Appunti di
Relatività Ristretta,
Meccanica Quantistica Relativistica
e
Relatività Generale

autore: Valeriano Barassi
(docente: *Roberto Pettorino*)

Anno Accademico 1992-93

"Non hai veramente capito qualcosa fino a quando non sei in grado di spiegarlo a tua nonna.".

– Albert Einstein

Indice

Note

Questo scritto è costituito dagli appunti raccolti durante il corso di Relatività tenuto dal prof. Roberto Pettorino *nell'Anno Accademico 1992-93 presso la Facoltà di Fisica dell'Università di Napoli "Federico II". Altri testi di riferimento per la parti descrittive sono il libro* "Filosofia della fisica", G. Boniolo, Bruno Mondadori Editore, 1997 *e alcuni estratti di Wikipedia.*

Queste note devono essere considerate piuttosto come una traccia per lo studio e certamente non una trattazione esaustiva dell'argomento. Diversi aspetti concettuali importanti della teoria della relatività sono però stati adeguatamente messi in evidenza. È consigliata una conoscenza preliminare della meccanica lagrangiana e hamiltoniana e della meccanica quantistica per la parte seconda.

Il corso è suddiviso in tre parti: Teoria della Relatività Ristretta, Meccanica Quantistica Relativistica, Teoria della Relatività Generale. Gli argomenti trattati nelle varie parti sono:

Teoria della Relatività Ristretta: *In questa parte affronteremo le trasformazioni di Lorentz, che costituiscono la base della teoria della relatività ristretta, sia dal punto di vista geometrico che a partire dall'invarianza del quadrintervallo. Una volta introdotto il concetto di quadrivettore e spaziotempo di Minkowski, passeremo a riformulare la meccanica in questo formalismo, utilizzando il principio variazionale sia per una particella libera che in interazione con il campo elettromagnetico. Questo ci condurrà da un lato ad estendere il formalismo ai sistemi continui, dall'altro a riformulare lo stesso elettromagnetismo classico in termini covarianti ed a scoprire che le equazioni di Maxwell sono le equazioni del moto del campo elettromagnetico. Indagheremo poi le conseguenze della richiesta di invarianza per traslazioni e rotazioni delle equazioni covarianti. Il capitolo conclusivo di questa parte è dedicato infine alla cinematica relativistica.*

Meccanica Quantistica Relativistica: *In questa parte affrontiamo la problematica della formulazione covariante della meccanica quantistica. I problemi che sorgono da una estensione troppo "intuitiva" della meccanica quantistica classica sono risolti dall'equazione di Dirac, che qui deriveremo dettagliatamente e di cui indagheremo le conseguenze. Tratteremo brevemente la teoria di Dirac dell'elettrone. Infine, considereremo le implicazioni sulla teoria dell'elettromagnetismo, mostrando come le equazioni di Maxwell siano in effetti le equazioni d'onda quantistiche del fotone.*

Teoria della Relatività Generale: *In questa parte affronteremo invece la teoria della Relatività Generale. Cominceremo con l'affrontare la base concettuale della teoria, il cosiddetto principio di equivalenza. Dopo aver generalizzato opportunamente i concetti matematici in modo da poterli utilizzare nell'ambito della geometria curva non euclidea, tratteremo il moto del punto materiale ed estenderemo la formulazione dell'elettromagnetismo nel campo gravitazionale tramite il metodo variazionale. Infine, sempre applicando il metodo variazionale, ricaveremo le cosiddette Equazioni di Einstein, che esprimono il moto della materia nel campo gravitazionale.*

Questo lavoro è distribuito liberamente nella speranza che possa essere utile nella preparazione degli esami, con la sola condizione che si citi il nome dell'autore e la fonte.

Parte I

Teoria della Relatività Ristretta

Capitolo 1

Principi generali

1.1 Fondamenti

Il principio di inerzia della meccanica classica presenta diverse astrazioni:

1. Un corpo lontano dall'azione di tutti gli altri corpi (corpo isolato). È un concetto nuovo.

2. L'esistenza di un sistema di riferimento assoluto.

È evidente che in un sistema di riferimento in moto – ad esempio circolare – una traiettoria dritta appare curvilinea: si imputa questo alla presenza di forze fittizie, nel caso specifico alla forza centrifuga.

Deve tuttavia valere il **Principio di relatività Galileiana**: *in sistemi di riferimento in moto uniforme l'uno rispetto all'altro la fisica deve essere la stessa*. Questo conduce ad un interessante problema: due osservatori determinano due diverse posizioni per il medesimo oggetto mobile che si trova in una certa posizione in un certo istante. Per poter correlare le due determinazioni, queste devono venire eseguite nel medesimo istante. I due osservatori si devono quindi scambiare un segnale per accordarsi quando fare la misura e il segnale deve trasmettersi istantaneamente.

Se il segnale si trasmette con velocità finita e conosciuta, i due osservatori, prima di allontanarsi l'uno dall'altro, possono sincronizzare i loro orologi, ma allora si deve supporre che il movimento degli orologi non alteri il sincronismo, cosa che si può verificare scambiando dei segnali, ma allora siamo daccapo.

Già Galilei comprese il problema e fece un tentativo per misurare la velocità della luce, che trovò rapidissima, per cui decise che si poteva trascurare il tempo di propagazione del segnale.

In questo consiste l'approssimazione della relatività di Galilei, validissima peraltro in situazioni comuni, ed è appunto la critica del concetto di sincronia che ha portato alla relatività di Einstein.

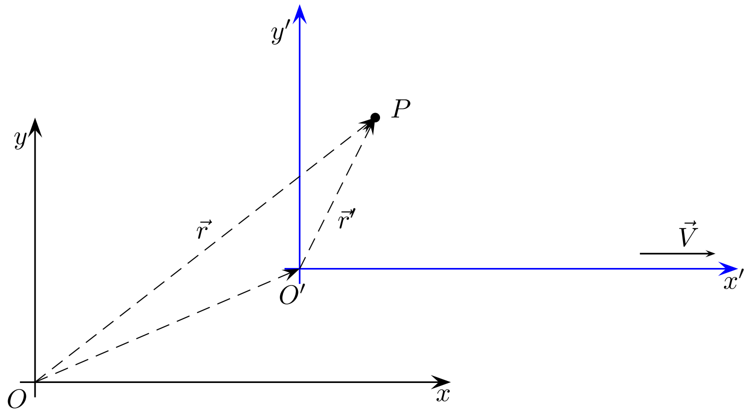
La relatività galileiana assume quindi che la velocità di propagazione dell'interazione sia infinita: questo è sempre assunto implicitamente in meccanica classica quando, ad esempio, si definiscono dei potenziali dipendenti dal punto.

Il principio di relatività galileiana implica quindi che non esista un sistema di riferimento privilegiato: mentre si perde quindi la nozione di isolocalità, resta ancora il concetto di un tempo assoluto e uniforme.

Galilei mise a punto delle regole di trasformazione dette *trasformazioni galileiane* che permettevano di spostare le analisi da un osservatore ad un altro.

Quando effettuano le loro misure, i due osservatori sono lontani l'uno dall'altro e, siccome le loro osservazioni devono essere fatte nello stesso istante, devono scambiarsi dei segnali. Come detto, Galilei era perfettamente conscio di tale problema tanto che nel 1667 provò a misurare la velocità della luce fra due osservatori che facevano segnali con una lanterna. Ne dedusse che la velocità della luce è rapidissima e archiviò la questione come irrilevante ai fini pratici. Si noti come questo implichi che la quarta coordinata, il tempo, sia la stessa in entrambi i sistemi inerziali: altrimenti detto, nell'ambito della meccanica classica tutti gli orologi marciano con lo stesso ritmo e di conseguenza gli intervalli temporali fra due eventi successivi saranno gli stessi per entrambi gli osservatori. Benché questa ipotesi sia naturale e logica, vedremo fra breve che risulta essere non corretta quando si ha a che fare con situazioni in cui la velocità è confrontabile con la velocità della luce.

Siano dati due sistemi di riferimento in moto relativo uniforme:



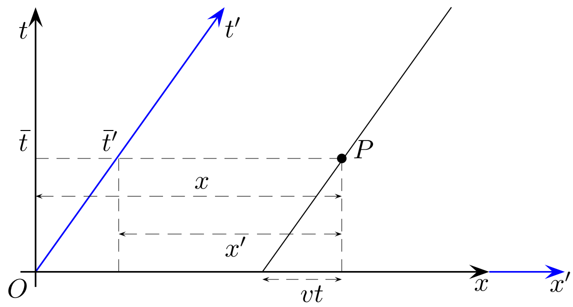
Le trasformazioni di Galileo per sistemi in moto relativo uniforme a velocità $\vec{V}$ risultano le seguenti:

$$\begin{cases} \vec{r} &= \vec{r}' + \overrightarrow{OO'} \\ \vec{v} &= \vec{v}' + \vec{V} \\ \vec{a} &= \vec{a}' + \vec{A} \rightarrow \vec{a} = \vec{a}' \end{cases}$$

da cui $\vec{r}' = \vec{r} - \overrightarrow{OO'} = \vec{r} - \vec{V}t$ e se si considera il moto lungo x ($y' = y$, $z' = z$):

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

Questo viene rappresentato tradizionalmente in un diagramma $x - t$, i cosiddetti diagrammi orari:



dall'interpretazione geometrica che ne dà il diagramma, discende immediatamente che $x' = x - vt$. In relatività Galileiana quindi il passaggio da un sistema di riferimento all'altro inclina l'asse t ma lascia l'asse x inalterato. I sistemi galileiani sono dunque rappresentati dai sistemi ad asse x fisso e t qualsiasi. Un punto P in questo piano è detto *evento*.

Notiamo che a causa dell'universalità del tempo nell'approccio galileiano, lungo un asse parallelo a x deve risultare lo stesso tempo t, pertanto è l'unità di misura deve variare.

Si sa attualmente – grazie all'esperienza di Michelson–Morley – che la velocità di propagazione dei segnali elettromagnetici, quindi anche della luce, non è infinita: in particolare vale esattamente $c = 299\,792\,458\, m/s$([1]) ed è costante in ogni sistema di riferimento. Questo elimina il ruolo assoluto che aveva il tempo nelle trasformazioni galileiane.

[1] A causa delle precedenti definizioni di metro, mano a mano che le misure si facevano più precise il valore della velocità della luce variava. La Conferenza Generale dei Pesi e Misure nel 1983 ha deciso pertanto di definire il metro come la distanza percorsa dalla luce *esattamente* in $1/299\,792\,458\ s$, soppiantando la vecchia definizione del 1960, ovvero $1\,650\,763.73$ volte la lunghezza d'onda della transizione fra i livelli $2p^{10}$ e $5d^5$ del ^{86}Kr. Questa definizione è ritenuta la più soddisfacente, perché basata sull'universalità e indipendenza di c.

L'esperienza di Michelson–Morley

La fisica nel XIX secolo postulava che le onde (luminose, sonore, etc.) dovessero avere un mezzo che consentiva la loro propagazione. Nel caso della luce si era ipotizzata l'esistenza di un "etere luminifero" come mezzo di propagazione. Albert A. Michelson decise di provare a misurare la velocità della luce per vedere se si trovava traccia del vento d'etere, usando a tale scopo uno strumento da lui stesso ideato che successivamente prese il nome di interferometro di Michelson.

Tuttavia il suo apparecchio prototipale non aveva la precisione sufficiente per escludere con certezza l'esistenza del movimento nell'etere. Per questo decise di effettuare esperimenti più precisi e, nel 1887, si mise in contatto con Edward Morley, che offrì il suo seminterrato per il nuovo esperimento. A tale scopo venne utilizzato un interferometro montato su una lastra di pietra quadrata di 15 cm di lato e circa 5 cm di spessore. Per eliminare le vibrazioni la lastra veniva fatta galleggiare su mercurio liquido, accorgimento che permetteva inoltre di mantenere la lastra orizzontale e di farla girare attorno ad un perno centrale. Un sistema di specchi inviava il raggio di luce per un percorso di otto viaggi di andata e ritorno allo scopo di rendere il viaggio del raggio di luce più lungo possibile.

Se la velocità di propagazione della luce nei due bracci dell'interferometro è diversa a causa del vento dell'etere, i due fasci di luce impiegano un tempo diverso per tornare a incontrarsi al centro e quindi le oscillazioni nei due fasci presentano una differenza di fase. Questo provoca la formazione di frange chiare e scure, che dovrebbero spostarsi variando l'orientamento dello strumento rispetto al vento dell'etere. La differenza attesa nei tempi impiegati dalla luce per percorrere i bracci dell'interferometro parallelo e perpendicolare al vento dell'etere si calcola facilmente.

Anche con il nuovo esperimento non si trovò traccia di un vento d'etere, la velocità della luce risultava indipendente dalla direzione e di poco inferiore a $300\,000\ km/s$. La cosa non accadde neanche ripetendo l'esperimento a distanza di tempo e di luogo.

Con questi esperimenti si dimostra il fallimento della legge di composizione galileiana delle velocità nel caso della luce, appunto perché la luce non viene trascinata da nessun mezzo fisico. Due spiegazioni sono possibili al fallimento dell'esperienza di Michelson–Morley:

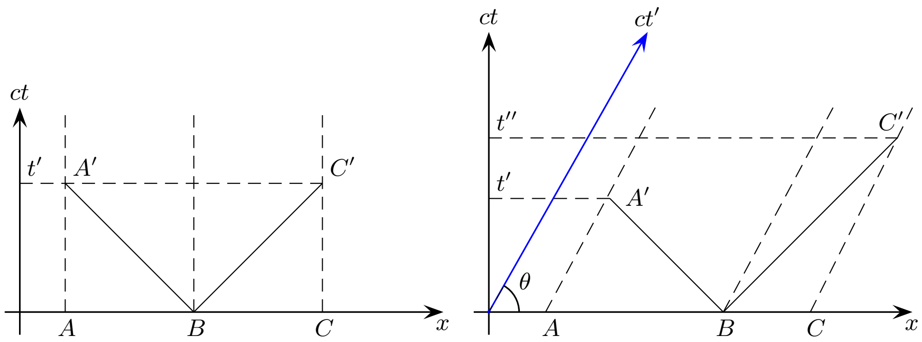
- il braccio dell'interferometro nella direzione del moto dell'etere si accorcia (contrazione di Fitzgerald), in modo da compensare la differenza nei tempi di percorrenza,
- la velocità della luce è la medesima in tutte le direzioni

Einstein accetta la seconda soluzione, che va considerata come una riprova dell'isotropia dello spazio per tutti gli osservatori. La spiegazione di tale risultato secondo Einstein è che non vi è nessun etere.

La conclusione, che *la velocità della luce è indipendente dal moto della sorgente e dell'osservatore*, fu l'ipotesi da cui partì Einstein per sviluppare la teoria della relatività ristretta.[2]

[2]Come abbiamo evidenziato nel testo, era possibile dare due interpretazioni del fallimento dell'esperimento di Michelson–Morley. Si può dimostrare che le due affermazioni sono in realtà matematicamente equivalenti: infatti, l'ipotesi che il braccio dell'interferometro si contrae nella direzione del moto porta alle leggi di trasformazione di Lorentz, le quali implicano a loro volta una velocità di propagazione della luce costante. D'altra parte, l'ipotesi della velocità della luce costante implica che la legge di trasformazione fra i sistemi di riferimento sia quella Lorentz e non quella di Galilei.
Storicamente, il problema essenziale era da un lato quello di dimostrare l'esistenza o meno dell'etere, dall'altro quello di comprendere perché nelle leggi di Maxwell – che sono svincolate da qualunque sistema di riferimento – compariva la velocità di propagazione della luce. Non si capiva, dalle equazioni di Maxwell, come la velocità di propagazione delle onde elettromagnetiche potesse essere indipendente dal sistema di riferimento, in modo che le leggi dell'elettromagnetismo rimanessero invariate nel passaggio fra sistemi di riferimento inerziali. Einstein era già al corrente di alcuni esperimenti che parevano dimostrare che la velocità della luce fosse effettivamente costante e fece il grande balzo concettuale di *postulare* che la velocità della luce fosse effettivamente la stessa indipendentemente dal moto della sorgente o dell'osservatore. Einstein notò che la velocità è semplicemente una misura di spazio diviso una misura di tempo. Per giustificare la costanza della velocità della luce, che va contro il senso comune espresso dalla composizione galileiana delle velocità, Einstein dedusse che per un osservatore in moto la misura dello spazio e del tempo dovevano modificarsi rispetto ad un osservatore in quiete, in un modo particolare tale da preservare il loro rapporto.
In altri termini, un osservatore in quiete vede un raggio luminoso viaggiare a velocità c. Un osservatore che sia in moto a velocità $0.5\,c$ che insegue un raggio di luce lungo la stessa direzione, non lo vede allontanarsi da lui a $c/2$, quando cercherà infatti di misurare la velocità della luce facendo una misura di distanza e di tempo, le sue misure risultaranno deformate (rispetto all'osservatore in quiete) in maniera tale che il loro rapporto sia sempre c.
Quindi, secondo Einstein, l'esperimento di Michelson–Morley *doveva* fallire perché la velocità della luce non si può comporre secondo le leggi di galileiane. Lo sviluppo di questa posizione ha dato origine alla teoria della relatività ristretta. Seguendo l'approccio storicamente seguito da Einstein, mostreremo quindi come da questa ipotesi discendano le trasformazioni di Lorentz e da questa la contrazione delle lunghezze.

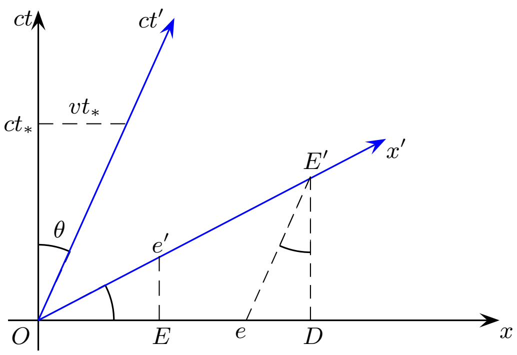
Si supponga che da un punto B in un sistema di riferimento fermo partano due segnali e che raggiungano i punti A e C contemporaneamente. In un sistema in moto rispetto al primo questi segnali *non potranno arrivare simultaneamente a causa della velocità finita di propagazione*. In termini di diagrammi spazio-tempo,[3] questo risulta evidente:



dove $\tan\theta = c/v$. Risulta chiaro che nel secondo diagramma – quello del sistema in moto – i due eventi non possono essere simultanei. Perché siano tali si deve ruotare l'asse x in modo che sia parallelo al segmento $\overline{A'C'}$: in pratica, la trasformazione deve coinvolgere sia x che t.

Vediamo ora come sia possibile ricavare le leggi di trasformazione fra questi due sistemi (leggi di Lorentz) per via geometrica in base alle richieste precedenti.

Il sistema in moto non è più un sistema ortogonale, tuttavia la richiesta che l'asse x' sia parallelo al segmento $\overline{A'C'}$ implica che anche nel nuovo sistema i raggi di luce devono essere delle bisettrici. Le unità di misura devono essere diverse, mentre i due angoli $\widehat{x\,x'}$ e $\widehat{ct\,ct'}$ sono uguali.



In questa figura, $\overline{OE}$ rappresenta l'unità di misura su x, nel sistema S; e' è l'unità di misura di S vista da S'; e è l'unità di misura $\overline{OE'}$ di S' vista nel sistema S. Inoltre qui $\tan\theta = v/c$.

La fisica vista nei due sistemi deve essere la stessa, questo implica che i rapporti delle unità di misura deve soddisfare:

$$\boxed{\frac{e}{E} = \frac{e'}{E'}} \tag{1.1}$$

Il triangolo $\overset{\triangle}{Oe'E}$ è evidentemente rettangolo. Questo permette di scrivere:[4]

$$e'^2 = E^2 + E^2\tan^2\theta = E^2 + E^2\frac{v^2}{c^2}$$

ovvero:

$$\boxed{e'^2 = E^2\left(1 + \frac{v^2}{c^2}\right)} \tag{1.2}$$

[3] A differenza dei diagrammi orari qui sulle ordinate è presente la quantità ct. In questo modo un raggio di luce è rappresentato sempre dalle bisettrici.

[4] Si sfrutta banalmente il teorema di Pitagora, unito al fatto che in un triangolo rettangolo la misura di un cateto è uguale a quella dell'altro cateto per la tangente dell'angolo opposto al primo. In questo caso $\overline{e'E} = E\tan\theta$.

Per quello che riguarda il segmento $\overline{Oe}$:[5]

$$e = D - \overline{eD} = D - E'D\tan\theta = D - \overline{E'D}\frac{v}{c} = D - D\frac{v^2}{c^2} = D\left(1 - \frac{v^2}{c^2}\right)$$

Per il segmento $\overline{OE'}$, utilizzando sempre la stessa relazione geometrica fra i cateti e la tangente:

$$E'^2 = D^2 + D^2\frac{v^2}{c^2} = D^2\left(1 + \frac{v^2}{c^2}\right)$$

sostituendo in questa relazione la $D = e/\left(1 - \frac{v^2}{c^2}\right)$ appena ricavata:

$$\boxed{E'^2 = \frac{\left(1 + \frac{v^2}{c^2}\right)}{\left(1 - \frac{v^2}{c^2}\right)^2}e^2} \tag{1.3}$$

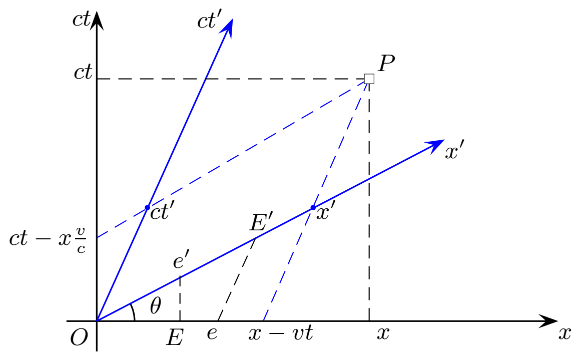
Avendo finalmente ottenuto e' in funzione di E – eq. (1.2) – e E' in funzione di e – eq. (1.3) –, si può imporre la (1.1), calcolando innanzitutto il rapporto fra le due:

$$\frac{e'^2}{E'^2} = \frac{E^2\left(1 + \frac{v^2}{c^2}\right)}{e^2\left(1 + \frac{v^2}{c^2}\right)}\left(1 - \frac{v^2}{c^2}\right)^2 \Rightarrow \frac{E^2}{e^2}\left(1 - \frac{v^2}{c^2}\right)^2 = \frac{e'^2}{E'^2}$$

dal che si ricava immediatamente la *legge con cui devono trasformare le unità di misura*:

$$\boxed{e = E\left(1 - \frac{v^2}{c^2}\right)} \tag{1.4}$$

Sulla base di quanto ricavato sulle unità di misura, cerchiamo ora la forma della trasformazione di coordinate fra due sistemi di riferimento tale da preservare la (1.1).



In questo diagramma, E' è l'unità di misura nel sistema non ortogonale S' e E quella in S, per cui risulta:

$$\frac{\overline{Ox'}}{E'} = x' \qquad \frac{\overline{O(x-vt)}}{E} = (x - vt)$$

da cui:

$$\frac{x'}{x - vt} = \frac{\overline{Ox'}}{\overline{O(x-vt)}}\frac{E}{E'}$$

Siccome i due triangoli $\overset{\triangle}{Ox'(x-vt)}$ e $\overset{\triangle}{OE'e}$ sono proporzionali, risulta:

$$\frac{\overline{Ox'}}{\overline{O(x-vt)}} = \frac{E'}{e}$$

[5] Qui $\overline{eD} = \overline{E'D}\tan\theta$ e $\overline{E'D} = D\tan\theta$.

che sostituita nella precedente permette di scrivere:

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

dove nell'ultimo passaggio abbiamo utilizzato la (1.4). Ne segue allora:

$$\boxed{x' = \frac{1}{\sqrt{1 - \beta^2}}(x - vt) \qquad \beta \equiv \frac{v}{c}}$$

Questa rappresenta l'essenza della trasformazione di Lorentz. È importante notare come in questa formula $v = c$ rappresenti un punto di discontinuità di seconda specie, in quanto il limite per $v \to c$ risulta infinito.[6]

1.2 Quadrintervallo. Trasformazioni di Lorentz

Ricaveremo ora le trasformazioni di Lorentz in maniera completa utilizzando un approccio differente.

