del moto tramite il principio variazionale di minima azione. Particelle che vengono create e distrutte sono descritte da linee di mondo finite, utilizzate nelle rappresentazioni grafiche dei diagrammi di Feynman.

4 Trasformazioni di Lorentz e formalismo tensoriale

Possiamo scrivere la trasformazione di Lorentz (2) come

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{20}$$

dove

$$\beta \equiv \frac{v}{c}, \qquad \gamma \equiv \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{21}$$

con $0 \leq \beta < 1$ e $1 \leq \gamma < \infty$ per sistemi di riferimento inerziali fisicamente realizzabili.

Queste trasformazioni possono essere riscritte in modo compatto usando varie notazioni

$$x' = \Lambda x \tag{22}$$

$$x'^{\mu} = \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{23}$$

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{24}$$

dove Λ indica la matrice della trasformazione di Lorentz riportata in (20) e $\Lambda^{\mu}{}_{\nu}$ i corrispondenti elementi di matrice

$$\Lambda = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix} . \tag{25}$$

La notazione matriciale (22) è comoda quando si hanno solo vettori e matrici. Nella (24) è stata usata la convenzione di Einstein, secondo cui gli indici ripetuti due volte sono considerati automaticamente sommati su tutti i loro possibili valori (in questo caso l'indice $\nu = 0, 1, 2, 3$). Come descritto sopra, $\Lambda^\mu{}_\nu$ indica al variare di μ e ν gli elementi della matrice Λ. In particolare $\Lambda^\mu{}_\nu$ con μ e ν fissato indica l'elemento in riga μ e in colonna ν. Dunque il primo indice (posto convenzionalmente in alto) indica l'indice di riga e il secondo indice (posto convenzionalmente in basso) indica l'indice di colonna. Inoltre, nella convenzione di Einstein in cui due indici ripetuti sono da considerarsi sommati, si richiede sempre che un indice sia in alto ed uno in basso (questa convenzione è molto utile poichè, come vedremo, garantisce a vista l'invarianza sotto un punto di vista gruppale).

Il quadrato della distanza minkowskiana s^2, che come abbiamo visto è un invariante di Lorentz, può essere scritto nei modi seguenti

$$\begin{aligned} s^2 = -c^2t^2 + x^2 + y^2 + z^2 = \begin{pmatrix} ct, & x, & y, & z \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \\ = x^T \eta \, x = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} x^\mu \eta_{\mu\nu} x^\nu = x^\mu \eta_{\mu\nu} x^\nu = \eta_{\mu\nu} x^\mu x^\nu \end{aligned} \tag{26}$$

dove si è introdotta la <u>metrica di Minkowski</u>, cioè la matrice η con componenti $\eta_{\mu\nu}$

$$\eta = \begin{pmatrix} \eta_{00} & \eta_{01} & \eta_{02} & \eta_{03} \\ \eta_{10} & \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{20} & \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{30} & \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{27}$$

che ci permette di valutare il modulo quadro dei quadrivettori. Dunque $\eta_{00} = -1$, $\eta_{11} = \eta_{22} = \eta_{33} = 1$, mentre $\eta_{\mu\nu} = 0$ se $\mu \neq \nu$. Questo quadrato della lunghezza minkowskiana generalizza il concetto di modulo quadro di un vettore in spazi euclidei. Da notare come gli indice di riga e quelli di colonna sono posizionati entrambi in basso. Inoltre, poiché la matrice η è simmetrica, si possono scambiare righe con colonne lasciando la matrice invariata ($\eta^T = \eta$), e quindi vale $\eta_{\mu\nu} = \eta_{\nu\mu}$, come facilmente verificabile.

In generale le trasformazioni di Lorentz sono per definizione tutte quelle che lasciano invariata la distanza minkowskiana. Vediamo di esprimere questa definizione in equazioni. Usando la notazione matriciale

$$\begin{aligned} s^2 &= x^T \eta \, x \qquad \text{(definizione del quadrato della distanza)} \\ x' &= \Lambda \, x \qquad \text{(trasformazione di Lorentz arbitraria)} \end{aligned} \tag{28}$$

per cui

$$s^2 = s'^2 \qquad \Rightarrow \qquad x^T \eta \, x = x'^T \eta \, x' = x^T \Lambda^T \eta \, \Lambda \, x \qquad \Rightarrow \qquad \Lambda^T \eta \, \Lambda = \eta \tag{29}$$

dove l'ultima equazione segue dal fatto che la relazione precedente deve valere per ogni quandrivettore x. Dunque tutte le trasformazioni di Lorentz possibili sono quelle definite da matrici Λ tali che

$$\Lambda^T \eta \, \Lambda = \eta \tag{30}$$

dove η è la metrica di Minkowski. In notazione tensoriale questa equazione si scrive come

$$\eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \eta_{\alpha\beta} \; . \tag{31}$$

Questa relazione può essere verificata riprendo i passaggi fatti in (29) usano le componenti dei quadrivettori e della metrica

$$s^2 = \eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} (\Lambda^\mu{}_\alpha x^\alpha)(\Lambda^\nu{}_\beta x^\beta) = \eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta x^\alpha x^\beta = \eta_{\alpha\beta} x^\alpha x^\beta = s^2 \tag{32}$$

da cui segue la (31).

Esercizio 2: *Verifcare che la matrice in (20) soddisfa la relazione (30).*

L'insieme delle trasformazioni di Lorentz formano un gruppo: *il gruppo di Lorentz*, indicato spesso con $O(3,1)$

$$O(3,1) = \{\Lambda, \text{matrici reali } 4 \times 4 \mid \Lambda^T \eta \Lambda = \eta\}$$

L'analisi della proprietà fondamentale (30) che caratterizza il gruppo di Lorentz permette di dedurre che tale gruppo, cioè l'insieme di tutte le matrici che definiscono trasformazioni di Lorentz (si dimostri che tale insieme soddisfa agli assiomi che definiscono un gruppo), può essere parametrizzato da 6 variabili: 3 angoli che definiscono la rotazione degli assi cartesiani spaziali più le 3 componenti della velocità $\vec{v}$ relativa tra i due sistemi di riferimento inerziali. Ci sono inoltre trasformazioni discrete, inversione spaziale (o parità) ed inversione temporale, che verranno discusse più avanti.

Nota tecnica: Abbiamo detto che $\Lambda^\mu{}_\nu$ indica, al variare di μ e ν, gli elementi della matrice Λ, ed in particolare indica l'elemento in riga μ e in colonna ν, come esemplificato in (25). Dunque il primo indice (posto convenzionalmente in alto) indica l'indice di riga e il secondo indice (posto convenzionalmente in basso) indica l'indice di colonna. Nel prodotto di due matrici, $A = BC$, l'elemento $A^\mu{}_\nu$ è dato dal prodotto vettoriale della riga μ di B con la colonna ν di C, e quindi da

$$A^\mu{}_\nu = B^\mu{}_\lambda C^\lambda{}_\nu \tag{33}$$

dove la somma in λ (indice ripetuto due volte e quindi sommato) produce il prodotto vettoriale. In alcune situazioni, ad esempio nella descrizione delle rotazioni, non è necessario distinguere tra indici in alto ed indici in basso. Per esempio gli elementi della matrice di rotazione R in uno spazio tridimensionale possono essere indicati da R^{ij}, con $i, j = 1, 2, 3$, per cui le componenti di un vettore si trasformano come $x'^i = R^{ij} x^j$. Il prodotto di matrici $R = PQ$ è descritto dagli elementi di matrice $R^{ij} = P^{ik} Q^{kj}$ (indice k ripetuto due volte e quindi sommato automaticamente da 1 a 3). Infine si noti che il trasposto R^T di R ha elementi di matrice $(R^T)^{ij} = R^{ji}$, si devono infatti scambiare le righe con le colonne, e ad esempio il prodotto $U = R^T S$ è dato da $U^{ij} = (R^T)^{ik} S^{kj} = R^{ki} S^{kj}$.