Iniziamo con il definire il quadrintervallo (distanza in 4 dimensioni).

Siano dati due punti nello spazio quadridimensionale (eventi) $P_1 = [x_1, y_1, z_1, t_1]$ e $P_2 = [x_2, y_2, z_2, t_2]$. Nella geometria tridimensionale euclidea, la distanza spaziale fra P_1 e P_2 è data da:

$$\overline{P_1P_2}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$$

mentre in termini di coordinata temporale la distanza percorsa da P_1 a P_2 è $v(t_1 - t_2)$. Per un raggio di luce questa distanza è quindi data da $c(t_1 - t_2)$.

Definiamo *quadrintervallo* la quantità:

$$s_{12}^2 = c^2(t_2 - t_1)^2 - (\vec{r_2} - \vec{r_1})^2$$

risulta immediato vedere che per un raggio di luce il quadrintervallo è sempre nullo per definizione:

$$c^2(t_2 - t_1)^2 = (\vec{r_2} - \vec{r_1})^2 \quad \Rightarrow \quad s_{12}^2 = 0$$

notiamo che a causa del segno meno il quadrintervallo così definito non è necessariamente positivo. Per infinitesimi risulta:

$$ds^2 = c^2dt^2 - d\vec{r}^2$$

L'ipotesi relativistica di invarianza della velocità della luce nei diversi sistemi di riferimento implica che il quadrintervallo di un raggio di luce sia sempre nullo in *ogni* sistema di riferimento, per cui gli infinitesimi – essendo dello stesso ordine – devono essere proporzionali per un fattore che al più dipende dal modulo della velocità relativa. Non può infatti dipendere né dalla posizione, né dal tempo perché tutti i punti dello spazio-tempo si assumono indistinguibili. Questo fattore non può neanche dipendere dalla direzione della velocità, perché si assume l'isotropia dello spazio. La relazione fra i quadrintervalli in tre sistemi S, S' e S'' pertanto deve essere:

$$ds^2 = a(v_1)ds'^2 \qquad ds^2 = a(v_2)ds''^2$$

ne consegue che l'analoga relazione fra i sistemi S' e S'' deve essere:

$$ds'^2 = a(|\vec{v_1} - \vec{v_2}|)ds''^2 \Rightarrow \frac{ds''^2}{ds'^2} = \frac{a(v_2)}{a(v_1)} = \frac{1}{a(|\vec{v_1} - \vec{v_2}|)} \equiv A$$

[6] Per pignoleria, il punto $v = c$ non appartiene al dominio di definizione della funzione rappresentata dalla trasformazione di Lorentz, ne rappresenta invece un punto di accumulazione. Da un punto di vista matematico formale, non si dovrebbe quindi parlare di discontinuità. Quello che è importante dal punto di vista fisico, tuttavia, è che $v = c$ non è un valore lecito e che il fattore di Lorentz β diverge per $v \to c$.

L'indistinguibilità dei sistemi di riferimento (che proviene dal principio di relatività Galileiana) implica che A sia effettivamente una costante e uguale a 1 per tutti i sistemi di riferimento. Infatti, per i due sistemi S' e S'' si può scrivere:

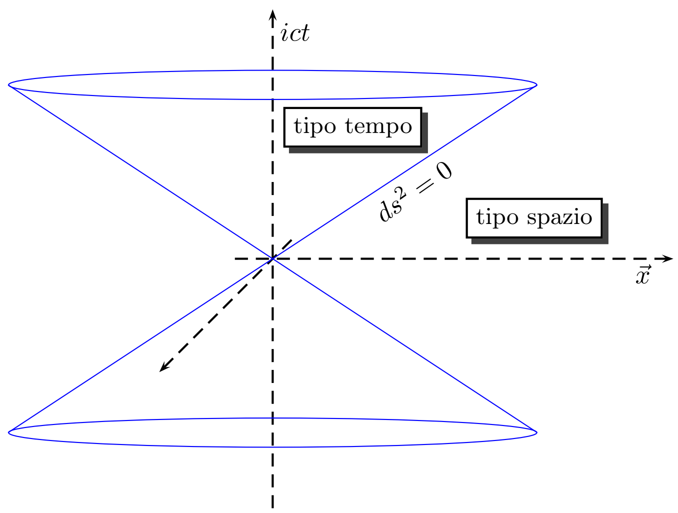
$$\begin{cases} ds'^2 = A'' ds''^2 \\ ds''^2 = A' ds'^2 \end{cases} \Rightarrow A' = \frac{1}{A''} \Rightarrow A' = A'' = 1$$

Da questo consegue che *il quadrintervallo è invariante*:

$$ds^2 = ds'^2 \tag{1.5}$$

In maniera analoga il quadrintervallo non infinitesimo è anch'esso un invariante.

Abbiamo già notato come il quadrintervallo non sia definito positivo. Studiamone quindi ora il significato con riferimento alla figura, in cui le tre dimensioni spaziali sono "appiattite" in un piano – per esigenze di disegno – e la terza rappresenta la coordinata temporale:



Il luogo geometrico costituito dai punti per cui $ds^2 = 0$, costituito dai raggi di luce, è un cono le cui falde sono a 45°. In questo caso si parla di quadrintervalli di *tipo luce*. Nella regione all'interno del cono risulta $x < ct$, pertanto $(c\Delta t)^2 - \Delta x^2 > 0$. In questa regione si può trovare al massimo un sistema di riferimento in cui $\Delta x = 0$, ma deve allora restare $c\Delta t > 0$. Questi punti rappresentano eventi avvenuti nello stesso luogo, ma in tempi diversi: può esistere quindi connessione causale. Si parla in questo caso di quadrintervallo di *tipo tempo*. Nella zona all'esterno del cono invece $(c\Delta t)^2 - \Delta x^2 < 0$, cioé si può trovare un sistema di riferimento in cui $c\Delta t = 0$ e quindi $\Delta x^2 < 0$. Questi sono avvenimenti che avvengono nello stesso istante, ma in posti diversi.[7] In questa regione la distanza spaziale è superiore a ct, pertanto nessuna informazione può viaggiare fra questi punti: si parla di quadrintervallo di *tipo spazio*. In questa regione non è possibile stabilire un ordinamento temporale ed un nesso di causalità, l'istante dipende infatti completamente dal sistema di riferimento.

Nel formalismo quadridimensionale questo diagramma è chiamato *cono di luce*.

Introduciamo ora l'importante nozione di *tempo proprio*. Si consideri una particella e si ipotizzi un sistema di riferimento S' centrato sulla particella istante per istante, cioé un sistema che si muova in maniera solidale alla particella ed alla sua velocità. Il *tempo proprio* è il tempo misurato dalla particella nel sistema di riferimento solidale S'. Allora, dall'invarianza del quadrintervallo:

$$ds'^2 = ds^2 \Rightarrow c^2 dt'^2 = c^2 dt^2 - d\vec{r}^2 \Rightarrow dt'^2 = dt^2 \left[1 - \frac{d\vec{r}^2}{c^2 dt^2}\right]$$

e cioé:

$$\boxed{dt' = dt\sqrt{1 - \beta^2}}$$

[7] In maniera suggestiva, questa regione viene talvolta chiamata *altrove*.

e quindi l'incremento temporale del sistema in moto solidale alla particella è inferiore rispetto a quello di un osservatore nel sistema fisso che guarda il moto. È il cosiddetto fenomeno della *dilatazione dei tempi*: il tempo appare scorrere più lentamente per un osservatore in moto rispetto ad uno fisso.

Il fenomeno è *simmetrico* nel senso che rispetto all'osservatore sul sistema in moto, è l'altro sistema che invece si muove di moto relativo con velocità $-\vec{v}$ e che quindi riscontra una dilatazione temporale, numericamente uguale a quella del primo osservatore; ma lo stesso fenomeno *non è invertibile*, perché per poter confrontare il risultato delle due misurazioni temporali si richiede in linea di principio che un osservatore torni indietro per incontrarsi con l'altro. Ma questo comporterebbe delle accelerazioni e pertanto il sistema considerato non sarebbe più inerziale: questo è sufficiente per rompere la simmetria e mostrare, in questo caso, che solo l'osservatore accelerato ha subito il fenomeno della dilatazione temporale. Questa è l'essenza del cosiddetto *paradosso dei gemelli* che sarà studiato in dettaglio nel paragrafo §1.3.

Introducendo la variabile complessa $\tau = ict$, il quadrintervallo assume la forma:

$$ds^2 = -\left(d\tau^2 + d\vec{r}^2\right)$$

che formalmente concide con una distanza euclidea. Si dice quindi che lo spazio di Minkowski è *pseudo-euclideo.*

Abbiamo mostrato che il quadrintervallo è invariante. Particolarmente importanti sono quindi le trasformazioni che rimangono invariate le distanze spazio-temporali, ovvero che preservano il quadrintervallo. In analogia alla geometria ordinaria, queste trasformazioni hanno la forma di rotazioni. In questo caso, però, vi sono 6 possibili rotazioni: xy, xz, $x\tau$, yz, $y\tau$, $z\tau$. Tre di esse corrispondono alle rotazioni ordinarie dello spazio tridimensionale (mantenendo fissi 2 assi fra cui τ), le nuove rotazioni miscelano le coordinate spaziali con quelle temporali e sono le più interessanti. Basta considerarne una sola, nel caso ad esempio il sistema scorra lungo x questa equivale ad una rotazione intorno agli assi $x\tau$. Applichiamo quindi una banale rotazione:

$$\begin{cases} x' = x\cos\varphi + \tau\sin\varphi \\ \tau' = -x\sin\varphi + \tau\cos\varphi \end{cases}$$

Supponendo s' il sistema solidale al moto, allora risulta $x' = 0$. Quindi:

$$\frac{x}{\tau} = -\tan\varphi \quad \Rightarrow \quad \frac{x}{ict} = \frac{v}{ic} \quad \Rightarrow \quad \tan\varphi = i\frac{v}{c}$$

da questa sono noti allora il seno e il coseno:

$$\sin\varphi = \frac{i\beta}{\sqrt{1-\beta^2}} \qquad \cos\varphi = \frac{1}{\sqrt{1-\beta^2}}$$

da cui, sostituendo nelle leggi di trasformazione sopra, si ottengono le **Trasformazioni di Lorentz** o **Boosts**:[8]

$$x' = \frac{1}{\sqrt{1-\beta^2}}(x - vt)$$
$$t' = \frac{1}{\sqrt{1-\beta^2}}\left(t - \frac{v}{c^2}x\right)$$

che nel limite di $c \to \infty$ restituiscono correttamente le leggi di trasformazione di Galileo. Pertanto *le leggi di trasformazione di Lorentz possono essere viste come le rotazioni quadridimensionali miste di un angolo immaginario nello spazio di Minkowski.*

[8]Per la precisione, un *Boosts* è una trasformazione fra due sistemi di riferimento con assi x, y e z paralleli e le cui origini coincidono all'istante iniziale e che si muovono lungo l'asse x (la cosiddetta *configurazione standard*). La più generale trasformazione di Lorentz coinvolge le rotazioni intorno ai tre assi. I Boosts sono rappresentati da matrici simmetriche – come quelle trattate qui – ma una trasformazione generale di Lorentz non è necessariamente simmetrica.

Tali relazioni possono ovviamente essere invertite cambiando semplicemente segno a v:

$$x = \frac{1}{\sqrt{1-\beta^2}}(x' + vt')$$
$$t = \frac{1}{\sqrt{1-\beta^2}}\left(t' + \frac{v}{c^2}x'\right)$$

ottenendo le componenti del sistema fisso nel sistema in moto. Da questa forma si derivano due interessanti fenomeni.

Uno è la dilatazione dei tempi già trovata. Si consideri infatti l'intervallo di tempo $t_A - t_B$ in un punto fisso $x_A = x_B$. Fisicamente, questo significa calcolare l'intervallo di tempo per un osservatore fisso nel sistema di riferimento e per un altro osservatore in moto relativo nel momento in cui passa per $x_A = x_B$. Applicando le relazioni precedenti, si trova:

$$t_A - t_B = \frac{1}{\sqrt{1-\beta^2}}(t'_A - t'_B) \quad \rightarrow \quad \Delta t = \Delta t' \sqrt{1-\beta^2}$$

che è esattamente la relazione trovata prima.

Il secondo fenomeno interessante si trova invece calcolando la distanza ad un tempo fisso, ovvero quando i due osservatori calcolano una distanza allo stesso tempo nel sistema solidale e nel sistema in moto. Si consideri a titolo di esempio un regolo i cui estremi sono x_A, x_B. In un determinato istante, quindi ponendo $t_A = t_B$, si ha che:

$$x_A - x_B = \frac{1}{\sqrt{1-\beta^2}}(x'_A - x'_B) \quad \rightarrow \quad x'_A - x'_B = (x_A - x_B)\sqrt{1-\beta^2}$$

dette quindi l e l_0 rispettivamente le lunghezze del regolo nel sistema mobile ed in quello fisso solidale, si ha:

$$\boxed{l = l_0\sqrt{1-\beta^2}}$$

cioé l'osservatore nel sistema di riferimento in moto misura una lunghezza inferiore rispetto a quella di un osservatore nel sistema di riferimento fisso. Tale fenomeno è noto sotto il nome di *contrazione delle lunghezze di Fitzgerald-Lorentz.*

Il fenomeno della dilatazione dei tempi e della contrazione delle lunghezze è suscettibile di una interpretazione puramente geometrica, su cui si baserà il paragrafo §1.3 per discutere l'approccio corretto alla relatività ristretta.

LEGGE DI TRASFORMAZIONE DELLE VELOCITÀ

Ricaviamo ora la legge di trasformazione delle velocità, sempre nell'ipotesi di moto lungo l'asse x. L'estensione al caso più generale è immediata e banale.

$$\begin{cases} dx &= \dfrac{1}{\sqrt{1-\beta^2}}(dx' + vdt') \\ dy &= dy' \\ dz &= dz' \\ dt &= \dfrac{1}{\sqrt{1-\beta^2}}\left(dt' + \frac{v}{c^2}dx'\right) \end{cases}$$

da cui:

$$\begin{cases} \dfrac{dx}{dt} &= \dfrac{dx' + vdt'}{dt' + \frac{v}{c^2}dx'} = \dfrac{v'_x + v}{1 + v\frac{v'_x}{c^2}} = v_x \\ \dfrac{dy}{dt} &= \dfrac{dy'\sqrt{1-\beta^2}}{dt' + \frac{v}{c^2}dx'} = \dfrac{v'_y\sqrt{1-\beta^2}}{1 + v\frac{v'_x}{c^2}} = v_y \\ \dfrac{dz}{dt} &= \dfrac{dz'\sqrt{1-\beta^2}}{dt' + \frac{v}{c^2}dx'} = \dfrac{v'_z\sqrt{1-\beta^2}}{1 + v\frac{v'_x}{c^2}} = v_z \end{cases} \tag{1.6}$$

Il limite $c \to \infty$ restituisce le trasformazioni di Galileo, tuttavia si vede come le componenti della velocità siano mischiate fra loro. Sviluppando in termini di v/c e v/c^2 (nell'ipotesi naturalmente di $v/c \ll 1$), si ricava:

$$\begin{cases} v_x &= (v_x + V)\left(1 - \dfrac{V}{c^2}v'_x\right) = v'_x + V - \dfrac{Vv'_x}{c^2} - \dfrac{V^2}{c^2}v'_x \\ v_y &= \left(1 - \dfrac{1}{2}\dfrac{V^2}{c^2}\right)v'_y\left(1 - \dfrac{V}{c^2}v'_x\right) = v'_y - v'_y\dfrac{V}{c^2}v'_x - \dfrac{1}{2}\dfrac{v^2}{c^2}v'_y + \dfrac{1}{2}\dfrac{V^3}{c^4}v'_xv'_y \\ v_z &= \left(1 - \dfrac{1}{2}\dfrac{V^2}{c^2}\right)v'_z\left(1 - \dfrac{V}{c^2}v'_x\right) = v'_z - v'_z\dfrac{V}{c^2}v'_x - \dfrac{1}{2}\dfrac{V^2}{c^2}v'_z + \dfrac{1}{2}\dfrac{V^3}{c^4}v'_xv'_z \end{cases}$$

notando che nel nostro caso $V = v_x$, in notazione vettoriale si può generalizzare come:

$$\boxed{\vec{v} = \vec{v}' - \left(\frac{\vec{V}\cdot\vec{v}'}{c^2}\right)\vec{v}' + \vec{V}}$$

anche in questo caso, quindi, le velocità si compongono per direzione, oltre che in modulo. Tale effetto è alla base del fenomeno dell'*aberrazione stellare*, cioé della variazione di angolo sotto cui è vista una stella a causa della velocità. Il fenomeno ottico si manifesta come una discrepanza tra la posizione apparente di una stella e la sua posizione vera. È causato dal moto dell'osservatore rispetto al cammino del fascio luminoso emesso dalla stella, ed è facilmente spiegabile tenendo conto che nell'intervallo di tempo necessario perché la luce dell'astro raggiunga l'osservatore, questi si è spostato rispetto alla propria posizione originaria.[9] Il moto dell'osservatore è la risultante del moto di rotazione della Terra, del suo moto di rivoluzione attorno al Sole e del moto del sistema solare nello spazio.

Per mostrare come si origina il fenomeno, si consideri a titolo di esempio la Terra ed il raggio di luce complanari nel piano xy. Allora utilizzando le (1.6):

$$\begin{cases} v\cos\theta &= \dfrac{v'_x + v}{1 + \frac{Vv'_x}{c^2}} = \dfrac{v'\cos\theta' + V}{1 + \frac{Vv'\cos\theta'}{c^2}} \\ v\sin\theta &= \dfrac{v'_y\sqrt{1-\beta^2}}{1 + v\frac{v'_x}{c^2}} = \dfrac{v'\sin\theta'\sqrt{1-\beta^2}}{1 + \frac{Vv'\cos\theta'}{c^2}} \end{cases}$$

per i raggi luminosi vale $v = c$ per cui:

$$\begin{cases} \cos\theta &= \dfrac{\cos\theta' + \frac{V}{c}}{1 + \frac{V\cos\theta'}{c}} \\ \sin\theta &= \dfrac{\sin\theta'\sqrt{1-\beta^2}}{1 + \frac{V\cos\theta'}{c}} \end{cases}$$

Per V/c piccolo si può trascurare il termine β al numeratore e sviluppare la componente orizzontale della velocità, che fornisce:

$$\sin\theta = \frac{\sin\theta'\sqrt{1-\beta^2}}{1 + \frac{V\cos\theta'}{c}} \Rightarrow \sin\theta\left(1 + \frac{V\cos\theta'}{c}\right) \approx \sin\theta' \Rightarrow$$

$$\sin\theta - \sin\theta' \approx -\frac{V}{c}\cos\theta'\sin\theta \Rightarrow \Delta\theta \approx \beta\sin\theta\cos\theta'$$

[9] L'aberrazione stellare è infatti una conseguenza diretta del fatto che la velocità della luce è finita.

1.3 La contrazione delle lunghezze, la dilatazione dei tempi e due paradossi della relatività ristretta

In questo paragrafo[10] tratteremo del modo corretto di interpretare la relatività ristretta e di due paradossi citati spesso come esempio del fatto che la relatività ristretta fa previsioni apparentemente incoerenti.

Il primo di questi paradossi riguarda lo spazio ed è il cosiddetto *paradosso dello sciatore* (noto anche come *paradosso del garage e della macchina*) e si formula generalmente nel seguente modo: si supponga una buca circolare di una certa larghezza l e degli sci di lunghezza l'. Nel sistema di riferimento della buca gli sci dello sciatore, che si avvicina con moto inerziale a velocità prossima a quella della luce, subiscono una contrazione di Lorentz nel senso della lunghezza pari a $\sqrt{1-\beta^2}$ tale da renderli più piccoli di l: lo sciatore deve pertanto cadere nella buca. Nel sistema di riferimento dello sciatore invece è la buca che si contrae dello stesso fattore, pertanto lo sciatore può passare senza cadere. Il paradosso consiste nel fatto che entrambe le descrizioni sono assolutamente corrette secondo la teoria della relatività, ma un effetto fisico (la caduta dello sciatore) non può dipendere dal sistema di riferimento inerziale scelto.

Il secondo paradosso riguarda il tempo ed è il cosiddetto *paradosso dei gemelli*, che in linea generale si formula nel seguente modo. Siano dati due gemelli: uno di questi resta sulla Terra (supposto ai fini di questo paradosso un riferimento inerziale in quiete) mentre l'altro parte su un'astronave che viaggia di moto inerziale a velocità prossima a c. Ognuno dei due gemelli vede scorrere il tempo proprio più lentamente di un fattore $\sqrt{1-\beta^2}$ rispetto all'altro a causa della simmetria dell'effetto di dilatazione dei tempi, ed il paradosso consiste nel fatto che se il gemello in viaggio compara il proprio tempo con quello dell'altro gemello (ad esempio ritornando sulla terra), ognuno dovrebbe vedere l'altro più vecchio di sé stesso, cosa naturalmente impossibile.

Il primo punto da affrontare è quello di comprendere se questi fenomeni relativistici sono fenomeni "fisici" o "apparenti" nel senso che andremo a discutere adesso. La questione è capire se questi abbiano una base fisica che li causano, se si tratta quindi di modificazioni delle interazioni elettromagnetiche o meccaniche dei corpi (in questo senso si parla di *spiegazione causale*), o se si tratti piuttosto di effetti che possono essere spiegati facendo ricorso alla sola struttura geometrica dello spaziotempo di Minkowski senza avere necessariamente una origine causale fisica (e in questo senso si parla di *spiegazione strutturale*). Attualmente, l'interpretazione causale è pressocché abbandonata perché non è effettivamente possibile trovare alcun meccanismo fisico che possa spiegare la contrazione di un corpo rigido o il rallentare del trascorrere del tempo (e di conseguenza anche di tutti i processi fisici e biologici associati) e si propende piuttosto per una spiegazione strutturale di questi effetti, che a parere dell'autore è l'unica corretta. L'approccio causale riemerge però implicitamente nel momento in cui si cerca di interpretare i paradossi relativistici in termini di causa-effetto fisico. Il classico esempio di questo tentativo è quando ci si chiede nel caso dello sciatore come sia possibile che uno sci lungo l' possa *materialmente* cadere in una buca $l < l'$. Si vedrà che questo approccio è quello sbagliato, mentre una spiegazione strutturale permette semplicemente di far sparire il paradosso stesso e mette in evidenza che il problema implicito in questo tentativo di interpretazione è il considerare simultanei eventi relativistici che in realtà non lo sono.

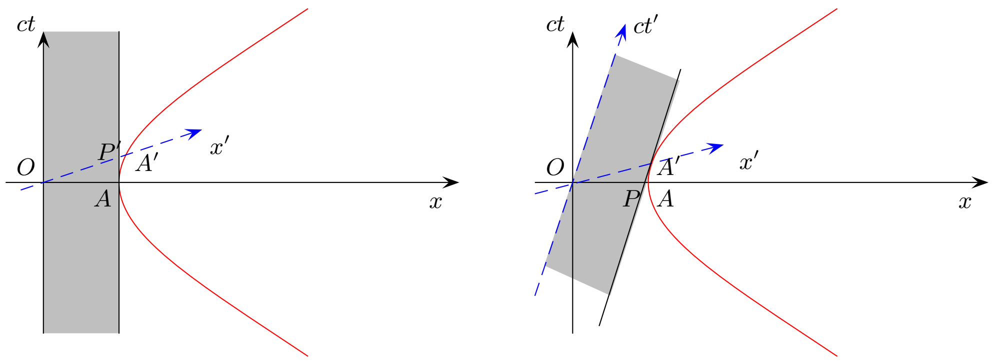
Vediamo quindi innanzitutto come è possibile giustificare la contrazione delle lunghezze e la dilatazione dei tempi in maniera strutturale.

Consideriamo un sistema di riferimento centrato su una sorgente luminosa. I raggi di luce sono rappresentati da linee di universo a 45°, mentre il quadrintervallo di un evento è dato da $s^2 = c^2t^2 - \vec{r}^2$ ed è un invariante in tutti i sistemi di riferimento. Questa forma quadratica rappresenta geometricamente un'iperbole equilatera, che è quindi anch'essa un invariante. Questo sistema di riferimento prende il nome di *sistema fondamentale* e le iperboli, chiamate *curve di calibrazione*, rappresentano in effetti il modo in cui variano le unità di misura.

[10] Questo paragrafo segue la trattazione riportata nel testo *"Filosofia della fisica", G. Boniolo, Bruno Mondadori Editore, 1997* per il paradosso dei gemelli e l'approccio generale alla questione dei paradossi e la spiegazione riportata in un articolo di Elio Fabri (Dipartimento di Fisica dell'università di Pisa) per il paradosso dello sciatore. L'articolo originale di Elio Fabri può essere trovato a questo indirizzo: *http://www.df.unipi.it/~fabri/sagredo/articoli/sciatore.pdf.*

Bisogna però fare attenzione al fatto che questa è un'iperbole molto particolare. In geometria euclidea, l'iperbole è definita come il luogo geometrico dei punti del piano in cui è costante il valore assoluto della differenza delle distanze da due punti fissi detti fuochi. Questo implica che la distanza dall'origine dei punti dell'iperbole non è invece una costante. Nella metrica di Minkowski, a causa al segno meno, questa iperbole rappresenta invece i punti del piano per cui la distanza dall'origine *in senso minkowskiano* è costante,[11] il che significa che anche in questo caso la distanza *puramente euclidea* di due punti dell'iperbole non è la stessa dall'origine. Mentre infatti la distanza nella geometria euclidea è definita da $\sqrt{dx^2 + dy^2 + dz^2}$, quella nello spazio di Minkowski è definita dalla relazione $\sqrt{c^2dt^2 - dx^2 - dy^2 - dz^2}$ e coincide proprio con il quadrintervallo, che è dunque invariante lungo l'iperbole. Si tratta in questo caso della distanza fra due eventi spaziotemporali, gli unici che hanno senso in relatività, mentre la distanza spaziale nello spazio di Minkowski è data sempre dalla forma euclidea $\sqrt{dx^2 + dy^2 + dz^2}$. Questa precisazione permette di capire come si originano la contrazione delle lunghezze e la dilatazione dei tempi.

Consideriamo la figura:[12]



Supponiamo di avere un regolo lungo $\overline{OA}$ in un sistema in quiete K. La linea di universo di questo regolo è espressa dalla banda grigia nella figura di sinistra compresa fra i due estremi O e A fermi. Consideriamo ora un sistema in moto K', il cui asse x' è indicato in figura e l'intersezione con l'iperbole invariante è indicata con A'. Ora, siccome i due punti giacciono sulla stessa iperbole (ovvero per il fatto che il quadrintervallo è invariante), la distanza $\overline{OA}$ nel sistema K è uguale alla distanza $\overline{OA'}$ nel sistema K'. Ma nel sistema in moto K' l'estremo del regolo è ancora P' (intersezione P' dell'asse delle ascisse x' con la linea di universo dell'estremo A) e questa distanza $\overline{OP'} < \overline{OA'}$, ovvero nel sistema in moto il regolo appare più corto rispetto al sistema in quiete.

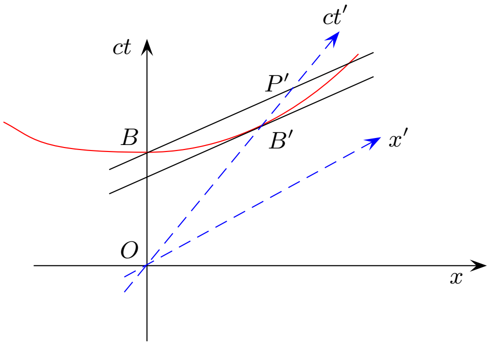
Il ragionamento è completamente invertibile come si vede dalla seconda figura. In questo caso, il regolo è in quiete nel sistema in moto K' e la sua linea di universo è dato dalla striscia inclinata in grigio, dove l'inclinazione dipende dalla velocità. Il regolo in questo sistema è lungo $\overline{OA'}$. Questa lunghezza corrisponde al segmento $\overline{OA}$ nel sistema in quiete K, ma la lunghezza del regolo in K è data dal segmento $\overline{OP} < \overline{OA}$, quindi anche nel sistema in quiete il regolo in moto appare più corto.

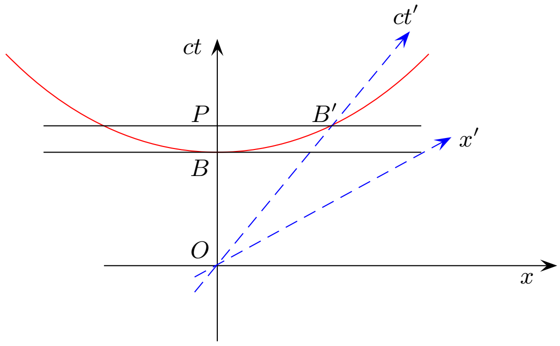
Possiamo costruire un approccio simile per mostrare come si origina la dilatazione dei tempi. Consideriamo la figura:[13]

[11] La distanza minkowskiana è data proprio dal quadrintervallo e questo è invariante.

[12] Le figure seguenti sono rifatte a partire dalle figure pubblicate su "Filosofia della fisica", G. Boniolo, Bruno Mondadori Editore, 1997.

[13] Le figure seguenti sono rifatte a partire dalle figure pubblicate su "Filosofia della fisica", G. Boniolo, Bruno Mondadori Editore, 1997.



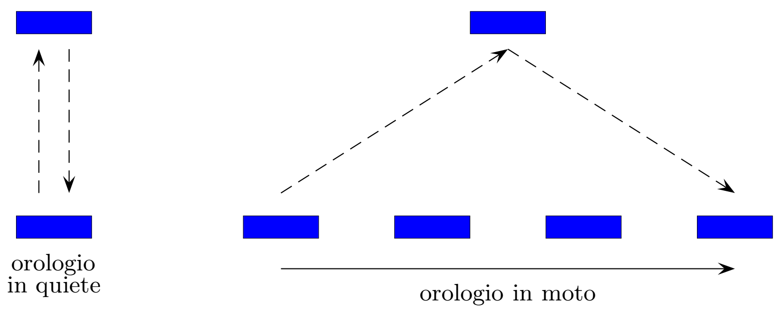


Consideriamo la prima figura e supponiamo di avere un orologio solidale con il sistema K in quiete e che misuri un intervallo di tempo $\overline{OB}$. Il punto B è un punto dell'iperbole invariante, per cui per un sistema K' in moto rispetto a K l'istante B corrisponde a B', ovvero l'intervallo $\overline{OB}$ è visto come $\overline{OB'}$. Ma in K' l'evento simultaneo a B è dato dall'evento in P'. Essendo $\overline{OB'} < \overline{OP'}$ ne risulta che l'orologio in K' ritarda rispetto a quello in K. Quando l'osservatore in K' misura B', l'evento P' ancora non si è verificato e quindi misura un tempo dilatato rispetto a K.

Ancora una volta il discorso è invertibile. Supponiamo un orologio solidale con il sistema K', dove misura un intervallo $\overline{OB'}$. Questo intervallo sarà lungo $\overline{OB}$ per un osservatore in K, ma per questo osservatore l'evento simultaneo con B' è P. Quindi l'intervallo $\overline{OB} < \overline{OP}$ e l'orologio in K ritarda rispetto a quello in K'. Quando l'osservatore in K misura B, l'evento P ancora non si è verificato e quindi sperimenta anche lui un tempo dilatato rispetto a K'

È quindi chiaro che questi fenomeni relativistici sono originati dal fatto che *una misura effettuata in K trasforma lungo l'iperbole dettata dal quadrintervallo e rappresenta come questa misura è vista in un sistema K', ma questo evento non è simultaneo con quello di partenza, che appare precedere (per misure spaziali) o seguire (per misure temporali) l'evento in K.*

Questi effetti derivano dall'ipotesi fondamentale della relatività ristretta: il fatto che *la velocità della luce è finita e la stessa in tutti i sistemi di riferimento*. Questo fatto può essere reso più intuitivo se si interpreta la dilatazione temporale in un altro modo. Consideriamo la figura seguente.



Supponiamo due osservatori distanti l l'uno dall'altro. Nel caso a sinistra l'osservatore in quiete misura la distanza $l = 1/2ct$, perché in realtà misura il tempo di andata e ritorno di un segnale luminoso. Nel caso invece l'orologio sia in moto, la maggiore lunghezza percorsa dai raggi di luce deve necessariamente essere compensata, dal punto di vista dell'osservatore in moto, da un rallentamento opportuno del tempo.

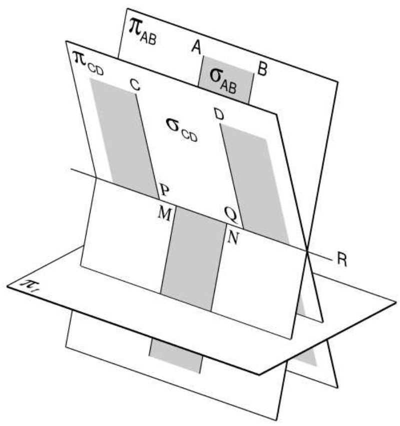
Come si vede, la relatività distrugge la nozione di contemporaneità (quello che tecnicamente si chiama *simultaneità degli eventi*) così come la si intende normalmente. Questo è una conseguenza diretta della finitezza della velocità della luce, in quanto quello che percepiamo come "qui" e "ora" è in realtà composto da informazioni che arrivano tanto più dal passato quanto più si guarda lontano.

L'esperienza comune però, contrariamente a quello che afferma la relatività, identifica come simultanei tutti questi eventi a distanza. Siccome è proprio questa intuizione comune che genera quelli che chiamiamo paradossi, è importante cercare di spiegare come questa "intuizione" nasce e stabilire quindi perchè la versione relativistica della simultaneità e quella percepita appaiono contradditorie. Questa spiegazione si basa su due fattori: il fatto che la velocità della luce è molto superiore alla velocità di propagazione degli altri segnali fisici e dalla limitazione del sistema nervoso umano di discriminare due eventi vicini nel tempo. Questo limite del sistema nervoso si posiziona a circa $1/100$ di secondo, in altri termini al di sotto di questo limite il sistema nervoso non riesce a individuare due eventi come temporalmente distinti. La luce in questo intervallo di tempo percorre circa 3000 km, quindi un essere umano non può fare altro che considerare ogni evento che si verifica all'interno di questo raggio come perfettamente simultaneo. Questa sfera comprende agevolmente ogni evento che possa *influenzare in pratica* il nostro comportamento "qui", e pertanto l'evoluzione ha portato l'essere umano ad identificare quello che "vede in questo momento" con quello che "è reale in questo momento". Inoltre, nel limite di $1/100$ di secondo la luce può viaggiare diverse volte fra l'osservatore e l'ambiente a lui circostante, fattore che contribuisce alla stabilità della percezione sensoriale e che tende quindi ancora una volta a consolidare l'identificazione del "vedo ora" con l'"esiste ora".

L'intuito porta quindi spesso ad applicare la teoria della relatività in maniera scorretta, in particolare quando non si tiene bene conto del ruolo della relatività della simultaneità anche nelle misure di spazio, dove è meno evidente. I cosiddetti paradossi si originano infatti quando in un contesto relativistico si introduce erroneamente il concetto di tempo assoluto, sia assumendo implicitamente un sistema di riferimento privilegiato, sia implicitamente non tenendo conto della relatività della simultaneità, che è poi proprio il concetto centrale della teoria.

Il paradosso dello sciatore

Bisogna notare preliminarmente che un punto è rappresentato in un diagramma spaziotemporale da una retta, e tutti i punti di un corpo rigido sono rappresentati da rette parallele. L'insieme dei punti sincronizzati a un determinato tempo t di questo sistema sono rappresentati da piani π_t perpendicolari all'asse temporale ct. Ogni sistema di riferimento avrà una famiglia di piani $\{\pi_t\}$ diversa e questo rappresenta geometricamente la relatività della simultaneità. Fatta questa premessa, illustriamo la situazione dello sciatore e della buca in un diagramma spaziotemporale.[14]



Supponiamo per generalità che sia la buca che gli sci siano in moto. Gli sci individuano una striscia definita dalle linee orarie degli estremi A e B e definiscono quindi un piano π_{AB}. In maniera analoga, la buca è definita dai suoi estremi C e D, che definiscono un piano π_{CD}. Nel riferimento in quiete K, un dato istante t definisce un piano π_t. Le intersezioni delle due strisce σ_{AB} e σ_{CD} con il piano π_t rappresentano le posizioni all'istante t della buca e degli sci e dalle condizioni del problema è evidente che queste sono parallele. Ne consegue che anche l'intersezione di π_{AB} e π_{CD} è una retta

[14] La figura è presa integralmente dall'articolo di di Elio Fabri (Dipartimento di Fisica dell'università di Pisa) sul paradosso dello sciatore. L'ottimo articolo di Elio Fabri può essere trovato a questo indirizzo: *http://www.df.unipi.it/~fabri/sagredo/articoli/sciatore.pdf.*

(indicata con R) ed è parallela al piano π_t. Questa intersezione rappresenta l'istante in cui gli sci e la buca si incrociano (sono quindi sulla stessa retta) e dimostra che il sistema K è quello più semplice per descrivere il paradosso. A ben vedere K non è l'unico sistema che gode di questa particolarità, ma vale per ogni sistema rappresentato da piani paralleli a R e in particolare quello in cui la buca è ferma.

Perché lo sci cada nella buca basta evidentemente che σ_{AB} non intersechi $\pi_{CD} - \sigma_{CD}$, ovvero che $MN \subset PQ$. I quattro punti M, N, P e Q sono eventi nello spaziotempo di Minkowski e pertanto le distanze sono invarianti, deve quindi risultare $\overline{MN} < \overline{PQ}$. *Questa condizione non dipende dal sistema di riferimento e pertanto non esiste alcun paradosso.* In particolare, nel riferimento K queste distanze coincidono con le distanze euclidee e quindi questo sistema risulta particolarmente comodo per "risolvere" il paradosso. *Quello che conta quindi è che nel sistema K gli sci siano più corti della buca.*

Nel sistema K' degli sci, il problema risulta invece particolarmente complesso da trattare. I piani π'_t di questo sistema non sono paralleli a R e di conseguenza le lunghezze calcolate in questo sistema non sono utilizzabili per risolvere il problema. Naturalmente, le distanze $\overline{MN}$ e $\overline{PQ}$ possono essere calcolate anche in K', ma in questo caso bisogna tenere presente che si riferiscono ad eventi *non simultanei* e pertanto il calcolo risulta più complesso. In definitiva si dimostra che nel sistema K' dello sciatore, qualora si tenga conto della non simultaneità della misura degli estremi dello sci dovuta al tempo di propagazione finito della luce, questo cade nella buca solo se risulta $\sqrt{1-\beta^2}\, l_{K'} < l_{\text{buca}}$.

Il paradosso dei gemelli

Per definire il problema chiaramente, evitiamo il più possibile di parlare di "gemelli" e "astronavi" e parliamo piuttosto di orologi e sincronizzazione, richiamandoci ai gemelli solo quando necessario per fare il parallelo con la formulazione classica. Chiamiamo quindi K il sistema di riferimento inerziale dell'osservatore G (quello a riposo, "che rimane sulla terra") e K' quello dell'osservatore G' in moto con velocità v. G e G' possono sincronizzare i loro orologi nell'istante iniziale t_0.

Data l'indistinguibilità dei sistemi di riferimenti inerziali, l'effetto di dilatazione temporale è simmetrico nei due sistemi. Per il sistema G', è il sistema G che si muove con velocità $-v$ e pertanto è il tempo proprio di G che sperimenta un effetto di dilatazione temporale per lo stesso fattore $\sqrt{1-\beta^2}$.

C'è accordo, in generale, sul fatto che questo effetto non può realmente essere simmetrico, nel senso che nel momento in cui si confronta il tempo trascorso per G e quello trascorso per G', per G' deve effettivamente essere trascorso meno tempo (nel linguaggio comune, "deve essere più giovane"). Questa asimmetria viene giustificata normalmente notando che il sistema in moto G' deve necessariamente invertire la rotta per poter confrontare le due misure e questo implica delle accelerazioni che richiedono l'utilizzo della relatività generale.

Questo non è strettamente vero, mentre è vero invece che il principio di invarianza galileiano vale solo fra sistemi inerziali e quindi non c'è alcun motivo per cui l'effetto di dilatazione temporale debba essere simmetrico anche per sistemi in moto relativo non inerziale. Spesso la relatività generale viene invocata ogni volta che intervengono delle accelerazioni, mentre questo non è necessario a priori. In realtà la relatività ristretta sa trattare benissimo le accelerazioni, quello che la relatività generale introduce – come vedremo più avanti nel corso – è il *principio di equivalenza*, ovvero l'assunzione che il campo gravitazionale possa essere localmente trattato come un sistema non inerziale. A causa delle masse gravitazionali, lo spaziotempo diventa curvo e solo in questo caso si è obbligati ad usare la relatività generale, che permette appunto di calcolare come si modifica la struttura geometrica dello spaziotempo a causa delle masse gravitazionali. Ogni volta che si ha a che fare con uno spazio tempo di Minkowski (ovvero piatto perché senza effetti gravitazionali), è sufficiente la relatività ristretta, che si parli di sistemi inerziali o no.

Parlare di accelerazioni e di relatività generale è quindi in realtà un modo un po' ingenuo di vedere questo paradosso. Il punto interessante è invece che esso può essere espresso anche utilizzando solo sistemi inerziali, dimostrando così che l'accelerazione non gioca alcun ruolo nella risoluzione di questo paradosso. Questo punto molto interessante è raramente messo in evidenza.

Una prima formulazione che prevede solo sistemi inerziali può essere chiamata dell'*universo cilindrico*. In questo caso si suppone semplicemente che in questo universo le rette siano linee chiuse. In maniera formalmente precisa, si suppone che lo spazio sia *minkowskiano con condizioni al con-*

torno spaziali periodiche, in modo che detto L il periodo si possa fare l'identificazione $x = x + L$ e similmente per le altre dimensioni spaziali. La coordinata temporale rimane invece non limitata per evitare curve causali chiuse. In questo modo si ottiene un universo la cui parte spaziale è uno *spazio compatto* ovvero, detto in termini più semplici, *chiuso e limitato*. Supponiamo anche assenza di materia, in modo da non essere costretti ad utilizzare la relatività generale.

In questo universo, sia K un sistema in quiete e K' un sistema in moto relativo inerziale. All'istante $t = t_0$ i due sistemi sincronizzano gli orologi che segnano così lo stesso tempo. K' è però in moto e registra quindi una dilatazione temporale rispetto al sistema K. Lo stesso accade per K, che misura una dilatazione temporale rispetto al sistema K' in moto relativo rispetto a lui, in quanto l'effetto in questo caso è realmente simmetrico. A causa della topologia cilindrica dello spazio, esiste però un istante $t_1 > t_0$ nel quale i due sistemi K e K' si incrociano di nuovo pur avendo viaggiato sempre di moto inerziale. Cosa misureranno i due osservatori G e G'?

La risoluzione del paradosso è in questo caso tecnicamente abbastanza complessa,[15] ma il punto essenziale è che in uno spazio compatto esiste effettivamente un sistema di riferimento privilegiato, ed è l'unico riferimento che può essere in quiete in questo spazio tempo: si tratta proprio del sistema in cui si definiscono le condizioni al contorno periodiche. In altri termini, per poter impostare le condizioni al contorno periodiche nel sottospazio spaziale (o qualunque altra condizione di compattezza) occorre stabilire necessariamente, anche se implicitamente, un sistema di riferimento che è in definitiva il sistema K del "gemello più vecchio".

Inoltre, a differenza di G, G' che si trova nel sistema in moto K' non può definire uno *spazio di quiete globale*, ovvero non può sincronizzare gli orologi nello spazio intorno a sé stesso. Questo è importante per poter definire un tempo associato all'osservatore in tutto lo spazio tempo e definire così una nozione di tempo anche lontano dall'osservatore.

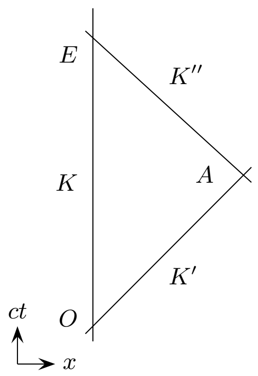
Se fosse possibile effettuare la sincronizzazione degli orologi in K', ogni gemello potrebbe utilizzare il proprio tempo globale per stimare l'età dell'altro e si cadrebbe in contraddizione. Ma questo è impossibile in uno spazio compatto, come vedremo subito.

Questa procedura di sincronizzazione su tutto lo spaziotempo (*procedura di sincronizzazione di Einstein*) prevede che l'osservatore G' posizioni degli osservatori nei vari punti dello spazio K' e che sincronizzi i vari orologi scambiando segnali luminosi. Ma nel caso di un universo compatto, l'osservatore G' ad un certo punto comincerà a ricevere i propri stessi segnali per resettare il proprio orologio di una quantità proporzionale a βL. In parole povere, G' è fuori sincrono con il proprio stesso tentativo di sincronizzazione. Ogni misura fatta in questo riferimento è quindi ambigua di una quantità pari a questo spostamento temporale e se ne deduce che non è possibile per G' definire una coordinata temporale globale che rispetti la procedura di sincronizzazione einsteniana. Per eventi lontani dalla sua linea di universo non è possibile associare loro una coordinata temporale sincronizzata con il tempo proprio di G', quindi non è possibile dire quanto tempo è passato e l'invarianza di Lorentz non è conservata.

In ultima analisi, i due sistemi di riferimento non sono simmetrici e pertanto l'effetto di dilatazione temporale non lo è: e siccome il tempo proprio di un osservatore in moto è sempre minore di un osservatore in quiete, all'istante t_1 G è effettivamente più vecchio di G'.

Una seconda formulazione del paradosso che non richiede accelerazioni prevede tre sistemi inerziali e lo spaziotempo piatto di Minkowski. Supponiamo ancora che all'instante t_0, nel punto O, un osservatore G nel sistema K (quello supposto in quiete) e uno in G' nel sistema K', sincronizzino gli orologi e poi si allontanino di moto inerziale a velocità $\vec{v}$. Ciascun osservatore vedrà quindi il proprio orologio rallentare rispetto all'altro. In un istante successivo t_1, nel punto A il sistema K' incrocia un altro sistema inerziale K'' in moto a velocità $-\vec{v}$, in questo istante un osservatore G'' su K'' sincronizza il proprio orologio con quello di G'. Il sistema K'' continua poi il suo moto fino ad incrociare il sistema K in E, dove confronta il suo orologio con quello di G che è rimasto nel sistema in quiete K. Ancora una volta, cosa misureranno i due osservatori G e G''?