NB. *La posizione degli indici in alto o in basso non è arbitraria, ma ha un significato ben preciso*: il quadrivettore posizione x^μ ha per definizione l'indice in alto, la metrica di Minkowski ha due indici in basso $\eta_{\mu\nu}$, gli elementi di matrice di una trasformazione di Lorentz sui quadrivettori x^μ ha l'indice di riga in alto e l'indice di colonna in basso, $\Lambda^\mu{}_\nu$. Partendo da queste definizioni, risulta utile definire un quadrivettore con l'indice in basso ("quadrivettore covariante") usando al metrica come

$$x_\mu \equiv \eta_{\mu\nu} x^\nu$$

da cui segue che si può riottenere il quadrivettore con l'indice in alto ("quadrivettore controvariante") usando l'inverso della metrica come

$$x^\mu = \eta^{\mu\nu} x_\nu \ , \qquad \eta^{\mu\nu} \equiv (\eta^{-1})^{\mu\nu} \ . \tag{34}$$

Per definizione, $\eta^{\mu\nu}$ indica gli elementi della matrice inversa della metrica, η^{-1}. Non è necessario specificare che si tratta della matrice inversa, poiché questo è deducibile dalla posizione in alto degli indici.

Queste definizioni sono autoconsistenti: in notazione tensoriale possiamo verificare che

$$x^\mu = \eta^{\mu\nu} x_\nu = \eta^{\mu\nu}(\eta_{\nu\lambda} x^\lambda) = (\eta^{\mu\nu}\eta_{\nu\lambda}) x^\lambda = \delta^\mu{}_\lambda x^\lambda = x^\mu \tag{35}$$

dove nel calcolo si è usata la delta di Kronecker $\delta^\mu{}_\lambda$ che corrisponde alle componenti della matrice identità (ricordare la convenzione di sommatoria di Einstein e notare che l'equazione matriciale $\eta^{-1}\eta = I$, dove I è la matrice identità, si scrive in componenti come $\eta^{\mu\lambda}\eta_{\lambda\nu} = \delta^\mu{}_\nu$. Similmente $Ix = x$ si scrive in componenti come $\delta^\mu{}_\nu x^\nu = x^\mu$). La delta di Kronecker rappresenta le componenti della matrice identità, che è una matrice simmetrica (si possono scambiare le righe con le colonne), per cui spesso è scritta come δ^μ_ν, senza rendere evidente quale sia l'indice di riga e quale quello di colonna.

In notazione matriciale, le x_μ sono quelle del quadri-vettore $\tilde{x} = \eta x$, da cui segue che $x = \eta^{-1}\tilde{x}$. Usando le componenti x_μ la tilde è sottintesa, in quanto la stessa informazione è contenuta nel fatto che l'indice è posto in basso.

Usando la definizione di x_μ si può anche riscrivere la (26) come

$$s^2 = x_\mu x^\mu \ . \tag{36}$$

L'operatore di derivata $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$ si comporta come un 4-vettore con l'indice in basso. Questo si può verificare calcolando la derivata dell'invariante s^2

$$\partial_\mu s^2 \equiv \frac{\partial s^2}{\partial x^\mu} = \frac{\partial}{\partial x^\mu}(\eta_{\nu\lambda} x^\nu x^\lambda) = 2\eta_{\mu\lambda} x^\lambda = 2x_\mu \ . \tag{37}$$

In generale si definiscono scalari, quadrivettori e quadritensori (o più brevemente vettori e tensori) quantità che si trasformano in modo ben preciso per trasformazioni di Lorentz

$$\begin{array}{rcll} s' &=& s & \text{(scalare)} \\ x'^\mu &=& \Lambda^\mu{}_\nu \, x^\nu & \text{(quadrivettore)} \\ F'^{\mu\nu} &=& \Lambda^\mu{}_\lambda \Lambda^\nu{}_\rho \, F^{\lambda\rho} & \text{(quadritensore di rango due)} \\ T'^{\mu\nu\lambda} &=& \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\lambda{}_\gamma \, T^{\alpha\beta\gamma} & \text{(quadritensore di rango tre)} \\ && \ldots & \end{array} \tag{38}$$

Esempio di uno scalare è la distanza Minkowskiana, di un vettore il quadrivettore coordinata o come vedremo più avanti il quadrimpulso, di un tensore a due indici il tensore campo elettromagnetico $F^{\mu\nu}$, tensore antisimmetrico ($F^{\mu\nu} = -F^{\nu\mu}$) con sei componenti indipendenti, che si può scrivere come (in unità gaussiane o in quelle di Heaviside–Lorentz)

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \tag{39}$$

Un modo di interpretare la (36) è dire che la trasformazione del quadrivettore x^μ è compensata dalla trasformazione del quadrivettore x_μ, per cui la "contrazione" degli indici in $x_\mu x^\mu$ produce uno scalare. Similmente dati vettori e tensori si deduce facilmente che ad esempio $A_\mu B_\nu F^{\mu\nu}$ è uno scalare, $B_\nu F^{\mu\nu}$ è uno quadrivettore controvariante, etc. In generale le posizioni degli indici in grandezze tensoriali indicano le proprietà di trasformazione sotto trasformazioni di Lorentz, ed indici contratti possono essere ignorati in quanto si comportano come uno scalare.

Tendo conto della metrica che abbassa e alza indici, possiamo scrivere la trasformazione di Lorentz di un vettore covariante (indici in basso) nella forma

$$x_\mu = \Lambda_\mu{}^\nu x_\nu \qquad \text{con} \qquad \Lambda_\mu{}^\nu = \eta_{\mu\alpha}\Lambda^\alpha{}_\beta \eta^{\beta\nu} = (\Lambda^{T,-1})_\mu{}^\nu \tag{40}$$

dove nella prima uguaglianza a destra si è abbassato il primo indice e alzato il secondo indice in $\Lambda^\alpha{}_\beta$, mentre la seconda uguaglianza segue dalla proprietà definente $\Lambda^T \eta \Lambda = \eta$ delle matrici di Lorentz (che possiamo riscrivere come $\Lambda^T \eta \Lambda \eta^{-1} = I$, con I matrice identità, da cui segue che $\eta \Lambda \eta^{-1} = \Lambda^{T,-1}$). Possiamo di nuovo verificare l'invarianza di $x_\mu x^\mu$

$$x_\mu x^\mu \qquad \to \qquad x'_\mu x'^\mu = \Lambda_\mu{}^\alpha x_\alpha \Lambda^\mu{}_\beta x^\beta = (\Lambda^{-1}\Lambda)^\alpha{}_\beta x_\alpha x^\beta = \delta^\alpha{}_\beta x_\alpha x^\beta = x_\alpha x^\alpha = x_\mu x^\mu \ . \tag{41}$$

Equazioni tensoriali sono quelle che eguagliano tensori dello stesso rango, e quindi oggetti con identiche proprietà di trasformazione. Queste assumono la stessa forma in tutti i sistemi di riferimento inerziali, e molto spesso contengono nel lato destro il tensore nullo. Ad esempio, metà delle equazioni di Maxwell nel vuoto possono essere scritte come

$$\partial_\nu F^{\mu\nu} = J^\mu$$

dove il quadrivettore $\partial_\nu F^{\mu\nu}$ è uguagliato al quadrivettore della densità di corrente di carica J^μ. Si noti che compare un solo indice libero (l'indice μ) che al suo variare descrive quattro equazioni distinte. Nel lato sinistro compare anche l'indice ν, compare due volte, una in alto ed una in basso, indicando una somma che produce un prodotto scalare invariante (per quanto riguarda questi indici). Entrambi i membri dell' equazione sono quadrivettori, e l'equazione assume la stessa forma in tutti i sistemi di riferimento collegati da trasformazioni di Lorentz (equazione invariante in forma).

Le proprietà geometriche dettate dalla metrica di Minkowski definiscono lo spazio-tempo di Minkowski, che rappresenta appunto un modello del nostro spazio-tempo fisico. Una scelta di quattro assi "cartesiani" lungo cui misurare le distanze x^μ dello spazio-tempo rappresenta un sistema di riferimento inerziale. Alcune definizioni e proprietà dello spazio-tempo di Minkowski sono descritte e riportate nella figura 5.

Oltre alla possibilità di cambiare il sistema di riferimento con trasformazioni di Lorentz è possibile fare una diversa scelta dell' origine del sistema di riferimento. Ciò corrisponde alla possibilità di operare 3 traslazioni spaziali ed 1 traslazione temporale. Quando si aggiunge questa ulteriore invarianza alle trasformazioni di Lorentz si ottiene un gruppo totale di trasformazioni a dieci parametri detto gruppo di Poincarè, sotto cui un punto dello spazio tempo si trasforma nel modo seguente

$$x^{\mu\prime} = \Lambda^\mu{}_\nu x^\nu + a^\mu \tag{42}$$

I dieci parametri corrispondono ai 4 parametri a^μ che definiscono una traslazione spazio-temporale più i 6 parametri del gruppo di Lorentz contenuti in una $\Lambda^\mu{}_\nu$ generica. Si può dimostarre che l'invarianza per traslazioni spazio-temporali è collegata alla conservazione dell'impulso e dell'energia. Similmente l'invarianza per trasformazioni di Lorentz comporta la conservazione di 6 quantità (che includono le 3 componenti del momento angolare).

Esercizio 3: *Utilizzare le leggi di trasformazione (38) per il campo elettromagnetico definito in (39) per derivare le seguenti leggi di trasformazione per cambio del sistema di riferimento specificato dalla trasformazione in (20)*

$$\begin{array}{ll} E_x{}' = E_x & B_x{}' = B_x \\ E_y{}' = \gamma(E_y - \beta B_z) & B_y{}' = \gamma(B_y + \beta E_z) \\ E_z{}' = \gamma(E_z + \beta B_y) & B_z{}' = \gamma(B_z - \beta E_y). \end{array} \tag{43}$$

Esercizio 4: *Verificare che se un tensore $T^{\mu\nu}$ è simmetrico (antisimmetrico), il suo trasformato rimane simmetrico (antisimmetrico).*
Esercizio 5: *Verificare che le equazioni di Maxwell si possono scrivere come ($\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$)*

$$\partial_\mu F^{\mu\nu} = -J^\nu \tag{44}$$

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \tag{45}$$

Si noti che se la sorgente J^μ si trasforma come un quadrivettore, allora è immediato verificare l'invarianza relativistica delle equazioni di Maxwell scritte in questa forma. Calcolare infine il valore dello scalare $F^{\mu\nu}F_{\mu\nu}$. Perchè è uno scalare?.
Esercizio 6: *Verificare che il secondo set di equazioni di Maxwell, cioè le equazioni senza sorgenti in (45), sono risolte automaticamente dall'introduzione di un quadripotenziale vettore A_μ definendo*

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

5 Trasformazioni discrete e gruppo di Lorentz proprio ed ortocrono

Il gruppo di Lorentz è spesso indicato con $O(3,1)$

$$O(3,1) = \{\Lambda, \text{matrici reali } 4\times 4 \mid \Lambda^T \eta\, \Lambda = \eta\}$$

che generalizza ad uno spaziotempo con tre direzioni spaziali ed una temporale il gruppo ortogonale delle rotazioni spaziali

$$O(3) = \{R, \text{matrici reali } 3\times 3 \mid R^T R = I\}\ .$$

Questi sono gruppi di Lie, con cui si intendono gruppi che dipendono in modo continuo da alcuni parametri. Per il gruppo di Lorentz questi parametri possono essere relazionati alle tre componenti della velocità relativa $\vec{v}$ con cui i due sistemi di riferimento si muovono in moto relativo uniforme, ed a tre angoli $\vec{\theta}$ che descrivono una eventuale rotazione degli assi spaziali. È quindi un gruppo a sei parametri. Se all'annullarsi di questi parametri si ha la trasformazione identità, variando in modo continuo questi parametri possiamo raggiungere tutte le matrici di Lorentz connesse all'identità. Tutte queste matrici, che per convenienza possiamo indicare con $\Lambda(\vec{v},\vec{\theta})$, hanno determinante $\det\Lambda(\vec{v},\vec{\theta}) = 1$ ed hanno la componente $\Lambda^0{}_0(\vec{v},\vec{\theta}) \geq 0$. In generale si deduce facilmente, calcolando il determinante della relazione definente le matrici del gruppo di Lorentz, che il determinante di una matrice di Lorentz λ può valere ± 1

$$\det(\Lambda^T \eta\, \Lambda) = \det(\eta) \qquad \rightarrow \qquad \det(\Lambda) = \pm 1 \; .$$

Inoltre, valutando la componente 00 della (31) si ottiene che

$$(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\Lambda^i{}_0)^2 \geq 1 \qquad \rightarrow \qquad \Lambda^0{}_0 \geq 1 \;\; \text{oppure} \;\; \Lambda^0{}_0 \leq -1 \; .$$

Poiché la trasformazione identità ha determinante unitario e $\Lambda^0{}_0 = 1$, per continuità le trasformazioni della parte connessa all'identità ha $\det\Lambda(\vec{v},\vec{\theta}) = 1$ e $\Lambda^0{}_0 \geq 1$. Queste matrici formano un sottogruppo, denominato *gruppo di Lorentz proprio ed ortocrono*, spesso indicato con $SO^{\uparrow}(3,1)$. Per invarianza relativistica solitamente si intende solo l'invarianza per trasformazioni appartenenti a questo sottogruppo.

Esistono poi trasformazioni discrete, inversione spaziale (o parità) ed inversione temporale, che non sono connesse all'identità, ma appartengono al gruppo di Lorentz $O(3,1)$.

L'inversione spaziale (indicata con P) è definita da

$$x^{\mu\prime} = P^\mu{}_\nu x^\nu \; , \qquad P^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

che evidentemente cambia l'orientamento degli assi spaziali. Appartiene al gruppo di Lorentz ($P^T \eta P = \eta$) ed ha $\det P = -1$.

Similmente l'inversione temporale (indicata con T) è definita da

$$x^{\mu\prime} = T^\mu{}_\nu x^\nu \; , \qquad T^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

cambia la direzione dell'asse temporale, appartiene al gruppo di Lorentz ($T^T \eta T = \eta$) ed ha $\det T = -1$.

Componendo le matrici del gruppo di Lorentz proprio ed ortocrono $SO^{\uparrow}(3,1)$ con le trasformazioni discrete P e T si ottengono gli elementi delle parti disconnesse dell'intero gruppo di Lorentz $O(3,1)$ che in totale ha quattro componenti disconnesse.

N.B. A questo punto è conveniente affrontare la teoria dei gruppi un po' più in generale (vedere appunti relativi).

6 Definizione relativistica di energia e momento: il quadrimomento

Abbiamo già analizzato il concetto di tempo proprio. In un sistema di riferimento inerziale arbitrario il tempo proprio infinitesimo di un oggetto in moto con velocità $\vec{v}$ può essere scritto come

$$d\tau = dt\,\sqrt{1-\frac{v^2}{c^2}}\ . \tag{46}$$

È un invariante relativistico (uno scalare), come risulta evidente se riscriviamo il tempo proprio come

$$d\tau = \frac{1}{c}\sqrt{-ds^2}\ . \tag{47}$$

Utilizzeremo questo parametro scalare per introdurre il concetto di quadrimomento (detto anche quadrimpulso) per particelle massive.

6.1 Particelle massive

Una particella nel suo moto descrive una traiettoria nello spazio-tempo. A questa traiettoria si dà il nome di linea di mondo (o linea d'universo), vedi figura 7.

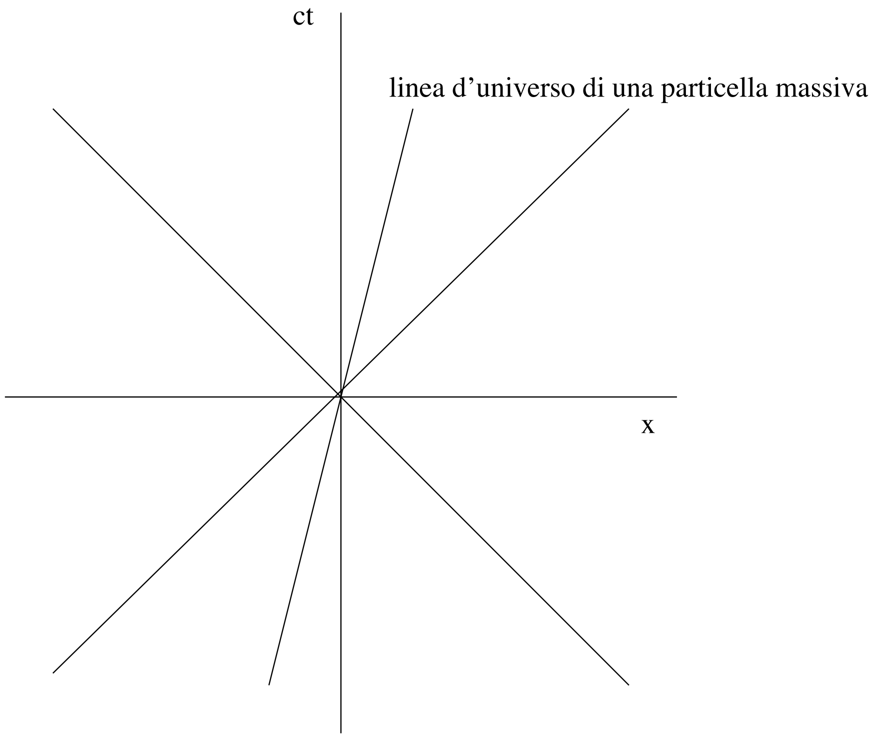


Figure 7: `La linea di universo di una particella massiva è contenuta all'interno del cono di luce.`