[15] Una discussione completa di questo argomento si trova in *John D. Barrow, Janna Levin, The Twin Paradox in Compact Spaces*, ***Phys.Rev. A63 (2001) 044104***, *http://arxiv.org/abs/gr-qc/0101014v1.*



La soluzione del paradosso si basa anche in questo caso sul fatto che i sistemi non sono equivalenti. Il punto essenziale in questo caso è che la retta $\overline{OE}$ non può mai trasformarsi nella spezzata $\overline{OAE}$ con una sola trasformazione di Lorentz, perché nello spazio di Minkowski le leggi di Lorentz trasformano sempre rette in rette. In altre parole, *l'effetto di dilatazione temporale è simmetrico solo fra* **due** *sistemi inerziali*, mentre la linea spezzata $\overline{OAE}$ si ottiene *applicando le trasformazioni di Lorentz in due punti diversi dello spaziotempo e questa composizione non è a sua volta una trasformazione di Lorentz*. Quindi l'insieme dei tre sistemi non è simmetrico e la dilatazione dei tempi a sua volta non è simmetrica. Ancora una volta, siccome il tempo proprio di un osservatore in moto è sempre minore di un osservatore in quiete, all'istante t_1 G è effettivamente più vecchio di G''.

Capitolo 2

Primo Intermezzo Matematico

2.1 Definizioni di base

Passeremo ora a impostare le basi del formalismo matematico adatto a trattare la relatività ristretta. Per fare ciò cominceremo innanzitutto a dare alcune definizioni di base che costituiranno l'ossatura di quanto diremo in seguito.

Abbiamo visto come un punto nello spazio di Minkowski sia definito da una quaterna di coordinate (x_1, x_2, x_3, x_4), che prende il nome di *evento*. Questo insieme di quattro componenti sarà indicato con x^α, e trasforma come:[1]

$$\begin{cases} x_1' &= \dfrac{1}{\sqrt{1-\beta^2}}(x_1 + i\beta x_4) \\ x_2' &= x_2 \\ x_3' &= x_3 \\ x_4' &= \dfrac{1}{\sqrt{1-\beta^2}}(x_4 - i\beta x_1) \end{cases}$$

La rappresentazione esplicita di queste coordinate è (ict, x, y, z). Essendo il differenziale delle coordinate dato da $(icdt, dx, dy, dz)$, il quadrintervallo può essere scritto nella forma $ds^2 = dx_\alpha dx^\alpha$.

Nota sulla notazione degli indici

Vale la pena sottolineare – per i più pignoli – che l'indice scritto in alto denota un quadrivettore controvariante, mentre l'indice in basso denota un quadrivettore covariante. La definizione dei due tipi di vettore è ritardata fino al paragrafo §11.2, in quanto nella I e II parte la distinzione non è essenziale: nel caso di coordinate piatte cartesiane - come quelle utilizzate in relatività speciale - le componenti covarianti e controvarianti coincidono.

In pratica, significa che fino a che non ci occuperemo della relatività generale, non è necessario preoccuparsi eccessivamente dell'innalzamento o abbassamento degli indici: questi sono mantenuti nella I e II parte solo per salvaguardare la convenzione sugli indici ripetuti spiegata nella nota successiva. Giusto per anticipare il discorso che seguirà, nel caso di coordinate piatte l'innalzamento o l'abbassamento di indice si effettua tramite una matrice diagonale con elementi unitari:

$$\eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

quindi in sostanza la notazione $dx_\alpha dx^\alpha$ indica in realtà $\eta_{\alpha\beta} dx^\alpha dx^\beta$.

[1] Sempre nell'ipotesi di un sistema di riferimento in moto lungo l'asse x.

Il segno relativo delle coordinate è conservato dalle trasformazioni di Lorentz e prende il nome di *segnatura*. È importante notare che la particolare segnatura in sé non è fondamentale: è piuttosto il segno relativo delle componenti che contiene il senso fisico. Si può quindi usare indifferentemente la segnatura $+---$ o $-+++$.[2] La segnatura adottata qui è $-+++$.

Un insieme A^α di quattro componenti che trasformano secondo le leggi di Lorentz si definisce *quadrivettore*. In notazione matriciale, più compatta:[3]

$$A'^\alpha = \Gamma^\alpha_\beta A^\beta \qquad \Gamma_{\alpha\beta} \equiv \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{i\beta}{\sqrt{1-\beta^2}} \\ 0 & 1 & 0 & o \\ 0 & 0 & 1 & 0 \\ -\frac{i\beta}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{1}{\sqrt{1-\beta^2}} \end{pmatrix}$$

la matrice qui denominata con $\Gamma_{\alpha\beta}$ prende il nome di *Matrice di Lorentz*.

Il prodotto scalare è definito come $A_\alpha B^\alpha$, diretta generalizzazione di quello già noto. Siccome risulta:

$$A'_\alpha B'^\alpha = \Gamma^\beta_\alpha A_\beta \Gamma^\alpha_\delta B^\delta$$

e vale la relazione sulle matrici di trasformazione di Lorentz:

$$\Gamma^\beta_\alpha \Gamma^\alpha_\delta = \left(\Gamma^T\right)^\beta_\alpha \Gamma^\alpha_\delta = \delta^\beta_\delta \Rightarrow \Gamma^\beta_\alpha A_\beta \Gamma^\alpha_\delta B^\delta = A_\beta B^\beta$$

vale – correttamente – la relazione:

$$A'_\alpha B'^\alpha = A_\alpha B^\alpha$$

ovvero, il prodotto scalare risulta invariante per trasformazioni di Lorentz. Questo è corretto, in quanto il prodotto scalare è un numero (uno scalare, appunto) e non può variare cambiando sistema di riferimento. Il modulo di un vettore invece non ha segno ben definito: $A_\alpha A^\alpha = A^2 \lesseqgtr 0$.

Possiamo definire un'altra quantità chiamata *Tensore*.

Si chiama tensore di *rango 2* un tensore che trasforma come il prodotto di due vettori, ovvero un oggetto che ha come matrice di trasformazione *formalmente* il "prodotto" di due matrici. Non esiste in questo caso somma sugli indici delle due matrici, quindi l'oggetto che regola la trasformazione (Jacobiano) presenta 16 componenti:

$$C'_{\alpha\beta} = \Gamma^\delta_\alpha \Gamma^\gamma_\beta C_{\delta\gamma}$$

Iterando la definizione, si otterranno i tensori di rango superiore. Riguardo ai tensori si possono definire le proprietà di:

$$\begin{aligned} C_{\alpha\beta} &= C_{\beta\alpha} && \textit{Simmetria} \\ C_{\alpha\beta} &= -C_{\beta\alpha} && \textit{Antisimmetria} \end{aligned}$$

In generale, un tensore sarà una combinazione delle due ovvero potrà essere simmetrico rispetto allo scambio di certi indici, ma antisimmetrico rispetto allo scambio di altri. È importante notare invece che la proprietà di simmetria o antisimmetria si conserva per trasformazioni. Il $\delta_{\alpha\beta}$ è un esempio di tensore numerico simmetrico, mentre un importante tensore antisimmetrico è il tensore:

$$\epsilon_{\alpha\beta\gamma\delta} = \begin{cases} 0 & \textit{due o più indici uguali} \\ 1 & \textit{permutazioni di indici pari} \\ -1 & \textit{permutazioni di indici dispari} \end{cases}$$

[2] A volte la coordinata temporale è la quarta e non la coordinata 0 (le coordinate spaziali essendo sempre la 1,2,3). In questo caso la segnatura è $---+$ o $+++-$. In pratica, questa è solo una questione di notazione e non cambia nulla nelle equazioni.

[3] Da questo momento sarà assunta la convenzione sugli indici ripetuti: se in una formula compare un indice covariante ed uno controvariante ripetuto, si sottintende allora una sommatoria (come l'indice β dell'esempio, che si intende sommato da 0 a 3: $\Gamma_{\alpha\beta}A^\beta \equiv \Gamma_{\alpha 0}A^0 + \Gamma_{\alpha 1}A^1 + \Gamma_{\alpha 2}A^2 + \Gamma_{\alpha 3}A^3$). Questi indici sono anche detti *muti*, perché la loro denominazione non influenza la fisica delle formule, ma sono presenti solo per ricordare, appunto, che è sottintesa una sommatoria sulle componenti. Questa convenzione è detta *convenzione di Einstein*.

il cui corrispondente tridimensionale è il tensore ϵ_{ijk}([4]) che definisce, ad esempio, il prodotto vettore:

$$\left[\vec{A} \times \vec{B}\right]_i = \epsilon_{ijk} C^{jk} \qquad C^{jk} = A^j B^k - A^k B^j$$

Di un tensore generico $F^{\gamma\delta}$ può essere definito un duale:

$$F^*_{\alpha\beta} = \frac{1}{6} \epsilon_{\alpha\beta\gamma\delta} F^{\gamma\delta}$$

Definizione di Tensore

Questa è chiaramente ben lontana dall'essere una definizione rigorosa di *tensore* e l'argomento richiederebbe un notevole approfondimento, che in genere è fatto in altri contesti. Per riassumere brevemente diciamo che, da un punto di vista fisico, un tensore è un oggetto molto generale definito intrinsecamente a partire da uno spazio vettoriale V (che può essere ad esempio lo spazio euclideo 3-dimensionale o lo spaziotempo 4-dimensionale) e quindi non dipendente da un particolare sistema di riferimento.

Rispetto ad un fissato sistema di riferimento, un vettore dello spazio è espresso come una sequenza di componenti numeriche (le sue coordinate), cioé una ennupla ordinata. Cambiando sistema di riferimento, lo stesso vettore è espresso con una sequenza diversa, secondo una legge ben precisa. Un tensore, espresso rispetto ad un particolare sistema di riferimento, è una più generale ennupla di numeri che può avere dimensione 1 (una sequenza), 2 o più alta (una matrice). Al mutare del sistema di riferimento le componenti di un tensore, come quelle di un vettore, sono modificate da leggi precise.

Questa nozione fisica di tensore come "oggetto non dipendente dal sistema di riferimento" è utile ad esprimere molte leggi fisiche, che per loro natura non dipendono dai sistemi di riferimento scelti. La nozione matematica di tensore è realizzata in modo più rigoroso tramite l'algebra lineare. Innanzitutto, nel linguaggio dell'algebra lineare un sistema di riferimento è una base e la legge di trasformazione è fornita dalla matrice di cambiamento di base. Inoltre, la definizione di tensore può essere data senza fare nessun riferimento ai sistemi di riferimento (cioé alle basi), usando le nozioni di applicazione multilineare e di spazio vettoriale duale.

La definizione di tensore che daremo ora è quella più intrinseca, perché non fa uso di basi, ed è la più usata in matematica. Una definizione alternativa, ampiamente usata in fisica, necessita di una base fissata.

Sia V uno spazio vettoriale di dimensione n su un campo K. Lo spazio duale V^* è lo spazio vettoriale formato da tutti i funzionali lineari $f : V \mapsto K$. Lo spazio V^* ha anch'esso dimensione n. Gli elementi di V e V^* sono chiamati rispettivamente vettori e covettori. Un tensore è una applicazione multilineare tale che:

$$T\colon \underbrace{V^* \times \cdots \times V^*}_{h} \times \underbrace{V \times \cdots \times V}_{k} \mapsto K$$

Un tensore T è quindi un'appplicazione che associa a k vettori $v_1, \dots v_k$ e h covettori $w_1 \dots w_h$ uno scalare $T(w_1 \dots w_h, v_1, \dots v_k)$. La multilinearità garantisce che la funzione sia lineare in ogni componente. L'ordine o tipo del tensore è la coppia (h, k).

Per gli scopi che ci prefiggiamo in questo corso quanto detto in questo capitolo sarà tuttavia sufficiente.

Il concetto di derivata in uno spazio di Minkowski non presenta alcuna differenza rispetto all'operatore già noto in $\mathbb{R}^3$ o più in generale in $\mathbb{R}^n$.

Per ciò che riguarda invece l'integrazione, in uno spazio quadridimensionale sono possibili 4 tipi di integrazione:

dx	1 dimensione	**Linea**
$d\sigma = dxdx'$	2 dimensioni	**Superficie**

essendo l'elemento di superficie definito come un prodotto vettore, esso è un vettore assiale a due indici del tipo:

$$df^{ij} = dx^i dx'^j - dx^j dx'^i$$

[4]In questo scritto si adotta la comoda convenzione di utilizzare le lettere greche per gli indici che vanno da 0 a 3 (o da 1 a 4) nello spazio-tempo quadridimensionale, mentre le lettere latine saranno utilizzate per indicare gli indici da 1 a 3 nello spazio tridimensionale euclideo galileiano o ristretti al sottospazio tridimensionale di uno quadridimensionale. In altre parole, gli indici latini variano sul solo spazio fisico, gli indici greci sull'intero spaziotempo.

in quattro dimensioni, questo è un oggetto antisimmetrico a due componenti (e si noti bene che questo non è un vettore). Il suo duale:

$$df^*_{\alpha\beta} = \frac{1}{6}\epsilon_{\alpha\beta\gamma\delta}f^{\gamma\delta}$$

rappresenta la superficie ortogonale a quella precedente.

$$d\Sigma = dxdx'dx'' \quad \text{3 dimensioni} \quad \textbf{Ipersuperficie}$$

In uno spazio tridimensionale è il prodotto scalare fra l'elemento di superficie ed il vettore ortogonale ad esso (elemento di volume). In quattro dimensioni, si costruisce come la generalizzazione di questo elemento di superficie:[5]

$$d\Sigma^{\alpha\beta\gamma} = \begin{vmatrix} dx^\alpha & dx'^\alpha & dx''^\alpha \\ dx^\beta & dx'^\beta & dx''^\beta \\ dx^\gamma & dx'^\gamma & dx''^\gamma \end{vmatrix}$$

In tal caso si può scrivere un quadrivettore completamente asimmetrico perpendicolare alla superficie:

$$C_\delta = \frac{1}{6}\epsilon_{\delta\alpha\beta\gamma}d\Sigma^{\alpha\beta\gamma}$$

che rappresenta la direzione perpendicolare allo spazio fisico tridimensionale. Infine:

$$d\Omega = dxdx'dx''dx''' \quad \text{4 dimensioni} \quad \textbf{Volume}$$

rappresenta l'elemento di volume quadridimensionale.

Si possono ovviamente generalizzare i teoremi di Gauss e Stokes. Infatti, per esempio, il teorema di Gauss che in tre dimensioni ha la forma:

$$\int_S a_i dS_i = \int_\Omega \partial a_i d\Omega$$

diventa in quattro dimensioni:[6]

$$\int_\Sigma A^\alpha dS_\alpha = \int_\Omega \partial_\alpha A^\alpha d\Omega$$

con $dS_\alpha = \frac{1}{6}\epsilon_{\alpha\beta\gamma\delta}dS^{\beta\gamma\delta}$, mentre:

$$\int_\Sigma A^\alpha dS_\alpha$$

rappresenta il flusso di A^α che attraversa l'ipersuperficie Σ.

Si può definire un quadrivettore *quadrivelocità* u^α come:

$$u^\alpha \equiv \frac{dx^\alpha}{ds}$$

che esplicitato per componenti ha la forma:

$$\begin{cases} \vec{u} &= \dfrac{d\vec{x}}{ds} = \dfrac{1}{c}\dfrac{1}{\sqrt{1-\beta^2}}\vec{v} \\ u_4 &= \dfrac{d(ict)}{ds} = \dfrac{i}{\sqrt{1-\beta^2}} \end{cases} \tag{2.1}$$

[5] In altri termini, l'elemento di ipersuperficie nello spazio quadridimensionale rappresenta l'elemento di volume nello spazio (fisico) tridimensionale.

[6] Spesso per abbreviare useremo nel seguito la notazione ∂_α per indicare la derivata rispetto a x^α, ovvero:

$$\partial_\alpha A_\beta \equiv \frac{\partial A_\beta}{\partial x^\alpha}$$

È immediato verificare che il modulo della quadrivelocità vale -1:

$$|u^\alpha| = u_\alpha u^\alpha = \vec{u}^2 + u_4^2 = \frac{\beta^2}{\sqrt{1-\beta^2}} - \frac{1}{\sqrt{1-\beta^2}} = -1$$

È possibile utilizzare per i quadrivettori anche una rappresentazione con matrici 2×2, che è analoga a quella con le matrici 4×4 $\Gamma_{\alpha\beta}$. Infatti se definiamo:

$$\hat{x} = \begin{pmatrix} x_4 - x_1 & x_2 - ix_3 \\ x_2 + ix_3 & x_4 + x_1 \end{pmatrix}$$

risulta $\det \hat{x} = c^2t^2 - x_1^2 - x_2^2 - x_3^2$, ovvero il determinante coincide proprio con il quadrintervallo. Le trasformazioni di Lorentz sono allora costituite dalle matrici che rimangono invariato il determinante, cioé le matrici che soddisfano:

$$\hat{x}' = g\hat{x}g^* \qquad \det g = 1$$

Se si considerano in particolare le trasformazioni tali che in più $g^* = g^{-1}$, queste conservano anche la traccia e costituiscono un sottoinsieme delle trasformazioni di Lorentz (le rotazioni). Non svilupperemo tuttavia ulteriormente questo punto.

Capitolo 3

Principio di Minima Azione e Sue Prime Applicazioni

3.1 Formulazione del principio

Introduciamo ora un metodo generale utile per ricavare l'equazione del moto e che sarà utilizzato molto spesso nel seguito: il cosiddetto *principio variazionale*. Questo principio consiste nel dire che l'"equazione del moto cercata" è quella che minimizza una determinata "quantità integrale", riconducendo così il problema del moto ad un problema di ricerca di minimo di un funzionale. In pratica, si permette a questa "quantità integrale" di variare e si impone la stazionarietà della derivata prima, ovvero la condizione di minimo. Questa condizione di minimo sarà l'equazione del moto del sistema descritto dalla "quantità integrale" cosiderata.

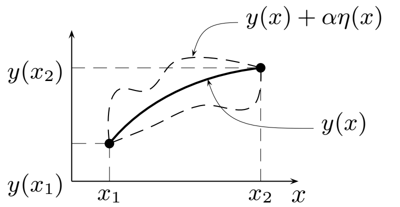
In meccanica, la "quantità integrale" è rappresentata dalla cosiddetta *azione*. In questo caso il principio variazionale prende il nome di *principio di minima azione*, ed asserisce che l'equazione del moto si ricava imponendo che l'azione abbia un estremo, ovvero che $\delta S = 0$. A tal fine in meccanica si definisce *integrale d'azione* il seguente funzionale della Lagrangiana:

$$S = \int_{t_0}^{t_1} L(x, \dot{x}, t)dt \tag{3.1}$$

Esporremo ora in maniera formalmente precisa il principio variazionale.

Sia $y(x)$ il percorso di una particella, con estremi x_1 e x_2. Supponiamo di parametrizzare una piccola variazione del percorso come segue:

$$\begin{aligned} y(x,\alpha) &= y(x) + \alpha\eta(x) \qquad && \alpha \text{ piccolo} \\ \eta(x_1) = \eta(x_2) &= 0 && y(x) = \left. y(x,\alpha)\right|_{\alpha=0} \end{aligned}$$



cioé consideriamo tutti i percorsi $y(x, \alpha)$ ottenibili tramite una piccola deformazione della curva (si considerano quindi una serie di cammini infinitamente vicini a quello che minimizza l'azione), mantenendo però gli estremi fissi ed imponendo che l'annullarsi del parametro α restituisca la curva

$y(x)$. Per fini pratici si può pensare $0 \le \alpha \le 1$. Allora il funzionale d'azione[1] assume la forma generica:

$$F \equiv F(x,\alpha) = \int_{x_1}^{x_2} f\left[y(x,\alpha); y'(x,\alpha); x\right] dx$$

e la condizione di minimo assume la forma:

$$\left.\frac{\partial}{\partial\alpha}F(x,\alpha)\right|_{\alpha=0} = 0$$

Tenendo presente la forma di $F(x,\alpha)$ e sviluppando la condizione di minimo:

$$\frac{\partial}{\partial\alpha}F(x,\alpha) = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial\alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial\alpha}\right] dx = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial\alpha} + \frac{\partial f}{\partial y'}\frac{\partial^2 y}{\partial x\partial\alpha}\right] dx$$

ed integrando per parti sulla variabile x:

$$\frac{\partial}{\partial\alpha}F(x,\alpha) = \left.\frac{\partial f}{\partial y'}\frac{\partial y}{\partial\alpha}\right|_{x_1}^{x_2} + \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y} - \frac{\partial}{\partial x}\frac{\partial f}{\partial y'}\right]\frac{\partial y}{\partial\alpha}dx$$

ponendo $\alpha = 0$ e tenendo presente che $\partial y/\partial\alpha = \eta(x)$ e che $\eta(x)$ si annulla agli estremi:

$$\frac{\partial}{\partial\alpha}F(x,\alpha) = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y} - \frac{\partial}{\partial x}\frac{\partial f}{\partial y'}\right]\left.\frac{\partial y}{\partial\alpha}\right|_{\alpha=0} dx$$

denotando $(\partial y/\partial\alpha)\delta\alpha|_{\alpha=0}$ con δy, si ricava:

$$\frac{\partial}{\partial\alpha}F(x,\alpha) = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y} - \frac{\partial}{\partial x}\frac{\partial f}{\partial y'}\right]\delta y dx$$

e dovendo valere $\delta F = 0$ indipendentemente dal valore di δy, deve risultare:

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0} \tag{3.2}$$

Questa è la forma generale delle equazioni del moto. Nel caso la funzione f coincida con la lagrangiana del sistema, $f(y(x), y'(x), x) \to L(q(t), \dot{q}(t), t)$, si riconosce facilmente in questa equazione l'equazione di Eulero-Lagrange, che è in effetti l'equazione del moto della meccanica:

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial\dot{q}} = 0$$

e per una particella libera $L = 1/2mv^2 \Rightarrow \frac{d}{dt}mv = 0$

Questa procedura è in realtà del tutto generale. Come esempio di applicazione, troveremo ora la funzione che rende minima la distanza fra due punti x_1 e x_2. In questo caso quindi il funzionale di azione è costituito dalla distanza, perché questa la quantità che vogliamo minimizzare per mezzo del principio:

$$F = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2}\sqrt{dx^2 + dy^2} = \int_{x_1}^{x_2}\sqrt{1 + y'^2}dx$$

e l'equazione (3.2) diventa:

$$-\frac{d}{dx}\frac{y'}{\sqrt{1+y'^2}} = 0 \quad\Rightarrow\quad \frac{y'}{\sqrt{1+y'^2}} = \text{cost.}$$

da cui risolvendo:

$$y' = \pm\frac{c}{1-c^2}$$

[1] Nella terminologia comune si parla di azione, o meglio di *funzionale di azione*, anche in casi più generali rispetto alla meccanica.

e cioé $y = ax + b$.([2]) Questa soluzione – quella che rende minima la distanza – è chiamata *geodetica*.