Possiamo parametrizzare questa linea di mondo in vari modi, ad esempio usando il tempo t come parametro e scrivere $\vec{x}(t)$. Un altro modo è quello di utilizzare il tempo proprio τ, che ha il vantaggio di essere uno scalare, ed indicare con $x^\mu(\tau)$ la linea di mondo della particella: per ogni valore del tempo proprio della particella τ le funzioni $x^\mu(\tau)$ ci dicono la posizione della particella nello spaziotempo nel sistema di riferimento scelto, cioè la posizione occupata $\vec{x}(\tau)$ al tempo $x^0(\tau)$.

Consideriamo dunque una una particella massiva che percorre la linea di mondo $x^\mu(\tau)$ parametrizzata tramite il tempo proprio. La quadrivelocità è per definizione il seguente quadrivettore

$$u^\mu(\tau) \equiv \frac{dx^\mu(\tau)}{d\tau} = \left(\frac{cdt(\tau)}{d\tau}, \frac{d\vec{x}(\tau)}{d\tau}\right) = \left(\frac{c}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}\right) \tag{48}$$

dove si è fatto uso dell'equazione (46). Questa quantità è un quadrivettore poichè $d\tau$ è uno scalare e dx^μ un quadrivettore. È immediato calcolarne il modulo quadrato

$$u^\mu u_\mu = -(u^0)^2 + \vec{u}\cdot\vec{u} = -c^2 \tag{49}$$

Poichè sappiamo che $u^\mu u_\mu$ è un invariante di Lorentz, per semplicità si sarebbe potuto effettuare il calcolo nel sistema di riferimento a riposo con la particella, in cui $u^\mu = (c, 0)$ come si evince dalla (48), e dunque $u^\mu u_\mu = -c^2$.

Il quadrimomento (o quadrimpulso) della particella è definito da

$$p^\mu = mu^\mu \tag{50}$$

dove m è la massa della particella. Questa massa è per definizione una proprietà scalare assegnata alla particella. A volte è detta massa invariante, massa a riposo o massa propria per differenziarla da eventuali altre definizioni (spesso inutili). Dunque anche p^μ è un quadrivettore, quindi per cambio del sistema di riferimento inerziale subisce una trasformazione di Lorentz della forma

$$p'^\mu = \Lambda^\mu{}_\nu p^\nu$$

Il modulo quadro di p^μ è facilmente calcolabile

$$p^\mu p_\mu = m^2 u^\mu u_\mu = -m^2c^2 \tag{51}$$

Se non agiscono forze sulla particella, la quadrivelocità ed il quadrimomento sono quadrivettori costanti. Familiarizziamo un pò con queste definizioni relativistiche per vedere come le usuali definizioni non-relativistiche per una particella libera sono generalizzate nella meccanica relativistica

$$p^\mu = m\frac{dx^\mu(\tau)}{d\tau} = \left(mc\frac{dt(\tau)}{d\tau}, m\frac{d\vec{x}(\tau)}{d\tau}\right) = \left(\frac{mc}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}\right) = \left(\frac{E}{c}, \vec{p}\right) = (p^0, \vec{p})\ . \tag{52}$$

Abbiamo identificato la componente p^0 con l'energia E associata alla particella (diviso c per immediate considerazioni dimensionali), mentre le componenti spaziali sono identificate con la definizione relativistica del momento spaziale. Giustifichiamo queste identificazioni analizzando il limite non-relativistico $v \ll c$.

<u>Energia E</u>: dalla (52) si ottiene

$$E = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}\ . \tag{53}$$

Per $v = 0$ si vede che la teoria della relatività assegna in modo naturale una energia a riposo proporzionale alla massa $E = mc^2$. Per $v \ll c$ possiamo sviluppare in serie di $\frac{v}{c}$ che è un numero piccolo rispetto ad 1

$$E = mc^2 \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = mc^2 \left(1 + \frac{1}{2}\frac{v^2}{c^2} + ...\right) = mc^2 + \frac{1}{2}mv^2 + ... \tag{54}$$

Questo ci fà vedere come la definizione non-relativistica di energia cinetica venga riprodotta nel limite di velocità basse rispetto a quella della luce. Alla luce di questo calcolo si può intuire la definizione appropriata di *energia cinetica* T nel caso relativistico

$$T = E - mc^2 \; . \tag{55}$$

Momento $\vec{p}$: dalla (52) si vede che nel limite $v \ll c$ si riottiene la definizione non-relativistica di momento lineare

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \rightarrow \qquad \vec{p} = m\vec{v} \; . \tag{56}$$

Si noti che particelle massive non possono raggiungere la velocità della luce: dovrebbero avere energia e momento infiniti! Dunque la velocità $v = c$ è una velocità limite teoricamente irraggiungibile per particelle massive, poichè nessun fenomeno fisico è in grado di cedere un'energia infinita ad una particella.

Queste definizioni relativistiche possono essere giustificate in maniera più rigorosa e derivate partendo da un principio d'azione che descriva la dinamica relativistica. L'azione corretta per una particella libera è proporzionale al tempo proprio integrato sulla linea di mondo della particella (così da garantire l'invarianza relativistica)

$$S[\vec{x}(t)] = -mc^2 \int dt \sqrt{1 - \frac{\dot{\vec{x}} \cdot \dot{\vec{x}}}{c^2}} \tag{57}$$

dove naturalmente $\vec{v} = \dot{\vec{x}} = \frac{d\vec{x}}{dt}$ descrive la velocità della particella. La costante di proporzionalità $(-mc^2)$ è stata scelta per ottenere il corretto limite non relativistico $L = \frac{mv^2}{2} - mc^2 + \cdots$.

Poiché la corrispondente lagrangiana $L = -mc^2\sqrt{1 - \frac{\dot{\vec{x}} \cdot \dot{\vec{x}}}{c^2}}$ non dipende dalla posizione $\vec{x}$, ma solo dalla velocità $\dot{\vec{x}}$, il momento coniugato

$$\vec{p} \equiv \frac{\partial L}{\partial \dot{\vec{x}}} = \frac{m\dot{\vec{x}}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{58}$$

è conservato (come conseguenza delle equazioni di Eulero-Lagrange $\frac{d\vec{p}}{dt} = 0$). Inoltre anche l'hamiltoniana

$$H = \vec{p} \cdot \dot{\vec{x}} - L = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{59}$$

che coincide con l'energia della particella è conservata.

Nel caso siano presente forze esterne, in una teoria relativistica le equazioni del moto possono essere scritte nella forma

$$\frac{dp^\mu}{d\tau} = f^\mu \tag{60}$$

dove f^μ è detto quadriforza, la generalizzazione relativistica del concetto di forza, che naturalmente si trasforma come un quadrivettore per trasformazioni di Lorentz.

6.2 Particelle con massa nulla

Abbiamo visto che particelle massive non possono raggiungere esattamente la velocità della luce. Però possono esistere particelle che viaggino alla velocità della luce purchè abbiano massa nulla. Questa interpretazione è consistente con la relatività ristretta. Infatti la relatività ristretta prevede che particelle che vanno alla velocità della luce siano costrette a viaggiare sempre a quella velocità, che infatti è necessariamente la stessa in tutti i sistemi di riferimento inerziali. Questa proprietà ci dice che non si può trovare un sistema di riferimento a riposo con particelle di massa nulla, e quindi non esiste il concetto di tempo proprio per tali particelle. Infatti dalla (19) si vede che il presunto tempo proprio dovrebbe essere nullo, e quindi sarebbe un tempo che non può scorrere e non può essere utilizzato per parametrizzare la linea d'universo di queste particelle. Infatti questa linea d'universo è una linea di tipo luce e deve necessariamente giacere sul cono di luce della particella stessa. In ogni caso possiamo ugualmente assegnare un quadrimomento alla particella con massa nulla. In tal caso l'invariante relativistico $p^\mu p_\mu$ si annulla e può essere utilizzato per ricavare una relazione tra energia ed impulso

$$p^\mu p_\mu = 0 \qquad \Rightarrow \qquad -\frac{E^2}{c^2} + \vec{p} \cdot \vec{p} = 0 \qquad \Rightarrow \qquad E = |\vec{p}| c\ . \tag{61}$$

Questa formula è valida approssimativamente anche per particelle che viaggiano a velocità molto vicine a quella della luce e che di conseguenza hanno energie molto maggiori della loro massa a riposo, $E \gg mc^2$ (particelle ultra-relativistiche. Si noti che il concetto di particella ultra-relativistica dipende dal sistema di riferimento scelto). Vediamo di derivare questa affermazione. Abbiamo visto che per una particella di massa m

$$p^\mu = \left(\frac{E}{c}, \vec{p}\right) = \left(\gamma mc, \gamma m\vec{v}\right) \tag{62}$$

da cui si ottiene la relazione

$$\vec{p} = \frac{E\vec{v}}{c^2}\ . \tag{63}$$

Nel limite di $v \to c$ si ottiene

$$|\vec{p}| = \frac{E}{c} \tag{64}$$

che coincide con quanto ottenuto in eq. (61) per particelle a massa nulla.

Concludiamo questa sezione ricordando che non è stato mai possibile interpretare in modo consistente, all'interno di teorie fisiche, i tachioni (ipotetiche particelle che possano viaggiare a velocità maggiori di c).

6.3 Legge di conservazione del quadrimomento

La definizione data sopra di quadrimomento per particelle massive e particelle senza massa è appropriata in quanto è proprio questa quantità che soddisfa a specifiche leggi di conservazione. Infatti si può dimostare che:
il quadrimomento totale è conservato dalle interazioni relativistiche invarianti per traslazioni spaziali e temporali (tutte le interazioni fondamentali lo sono)

$$P^\mu_{totale,iniziale} = P^\mu_{totale,finale}\ . \tag{65}$$

Dunque il quadrimomento totale si conserva per sistemi isolati in cui avvengono processi di urto regolati dalle interazioni fondamentali. Questo corrisponde a quattro leggi di conservazione (i quattro valori possibili dell'indice μ in (65): conservazione dell'energia e conservazione delle tre componenti del momento lineare.

Consideriamone immediatamente una applicazione: il processo di una particella di massa M a riposo che decade in due particelle di massa m_1 e m_2. Utilizziamo la conservazione del quadrimomento totale per calcolare le energie delle particelle finali. Il quadrimomento della particella iniziale è dato da

$$p^\mu = (Mc, 0)$$

mentre indichiamo con

$$p_1^\mu = (E_1/c, \vec{p}_1) \ , \qquad p_2^\mu = (E_2/c, \vec{p}_2)$$

i quadrimomenti delle particelle finali. La legge di conservazione è data da

$$p^\mu = p_1^\mu + p_2^\mu \tag{66}$$

che esplicitata si riduce a

$$\begin{aligned} &Mc^2 = E_1 + E_2 \\ &\vec{p}_1 + \vec{p}_2 = 0 \ . \end{aligned}$$

Si noti che, poiché $E_1 \geq m_1 c^2$ e $E_2 \geq m_2 c^2$, necessariamente $M \geq m_1 + m_2$ affiché il processo possa avvenire. Per proseguire consideriamo la seconda equazione che implica l'eguaglianza del quadrato dei momenti delle particelle finali

$$\vec{p}_1^{\,2} = \vec{p}_2^{\,2} \qquad \rightarrow \qquad E_1^2 - m_1^2 c^4 = E_2^2 - m_2^2 c^4$$

ed unita alla prima equazione (la conservazione dell'energia) implica che

$$E_1 = \frac{M^2 + m_1^2 - m_2^2}{2M} c^2 \ , \qquad E_2 = \frac{M^2 + m_2^2 - m_1^2}{2M} c^2 \ .$$

In particolare si deduce che le particelle finali sono monoenergetiche.

Il decadimento β del neutrone è invece un decadimento a tre corpi: $n \rightarrow p + e^- + \bar{\nu}_e$. Originariamente non si conosceva l'esistenza del neutrino, ma siccome l'elettrone finale non mostrava uno spettro monoenergetico, nel 1930 Pauli assumendo la conservazione dell'energia ipotizzò l'esistenza di una terza particella prodotta nel decadimento, il neutrino, poi scoperto sperimentalmente nel 1956.

Per investigare processi di decadimento, è comodo definire il concetto di massa invariante: in processi in cui nello stato finale emergono un numero N fissato di particelle, con relativo quadrimomento p_i^μ, con $i = 1, ..., N$, si definisce massa invariante del sistema la massa associata alla somma dei quadrimonenti delle N particelle

$$M = \frac{1}{c}\sqrt{-p^\mu p_\mu} \ , \qquad p^\mu = \sum_{i=1}^{N} p_i^\mu \tag{67}$$

che coincide con la massa di una eventuale particella che possa decadere nelle N particelle in questione. Il caso con $N = 2$ è quello usato con più frequenza.

Per descrivere processi di urto tra due particelle che vanno in due particelle, come in figura 7, risulta utile l'uso delle variabli di Mandelstam

$$\begin{aligned} s &= -(p_1+p_2)^2 \\ t &= -(p_2-p_3)^2 \\ u &= -(p_1-p_3)^2 \end{aligned} \tag{68}$$

dove nel lato destro si intende il modulo quadrato minkowskiano dei rispettivi quadrivettori. Il quadrivettore p_n è il quadrimomento delle particella n-esima con massa m_n ed è orientato come in figura 7.

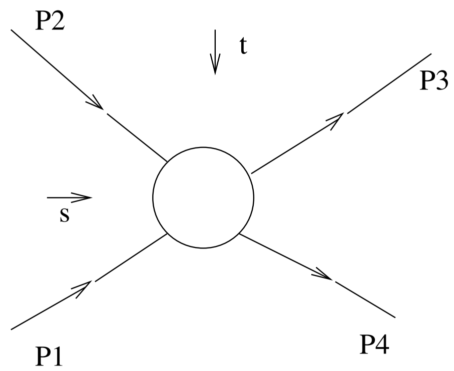


Figure 8: Un processo d'urto. Sono rappresentate le variabili di Mandelstam s e t.

Queste variabili di Mandelstam non sono indipendenti, ma si può mostare che (ponendo per semplicità $c = 1$)

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2 \tag{69}$$

dove $(p_1)^2 = -m_1^2$, $(p_2)^2 = -m_2^2$, etc. Per dimostare questa relazione occorre usare la conservazione del quadrimpulso totale, che nel presente caso assume la forma

$$p_1^\mu + p_2^\mu = p_3^\mu + p_4^\mu \ . \tag{70}$$

Se nel processo di urto sono prodotte ulteriori particelle le variabili t ed u perdono il loro significato immediato, mentre la variabile s continua ad essere estremamente utile: infatti $\sqrt{s}$ corrisponde all' energia totale presente nel sistema di riferimento "centro di massa", disponibile per la creazione di nuove particelle.

Per definizione il sistema di riferimento "centro di massa" (CM) è quello in cui il momento spaziale totale è nullo. Un altro sistema di riferimento utile quando si studia un processo di scattering di due particelle è il "sistema del laboratorio" (LAB), in cui una particella è a riposo (bersaglio) mentre l'altra particella possiede un momento non nullo (proiettile). L'uso di invariati relativistici è spesso comodo per studiare i processi di scattering. Ad esempio consideriamo il seguente esercizio

Esercizio 7: *Calcolare l'energia cinetica minima (energia di soglia) che un protone deve avere per interagire con un altro protone fermo e generare nell'urto una coppia protone–antiprotone*

$$p + p \quad \rightarrow \quad p + p + p + \bar{p} \ .$$

Per risolvere l'esercizio si consideri prima il sistema CM. L'energia minima si ha quando dopo l'urto tutte le particelle create sono ferme nel sistema CM, per cui il quadrimpulso totale finale in tale sistema vale $P_f'^{\mu} = (4m_p, \vec{0})$ ed ha come modulo quadro $P_f'^2 = -16m_p^2$. Nel sistema LAB il quadrimpulso iniziale totale vale $P_i^\mu = (m_p, \vec{0}) + (E, \vec{p}) = (m_p + E, \vec{p})$, ed ha come modulo quadro $P_i^2 = -2m_p^2 - 2m_p E$. Usando la conservazione del quadrimpulso totale, e l'invarianza relativistica del modulo quadro del quadrimpulso, possiamo scrivere $P_f'^2 = P_i^2$, per cui $E = 7m_p$. La corrispondente energia cinetica dal protone proiettile è quindi data da $T = E - m_p = 6m_p \sim 5.6\,\mathrm{GeV}$.