Quindi, *affinché il metodo variazionale fornisca un risultato (fisicamente) corretto è cruciale la ricerca di un appropriato funzionale di azione.*

Nel caso relativistico, l'azione per il punto libero deve rispettare l'invarianza relativistica e riprodurre per $c \to \infty$ l'andamento classico, deve essere perciò uno scalare.[3]

L'unico scalare disponibile è ds. Poniamo quindi che il funzionale di azione sia dato da:

$$S = \text{cost.} \times \int_1^2 ds$$

e vediamo se e come sia possibile riprodurre le leggi classiche del moto a partire da questa quantità. Innanzitutto, dalla definizione di tempo proprio si sa che $\frac{1}{c}\int ds$ è un massimo, dunque per renderlo esplicitamente un minimo si pone:

$$S = -k \int_1^2 ds$$

che servirà, fra l'altro, per riprodurre effettivamente il limite classico. Segue:

$$S = -k \int_1^2 ds = -k \int_{t_1}^{t_2} \sqrt{c^2 dt^2 - \vec{x}^2} = -k \int_{t_1}^{t_2} c\sqrt{1-\beta^2}dt \quad \Rightarrow L = -kc\sqrt{1-\beta^2}$$

che sulla base della definizione di integrale di azione ci permette di identificare la Lagrangiana come $L = -kc\sqrt{1-\beta^2}$. [4]

Il limite classico si deve ottenere per velocità $v \ll c$ e in questa ipotesi si può sviluppare la radice in termini di β:

$$\lim_{v\to 0} -kc\sqrt{1-\beta^2} \simeq -kc + \frac{1}{2}k\frac{v^2}{c}$$

Il limite classico si ritrova appunto ponendo $k = mc$, a meno di un fattore costante. In ogni caso, questo fattore costante non crea problemi nella scrittura delle equazioni del moto a partire dalla Lagrangiana, in quanto questa compare sempre sotto segno di derivazione. Si può quindi concludere che la Lagrangiana relativistica per un punto materiale è data da:

$$L = -mc^2\sqrt{1-\beta^2}$$

La quantità di moto è data dalla derivata della Lagrangiana rispetto alla variabile coniugata di $\vec{x}$:

$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{x}}} \quad \Rightarrow \quad p_i = \frac{\partial L}{\partial \dot{x}_i} = +mc^2 \frac{\frac{2v_i}{c^2}}{2\sqrt{1-\beta^2}} = \frac{mv_i}{\sqrt{1-\beta^2}}$$

si vede quindi che rispetto alla definizione classica di quantità di moto compare un fattore moltiplicativo $1/\sqrt{1-\beta^2}$. Di conseguenza, per un corpo in moto con il termine *massa* ci si deve riferire in realtà all'intero termine $m/\sqrt{1-\beta^2}$.

Per la quarta componente, che rappresenta l'energia:

$$\begin{aligned} E &= p_i\dot{x}_i - L = p_i\dot{x}_i + mc^2\sqrt{1-\beta^2} = \\ &= \frac{m\dot{x}^2}{\sqrt{1-\beta^2}} + mc^2\sqrt{1-\beta^2} = \frac{1}{\sqrt{1-\beta^2}}\left[mv^2 + mc^2 - mv^2\right] \end{aligned}$$

ovvero in definitiva:

$$E = \frac{mc^2}{\sqrt{1-\beta^2}}$$

[2]Un modo dispendioso per ritrovare il fatto che la distanza minima fra due punti è una retta.

[3]Infatti, l'azione classica è uno scalare.

[4]Si noti che questo non è rigorosamente corretto da un punto di vista formale: si è isolato il tempo dalle altre coordinate, perdendo così la covarianza a vista. Si parla di *covarianza a vista* quando in un'equazione le quattro coordinate spazio-temporali appaiono formalmente tutte sullo stesso piano. Quando questo non accade si dice che si è persa la covarianza a vista.

Il termine $m' = m/(\sqrt{1-\beta^2})$ prende il nome di *massa relativistica*.

Vale la pena notare che nel limite classico di $c \to \infty$ si ottiene ancora una volta:

$$\lim_{v\to 0} \frac{mc^2}{\sqrt{1-\beta^2}} \simeq mc^2\left(1+\frac{1}{2}\frac{v^2}{c^2}\right) = mc^2 + \frac{1}{2}mv^2$$

ovvero l'energia relativistica comprende, oltre al termine classico di energia cinetica, un termine aggiuntivo presente anche per masse non in moto. Questo termine deve essere considerato pertanto come una "energia a riposo" dovuta al semplice fatto di possedere una massa ed è pari a:

$$E = mc^2$$

La massa è quindi una forma di energia e deve essere possibile trasformare l'una nell'altra.[5]

Ricordando la definizione di quadrivelocità (2.1) si può scrivere:

$$\begin{cases} E &= -imc^2u_4 \\ \vec{p} &= cm\vec{u} \end{cases}$$

e questo permette di definire un quadrivettore *quadrimpulso*:

$$p^\alpha \equiv (\vec{p}, i\frac{E}{c}) \qquad p^\alpha = mcu^\alpha \tag{3.3}$$

Le trasformazioni di Lorentz "mescolano" quindi le componenti dell'energia e dell'impulso. È immediato ricavare la relazione:

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

che resta valida anche per particelle senza massa, nel qual caso fornisce $\vec{p} = E/c$.([6])

Il calcolo del modulo di p^α a partire dalle sue definizioni (3.3) fornisce:

$$p_\alpha p^\alpha = -m^2c^2 \qquad\qquad p_\alpha p^\alpha = -\frac{E^2}{c^2} + p^2$$

che eguagliate permettono di ricavare l'importantissima relazione della relatività ristretta:

$$\boxed{c^2p^2 - E^2 = -m^2c^4 \qquad (E^2 = p^2c^2 + m^2c^4)} \tag{3.4}$$

Il segno meno viene dalla particolare scelta della segnatura della metrica $+++-$.

[5] È interessante notare che nel momento in cui scriviamo $m/\sqrt{1-\beta^2}$ stiamo dicendo che aumentando la velocità del corpo in realtà aumenta *anche* la sua massa, più precisamente aumentando l'energia di un corpo se ne aumenta sia la velocità sia la massa. Il modo in cui questa diverge all'avvicinarsi alla velocità della luce è tale che questa costituisce un limite invalicabile perchè la massa del corpo tende a infinito quando la sua velocità tende a c (matematicamente, si tratta di un punto di discontinuità di seconda specie). Occorre quindi un'energia infinita perché un corpo dotato di massa possa raggiungere la velocità della luce.

[6] Occorre notare qui una finezza concettuale: in teoria della relatività ristretta abbiamo visto che la "massa" è sostanzialmente l'energia a riposo di un corpo ($E = mc^2$), ed è invariante in tutti i sistemi di riferimento inerziali. Una volta introdotto il concetto di quadrimpluso – che può essere derivato direttamente a partire dai postulati della teoria ed applicando il metodo variazionale – il concetto di "massa relativistica" $m/\sqrt{1-\beta^2}$ diventa in pratica superfluo. In questo senso diventa scorretto dire che "la massa di un corpo aumenta con la sua velocità", che deriva invece dall'aver esteso in maniera impropria il concetto classico di impulso. In effetti, la quantità relativisticamente invariante è il quadrimpulso $\vec{p} = m\vec{v}/(\sqrt{1-\beta^2})$, che può essere interpretato in due modi:

$$\vec{p} = \left(\frac{m}{\sqrt{1-\beta^2}}\right)\vec{v} \qquad \text{e} \qquad \vec{p} = \frac{1}{\sqrt{1-\beta^2}}(m\vec{v})$$

La differenza concettuale delle due interpretazioni è che nel primo caso si sta dicendo che la quantità di moto relativistica coincide con quella classica, a patto di introdurre una nuova massa (quella relativistica) che dipende dalla velocità, nel secondo caso si dice invece che la quantità di moto classica dipende dal particolare sistema di riferimento in cui viene misurata e pertanto la sua forma trasforma secondo le leggi di Lorentz, mentre la massa della particella (energia a riposo) è invariante e resta la stessa in ogni sistema.

L'equazione di Lagrange è data $\frac{d\vec{p}}{dt} = 0$, ma è un'equazione non covariante ed è solo su tre componenti. Per ricavare le equazioni del moto in forma quadridimensionale, consideriamo che $ds = \sqrt{-dx_\alpha dx^\alpha}$ e:

$$\delta S = -mc\int_1^2 \delta ds$$
$$\delta ds = \frac{(-2dx^\beta)\delta dx_\beta}{2\sqrt{-dx_\alpha dx^\alpha}} = -\frac{dx^\beta}{ds}\delta dx_\beta = -u^\beta \delta dx_\beta$$

da cui risulta:

$$\delta S = +mc\int_1^2 u^\beta \delta dx_\beta = \cancelto{0}{mcu^\beta \delta x_\beta\Big|_1^2} - mc\int_1^2 \frac{du^\beta}{ds}\delta x_\beta ds$$

avendo scambiato variazioni con integrazioni e integrato per parti nell'ultimo passaggio. La forma quadridimensionale covariante a vista delle equazioni del moto è pertanto:

$$mc\frac{du^\beta}{ds} = 0 \tag{3.5}$$

che infatti per componenti fornisce:

$$\begin{cases} mc\dfrac{d\vec{u}}{ds} = \dfrac{mc}{c\sqrt{1-\beta^2}}\dfrac{d\vec{u}}{dt} = \dfrac{m}{\sqrt{1-\beta^2}}\dfrac{d}{dt}\dfrac{\vec{v}}{\sqrt{1-\beta^2}} \Rightarrow \dfrac{d\vec{p}}{dt} & = 0 \\ \dfrac{mc}{c\sqrt{1-\beta^2}}\dfrac{d}{dt}\dfrac{i}{\sqrt{1-\beta^2}} & = 0 \end{cases}$$

Considerando il punto 1 fissato e il punto 2 generico, l'azione S sarà una funzione dell'estremo di integrazione:

$$\delta S(x) = mcu^\beta \delta x_\beta$$

e questo consente di definire:

$$\frac{\delta S(x)}{\delta x_\beta} = mcu^\beta$$

che corrisponde all'impulso classico. Infatti, esplicitando le componenti:

$$\begin{cases} \dfrac{\partial S}{\partial t} & = -E \\ \dfrac{\partial S}{\partial x_i} & = p_i \end{cases}$$

e:

$$\left(\frac{\partial S}{\partial x_i}\right)^2 + \left(\frac{\partial S}{\partial x_4}\right)^2 = -m^2c^2$$

che è equivalente a:

$$\left(\frac{\partial S}{\partial x_i}\right)^2 - \frac{1}{c^2}\left(\frac{\partial S}{\partial t}\right)^2 + m^2c^2 = 0$$

che per $c \to \infty$ si riduce all'equazione Hamilton–Jacobi, a parte il solito termine m^2c^2. Ci si può tuttavia liberare di questo fattore ridefinendo S' in modo che risulti:

$$\frac{\partial S'}{\partial t} = -E + m^2c^2$$

imponendo in pratica che:

$$S' = S + mc^2 t \tag{3.6}$$

3.2 Sviluppi relativi alla Lagrangiana ed all'Hamiltoniana

Abbiamo quindi visto che nella formulazione quadridimensionale risulta:

$$mc\frac{du^\alpha}{ds} \equiv \frac{dp^\alpha}{ds} = 0$$

ora:

$$S = \int_1^t L dt \Rightarrow S \equiv S(q(t), \dot{q}(t))$$

$$\delta S = \int_1^t \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} \right) dt$$

che, integrando il secondo termine per parti, fornisce:

$$\delta S = \frac{\partial L}{\partial \dot{q}}\delta q \bigg|_1^t + \int_1^t \left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right] \delta q dt$$

Notando ora che variare l'estremo superiore dell'integrale equivale a muoversi lungo un moto (in quanto sotto integrale abbiamo l'equazione del moto stesso), l'integrando risulta in realtà nullo proprio perché corrisponde all'equazione di Eulero-Lagrange. Quindi:

$$\delta S = \frac{\partial L}{\partial \dot{q}}\delta q$$

siccome la variazione δS può essere scritta anche banalmente come:

$$\delta S = \frac{\partial L}{\partial q}\delta q$$

ne consegue che:

$$p = \frac{\partial L}{\partial q}$$

Per l'energia risulta:

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial q}\dot{q}$$

e siccome per definizione $dS/dt = L$, si ricava l'importante relazione:

$$L = \frac{\partial S}{\partial t} + p\dot{q} \rightarrow \frac{\partial S}{\partial t} = L - p\dot{q} = -H$$

da cui:

$$E = -\frac{\partial S}{\partial t} \quad \text{ovvero} \quad \frac{\partial S}{\partial t} = -H$$

che permette proprio di ritrovare le equazioni di Hamilton-Jacobi.

Nel caso relativistico, $p_\alpha p^\alpha = -mc^2$ e $\vec{p}^2 = -E^2/c^2 + m^2c^2 = 0$, quindi:

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 - \frac{1}{c^2}\left(\frac{\partial S}{\partial t}\right)^2 + m^2c^2 = 0$$

che è una forma quadratica in d/dt. Per ritrovarsi nel limite classico bisogna allora imporre, come mostrato precedentemente, la (3.6), infatti:

$$\frac{\partial S}{\partial t} = \frac{\partial S'}{\partial t} - m^2c^2$$

mentre per quanto riguarda le derivazioni rispetto alle coordinate spaziali, S e S' sono uguali:

$$\left(\frac{\partial S'}{\partial x}\right)^2 + \left(\frac{\partial S'}{\partial y}\right)^2 + \left(\frac{\partial S'}{\partial z}\right)^2 - \frac{1}{c^2}\left(\frac{\partial S'}{\partial t}\right)^2 - m^2c^2 + 2m\frac{\partial S}{\partial t} + m^2c^2 = 0$$

che divisa per $2m$ fornisce:

$$\frac{p^2}{2m} - \frac{1}{c^2}\left(\frac{\partial S'}{\partial t}\right)^2 + \frac{\partial S}{\partial t} = 0$$

che nel limite classico restituisce l'equazione di Hamilton-Jacobi $\frac{p^2}{2m} + \frac{\partial S}{\partial t} = 0$.

Da questa e dalla definizione di quadrimpulso si può definire una forza in maniera banale:

$$p^\alpha = mc\frac{dx^\alpha}{ds} \qquad\qquad \frac{dp^\alpha}{ds} = 0$$

da cui la definizione:

$$F^\alpha = \frac{dp^\alpha}{ds} \tag{3.7}$$

Le componenti spaziali di questo quadrivettore sono:

$$F_i = \frac{dp_i}{ds} = \frac{1}{c\sqrt{1-\beta^2}}\frac{d}{dt}\frac{m\vec{v}}{\sqrt{1-\beta^2}}$$

dove la "forza classica"[7] sarebbe:

$$f_i = \frac{d}{dt}\frac{m\vec{v}}{\sqrt{1-\beta^2}} \equiv \frac{d}{dt}m'\vec{v}$$

per cui:

$$F_i = \frac{1}{c\sqrt{1-\beta^2}}f_i$$

Per trovare invece la componente temporale di $\frac{d}{ds}mcu^\alpha$ si calcola il prodotto scalare con u_α e si sfrutta il fatto che u^α e du^α/ds sono ortogonali. Si deduce quindi che $F^\alpha u_\alpha = F_i u_i + F_4 u_4 = 0$, ovvero deve accadere che:

$$F_4 \cdot \frac{i}{\sqrt{1-\beta^2}} = -\frac{1}{c\sqrt{1-\beta^2}} \cdot f_i \cdot \frac{v_i}{c\sqrt{1-\beta^2}}$$

cioé in definitiva:

$$\boxed{\begin{aligned} F_i &= \frac{1}{c\sqrt{1-\beta^2}}f_i \\ F_4 &= i\frac{\vec{f}\cdot\vec{v}}{c^2\sqrt{1-\beta^2}} \end{aligned}}$$

In relazione alle equazioni del moto, le componenti spaziali danno $\vec{F} = m\vec{a}$, mentre la quarta componente fornisce:

$$i\frac{\vec{f}\cdot\vec{v}}{c^2\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\beta^2}}\frac{1}{c}\frac{d}{dt}\left(\frac{imc^2}{\sqrt{1-\beta^2}}\right)$$

ovvero:

$$\vec{f}\cdot\vec{v} = \frac{d}{dt}\left(\frac{imc^2}{\sqrt{1-\beta^2}}\right)$$

ma $\vec{f}\cdot\vec{v}$ è il lavoro per unità di tempo, e quindi rappresenta il teorema delle forze vive in forma relativistica:

$$\frac{dT}{dt} = \frac{dE}{dt}$$

Pertanto, riassumendo, la (3.7) contiene sia la seconda legge della dinamica (nelle componenti spaziali) che il teorema delle forse vive (nella componente temporale).

[7]Classica è messo fra virgolette perché è la forma della forza che ci si attenderebbe da una estensione del concetto classico, inserendo semplicemente la definizione di massa relativistica al posto di quella classica.

3.3 Formulazione dell'interazione

Aggiungiamo ora un altro tassello alla costruzione teorica che stiamo conducendo qui. Si tratta di introdurre in maniera formale l'interazione elettromagnetica. La base di partenza sarà costituita dalle relazioni:

$$S = -mc\int ds = \int L_0 dt \qquad\qquad L_0 = -mc^2\sqrt{1-\beta^2}$$

dove L_0 rappresenta la lagrangiana libera senza interazione. Come è noto, l'interazione è descritta da un termine di potenziale $L = L_0 - U$. Nell'ipotesi che $U \equiv U(q)$ non dipenda da $\dot{q}$, l'equazione di Lagrange diventa:

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial (L_0 - U)}{\partial q} - \frac{d}{dt}\frac{\partial (L_0 - U)}{\partial \dot{q}} = -\frac{\partial U}{\partial q} - \frac{d}{dt}\frac{\partial L_0}{\partial \dot{q}} = 0$$

ovvero:

$$\frac{dp}{dt} = -\frac{\partial U}{\partial q}$$

In elettrodinamica d'altra parte è possibile definire un potenziale $U(q, \dot{q})$ per descrivere la forza di Coulomb $\vec{F} = e\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right)$ che dipende dalla velocità, e che implica quindi un termine aggiuntivo:

$$\frac{dp}{dt} = -\frac{\partial U}{\partial q} + \frac{d}{dt}\frac{\partial U}{\partial \dot{q}}$$

in cui il potenziale corrispondente alla forza di Coulomb è:

$$U = e\left(\Phi + \frac{1}{c}\vec{v}\cdot\vec{A}\right)$$

Questo fa comparire nell'azione un termine di interazione Udt che deve (e può) essere messo in forma covariante:

$$Udt = e\left(\Phi dt + \frac{1}{c}\frac{d\vec{x}}{dt}\cdot\vec{A}dt\right)$$

Ricordando che $x^\alpha = (\vec{x}, ict)$, si può definire direttamente un quadrivettore *quadripotenziale* $A_\alpha \equiv (\vec{A}, i\Phi)$ tale che il termine di interazione possa essere scritto in una forma covariante a vista:[8]

$$Udt = e\left[-(i\Phi)\frac{icdt}{c} - \frac{1}{c}d\vec{x}\cdot\vec{A}\right] \rightarrow Udt = -\frac{e}{c}A_\alpha dx^\alpha$$

questo termine può essere inserito nell'azione in forma covariante:

$$\boxed{S = \int_1^2 \left[-mc\, ds + \frac{e}{c}A_\alpha dx^\alpha\right]} \tag{3.8}$$

Notare che siccome il termine di interazione è uno scalare, esso è invariante per trasformazioni.

Secondo il principio di minima azione, per ricavare ora le equazioni del moto di una particella carica in un campo elettromagnetico si deve imporre $\delta S = 0$:[9]

$$\begin{aligned}\delta S &= \int_1^2 \left[-mc\delta ds + \frac{e}{c}\frac{\partial A_\alpha}{\partial x^\beta}\delta dx^\beta dx^\alpha + \frac{e}{c}A_\alpha \delta dx^\alpha\right] = \\ &= \int_1^2 \left[\frac{e}{c}\frac{\partial A_\alpha}{\partial x^\beta}dx^\alpha \delta dx^\beta - \frac{mc}{2}\frac{2}{ds}dx^\alpha \delta dx_\alpha + \frac{e}{c}A_\alpha \delta dx^\alpha\right] =\end{aligned}$$

[8]Per essere formalmente precisi, dovremmo dimostrare che la quantità $A_\alpha = (\vec{A}, i\Phi)$ trasformi effettivamente secondo la matrice di Lorentz data in §2.1. Non ne faremo tuttavia la dimostrazione qui e prenderemo per buono il risultato.

[9]Ricordiamo che $ds = \sqrt{dx_\alpha dx^\alpha}$, quindi $\delta ds = \frac{1}{2\sqrt{dx_\alpha dx^\alpha}}2dx^\alpha \delta dx_\alpha$.

integrando ora per parti il II° ed il III° pezzo e considerando che le variazioni sono nulle agli estremi:

$$= \int_1^2 \left[\frac{e}{c}\frac{\partial A_\alpha}{\partial x^\beta} dx^\alpha \delta dx^\beta - mc\, du^\alpha \delta x_\alpha - \frac{e}{c} dA_\alpha \delta x^\alpha \right] =$$
$$= \int_1^2 \left[\frac{e}{c}\frac{\partial A_\beta}{\partial x^\alpha} dx^\beta - mc\, du^\alpha - \frac{e}{c} dA_\alpha \right] \delta x^\alpha =$$

(nell'ultimo passaggio si sono scambiati gli indici α e β del primo pezzo, in quanto muti). Segue:

$$= \int_1^2 \left[\frac{e}{c}\frac{\partial A_\beta}{\partial x^\alpha}\frac{dx^\beta}{ds} - mc\frac{du^\alpha}{ds} - \frac{e}{c}\frac{\partial A_\alpha}{\partial x^\beta}\frac{dx^\beta}{ds} \right] \delta x^\alpha\, ds =$$
$$= \int_1^2 \left[-mc\frac{du^\alpha}{ds} + \frac{e}{c}\left(\frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta} \right)\frac{dx^\beta}{ds} \right] \delta x^\alpha\, ds$$

dove abbiamo espresso fra l'altro il differenziale totale dA_α come $\frac{\partial A_\alpha}{\partial x^\beta}\frac{dx^\beta}{ds}$. Ora, questa equazione deve valere per ogni variazione δx^α, pertanto deve essere nullo l'integrando:

$$mc\frac{du^\alpha}{ds} = \frac{e}{c} F_{\alpha\beta} u^\beta$$

dove è stato definito il *tensore elettromagnetico* $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$. La quarta componente restituisce anche in questo caso il teorema delle forze vive. L'impulso può ora essere ricavato come segue:

$$p^\alpha = \frac{\partial S}{\partial \dot{q}_\alpha} \quad \to \quad \vec{P} = \vec{p} - \frac{e}{c}\vec{A}$$

che è la forma invariante di Gauge. La trasformazione stessa di Gauge è "ambigua", infatti può essere aggiunto un termine scalare tale che:

$$A_\mu \to A_\mu + \partial_\mu \Lambda \Rightarrow \begin{cases} \vec{A} \to \vec{A} + \vec{\nabla}\Lambda \\ \Phi \to \Phi - \dfrac{\partial \Lambda}{\partial t} \end{cases}$$

La trasformazione di Gauge dell'impulso aggiunge quindi all'azione il termine:

$$\int_1^2 \frac{e}{c}\partial_\alpha \Lambda dx^\alpha$$

che è però un differenziale totale, scrivibile quindi nella forma $\Lambda_2 - \Lambda_1$: pertanto non contribuisce nella variazione dell'azione.[10] In maniera analoga si ricava $E = \frac{\vec{p}^2}{2m}$.