Esercizio 8: *Verificare l'equazione (69).*

7 Appendice

Equazioni di Maxwell

Svolgiamo esplicitamente l'esercizio n. 5 facendo uso delle unità di misura di Heaviside-Lorentz con $c = 1$ ($\epsilon_0 = \mu_0 = 1$). Ricordiamo che le equazioni di Maxwell possono essere scritte in forma vettoriale (che tiene conto delle proprietà di simmetria sotto rotazioni) come

$$\vec{\nabla} \cdot \vec{E} = \rho \ , \quad \vec{\nabla} \times \vec{B} - \partial_t \vec{E} = \vec{J}$$
$$\vec{\nabla} \cdot \vec{B} = 0 \ , \quad \vec{\nabla} \times \vec{E} + \partial_t \vec{B} = 0 \ .$$

Consideriamo ore le quattro equazioni di Maxwell con sorgenti, scritte in forma covariante,

$$\partial_\mu F^{\mu\nu} = -J^\nu$$

dove

$$\begin{aligned} F^{\mu\nu} &= \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \\ J^\mu &= (J^0, \vec{J}) = (\rho, \vec{J}) \end{aligned} \tag{71}$$

e valutiamo l'equazione per i diversi valori dell'indice libero $\nu = (0, i)$ con $i = (1, 2, 3)$.

Per $\nu = 0$ abbiamo

$$\begin{aligned} \partial_\mu F^{\mu 0} &= \partial_0 F^{00} + \partial_i F^{i0} = \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} = -(\partial_1 E_1 + \partial_2 E_2 + \partial_3 E_3) \\ &= -\vec{\nabla} \cdot \vec{E} = -J^0 = -\rho \ . \end{aligned}$$

Riconosciamo questa come l'equazione di Gauss

$$\vec{\nabla} \cdot \vec{E} = \rho \ .$$

Per $\nu = 1$ abbiamo

$$\begin{aligned} \partial_\mu F^{\mu 1} &= \partial_0 F^{01} + \partial_i F^{i1} = \partial_0 E_1 + \partial_2 F^{21} + \partial_3 F^{31} = \partial_t E_1 - \partial_2 B_3 + \partial_3 B_2 \\ &= (\partial_t \vec{E} - \vec{\nabla} \times \vec{B})_1 = -J_1 \end{aligned}$$

che corrisponde alla prima componente dell'equazione vettoriale

$$\vec{\nabla} \times \vec{B} - \partial_t \vec{E} = \vec{J} \ .$$

Un calcolo simile con $\nu = 2, 3$ genera le altre componenti di questa equazione vettoriale.

Analogamente si può procedere con le equazioni di Maxwell senza sorgenti

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \tag{72}$$

dove, abbassando gli indici con la metrica di Minkowski, si ha

$$F_{\mu\nu} = \eta_{\mu\lambda}\eta_{\nu\sigma}F^{\lambda\sigma} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{73}$$

cioè $F_{0i} = -F^{0i}$, $F_{ij} = F^{ij}$ (basta ricordare che abbassando un indice temporale si cambia segno, mentre non succede nulla abbassando un indice spaziale). Valutando l'eq. (72) con $(\mu, \nu, \lambda) = (1, 2, 3)$ si ha

$$\partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12} = \partial_1 B_1 + \partial_2 B_2 + \partial_3 B_3 = 0$$

ed otteniamo l'equazione

$$\vec{\nabla} \cdot \vec{B} = 0 \, .$$

Similmente con $(\mu, \nu, \lambda) = (0, 1, 2)$ si ha

$$\partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} = \partial_t B_3 + \partial_1 E_2 - \partial_2 E_1 = 0$$

che corrisponde alla terza componente dell'equazione vettoriale

$$\vec{\nabla} \times \vec{E} + \partial_t \vec{B} = 0$$

e così via. Le altre componenti contenute in (72) non dicono nulla di nuovo, il tensore è infatti completamente antisimmetrico e le altre componenti non sono indipendenti.

Infine, usando (71) ed (73), è facile valutare

$$F^{\mu\nu} F_{\mu\nu} = 2(\vec{B}^2 - \vec{E}^2) \tag{74}$$

che è manifestamente uno scalare perchè tutti gli indici sono sommati due a due nel modo corretto per formare un invariante.

Consideriamo ora la soluzione dell' esercizio n. 6. L'equazione (72) può essere interpretata come una condizione di integrabilità che ha come soluzione

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

Infatti sostituendo questa relazione in (72) si ottiene

$$\partial_\mu(\partial_\nu A_\lambda - \partial_\lambda A_\nu) + \partial_\nu(\partial_\lambda A_\mu - \partial_\mu A_\lambda) + \partial_\lambda(\partial_\mu A_\nu - \partial_\nu A_\mu) = 0$$

dove i termini si cancellano due a due perchè le derivate commutano. Identificando $A^\mu = (A^0, \vec{A}) = (\Phi, \vec{A})$, dove Φ è il potenziale scalare elettrico ed $\vec{A}$ il potenziale vettore, e quindi $A_\mu = (A_0, \vec{A}) = (-\Phi, \vec{A})$, possiamo verificare che

$$\begin{aligned} F_{i0} &= \partial_i A_0 - \partial_0 A_i = -\partial_i \Phi - \partial_t A_i = E_i \\ F_{ij} &= \partial_i A_j - \partial_j A_i = \epsilon_{ijk} B_k \end{aligned}$$

o, equivalentemente,

$$\begin{aligned} E_i &= F_{i0} \\ B_i &= \frac{1}{2}\epsilon_{ijk}F_{jk} \end{aligned}$$

in accordo con (73).

Simmetria di gauge

I potenziali A_μ introdotti sopra non sono univoci. Potremmo equivalentemente usare il quadripotenziale

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \Lambda(x) \tag{75}$$

dove $\Lambda(x)$ è una funzione arbitraria dello spaziotempo. Infatti si può verificare che i campi elettromagnetici rimangono invariati

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = \partial_\mu A_\nu + \partial_\mu \partial_\nu \Lambda - \partial_\nu A_\mu - \partial_\nu \partial_\mu \Lambda = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu} \ .$$

La trasformazione (75) è chiamata simmetria di gauge, ed i campi $F_{\mu\nu}$ sono invarianti di gauge. È una simmetria locale (le funzioni $\Lambda(x)$ dipendono in modo arbitrario dal punto dello spaziotempo) e sono riconducibili al gruppo $U(1)$, il gruppo delle fasi. Infatti possiamo scrivere la trasformazione (75) in modo equivalente come

$$A'_\mu(x) = A_\mu(x) - ie^{-i\Lambda(x)}\partial_\mu e^{i\Lambda(x)} \tag{76}$$

dove le fasi $e^{i\Lambda(x)} \in U(1)$ per ogni x. Si dice che l'elettromagnetismo è una teoria di gauge basata sul gruppo $U(1)$, ed A_μ è spesso chiamato potenziale di gauge.

La generalizzazione di questa teoria a gruppi più generali (in particolare a gruppi non abeliani) sta alla base della formulazione del Modello Standard delle particelle elementari, dove il gruppo di simmetria di gauge è dato dal prodotto di tre gruppi indipendenti, $SU(3) \times SU(2) \times U(1)$.

Azione di particella relativistica e forza di Lorentz

Abbiamo descritto il principio d'azione per una particella relativistica in (57). In particolare, abbiamo visto come l'azione debba essere proporzionale al tempo proprio, così da assicurare l'invarianza relativistica. Si può procedere anche in maniera manifestamente covariante (procedura che induce una simmetria locale associata al cambio di parametrizzazione della linea di mondo). Vediamone un po' i dettagli e sfruttiamo questa descrizione per introdurre l'accoppiamento di una particella carica al campo elettromagnetico e trovare l'espressione della forza di Lorentz.