Notiamo che il gradiente in forma quadrimensionale assume l'aspetto $\partial_\alpha \equiv \left(\vec{\nabla}, \frac{1}{ic}\frac{\partial}{\partial t}\right)$. Ne consegue che:

$$\partial_\alpha \partial^\alpha = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \equiv \Box$$

L'operatore $\Box$ prende il nome di *Dalambertiano.*

Troviamo ora le componenti del tensore $F_{\alpha\beta}$. Ricordando inannzitutto che:

$$\vec{B} = \vec{\nabla} \times \vec{A} \qquad \vec{E} = -\vec{\nabla}\Phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}$$

calcoliamo esplicitamente alcune componenti:

$$F_{12} = \frac{\partial}{\partial x_1}A_2 - \frac{\partial}{\partial x_2}A_1 = \frac{\partial}{\partial x}A_y - \frac{\partial}{\partial y}A_x = (\vec{\nabla} \times \vec{A})_z \equiv B_z$$
$$F_{13} = \frac{\partial}{\partial x_1}A_3 - \frac{\partial}{\partial x_3}A_1 = \frac{\partial}{\partial x}A_z - \frac{\partial}{\partial z}A_x = -(\vec{\nabla} \times \vec{A})_x \equiv -B_y$$
$$F_{14} = \frac{\partial}{\partial x_1}A_4 - \frac{\partial}{\partial x_4}A_1 = \frac{\partial}{\partial x}i\Phi - \frac{1}{ic}\frac{\partial}{\partial t}A_x = -i\left(-\vec{\nabla}\Phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \right)_x \equiv -iE_x$$

[10] Questo aspetto è legato alla costanza della carica elettrica.

in definitiva la forma del tensore risulta la seguente:

$$F_{\alpha\beta} = \begin{pmatrix} 0 & B_z & -B_y & -iE_x \\ -B_z & 0 & B_x & -iE_y \\ B_y & -B_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{pmatrix} \tag{3.9}$$

come si può notare, il tensore è antisimmetrico e dipende solo dai campi (e quindi non dai potenziali). Essendo un tensore del II° ordine, esso trasforma secondo la legge: $F'_{\alpha\beta} = \Gamma^{\mu}_{\alpha}\Gamma^{\nu}_{\beta}F_{\mu\nu}$.

Studiamo ora più in dettaglio le trasformazioni delle varie componenti applicando la matrice di Lorentz data in §2.1. Se nella trasformazione non sono coinvolti gli indici 1 e 4 (ovvero solo per le componenti $F_{23} = -F_{32}$), risulta $F_{23} = F'_{23}$, e le componenti non si trasformano. Le componenti miste F_{12}, F_{13}, F_{42}, F_{43} trasformano invece come dei quadrivettori. Le componenti $F_{14} = -F_{41}$ si vede che non trasformano ($F_{14} = F'_{14}$). Infine, $F_{11} = F_{22} = F_{33} = F_{44} = 0$.

A partire dal tensore $F_{\alpha\beta}$ si possono costruire due quantità invarianti. La prima è:

$$F^2 = F_{\alpha\beta}F^{\alpha\beta} = 2B^2 - 2E^2 = 2(B^2 - E^2)$$

quindi, sebbene per trasformazioni $F_{\alpha\beta}$ vari (e varino quindi i campi $\vec{E}$ e $\vec{B}$), la differenza $B^2 - E^2$ è costante in ogni sistema di riferimento.

L'altra quantità invariante è legata al tensore duale di $F_{\alpha\beta}$:

$$\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\alpha\beta}$$

si tratta ancora di un tensore antisimmetrico in μ e ν, ed è un tensore in cui $\vec{E}$ e $\vec{B}$ risultano scambiati:[11]

$$\tilde{F}_{12} = \frac{1}{2}\epsilon_{1234}F^{34} - \frac{1}{2}\epsilon_{1243}F^{43} = 2F_{34}$$

cioé la componente E_z va in B_z e viceversa. Siccome risulta:

$$\tilde{F}_{\alpha\beta}F^{\alpha\beta} \propto \vec{E}\cdot\vec{B}$$

anche $\vec{E}\cdot\vec{B}$ è costante in ogni sistema di riferimento.

3.4 Riscrittura covariante delle equazioni di Maxwell

Abbiamo visto quindi nella (3.8) che l'azione di una particella libera e soggetta alla forza elettromagnetica è data da:

$$S = \int_1^2 \left[-mc\,ds + \frac{e}{c}A_\alpha dx^\alpha\right]$$

e che risulta:

$$\delta S = 0 \quad \Rightarrow \quad mc\frac{du^\alpha}{ds} = \frac{e}{c}F_{\alpha\beta}u^\beta$$

Mostreremo ora che le equazioni di Maxwell sono facilmente riscrivibili in forma covariante in termini del tensore $F_{\alpha\beta}$.

In realtà potremo spingerci ben oltre: arriveremo infatti a mostrare che le equazioni di Maxwell sono proprio le equazioni del moto del campo elettromagnetico. Per mostrare ciò, però, sarà necessario trovare una azione "appropriata" che tenga conto del fatto che campo elettromagnetico è un sistema lagrangiano ad infiniti gradi di libertà. Questo ci porterà nel capitolo §4 a studiare prima un formalismo adatto per i sistemi continui per poi giungere finalmente ad un termine per l'azione per il campo elettromagnetico in §4.3.

[11]Si noti en passant la relazione fra il fatto che il tensore duale rappresenta la superficie ortogonale a quella del tensore dato ed il fatto che i campi $\vec{E}$ e $\vec{B}$ sono ortogonali fra loro.

Iniziamo quindi con l'esplicitare la forma covariante delle equazioni di Maxwell senza le sorgenti:

$$\vec{\nabla} \cdot \vec{B} = 0 \qquad\qquad \vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$$

Dalla definizione del tensore elettromagnetico (3.9), queste possono essere riscritte facilmente come:

$$\frac{\partial F_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial F_{\beta\gamma}}{\partial x^\alpha} + \frac{\partial F_{\gamma\alpha}}{\partial x^\beta} = 0$$

notiamo ora che questa è una combinazione completamente antisimmetrica: il tensore $F_{\mu\nu}$ è antisimmetrico e se due indici sono uguali, allora si riduce all'identità $0 = 0$. Questa relazione può quindi essere messa in forma compatta utilizzando il tensore completamente antisimmetrico $\epsilon_{\alpha\beta\gamma\delta}$, assumendo la forma:

$$\boxed{\epsilon_{\alpha\beta\gamma\delta}\, \partial^\beta F^{\gamma\delta} = 0} \tag{3.10}$$

Notiamo anche che introducendo il tensore duale $\tilde{F}^{\alpha\beta}$, queste possono essere scritte come quadridivergenza di questo tensore.

Consideriamo ora invece le due equazioni che contengono le sorgenti:

$$\vec{\nabla} \cdot \vec{E} = 4\pi\varrho \qquad\qquad \vec{\nabla} \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c}\vec{\jmath}$$

dove, poiché si parla di punti materiali, le densità sono definite come segue:

$$\begin{cases} \varrho(\vec{r}) & = \displaystyle\sum_{i=1}^{n} e_i \delta(\vec{r} - \vec{r}_i) \\ \vec{\jmath} = \varrho\vec{v} & = \displaystyle\sum_{i=1}^{n} e_i \delta(\vec{r} - \vec{r}_i)\vec{v}_i \end{cases}$$

In questo caso, l'equazione di continuità fornisce:

$$\frac{\partial \varrho}{\partial t} = \sum_{i=1}^{n} \frac{\partial \varrho}{\partial \vec{r}_i} \cdot \frac{\partial \vec{r}_i}{\partial t} = -\sum_{i=1}^{n} \frac{\partial \varrho}{\partial \vec{r}_i} \cdot \vec{v}_i = -\sum_{i=1}^{n} \left(\vec{\nabla} \cdot \varrho\right) \cdot \vec{v}_i = -\sum_{i=1}^{n} \vec{\nabla} \cdot (\varrho \vec{v}_i)$$

da cui, se $\vec{v}_i$ non dipende da $\vec{r}_i$ discende la classica equazione di continuità:

$$\frac{\partial \varrho}{\partial t} + \vec{\nabla} \cdot \vec{\jmath} = 0$$

ora, considerando che per una carica infinitesima vale $de = \varrho dV$ e che de deve essere invariante di Lorentz,[12] ne risulta che la densità ϱ deve trasformare come:

$$de\, dx^\alpha = \varrho dV dx^\alpha = \varrho dV \frac{dx^\alpha}{dt} dt = \varrho \frac{1}{ic}\frac{dx^\alpha}{dt} d(icdt) = \varrho \frac{dx^\alpha}{dx^4} d\Omega$$

$\varrho\frac{dx^\alpha}{dt}$ deve pertanto trasformare come un quadrivettore, il che permette di definire il *quadrivettore corrente*:[13]

$$j^\mu \equiv (\vec{\jmath}, ic\varrho)$$

L'equazione di continuità diventa in questo caso semplicemente:

$$ic\frac{\partial \varrho}{\partial ict} + \vec{\nabla} \cdot \vec{\jmath} = 0$$

ovvero in forma covariante a vista:

$$\boxed{\partial_\mu j^\mu = 0}$$

[12] La carica non può cambiare se si cambia sistema di riferimento.

[13] Ancora una volta, per essere formalmente precisi dovremmo dimostrare che la quantità $j^\mu = (\vec{\jmath}, ic\varrho)$ trasforma realmente secondo la matrice di Lorentz. Ancora una volta, qui prenderemo per buono il risultato.

Anche le equazioni di Maxwell con le correnti possono allora essere scritte nella forma covariante a vista:

$$\boxed{\partial_\nu F^{\mu\nu} = \frac{4\pi}{c} j^\mu} \tag{3.11}$$

che insieme alla (3.10) costituiscono le equazioni di Maxwell in forma covariante.

Occorre ora trovare una forma per l'azione completa da associare al campo elettromagnetico.

Siccome $e = \int \varrho dV$, il termine di interazione elettromagnetica contenuto nella (3.8) si può riscrivere nella forma:

$$\int \frac{e}{c} A_\mu dx^\mu = \iint \frac{\varrho(\vec{r})}{c} dV A_\mu dx^\mu = \iint \frac{\varrho(\vec{r})}{ic^2} A_\mu \frac{dx^\mu}{dt} d(ict) dV = \int \frac{1}{ic^2} A_\mu j^\mu d\Omega$$

L'azione ottenuta ora per una particella libera in un campo elettromagnetico comprende il termine di moto della particella ed il termine di interazione della particella con il campo. L'azione completa – che ci permetterà di ricavare completamente le equazioni del moto – dovrà includere anche un termine che coinvolge il *campo in sé*, ovvero un termine che restituisca le equazioni di Maxwell come equazioni del moto. L'azione finale sarà quindi costituita da tre pezzi che permettono di ritrovare contemporaneamente, una volta applicato il metodo variazionale, le equazioni del moto della particella, le equazioni di Maxwell ed il termine di interazione fra campo e particella.

Come detto più sopra, per ottenere ciò bisogna però prima dire qualcosa sui sistemi continui.

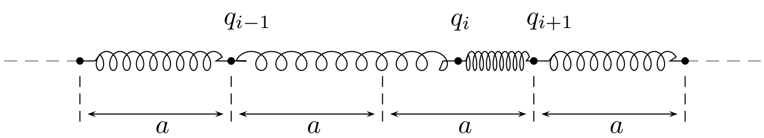
Capitolo 4

Formalismo per Sistemi Continui. Azione del campo elettromagnetico

4.1 Formalismo Lagrangiano

Lo scopo di questo paragrafo è di partire da un sistema discreto (avente anche una utilità pratica) e vedere come si modifica il formalismo lagrangiano quando se ne fa il limite al continuo. Questo ci permetterà di impostare il formalismo necessario per trattare il campo elettromagnetico, che è matematicamente equivalente ad un sistema ad infiniti gradi di libertà.

Il sistema utilizzato qui come base di partenza è la cosiddetta *catena lineare*. Si prenda un insieme di N particelle accoppiate da molle, la cui distanza di riposo sia a. L'eccitazione è rappresentata dallo spostamento di una particella dall'equilibrio, che si ripercuote sull'intero sistema.



Diciamo q_i lo spostamento dell'i-sima particella dall'equilibrio. Allora supponendo che le masse siano tutte uguali e che $q_{i+N} = q_i$ (condizione di periodicità, gli estremi si identificano), la Lagrangiana è data da:

$$L = T - U \qquad\qquad T = \sum_{i=1}^{N} \frac{1}{2} m\dot{q}_i^2$$

il termine di energia cinetica non pone problema, essendo la somma su tutti gli elementi della catena di un termine $1/2m\dot{q}^2$. Il termine di energia potenziale, notando che $\Delta x = q_i - q_{i-1}$, può invece essere scritto nella forma:

$$U = \sum_{i=1}^{N} \frac{1}{2} k(q_{i+1} - q_i)^2$$

e dove i singoli termini sono della forma:

$$U_i = \frac{1}{2} k(q_{i+1} - q_i)^2 + \frac{1}{2} k(q_i - q_{i-1})^2$$

d'altra parte:

$$-F_i = \frac{\partial U}{\partial q_i} = -k(q_{i+1} - q_i) + k(q_i - q_{i-1})$$

sono proprio le forze agenti sul singolo oscillatore q_i.[1] Allora la lagrangiana risulta:

$$L = \sum_{i=1}^{N} \left[\frac{1}{2} m \dot{q}_i^2 - \frac{1}{2} k (q_{i+1} - q_i)^2 \right]$$

e le equazioni del moto di ogni singolo oscillatore sono date dalla serie di N equazioni di Lagrange:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \qquad \forall i \in 1, \ldots, N$$

e infatti, esplicitando le equazioni di Lagrange si ottiene la classica $F_i = ma_i$:

$$\begin{cases} \dfrac{\partial L}{\partial q_i} & = k(q_{i+1} - q_i) - k(q_i - q_{i-1}) \\ \dfrac{d}{dt} \dfrac{\partial L}{\partial \dot{q}_i} & = m \dfrac{d}{dt} \dot{q}_i = ma_i \end{cases} \Rightarrow k(q_{i+1} - q_i) - k(q_i - q_{i-1}) = ma_i$$

Ora passiamo dal formalismo discreto a quello continuo. Fare questo significa mandare la distanza fra le particelle a zero $a \to 0$ in modo che risulti $Na \to l$, con l lunghezza della catena lineare. La sommatoria si trasforma dunque in un integrale: $\sum \cdot a \to \int dx$. Riscriviamo innanzitutto la lagrangiana in una forma più adatta per effettuare questo limite:

$$L = \sum_{i=1}^{N} a \left[\frac{1}{2} \frac{m}{a} \dot{q}_i^2 - \frac{1}{2} \frac{ka}{a^2} (q_{i+1} - q_i)^2 \right]$$

quindi eseguiamo il limite $a \to 0$ e denotiamo con $m/a \equiv \mu$ la densità lineare e con $ka \equiv Y$ il modulo di elasticità Young.

Per quanto riguarda il termine $(q_{i+1} - q_i)^2/a^2$, se $a \to 0$, allora $q_{i+1} \to q_i$. Il numero di elementi q_i per unità di lunghezza quindi si infittisce, tendendo ad infinito. In pratica, si tende ad una funzione continua di x: $q_i \to q(x)$. In termini formali, il termine diviene un vero e proprio rapporto incrementale e restituisce quindi una derivata:

$$\lim_{a \to 0} \frac{(q_{i+1} - q_i)^2}{a^2} = \left[\frac{\partial}{\partial x} q(x) \right]^2$$

e quindi la Lagrangiana si trasforma nel modo seguente:

$$L = \int_0^l \left[\frac{1}{2} \mu \dot{q}^2(x) - \frac{1}{2} Y \left(\frac{\partial q(x)}{\partial x} \right)^2 \right] dx$$

ovvero:

$$L = \int_0^l \mathcal{L}(x) dx \tag{4.1a}$$

$$\mathcal{L}(x) \equiv \frac{1}{2} \mu \dot{q}^2(x) - \frac{1}{2} Y \left(\frac{\partial q(x)}{\partial x} \right)^2 \tag{4.1b}$$

La $\mathcal{L}(x)$ si presenta quindi a tutti gli effetti come una densità e pertanto prende il nome di *densità di Lagrangiana*

Ricordiamo che le equazioni del moto possono scriversi come segue:

$$\frac{ak(q_{i+1} - q_i) - ak(q_i - q_{i-1})}{a^2} = \frac{m}{a} \frac{d}{dt} \dot{q}_i$$

[1] Ogni oscillatore è soggetto alle forze dei due oscillatori vicini. Quando un oscillatore è fuori della sua posizione di equilibrio, esso risente pertanto di due forze concordi che tendono a riportarlo nella posizione di equilibrio.

nel limite di $a \to 0$ il primo termine fornisce:

$$\lim_{a\to 0}\left[\frac{ak(q_{i+1}-q_i)-ak(q_i-q_{i-1})}{a^2}\right] = Y\lim_{a\to 0}\frac{\frac{q_{i+1}-q_i}{a}-\frac{q_i-q_{i-1}}{a}}{a}$$

dunque l'equazione del moto nel limite del continuo diventa:

$$Y\frac{\partial^2 q}{\partial x^2}-\mu\frac{\partial^2 q}{\partial t^2}=0$$

che è essenzialmente l'equazione delle onde che si propagano lungo una corda con velocità data da $v=\sqrt{Y/\mu}$.

Da N equazioni del moto in t, siamo quindi arrivati tramite il passaggio al limite del continuo ad una sola equazione differenziale in x e t.

Ricaviamo ora l'equazione del moto esplicitamente tramite variazioni dell'integrale di azione. Scriviamo la forma più generale dell'azione, partendo dalla definizione (3.1):

$$S=\int L(q,\dot{q},t)dt=\iint \mathcal{L}(q(x),\partial q/\partial x,\dot{q}(x))dxdt$$

da cui calcolando le variazioni minime:

$$\delta S=\iint\left[\frac{\partial\mathcal{L}}{\partial q}\delta q+\frac{\partial\mathcal{L}}{\partial\dot{q}}\delta\left(\frac{\partial q}{\partial t}\right)+\frac{\partial\mathcal{L}}{\partial\frac{\partial q}{\partial x}}\delta\left(\frac{\partial q}{\partial x}\right)\right]dxdt$$

che integrata per parti su $\partial\mathcal{L}/\partial\left(\frac{\partial q}{\partial x}\right)$ e $\partial\mathcal{L}/\partial\dot{q}$ fornisce:[2]

$$\delta S=\iint\left[\frac{\partial\mathcal{L}}{\partial q}-\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}}-\frac{d}{dx}\frac{\partial\mathcal{L}}{\partial\frac{\partial q}{\partial x}}\right]\delta q dxdt$$

e quindi:

$$\delta S=0\Leftrightarrow\frac{\partial\mathcal{L}}{\partial q}-\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}}-\frac{d}{dx}\frac{\partial\mathcal{L}}{\partial\frac{\partial q}{\partial x}}=0$$

Applicando ora l'equazione appena trovata alla densità di Lagrangiana $\mathcal{L}$ in (4.1b):

$$\frac{\partial\mathcal{L}}{\partial q}=0\to-\frac{d}{dt}\mu\dot{q}+\frac{d}{dx}Y\frac{\partial q}{\partial x}=0\to-\mu\ddot{q}+Y\frac{\partial^2 q}{\partial x^2}=0$$

che è esattamente l'equazione delle onde già trovata sopra. Notiamo che l'essere passati dalla Lagrangiana alla densità Lagrangiana ha comportato, fra l'altro, di far comparire sullo stesso piano le derivate rispetto a x e t.

Al fine di generalizzare l'equazione di variazione ricavata sopra, introduciamo il formalismo delle *derivate funzionali* (che saranno indicate con $\delta/\delta f(x)$). Definiamo quindi:

$$\begin{cases}\dfrac{\delta q(x')}{\delta q(x)} & =\delta(x'-x)\\[2ex] \dfrac{\delta}{\delta q(x)}\dfrac{\partial q(x')}{\partial x'} & =\dfrac{\partial}{\partial x'}\delta(x'-x)\end{cases}$$

In tale formalismo possiamo allora scrivere:

$$\frac{\delta\mathcal{L}}{\delta q(x)}=\int\left[\frac{\partial\mathcal{L}}{\partial q(x')}\frac{\delta q(x')}{\delta q(x)}+\frac{\partial\mathcal{L}}{\partial\frac{\partial q(x')}{\partial x'}}\frac{\delta}{\delta q(x)}\frac{\partial q(x')}{\partial x'}\right]dx'=\int\left[\frac{\partial\mathcal{L}}{\partial q(x')}-\frac{d}{dx'}\frac{\partial\mathcal{L}}{\partial\frac{\partial q(x')}{\partial x'}}\right]\frac{\delta q(x')}{\delta q(x)}dx'$$

[2] I termini estratti fuori integrale sono evidentemente nulli, in quanto tutte le variazioni sono nulle agli estremi.

in definitiva quindi:

$$\frac{\delta \mathcal{L}}{\delta q(x)} = \frac{\partial \mathcal{L}}{\partial q(x)} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \frac{\partial q(x)}{\partial x}}$$
$$\frac{\delta \mathcal{L}}{\delta \dot{q}(x)} = \int \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{\delta \dot{q}(x)}{\delta \dot{q}(x)}dx = \frac{\partial \mathcal{L}}{\partial \dot{q}(x)}$$

e l'equazione di Eulero-Lagrange diviene nella sua forma più generale:

$$\frac{\delta \mathcal{L}}{\delta q(x)} - \frac{d}{dt}\frac{\delta \mathcal{L}}{\delta \dot{q}(x)} = 0$$

che ha un'espressione formalmente analoga a quella già nota.

Nel caso tridimensionale, compaiono evidentemente le derivate rispetto alle tre coordinate. In generale quindi:

$$\mathcal{L} \equiv \mathcal{L}(q_j(x,y,z,t), \partial_\mu q_j(x,y,z,t), t)$$

dove l'indice j è l'indice che identifica i campi da cui dipende $\mathcal{L}$. L'equazione di Eulero-Lagrange assume quindi la forma (moltiplicando e dividento la derivata rispetto al tempo per ic):

$$\frac{\partial \mathcal{L}}{\partial q(\vec{r})} - \frac{d}{icdt}\frac{\partial \mathcal{L}}{\partial \frac{\partial q}{ic\partial t}} - \frac{d}{d\vec{r}}\frac{\partial \mathcal{L}}{\partial \frac{\partial q}{\partial \vec{r}}}$$

che può immediatamente essere messa in forma quadridimensionale covariante a vista:

$$\frac{\delta \mathcal{L}}{\delta q(\vec{r})} - \partial_\mu \frac{\delta \mathcal{L}}{\delta\, \partial_\mu q(\vec{r})} = 0$$

4.2 Formalismo Hamiltoniano

Passiamo ora ad estendere quanto visto al formalismo Hamiltoniano, impostando anche qui inizialmente il formalismo per il caso discreto.

Abbiamo visto che nel caso della catena lineare la Lagrangiana "discreta" risulta:

$$L = \sum_{i=1}^{N}\left[\frac{1}{2}m\dot{q}_i^2 - \frac{1}{2}k(q_{i+1}-q_i)^2\right]$$

come è noto, la variabile coniugata della q_i è data da $\partial L/\partial \dot{q}_i = m\dot{q}_i$ e l'Hamiltoniana, secondo le "regole classiche" è data da:

$$H = \sum_{i=1}^{N}\dot{q}_i p_i - L = \sum_{i=1}^{N} a\cdot\dot{q}_i\cdot\frac{p_i}{a} - \sum_{i=1}^{N} a\frac{L}{a}$$

il limite per il sistema continuo prevede naturalmente ancora di effettuare il limite $a \to 0$, $Na \to l$. Il passaggio al limite $\lim_{a\to 0} p_i/a \equiv \pi_i(x)$ restituisce invece una quantità denominata *densità di impulso*.[3]

Il passaggio al limite completo per i sistemi continui fornisce quindi, in analogia con il formalismo lagrangiano, una *densità di Hamiltoniana*. Infatti si ottiene, molto banalmente:

$$\lim_{a\to 0}\sum_{i=1}^{N}\left[\dot{q}_i\cdot\frac{p_i}{a} - \frac{L}{a}\right]a = \int \left[\dot{q}(x)\pi(x) - \mathcal{L}(x)\right]dx$$

[3]Questo naturalmente perché, in analogia ai casi precedenti, per ottenere l'impulso totale del sistema occorre integrare su x: $p = \int \pi(x)dx$.

ovvero:

$$H = \int [\dot{q}(x)\pi(x) - \mathcal{L}(x)]\, dx \tag{4.2a}$$

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{q}(x)} \tag{4.2b}$$

che permette di definire la *densità di Hamiltoniana* $\mathcal{H}$:

$$H = \int \mathcal{H}(x)dx \tag{4.3a}$$

$$\mathcal{H}(x) \equiv \dot{q}(x)\pi(x) - \mathcal{L}(x) \tag{4.3b}$$

con la densità di impulso definita dalla (4.2b) e generalizzabile in termini di derivate funzionali come $\delta\mathcal{L}/\delta\dot{q}(x)$.

Si applichi ora il metodo variazionale all'Hamiltoniana così definita.

$$\begin{aligned}
\delta H &= \int \left[\frac{\partial \mathcal{H}}{\partial q}\delta q + \frac{\partial \mathcal{H}}{\partial \pi}d\pi + \frac{\partial \mathcal{H}}{\partial \frac{\partial q}{\partial x}} d\left(\frac{\partial q}{\partial x}\right) + \frac{\partial \mathcal{H}}{\partial t}dt\right] dx = \\
&= \int \left[\frac{\partial \mathcal{H}}{\partial q}\delta q - \frac{d}{dx}\frac{\partial \mathcal{H}}{\partial \frac{\partial q}{\partial x}} dq + \frac{\partial \mathcal{H}}{\partial \pi}d\pi + \frac{\partial \mathcal{H}}{\partial t}dt\right] dx = \\
&= \int \left[\frac{\delta \mathcal{H}}{\delta q}dq + \frac{\delta \mathcal{H}}{\delta \pi}d\pi + \frac{\delta \mathcal{H}}{\delta t}dt\right]
\end{aligned}$$

Ma della definizione precedente (4.3b):

$$\delta H = \int \left[\pi d\dot{q} + \dot{q}d\pi - \frac{\delta \mathcal{L}}{\delta q}dq - \frac{\delta \mathcal{L}}{\delta \dot{q}}d\dot{q} - \frac{\delta \mathcal{L}}{\delta t}dt\right] dx$$

tenendo presente che $\pi d\dot{q} - \frac{\delta\mathcal{L}}{\delta\dot{q}}d\dot{q} = 0$ per la (4.2b). Dal confronto delle due espressioni, uguagliando i coefficienti dei differenziali, si ricavano le seguenti relazioni per le variabili coniugate:[4]

$$\begin{aligned}
\dot{q} &= \frac{\delta \mathcal{H}}{\delta \pi} \\
\dot{\pi} &= \frac{\delta \mathcal{H}}{\delta t}
\end{aligned}$$

4.3 Formulazione lagrangiana e hamiltoniana del campo. Azione del campo elettromagnetico

Sia dato ora un campo $\phi(\vec{x}, t)$. Come ormai sappiamo la Lagrangiana, l'azione e le equazioni di Eulero–Lagrange sono date da:

$$L = L(\phi, \partial_\mu\phi, t) \qquad S = \int_1^2 \int \mathcal{L} d^3x dt \qquad \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} = 0$$

Nel caso la Lagrangiana dipenda da più campi, si avranno tante equazioni di Lagrange quanti sono i campi coinvolti. Nel caso di campi vettoriali $A_\mu \equiv (\vec{A}, i\Phi)$, la Lagrangiana è funzione di 4 componenti, come se fossero 4 campi scalari indipendenti $L(A_\mu, \partial_\mu A_\nu, t)$. Nel caso particolare del campo elettromagnetico, le variabili dinamiche sono proprio i potenziali e devono essere soddisfatte – nel caso di assenza delle sorgenti – le equazioni per il campo libero (3.10) e le (3.11) senza sorgenti, ovvero $\partial_\nu F^{\mu\nu} = 0$.

Il procedimento che seguiremo ora è il seguente:

[4] Per ricavare la seconda, si osservi che $\frac{\delta\mathcal{L}}{\delta q} = \frac{d}{dt}\pi$.

- scrivere le equazioni del moto come variazione dell'azione, in modo che l'annullamento del δS implichi l'equazione del moto stessa,
- esplicitare le variazioni dei campi nella scrittura del δS,
- risalire dalla variazione dei campi (e quindi dal δS) alla forma che deve avere l'azione.

questo procedimento, che permette di trovare l'azione note le equazioni del moto, è di validità del tutto generale e sarà utilizzato molto spesso nel seguito.

In pratica, in questo caso le equazioni del moto ricercate sono $\partial_\nu F^{\mu\nu} = 0$:

$$\delta S = k \int_1^2 \partial_\nu F^{\mu\nu} \delta A_\mu d^3x dt = -k \int_{t_1}^{t_2} F^{\mu\nu} \delta \partial_\nu A_\mu d^3x dt = -k \int_{t_1}^{t_2} F^{\mu\nu} \delta \frac{[\partial_\nu A_\mu - \partial_\mu A_\nu]}{2} d\Omega =$$
$$= \frac{k}{2} \int_{t_1}^{t_2} F^{\mu\nu} \delta F_{\mu\nu} d\Omega$$

da cui si ricava immediatamente la forma dell'azione:

$$S = k \int_{t_1}^{t_2} F^{\mu\nu} F_{\mu\nu} d\Omega$$

il che implica che il termine della Lagrangiana associato al "moto" del campo è dato da:

$$L = kF^{\mu\nu} F_{\mu\nu}$$

Al termine di azione (3.8) già trovato per una particella *in* un campo magnetico:

$$S = \int \left[-mc\, ds + \frac{e}{c} A_\mu dx^\mu \right] = -\int mc\, ds + \int \frac{\varrho \vec{x}}{c} A_\mu \frac{d\vec{x}}{dt} dt = -\int mc\, ds + \int \frac{1}{ic^2} j^\mu A_\mu d\Omega$$

deve dunque essere aggiunto il termine appena ricavato, ottenendo l'azione totale:

$$S = -\int mc\, ds + \int \left[\frac{1}{ic^2} j^\mu A_\mu + ikF^{\mu\nu} F_{\mu\nu} \right] d\Omega$$

Trascurando ora la parte relativa alla particella libera già nota, applichiamo il metodo variazionale alla parte di Lagrangiana relativa al campo elettromagnetico. Questo ci permetterà di trovare il valore corretto per la costante k utilizzando le equazioni di Maxwell con le sorgenti (3.11):[5]

$$\delta S = \int \left[\frac{1}{ic^2} j^\mu \delta A_\mu + ik2F^{\mu\nu} \delta F_{\mu\nu} \right] d\Omega = \int \left[\frac{1}{ic^2} j_\mu \delta A_\mu + i4kF^{\mu\nu} \partial_\nu \delta A_\mu \right] d\Omega =$$
$$= \int \left[\frac{1}{ic^2} j^\mu \delta A_\mu + i4k \partial_\nu F^{\mu\nu} \delta A_\mu \right] d\Omega = \int \left[-\frac{1}{c^2} j_\mu + 4k \partial_\nu F_{\mu\nu} \right] i\delta A_\mu d\Omega$$

Ora, imporre che la variazione δS dell'azione sia nulla, significa imporre che:

$$\partial_\nu F^{\mu\nu} = \frac{1}{4kc^2} j^\mu$$

e dal confronto con la (3.11) si ricava il valore della costante $k = 1/16\pi c$. Il termine della densità di lagrangiana relativa al "moto" del campo è pertanto:

$$\mathcal{L} = \frac{1}{16\pi c} F^{\mu\nu} F_{\mu\nu}$$

il quale – ricordando che $F^{\mu\nu} F_{\mu\nu}$ è un invariante che vale $2(B^2 - E^2)$ – si può riscrivere come $\mathcal{L} = -1/8\pi(B^2 - E^2)$. L'azione completa per una particella interagente con un campo elettromagnetico è quindi:

$$\boxed{S = -\int mc\, ds + \int \left[\frac{1}{ic^2} j^\mu A_\mu + \frac{i}{16\pi c} F^{\mu\nu} F_{\mu\nu} \right] d\Omega}$$

[5] Nel secondo passaggio si è sfruttata la relazione $\delta F_{\mu\nu} = \delta(\partial_\mu A_\nu - \partial_\nu A_\mu) = \partial_\mu \delta A\nu - \partial_\nu \delta A_\mu$.

che comprende in sé le equazioni del moto della particella, le equazioni del campo (Maxwell) e il termine di interazione fra i due.

Al fine di impostare il formalismo Hamiltoniano, occorre trovare la variabile coniugata con A_μ. Detta questa π_μ:

$$\pi_\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} = \frac{\partial \mathcal{L}}{\partial \frac{\partial A_\mu}{\partial t}}$$

la cui quarta componente fornisce:

$$\pi_4 = \frac{\partial \mathcal{L}}{\partial \frac{\partial A_4}{\partial t}} \equiv \frac{1}{ic}\frac{\partial \mathcal{L}}{\partial\, \partial_4 A_4}$$

infatti $\partial/\partial t \equiv ic\partial_4$. Siccome risulta $F_{44} = 0$ $(\mu = \nu = 4)$ ([6]), ne consegue che $\pi_4 = 0$. Non esiste quindi una variabile coniugata con A_4, ed in effetti queste componenti non sono indipendenti per l'arbitrarietà della scelta di Gauge.[7]

Per le altre componenti:

$$\pi_k = \frac{\partial \mathcal{L}}{\partial \frac{\partial A_k}{\partial t}} \qquad k = 1, 2, 3$$

ovvero:

$$\frac{\partial \mathcal{L}}{ic\partial \frac{\partial A_k}{\partial x_4}} = \frac{1}{ic}\frac{\partial \mathcal{L}}{\partial F_{\mu\nu}}\frac{\partial F_{\mu\nu}}{\partial\, \partial_4 A_k} = \frac{1}{16ic\pi}2F_{\mu\nu}\frac{\partial F_{\mu\nu}}{\partial\, \partial_4 A_k}$$

siccome vale:

$$\frac{\partial F_{\mu\nu}}{\partial\, \partial_4 A_k} = [\delta_{4\mu}\delta_{k\mu} - \delta_{k\mu}\delta_{4\mu}]$$

segue:

$$\pi_k = -\frac{1}{16ic\pi}2F_{\mu\nu}[\delta_{4\mu}\delta_{k\mu} - \delta_{k\mu}\delta_{4\mu}] = \frac{1}{ic}\frac{1}{8\pi}2F_{4k} = -\frac{1}{4\pi c}E_k$$

cioé il momento coniugato della variabile A_k è proprio la componente E_k del campo elettrico lungo $\hat{k}$. Le variabili hamiltoniane sono quindi:

$$\begin{cases} A_\mu \\ \pi_4 \equiv 0 \\ \pi_k = -\dfrac{1}{4\pi c}E_k \end{cases}$$

Si può ora scrivere la densità di Hamiltoniana utilizzando la (4.3b):

$$\mathcal{H} = \pi_\mu \frac{\partial}{\partial t}A_\mu - \mathcal{L} \equiv \frac{\partial}{\partial t}A_k \cdot \pi_k - \mathcal{L}$$

che tenendo presente la seguente:

$$\frac{\partial}{\partial t}A_k = -e\partial_k\Phi - eE_k \qquad \left(\vec{E} = -\vec{\nabla}\Phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}\right)$$

assume la forma:

$$\mathcal{H} = e(\partial_k\Phi + E_k)\frac{1}{4\pi c}E_k - \mathcal{L} = \frac{1}{4\pi}E^2 + \frac{1}{8\pi}(B^2 - E^2)$$

da cui la forma definitiva della densità Hamiltoniana:

$$\mathcal{H} = \frac{1}{8\pi}(B^2 + E^2) + \frac{\vec{E}\cdot\vec{\nabla}\Phi}{4\pi} \tag{4.4}$$

[6] Il calcolo esplicito:

$$\frac{\partial \mathcal{L}}{\partial\, \partial_4 A_4} = \frac{\partial}{\partial\, \partial_4 A_4}F_{\mu\nu}F_{\mu\nu} = \frac{\partial}{\partial\, \partial_4 A_4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) = \frac{\partial}{\partial\, \partial_4 A_4}(\partial_4 A_4 - \partial_4 A_4)(\partial_4 A_4 - \partial_4 A_4) \equiv 0$$

[7] È un problema, questo, che emerge nella quantizzazione a causa delle parentesi di commutazione fra momenti coniugati; tuttavia si può fissare una Gauge particolare.

L'Hamiltoniana si ottiene naturalmente integrando la relazione precedente. Integrando per parti il secondo termine della (4.4) si ottiene un termine $\vec{\nabla}\cdot\vec{E}$, che è nullo in virtù delle equazioni di Maxwell, pertanto l'Hamiltoniana del campo elettromagnetico è data in definitiva da:

$$H = \int \frac{B^2 + E^2}{8\pi} d\vec{x}$$

Consideriamo ora le seconde equazioni di Maxwell ma senza le sorgenti (3.11):

$$\partial_\nu F^{\mu\nu} = \partial_\nu[\partial^\mu A^\nu - \partial^\nu A^\mu] = \partial_\nu\partial^\mu A^\nu - \partial_\nu\partial^\nu A^\mu = 0$$

ma $\partial_\nu\partial^\nu \equiv \Box$ è il dalambertiano in forma quadridimensionale, pertanto la seconda coppia di equazioni di Maxwell in termini del potenziale A_μ è data da:

$$\Box A^\mu - \partial^\mu(\partial_\nu A^\nu) = 0$$

che è esattamente un'equazione delle onde scritta per il potenziale, cosa resa ancora più evidente dalla scelta della *Gauge di Lorentz* $\partial_\nu A^\nu = 0$ che la rende covariante a vista e mette le quattro componenti sullo stesso piano.

Se si opera con formalismo hamiltoniano su tutte e quattro le componenti con la scelta di Gauge fatta, si deve aggiungere un termine proporzionale alla scelta di Gauge stessa. La Lagrangiana che si può prendere in considerazione è la cosiddetta *Lagrangiana di Fermi*:

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu}F^{\mu\nu} - \frac{1}{8\pi}(\partial_\mu A^\mu)^2$$

prima di ricercare l'Hamiltoniana corrispondente, conviene sviluppare ulteriormente questa Lagrangiana:[8]

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{16\pi}\left[(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu)\right] - \frac{1}{8\pi}(\partial_\mu A^\mu)^2 \\
&= -\frac{1}{16\pi}\left[\partial_\mu A_\nu \partial^\mu A^\nu - \partial_\mu A_\nu \partial^\nu A^\mu - \partial_\nu A_\mu \partial^\mu A^\nu + \partial_\nu A_\mu \partial^\nu A^\mu\right] - \frac{1}{8\pi}(\partial_\mu A^\mu)^2 \\
&= -\frac{1}{16\pi}\left[2(\partial_\mu A_\nu)^2 - 2(\partial_\mu A_\nu \partial^\nu A^\mu) + 2(\partial_\mu A^\mu)^2\right] = \\
&= -\frac{1}{8\pi}\left[(\partial_\mu A_\nu)^2 - \partial_\mu(A_\nu \partial^\nu A^\mu) + A_\nu \partial_\mu \partial^\nu A^\mu + (\partial_\mu A^\mu)^2\right] = \\
&= -\frac{1}{8\pi}\left[(\partial_\mu A_\nu)^2 - \partial_\mu(A_\nu \partial^\nu A^\mu) + A_\nu \partial_\nu(\partial_\mu A^\mu) + (\partial_\mu A^\mu)^2\right]
\end{aligned}$$

il termine di divergenza totale $\partial_\mu(A_\nu\partial^\nu A^\mu)$ non dà contributi all'Hamiltoniana e si può pertanto trascurare. I due termini in $\partial_\mu A^\mu$ sono nulli per la scelta della Gauge di Lorentz. Pertanto:

$$\mathcal{L} = -\frac{1}{8\pi}\left[(\partial_\mu A_\nu)^2 + A_\nu\partial_\nu\chi + \chi^2\right] \qquad \chi \equiv \partial_\mu A^\mu$$

I momenti coniugati sono pertanto:

$$\pi_\mu = \frac{\partial\mathcal{L}}{\partial\frac{\partial A_\mu}{\partial t}} = \frac{1}{ic}\frac{\partial\mathcal{L}}{\partial\,\partial_4 A_\mu}$$

i quali, una volta imposta la condizione $\chi = 0$, forniscono:[9]

$$\pi_\mu = \frac{1}{ic}\frac{\partial\mathcal{L}}{\partial\,\partial_4 A_\mu} = -\frac{1}{ic}\frac{1}{8\pi}2\cdot\partial_4 A_\mu = -\frac{1}{ic}\frac{1}{4\pi}\cdot\partial_4 A_\mu \to \partial_4 A_\mu = -ic4\pi\,\pi_\mu$$

[8] Nel quarto passaggio si è integrato per parti.

[9] Si noti, incidentalmente, che l'equazione di Eulero scritta rispetto a χ fornisce:

$$\frac{\partial\mathcal{L}}{\partial\chi} = \partial_\mu\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\chi} = 2\chi - \partial_\mu A_\mu = 0 \Rightarrow \partial_\mu A_\mu = 0 \quad (\chi \equiv \partial_\mu A_\mu)$$

La densità di Hamiltoniana assume quindi la forma:

$$\mathcal{H} = \pi_\mu \frac{\partial A_\mu}{\partial t} - \mathcal{L} = \pi_\mu \cdot ic\partial_4 A^\mu - \mathcal{L}$$

ed essendo $\partial_4 A^\mu = -ic4\pi\, \pi^\mu$:

$$\begin{aligned}\mathcal{H} &= -4\pi c^2 \pi_\mu \pi^\mu + \frac{1}{8\pi}(\partial_4 A_\nu)^2 + \frac{1}{8\pi}(\partial_i A_\nu)^2 \\ &= -4\pi c^2 \pi_\mu \pi^\mu + \frac{1}{8\pi}\left(16\pi^2 c^2 \pi_\mu \pi^\mu\right) + \frac{1}{8\pi}(\partial_i A_\nu)^2 \\ &= -4\pi c^2 \pi_\mu \pi^\mu + 2\pi c^2 \pi_\mu \pi^\mu + \frac{1}{8\pi}(\partial_i A_\nu)^2\end{aligned}$$

ovvero:

$$\mathcal{H} = -2\pi c^2 \pi_\mu \pi^\mu + \frac{1}{8\pi}(\partial_i A_\nu)^2$$

quindi, conservando la formulazione covariante a vista, l'Hamiltoniana non è più definita positiva. Questo è il problema della quantizzazione in forma covariante: l'Hamiltoniana non definita positiva implica che esistano degli stati quantistici ad energia minore di zero, cosa che è tuttavia esclusa dalla condizione $\partial_\mu A^\mu = 0$.

Capitolo 5

Invarianza per Trasformazioni del Sistema di Riferimento

Vedremo ora cosa comporta richiedere che le equazioni della relatività ristretta siano invarianti per trasformazioni di coordinate. Vedremo in effetti che l'invarianza per traslazioni del sistema di riferimenti implica la presenza di un tensore che si conserva, che sarà chiamato *tensore energia–impulso*.

Troveremo poi la forma esplicita di questo tensore e dell'azione in tre casi particolarmente interessanti: il campo elettromagnetico, l'equazione di Klein-Gordon e l'equazione di Schrödinger.

Infine, vedremo le conseguenze dell'invarianza sotto rotazioni spazio-temporali di Lorentz. Questo caso ci condurrà alla scoperta di un altro tensore conservato, che sarà chiamato *tensore momento angolare*.

5.1 Invarianza per Traslazioni. Tensore Energia–Impulso

Supponiamo di avere una Lagrangiana dipendente da un campo scalare:

$$L = \int \mathcal{L}(\Phi, \partial_\mu \Phi) d^3x \qquad \Phi(x_i) \text{ scalare}$$

ed effettuiamo una traslazione delle quadricoordinate $x^\mu \to x'^\mu = x^\mu + \delta x^\mu$. Dire che il sistema è invariante sotto questa trasformazione di coordinate significa dire che la variazione indotta dalla trasformazione sul campo Φ induce a sua volta una variazione della Lagrangiana tale che questa variazione sia un termine costante e che quindi non influirà sulle equazioni del moto di Eulero-Lagrange, che comportano delle derivazioni.

La variazione indotta su $\mathcal{L}$ dal campo Φ è data da:

$$\delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \Phi}\delta\Phi + \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi}\delta\partial_\mu\Phi = \frac{\partial \mathcal{L}}{\partial \Phi}\delta\Phi + \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi}\delta\Phi\right] - \partial_\mu \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi}\delta\Phi$$

come si vede, la variazione $\delta\mathcal{L}$ indotta da $\delta\Phi$ è composta da un pezzo che è l'equazione del moto e da una divergenza totale:

$$\underbrace{\left(\frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi}\right)}_{\textit{equazione del moto}}\delta\Phi + \underbrace{\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi}\delta\Phi\right)}_{\textit{divergenza totale}}$$

Detta Λ^μ la variazione della Lagrangiana, lungo un moto questa risulta proprio:

$$\partial_\mu \Lambda^\mu = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi}\delta\Phi\right) \tag{5.1}$$

ed è effettivamente una costante, essendo una divergenza totale. Si danno ora due casi possibili:

- $\Lambda_\mu = 0$: In tal caso la Lagrangiana rimane completamente invariata per trasformazioni di coordinate. Si parla in questo caso di *Simmetria Interna*,[1]
- $\Lambda_\mu \neq 0$: Si parla in tal caso di *Simmetria Spazio-temporale.*

Del tutto in generale la variazione dipende dal campo: $\Lambda \equiv \Lambda(\Phi)$.

Generalizzando, se la Lagrangiana dipende da campi complessi: $\mathcal{L}(\Phi, \Phi^*, \partial_\mu\Phi, \partial_\mu\Phi^*)$, si possono effettuare delle trasformazioni di fase $\Phi' = e^{i\alpha}\Phi$ tali da lasciare la Lagrangiana invariata.

Se invece la variazione indotta non è esprimibile come una divergenza totale, allora la Lagrangiana non è simmetrica e non risulta invariante per la data trasformazione di coordinate.

Ora, è noto che ad ogni simmetria è associato un quadrivettore "corrente" che si conserva.[2] La quantità che si conserva è proprio il termine di divergenza trovato: infatti, in caso di invarianza della Lagrangiana, a causa della (5.1) deve valere banalmente:

$$\partial_\mu \left(\Lambda^\mu - \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi} \delta\Phi \right) = 0$$

per cui risulta evidente che la quantità conservata sia esattamente $\partial\mathcal{L}/\partial\,\partial_\mu\Phi$. Se da questa inoltre si può ricavare un quadrivettore corrente j_μ che si conserva ($\partial_\mu j_\mu = 0$), allora:

$$\begin{aligned}
\int_{t_1}^{t_2} \int \partial_\mu j^\mu d^3x dt &= \int_{t_1}^{t_2} \left[\int \partial_4 j_4 d^3x + \int \partial_i j^i d^3x \right] = \\
&= \int_{t_1}^{t_2} \left[\frac{1}{ic}\frac{\partial}{\partial t} \int j_4 d^3x + \textit{integrale di superficie} = 0 \right] dt \rightarrow \\
&\rightarrow \left[\frac{1}{ic} \int j_4 d^3x \right]_{t_1}^{t_2} = 0
\end{aligned}$$

esiste quindi una "carica" conservata data esattamente da $Q = \int j_4 d^3x$. Più in generale, la carica conservata può essere espressa coma una ipersuperficie (integrale sulle coordinate spaziali): $Q = \int j_\alpha dS^\alpha$.