Indichiamo con

$$d\tau = \frac{1}{c}\sqrt{-dx^\mu dx_\mu} \tag{77}$$

il tempo proprio infinitesimo della particella che percorre una distanza infinitesima dx^μ nello spaziotempo. Possiamo riscrivere l'azione (57) nella forma

$$S = -mc^2 \int d\tau = -mc \int \sqrt{-dx^\mu dx_\mu} = -mc \int d\lambda \sqrt{-\frac{dx^\mu}{d\lambda}\frac{dx_\mu}{d\lambda}} \tag{78}$$

dove λ è un parametro arbitrario che descrive la linea di mondo della particella. Può essere scelto a piacimento (invarianza locale o di gauge della particella relativistica[1]). Questa è l'azione manifestamente invariante a cui accennavamo sopra, con la sua simmetria di gauge. La scelta del parametro che abbiamo fatto precedentemente in (57) è usare $\lambda = t$, il tempo del sistema di riferimento scelto.

Scegliamo ora di usare lo stesso tempo proprio τ per parametrizzare la linea di mondo $x^\mu(\tau)$. Indicando con $\dot{x}^\mu = \frac{dx^\mu}{d\tau}$ riscriviamo l'azione come

$$S[x^\mu] = -mc \int d\tau \sqrt{-\dot{x}^\mu \dot{x}_\mu} \tag{79}$$

interpretata come funzionale delle quattro funzioni $x^\mu(\tau)$. L'accoppiamento al campo elettromagnetico è ottenuto usando il 4-potenziale $A^\mu = (A^0, \vec{A}) = (\Phi, \vec{A})$ ed introducendo il termine d'interazione nella lagrangiana

$$L_{int} = \frac{e}{c} A_\mu(x(\tau)) \dot{x}^\mu(\tau) \tag{80}$$

che è ovviamente uno scalare, e quindi consistente con l'invarianza di Lorentz. La costante e rappresenta la carica elettrica della particella. L'invarianza di Lorentz ci ha permesso con poco sforzo di trovare il termine di interazione corretto!

Abbiamo quindi trovato l'azione totale

$$S[x^\mu] = \int d\tau \left[-mc\sqrt{-\dot{x}^\mu \dot{x}_\mu} + \frac{e}{c} A_\mu(x) \dot{x}^\mu \right] \tag{81}$$

da cui seguono le equazioni del moto (imponendo $\delta S = 0$). Troviamo

$$\frac{dp^\mu}{d\tau} = \frac{e}{c} F^{\mu\nu} \dot{x}_\nu \tag{82}$$

dove

$$p^\mu = \frac{mc\dot{x}^\mu}{\sqrt{-\dot{x}^2}} = m\dot{x}^\mu$$

è il momento coniugato ad x^μ (la seconda espressione segue dalla definizione di tempo proprio). La parte spaziale della (82) contiene l'usuale forza di Lorentz. Se torniamo ad usare il parametro t piuttosto che τ, per confrontare con espressioni note da corsi precedenti, troviamo l'espressione della forza di Lorentz

$$\begin{aligned} \frac{dp^i}{dt} &= \frac{e}{c} F^{i0} \frac{dx_0}{dt} + \frac{e}{c} F^{ij} \frac{dx_j}{dt} \\ &= eE^i + \frac{e}{c} \epsilon^{ijk} B^k \frac{dx_j}{dt} = eE^i + \frac{e}{c} \epsilon^{ijk} \frac{dx_j}{dt} B^k \\ &= eE^i + \frac{e}{c} (\vec{v} \times \vec{B})^i \, . \end{aligned} \tag{83}$$

Formalismo hamiltoniano e sostituzione minimale

[1] Il cui studio dettagliato è molto interessante, utile anche allo studio delle stringhe relativistiche, che sono alla base delle moderne teorie di stringa.

È noto che nel formalismo hamiltoniano, l'accoppiamento di una particella carica ad un campo magnetico è ottenuta tramite la sostituzione minimale $\vec{p} \to \vec{p} - \frac{e}{c}\vec{A}(x)$ operata nell'hamiltoniana

$$H = \frac{\vec{p}^2}{2m} \quad \to \quad H = \frac{(\vec{p} - \frac{e}{c}\vec{A})^2}{2m} \ . \tag{84}$$

Questo può essere dedotto studiando la formulazione hamitoniana del formalismo lagrangiano. Si è soliti indicare

$$\pi_i = p_i - \frac{e}{c}A_i$$

chiamato momento covariante (covariante per simmetria di gauge $U(1)$).

La generalizzazione relativistica dell'accoppiamento minimale è evidentemente data da

$$p_\mu \quad \to \quad p_\mu - \frac{e}{c}A_\mu(x) \ . \tag{85}$$

Campi di gauge abeliani e non-abeliani

Descriviamo brevemente la generalizzazione dei campi elettromagnetici, che pensiamo come associati ad una teoria di gauge $U(1)$, a campi di gauge non-abeliani. In meccanica quantistica si vedrà come il momento $\vec{p}$ può essere associato ad un operatore differenziale

$$\vec{p} \to -i\hbar\vec{\nabla}$$

cosicché la componente $p_x \to -i\hbar\frac{\partial}{\partial x}$, etc. Per il momento pensiamo a questa relazione come ad una sostituzione formale, senza considerarne le implicazioni fisiche della meccanica quantistica.

Usiamo per semplicità unità di misura naturali, con $\hbar = c = 1$, e notiamo che dal momento covariante che emerge dalla sostituzione minimale si ottiene la cosiddetta derivata covariante $D_i = \partial_i - ieA_i$

$$p_i \to -i\partial_i \ , \qquad \pi_i = p_i - eA_i \to -i(\partial_i - ieA_i) = -iD_i \ . \tag{86}$$

Le derivate covarianti non commutano e danno origine al campo magnetico

$$[D_i, D_j] = -iF_{ij} \ . \tag{87}$$

Infatti il calcolo esplicito produce

$$\begin{aligned}
[D_i, D_j] &= [\partial_i - ieA_i(x), \partial_j - ieA_j(x)] = -ie[\partial_i, A_j(x)] - ie[A_i(x), \partial_j] \\
&= -ie([\partial_i, A_j(x)] - [\partial_j, A_i(x)]) = -ie(\partial_i A_j - \partial_j A_i) \\
&= -ieF_{ij} = -ie\epsilon_{ijk}B^k
\end{aligned} \tag{88}$$

Il campo magnetico è convenientemente ottenuto dal commutatore delle derivate covarianti. Questo è analogo al fatto che il momento covariante ha con se stesso parentesi di Poisson non nulle, e infatti queste parentesi sono proporzionali al campo magnetico

$$\begin{aligned}
\{\pi_i, \pi_j\} &= \{p_i - eA_i(x), p_j - eA_j(x)\} = -e\{p_i, A_j(x)\} - e\{A_i(x), p_j\} \\
&= e(\partial_i A_j - \partial_j A_i) = eF_{ij} \ .
\end{aligned} \tag{89}$$

Procedendo in modo simile, nella generalizzazione relativistica 4-dimensionale

$$p_\mu \to -i\partial_\mu \ , \qquad p_\mu - eA_\mu \to -i(\partial_\mu - ieA_\mu) = -iD_\mu \ . \tag{90}$$

abbiamo che il campo elettromagnetico $F_{\mu\nu}$ emerge come commutatore delle derivate covarianti $D_\mu = \partial_\mu - ieA_\mu$

$$[D_\mu, D_\nu] = -i(\partial_\mu A_\nu - \partial_\nu A_\mu) = -ieF_{\mu\nu} \ . \tag{91}$$

Abbiamo già descritto come il campo di gauge A_μ sia associato alla simmetria di gauge $U(1)$, vedi eq. (76),

$$A_\mu(x) \quad \to \quad A'_\mu(x) = A_\mu(x) - iU^{-1}(x)\partial_\mu U(x) \tag{92}$$

dove $U(x) = e^{i\Lambda(x)}$ è un elemento del gruppo $U(1)$ per ogni punto x^μ dello spaziotempo. Usando l'identità $U^{-1}\partial_\mu U = -(\partial_\mu U^{-1})U$, e il fatto che le fasi commutano, possiamo riscrivere questa trasformazione di gauge come

$$A_\mu(x) \quad \to \quad A'_\mu(x) = U(x)A_\mu(x)U^{-1}(x) + iU(x)\partial_\mu U^{-1}(x) \ . \tag{93}$$

Una proprietà caratterizzante le derivate covarianti è quella di produrre tensori da tensori, cioè preservare il carattere tensoriale quando agisce su tensori. Per esplicitare questa affermazione, iniziamo preliminarmente a semplificare la notazione assorbendo la costante e nella definizione di A_μ (o equivalentemente ponendo $e = 1$). Consideriano ora una funzione $\phi(x)$ che si trasformi nella rappresentazione fondamentale di $U(1)$, interpretato come gruppo di gauge,

$$\phi(x) \quad \to \quad \phi'(x) = e^{i\Lambda(x)}\phi(x) \tag{94}$$

Questo è un tensore (infatti un "vettore" della rappresentazione definente). La sua derivata non è un tensore, poichè si trasforma in modo diverso

$$\partial_\mu\phi(x) \quad \to \quad \partial_\mu\phi'(x) = e^{i\Lambda(x)}\partial_\mu\phi(x) + i(\partial_\mu\Lambda(x))e^{i\Lambda(x)}\partial_\mu\phi(x) \tag{95}$$

La sua derivata covariante invece è un tensore

$$D_\mu\phi \quad \to \quad D'_\mu\phi' = e^{i\Lambda}D_\mu\phi \tag{96}$$

come mostrato dal calcolo esplicito

$$D'_\mu\phi' = (\partial_\mu - iA'_\mu)\phi' = (\partial_\mu - iA_\mu - i\partial_\mu\Lambda)e^{i\Lambda}\phi = e^{i\Lambda}(\partial_\mu - iA_\mu)\phi = e^{i\Lambda}D_\mu\phi \ . \tag{97}$$

Più in generale, su un tensore di carica q

$$\phi_q(x) \quad \to \quad \phi'_q(x) = e^{iq\Lambda(x)}\phi(x) \tag{98}$$

agisce la derivata covariante

$$D_\mu = \partial_\mu - iA_\mu Q \tag{99}$$

dove Q è il generatore di $U(1)$, con algebra di Lie abeliana $[Q, Q] = 0$, nella rappresentazione del tensore su cui agisce (e quindi $Q = q$). Si può quindi verificare

$$D_\mu\phi_q \quad \to \quad D'_\mu\phi'_q = e^{iq\Lambda}D_\mu\phi_q \tag{100}$$

e quindi anche $D_\mu\phi_q$ è un tensore di carica q.

Come generalizzare questa struttura a gruppi non-abeliani?

Per vederlo, consideriamo l'estensione ad $SU(N)$ (o più in generale ad un gruppo di Lie compatto). La derivata covariante per definizione produce tensori quando applicata a tensori. Per definizione, un tensore si trasforma in una rappresentazione del gruppo di gauge $SU(N)$. La derivata covariante è definita da

$$D_\mu = \partial_\mu + W_\mu(x) \tag{101}$$

dove $W_\mu(x)$ è il "potenziale di gauge", che opera trasformazioni infinitesime associate al gruppo (per compensare un termine non voluto nella trasformazione della derivata semplice), e che quindi assume valori nell'algebra di Lie del gruppo, ed è formato dalle matrici dell'algebra di Lie nella rappresentazione del tensore su cui agisce. Sviluppiamo il potenziale $W_\mu(x)$ in termini dei generatori T^a dell'algebra di Lie come segue

$$W_\mu(x) \equiv -iW_\mu^a(x)T^a \ . \tag{102}$$

Questa relazione definisce i campi $W_\mu^a(x)$, funzioni dello spazio tempo per ogni valore degli indici a e μ fissati. Dalla richiesta di covarianza

$$\begin{array}{rcl} \psi(x) & \rightarrow & \psi'(x) = U(x)\psi(x) \\ D_\mu\psi(x) & \rightarrow & D'_\mu\psi'(x) = U(x)D_\mu\psi(x) \end{array} \tag{103}$$

si ottene la legge di trasformazione dei potenziali di gauge

$$W_\mu(x) \ \rightarrow \ W_\mu(x)' = U(x)W_\mu(x)U^{-1}(x) + U(x)\partial_\mu U^{-1}(x) \ . \tag{104}$$

Infatti, imponendo che $D'_\mu\psi' = UD_\mu\psi$, possiamo calcolare

$$\begin{aligned} D'_\mu\psi' &= (\partial_\mu + W'_\mu)\psi' = U(\partial_\mu + W_\mu)\psi \\ &= U(\partial_\mu + W_\mu)U^{-1}U\psi \\ &= U(\partial_\mu + W_\mu)U^{-1}\psi' = \partial_\mu\psi' + U[(\partial_\mu + W_\mu)U^{-1}]\psi' \end{aligned} \tag{105}$$

da cui segue la legge di trasformazione riportata sopra.

Le derivate covarianti in generale non commutano. Questo permette di definire il tensore di "curvatura" o tensore campo di forza ("field strength") nel seguente modo

$$[D_\mu, D_\nu] = F_{\mu\nu} \tag{106}$$

da cui

$$F_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + [W_\mu, W_\nu] \tag{107}$$

dove compare un termine quadratico dovuto ai commutatori del gruppo non abeliano (che si annulla per un gruppo abeliano come $U(1)$). È facile vedere che il tensore campo di forza si trasforma in modo covariante (nella cosiddetta rappresentazione aggiunta)

$$F_{\mu\nu} \ \rightarrow \ F'_{\mu\nu} = UF_{\mu\nu}U^{-1} \tag{108}$$

che segue dalla covarianza della (106). Esplicitando la (107), in termini delle funzioni componenti $W_\mu(x) = -iW_\mu^a(x)T^a$ e $F_{\mu\nu}(x) = -iF_{\mu\nu}^a(x)T^a$ trovando

$$F_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + f^{abc}W_\mu^b W_\nu^c \tag{109}$$

in cui compaiono esplicitamente le costanti di struttura del gruppo non abeliano.

In applicazioni al modello standard, si usa as esempio il gruppo $SU(3)$ per la teoria di gauge associata alle interazioni forti, il gruppo $SU(3)$ ha 8 generatori, ed i campi W_μ^a sono associati ai gluoni (i gluoni sono i quanti del campo d'onda W_μ^a, proprio come i fotoni sono interpretati come i quanti del campo A_μ dell'elettromagnetismo). Il tensore $F_{\mu\nu}^a$ costituisce il campo cromodinamico, che generalizza il tensore di Faraday del campo elettromagnetico alle teorie non abeliane.

Questi accenni preliminari sono serviti per illustrare l'uso cruciale che si fa della teoria dei gruppi nella descrizione delle forze nucleari e subnucleari.

Concludiamo con la scrittura estesa, con tutti gli indici espliciti, della derivata covariante di $SU(3)$ che agisce su un campo d'onda che si trasforma nella rappresentazione fondamentale **3** del gruppo. Denotiamo quindi il campo d'onda (che possiamo pensare associato ad un quark che ha tre "colori") con $q^\alpha(x)$ con $\alpha = 1, 2, 3$, oppure $\alpha = rosso,\ verde,\ blu$. La derivata covariante si esplicita come

$$D_\mu q(x) \quad \to \quad \partial_\mu q^\alpha(x) - iW_\mu^a(x)(T_F^a)^\alpha{}_\beta q^\beta(x)$$

dove T_F^a sono i generatori nella rappresentazione fondamentale **3** del gruppo, e sono quindi matrici 3×3, che naturalmente soddisfano l'agebra di Lie del gruppo, $[T_F^a, T_F^b] = if^{abc}T_F^c$, con f^{abc} le costanti di struttura di $SU(3)$. La loro espressione esplicita si ha tramite le matrici di Gell-Mann, $T_F^a = \frac{1}{2}\lambda^a$. I campi W_μ^a hanno un'indice $a = 1, ..., 8$, che indica che il suo colore appartiene alla **8**, la rappresentazione aggiunta di $SU(3)$, per cui si dice che i gluoni hanno otto colori (otto combinazioni diverse di colore, che si mescolano tra di loro con le trasformazioni di $SU(3)$ nella rappresentazione **8**).