Consideriamo ora le traslazioni più in dettaglio. Se la variazione indotta sul campo Φ è tale che:

$$x^\mu \rightarrow x'^\mu = x^\mu + \delta x^\mu \qquad \Rightarrow \qquad \Phi \rightarrow \Phi' : \Phi'(x') = \Phi(x')$$

allora:

$$\delta\Phi(x) = \Phi'(x') - \Phi(x) = \Phi(x') - \Phi(x) = \Phi(x - \delta x) - \Phi(x) = -\partial_\mu\Phi(x)\delta x^\mu$$

per costruire la corrente conservata associata consideriamo quindi:[3]

$$\delta\mathcal{L} = \mathcal{L}'(x) - \mathcal{L}(x) = \partial_\mu\mathcal{L}\delta x^\mu$$

ricordando che sussiste anche la relazione (5.1) e che $\delta\Phi(x) = -\partial_\mu\Phi(x)\delta x^\mu$:

$$\begin{aligned}
-\partial_\mu\mathcal{L}\delta x^\mu = \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\Phi}(-\partial_\nu\Phi\delta x^\nu) \right] &\rightarrow \\
\partial_\mu \left[\mathcal{L}\delta x^\mu - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\Phi}\partial_\nu\Phi\delta x^\nu \right] = 0 &\rightarrow \\
\partial_\mu \left[\left(\mathcal{L}\delta^\mu_\nu - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\Phi}\partial_\nu\Phi \right) \delta x^\nu \right] = 0 &
\end{aligned}$$

ovvero:

$$\partial_\mu \left[\mathcal{L}\delta^\mu_\nu - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\Phi}\partial_\nu\Phi \right] \delta x^\nu = 0 \qquad \forall \delta x^\nu$$

[1] È il caso, ad esempio, delle trasformazioni di fase.

[2] Niente altro che il teorema di Noether: esso sancisce un legame tra l'invarianza di una certa quantità sotto trasformazioni di uno o più campi e una legge di conservazione di una corrente, detta appunto corrente di Noether.

[3] La lagrangiana è uno scalare.

Questa quantità risulta essere un tensore del II° ordine, che sarà denotato con $T_{\mu\nu}$ e chiamato *Tensore Energia–Impulso*. Quindi possiamo scrivere:

$$\boxed{\partial_\mu T^{\mu\nu} = 0}$$

con la posizione:

$$\boxed{T^{\mu\nu} \equiv \mathcal{L}\delta^{\mu\nu} - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\Phi}\partial^\nu\Phi} \qquad \textbf{(Tensore Energia–Impulso)}$$

Questo tensore porta con sé quattro correnti conservate, le cui cariche associate sono:

$$\begin{cases} Q_4 = \int T^{44} d^3x \\ Q_i = \int T^{4i} d^3x \end{cases}$$

Per la quarta componente:

$$Q_4 = \int \left[\mathcal{L} - \frac{\partial\mathcal{L}}{\frac{1}{ic}\partial\left(\frac{\partial\Phi}{\partial t}\right)}\frac{1}{ic}\frac{\partial}{\partial t}\Phi\right] d^3x = \int \left[\mathcal{L} - \frac{\partial\mathcal{L}}{\partial\dot\Phi}\dot\Phi\right] d^3x = \int \left[\mathcal{L} - \pi\dot\Phi\right] d^3x$$

infatti $\partial\mathcal{L}/\partial\dot\Phi = \pi$ è la densità di impulso associata alla variabile dinamica Φ. La quarta carica conservata è quindi proprio l'Hamiltoniana del sistema:

$$Q_4 = -\int \mathcal{H} d^3x$$

Le altre cariche forniscono:

$$Q_i = \int T^{4i} d^3x = \int \left[\mathcal{L} - \frac{\partial\mathcal{L}}{\frac{1}{ic}\partial\left(\frac{\partial\Phi}{\partial t}\right)}\partial^i\Phi\right] d^3x = ic\int \pi\partial_i\Phi d^3x$$

Questa quantità prende il nome di *impulso associato al campo* Φ. Siccome per una particella vale $p^\mu \equiv (\vec{p}, iE/c)$, ne risulta $P^\alpha = -iQ^\alpha/c$, con l'impulso associato dato da: $P^\nu = -i/c\int T^{4\nu}d^3x$. Siccome il tensore si conserva, le cariche conservate corrispondono alla conservazione dell'impulso e dell'energia.

Il tensore Energia-Impulso così definito non ha in realtà proprietà definite di simmetria, mentre per definire il momento angolare sarà necessario un tensore simmetrico, come mostrato più avanti in §5.3.

Iniziamo col notare che il tensore $T^{\mu\nu}$ non è univocamente definito, in quanto per trasformazioni della lagrangiana del tipo $\mathcal{L} \to \mathcal{L} + \partial_\mu\Lambda^\mu(\Phi)$ esso varia. Infatti:

$$\begin{aligned} T'^{\mu\nu} &= \mathcal{L}\delta^{\mu\nu} + \delta^{\mu\nu}\partial_\alpha\Lambda^\alpha - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\Phi}\partial^\nu\Phi - \frac{\partial(\partial_\alpha\Lambda^\alpha)}{\partial\,\partial_\mu\Phi}\partial^\nu\Phi = \\ &= T^{\mu\nu} + \delta^{\mu\nu}\partial_\alpha\Lambda^\alpha - \frac{\partial}{\partial\,\partial_\mu\Phi}\left(\frac{\partial\Lambda^\alpha}{\partial\Phi}\partial_\alpha\Phi\right)\partial^\nu\Phi = \\ &= T^{\mu\nu} + \delta^{\mu\nu}\partial_\alpha\Lambda^\alpha - \frac{\partial\Lambda^\alpha}{\partial\Phi}\delta^\mu_\alpha\partial^\nu\Phi \end{aligned}$$

ovvero in definitiva:

$$T'^{\mu\nu} = T^{\mu\nu} + \delta^{\mu\nu}\partial_\alpha\Lambda^\alpha - \frac{\partial\Lambda^\mu}{\partial\Phi}\partial^\nu\Phi$$

Il tensore è quindi variato. Vale la pena notare esplicitamente che essendo le correnti conservate espresse da:

$$\begin{cases} T'_{44} &= T_{44} + \partial_\alpha\Lambda^\alpha - \dfrac{\partial\Lambda^4}{\partial\Phi}\partial_4\Phi = T_{44} + \partial_i\Lambda^i \\ T'_{4i} &= T_{4i} + \partial_\alpha\Lambda^\alpha - \dfrac{\partial\Lambda^4}{\partial\Phi}\partial_i\Phi = T_{44} - \partial_i\Lambda^i \end{cases}$$

esse sono variate per delle divergenze totali e pertanto il loro integrale (le cariche conservate) non varia.

Per altra via, considerando la relazione $\partial_\mu T^{\mu\nu} = 0$ si nota che aggiungendo una divergenza di una quantità completamente antisimmetrica $\partial_\alpha \psi^{\mu\nu\alpha}$, si ricava:

$$\partial_\mu T'^{\mu\nu} = \partial_\mu T^{\mu\nu} + \partial_\mu \partial_\alpha \psi^{\mu\nu\alpha}$$

e quindi a livello di correnti:

$$\int j^4 d^3x = \int j^\mu dS_\mu$$

$$P^\nu = -\frac{i}{c}\int T^{4\nu} d^3x = -\frac{i}{c}\int T^{\mu\nu} dS_\mu \rightarrow P'^\nu = P^\nu - \frac{i}{c}\int \partial_\alpha \psi^{\mu\nu\alpha} dS_\mu$$

considerando solo la parte variata e ricordando che $\psi^{\mu\nu\alpha}$ è supposta antisimmetrica:

$$\int \partial_\alpha \psi^{\mu\nu\alpha} dS_\mu = \frac{1}{2}\int [\partial_\alpha \psi^{\mu\nu\alpha} dS_\mu - \partial_\mu \psi^{\alpha\nu\mu} dS_\alpha] = \frac{1}{2}\int \psi^{\mu\nu\alpha} df_{\mu\alpha}$$

dove $df_{\mu\alpha}$ è l'elemento di superficie in 4 dimensioni. Quindi il termine è nullo. In definitiva, aggiungendo un termine completamente antisimmetrico alla definizione di tensore Energia-Impulso data sopra, questo si conserva per trasformazioni di Gauge ma assume in più proprietà di simmetria definite, diventa cioé antisimmetrico.

5.2 Densità di Lagrangiana e Tensore energia–impulso in alcuni casi particolari

Precederemo ora a calcolare il tensore energia–impulso e la densità di Lagrangiana in tre casi interessanti.

5.2.1 Equazioni di Maxwell

Consideriamo ora la densità di Lagrangiana associata al campo elettromagnetico:

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu}$$

Cerchiamo il tensore energia-impulso per il campo elettromagnetico:

$$\delta A_\mu = -\delta x^\nu \partial_\nu A_\mu$$

ora:

$$\frac{\partial \mathcal{L}}{\partial\, \partial_\nu A_\mu} = -\frac{1}{16\pi} 2F_{\alpha\beta} \frac{F^{\alpha\beta}}{\partial_\nu A_\mu} = -\frac{1}{8\pi}[\delta^{\alpha\nu}\delta^{\beta\mu} - \delta^{\alpha\mu}\delta^{\beta\nu}]F_{\alpha\beta} \underbrace{=}_{\text{antisimmetria}} -\frac{1}{4\pi}F^{\nu\mu}$$

quindi:

$$T^{\mu\nu} = -\frac{1}{16\pi}\delta^{\mu\nu} F_{\alpha\beta}F^{\alpha\beta} - \frac{\partial \mathcal{L}}{\partial\, \partial_\nu A_\mu} = -\frac{1}{16\pi}\delta^{\mu\nu} F^2_{\alpha\beta} + \frac{1}{4\pi} F^{\mu\alpha}\partial^\nu A_\alpha$$

Questo oggetto non è simmetrico per scambio di indici $\mu \leftrightarrow \nu$: sebbene $F^{\mu\nu}$ in sé non cambi per trasformazioni di Gauge, il termine $\partial_\nu A_\alpha$ fa dipendere $T^{\mu\nu}$ dalla scelta di Gauge. Occorrerebbe, per evitare questo, aggiungere un termine del tipo $-1/4\pi F^{\mu\alpha}\partial_\alpha A^\nu$, ovvero modificare il tensore in modo da produrre questo termine. Ricordando che la divergenza di $F^{\mu\nu}$ è nulla,[4] il termine necessario può essere riscritto come $\partial_\alpha(F^{\mu\alpha}A^\nu)$. Ma $F^{\mu\alpha}A^\nu$ è antisimmetrico e non cambia le equazioni fisiche, pertanto può effettivamente essere aggiunto senza problemi:

$$T^{\mu\nu} = -\frac{1}{16\pi}\delta^{\mu\nu} F^2 + \frac{1}{4\pi} F^{\mu\alpha} F^\nu_\alpha$$

[4]Vedi le equazioni del moto (3.11) in assenza di sorgenti.

e per l'antisimmetria

$$T^{\mu\nu} = -\frac{1}{16\pi}\delta^{\mu\nu}F^2 - \frac{1}{4\pi}F^{\mu\alpha}F^{\nu}_{\alpha}$$

Le cariche conservate sono date da:

$$\int T^{44}d^3x = \int \left[-\frac{1}{16\pi}2(H^2 - E^2) - \frac{1}{4\pi}E^2\right] d^3x = \int \left[-\frac{1}{8\pi}H^2 - \frac{1}{8\pi}E^2\right] d^3x =$$
$$= -\frac{1}{8\pi}\int [H^2 + E^2]d^3x$$

cioé proprio la densità di energia. Le componenti spaziali della corrente forniscono:

$$P^i = -\frac{i}{c}\int T^{4i}d^3x = -\frac{i}{c}\int \left[-\frac{1}{4\pi}F^{4\alpha}F^i_{\alpha}\right] d^3x = -\frac{i}{c}\int i[\vec{E}\times\vec{H}]d^3x$$

infatti, ad esempio per la seconda componente:

$$P^2 = -\frac{i}{c}\int T^{42}d^3x = -\frac{i}{c}\int [iE_xH_z - iE_zH_x]\, d^3x = -\frac{i}{c}\int \left[i\left(\vec{E}\times\vec{H}\right)_y\right] d^3x$$

e quindi la corrente conservata risulta legata al vettore di Poynting.

5.2.2 Equazione di Klein–Gordon

Cerchiamo ora la densità di Lagrangiana ed il tensore energia-impulso che hanno l'equazione di Klein-Gordon come equazione del moto:

$$\Box\,\Phi + m^2\Phi = 0 \qquad\qquad \Phi \text{ Campo Scalare}$$

ovvero:

$$\delta S = \iint \left[\frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2} - \nabla^2\Phi + m^2\Phi\right]\delta\Phi d^3x =$$
$$= \iint_{t_1}^{t_2}\left[\frac{1}{c^2}\frac{\partial\Phi}{\partial t}\delta\left(\frac{\partial\Phi}{\partial t}\right) + \partial_i\Phi\delta\partial_i\Phi + m^2\Phi\delta\Phi\right] dtd^3x =$$
$$= \iint_{t_1}^{t_2}\delta\left[-\frac{1}{2c^2}\left(\frac{\partial\Phi}{\partial t}\right)^2 + \frac{1}{2}(\partial_i\Phi)^2 + m^2\Phi^2\right] dtd^3x$$

e quindi la densità di lagrangiana che genera come equazione del moto l'equazione di Klein-Gordon è:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi + \frac{1}{2}m^2\Phi^2$$

ed il relativo tensore energia-impulso è:

$$T^{\mu\nu} = \mathcal{L}\delta^{\mu\nu} - \partial^\mu\Phi\partial^\nu\Phi$$

il quale è simmetrico in μ e ν grazie alla simmetria della densità di Lagrangiana.

5.2.3 Equazione di Schrödinger

Consideriamo ora il caso in cui l'equazione del moto sia costituita dall'equazione di Schrödinger (*campo di materia*):

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

Questo è un campo complesso, per cui la densità di Lagrangiana è espressa in termini delle coordinate $\mathcal{L} = \mathcal{L}(\psi, \psi^*, \partial_\mu\psi, \partial_\mu\psi^*)$. L'equazione di Schrödinger è riprodotta se:

$$\mathcal{L} = \frac{\hbar^2}{2m}\nabla\psi(\vec{x},t)\nabla\psi^*(\vec{x},t) + U(\vec{x})\psi(\vec{x},t)\psi^*(\vec{x},t) + \frac{1}{2}\frac{\hbar}{i}\left[\psi^*(\vec{x},t)\dot{\psi}(\vec{x},t) - \psi(\vec{x},t)\dot{\psi}^*(\vec{x},t)\right]$$

infatti:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \psi^*} &= U(\vec{x})\psi(\vec{x},t) + \frac{1}{2}\frac{\hbar}{i}\dot{\psi}(\vec{x},t) \\
\frac{\partial \mathcal{L}}{\partial\, \partial_i \psi^*} &= \frac{\hbar^2}{2m}\nabla\psi(\vec{x},t) \\
\frac{\partial \mathcal{L}}{\partial \dot{\psi}^*} &= \frac{1}{2}\frac{\hbar}{i}\psi(\vec{x},t)
\end{aligned}$$

e l'equazione del moto rispetto a $\psi^*(\vec{x},t)$ è allora:

$$\frac{\partial \mathcal{L}}{\partial \psi^*} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \psi^*} = 0 \Rightarrow U(\vec{x})\psi(\vec{x},t) + \frac{1}{2}\frac{\hbar}{i}\dot{\psi}(\vec{x},t) - \frac{\hbar^2}{2m}\nabla^2\psi(\vec{x},t) + \frac{1}{2}\frac{\hbar}{i}\psi(\vec{x},t)$$

L'impulso coniugato fornisce:

$$\pi^*_\mu = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^*} = \frac{1}{2}\frac{\hbar}{i}\psi(\vec{x},t)$$

cioé il momento coniugato di $\psi^*(\vec{x},t)$ è $\psi(\vec{x},t)$. Le variabili coniugate sono quindi proprio $\psi^*(\vec{x},t)$ e $\psi(\vec{x},t)$.

5.3 Invarianza per Rotazioni spazio–temporali di Lorentz

Consideriamo ora l'invarianza sotto rotazioni spazio-temporali di Lorentz:

$$x'^\mu = \Lambda^\mu_\nu x^\nu \qquad\qquad \Lambda_{\mu\nu} = \text{matrice di Lorentz}$$

che per una trasformazione infinitesima si scrive:[5]

$$x'^\mu = x^\mu + \epsilon\omega^\mu_\nu x^\nu$$

Sapendo che $x'_\mu x'^\mu = x^2$ ed è un invariante, e denotando $\epsilon_{\mu\nu} \equiv \epsilon\omega_{\mu\nu}$:

$$\begin{aligned}
x'_\mu x'^\mu &= x^\rho[\delta_{\mu\rho} + \epsilon_{\mu\rho}]x^\sigma[\delta_{\nu\sigma} + \epsilon_{\nu\sigma}]\delta^{\mu\nu} = \\
&= x^\rho x^\sigma[\delta_{\mu\rho}\delta_{\nu\sigma} + \delta_{\mu\rho}\epsilon_{\nu\sigma} + \delta_{\nu\sigma}\epsilon_{\mu\rho} + o(\epsilon^2)]\delta^{\mu\nu}
\end{aligned}$$

da cui, dovendo essere invariante:

$$\delta_{\mu\rho}\epsilon_{\nu\sigma} + \delta_{\nu\sigma}\epsilon_{\mu\rho} = 0 \Rightarrow \epsilon_{\rho\sigma} + \epsilon_{\sigma\rho} = 0$$

ovvero, in definitiva:

$$\epsilon(\omega_{\rho\sigma} + \omega_{\sigma\rho}) = 0$$

e quindi $\omega_{\rho\sigma}$ deve essere antisimmetrico.

Se abbiamo a che fare con un campo puramente scalare:

$$\delta\phi(x) = \phi'(x) - \phi(x) = \phi(x - \epsilon\omega_{\mu\nu}x^\nu) - \phi(x) = (\partial_\mu\phi)\epsilon\omega^\mu_\nu x^\nu$$

la quantità conservata è come al solito:

$$\delta\mathcal{L} = \partial_\mu\left[\frac{\partial \mathcal{L}}{\partial\, \partial_\mu\phi}\delta\phi\right]$$

ma:

$$\delta\mathcal{L} = -\partial_\mu\mathcal{L}\cdot\epsilon\omega^\mu_\nu x^\nu = -\partial_\mu\left[\mathcal{L}\cdot\epsilon\omega^\mu_\nu x^\nu\right] + \mathcal{L}\cdot\partial_\mu\epsilon\omega^\mu_\nu x^\nu$$

[5] La trasformazione più generale del tipo di Lorentz è $x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu$, che costituisce il cosiddetto *Gruppo di Poincaré*, mentre la più generale possibile è $x'^\mu = f_\mu(x)$.

da cui per l'antisimmetria di $\omega_{\mu\nu}$:

$$\delta\mathcal{L} = -\partial_\mu \left[\mathcal{L} \cdot \epsilon\omega^\mu_\nu x^\nu\right]$$

pertanto:

$$\begin{aligned}
\partial_\mu \left[\mathcal{L} \cdot \epsilon\omega^\mu_\nu x^\nu - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi}(\partial_\nu\phi)\epsilon\omega^\nu_\rho x^\rho\right] &= 0 \Rightarrow \\
\epsilon\partial_\mu\omega^\rho_\nu x^\nu \left[\mathcal{L}\delta^\mu_\rho - \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi}\partial_\rho\phi\right] &= 0 \Rightarrow \\
= \epsilon\partial_\mu[\omega^\rho_\nu x^\nu T^\mu_\rho] &= 0 \Rightarrow \\
= \epsilon\omega^\rho_\nu\partial_\mu[x^\nu T^\mu_\rho] &= 0
\end{aligned}$$

ricordando ora che $\omega_{\rho\nu}$ è antisimmetrico, la relazione si può riscrivere come:

$$\epsilon\omega^\rho_\nu\partial_\mu\frac{1}{2}[x^\nu T^\mu_\rho - x_\rho T^\mu_\nu] = 0$$

che, una volta introdotto un *tensore momento angolare* a tre indici definito come:

$$\boxed{m^{\nu\mu\rho} \equiv \frac{1}{2}[x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu}]} \qquad \textbf{(Tensore Momento Angolare)}$$

permette di scrivere la quantità conservata nella forma:

$$\boxed{\partial_\mu m^{\nu\mu\rho} = 0}$$

Riconsideriamo ora la trasformazione infinitesima di Lorentz.

$$\delta S = \sum_i p_\mu\delta x^\mu \qquad\qquad p_\mu = \frac{\delta S}{\delta x^\mu} = \frac{\partial S}{\partial x^\mu}$$

quindi:

$$\delta S = \sum_i p^\mu\delta\omega_{\mu\nu}x^\nu = \sum_i \frac{1}{2}(p^\mu x^\nu - p^\nu x^\mu)\delta\omega_{\mu\nu}$$

la quantità conservata è allora:

$$\frac{\partial S}{\partial\omega_{\mu\nu}} = \frac{1}{2}\sum_i(p^\mu x^\nu - p^\nu x^\mu)$$

che è antisimmetrica. A ben guardare, le componenti spaziali sono proprio le componenti del momento angolare:

$$M_{ij} = x_i p_j - x_j p_i \qquad\qquad (\vec{M} = \sum_i(\vec{r}_i \times \vec{p}_i)\cdot\hat{e}_i \quad)$$

la quarta componente fornisce invece:

$$M_{4j} = \sum\left(ictp_j - x_j\frac{iE}{c}\right) = ic\sum(tp_j - x_j\frac{E}{c^2})$$

e deve ovviamente conservarsi anche questa. Allora risulta:

$$\sum\left[tp_j - x_j\frac{E}{c^2}\right] = \text{cost.} \quad \rightarrow \quad \sum\left[t\vec{p} - \vec{r}\frac{E}{c^2}\right] = \text{cost.}$$

ma $\sum E$ è costante e pertanto si può dividere per questa quantità:

$$\frac{\sum t\vec{p}}{\sum E} - \frac{1}{c^2}\frac{\sum\vec{r}E}{\sum E} = \text{cost.} \Rightarrow \frac{\sum t\vec{p}}{\sum E} - \frac{1}{c^2}\vec{R} = \text{cost.}$$

con $\vec{R}$ coordinate del baricentro. Siccome t è costante e $\vec{p}/E = \vec{V}/c^2$, la sommatoria fornisce una sorta di "velocità totale":

$$\frac{1}{c^2}\vec{V}t - \frac{1}{c^2}\vec{R} = \text{cost.}$$

e questa è in effetti l'equazione del centro di massa.

Fin qua, quello che riguarda i punti materiali.

Applichiamo ora l'invarianza per rotazioni alla densità di Lagrangiana. Sappiamo che nel caso questa dipenda da un campo scalare Φ:

$$\delta\mathcal{L} = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi} \delta\Phi \right]$$

ma applicando direttamente la definizione di variazione alla Lagrangiana:

$$\delta\mathcal{L} = \mathcal{L}' - \mathcal{L} = \mathcal{L}(x - \epsilon\omega^\mu_\nu x^\nu) - \mathcal{L} = \epsilon \frac{\partial \mathcal{L}}{\partial x^\mu} \omega^\mu_\nu x^\nu$$

ma $\omega_{\mu\nu}$ non dipende dalle x, pertanto: $\delta\mathcal{L} = -\epsilon\omega^\mu_\nu \partial_\mu(\mathcal{L}x^\nu)$ è una divergenza. Per le correnti conservate consideriamo:

$$\delta\mathcal{L} = \partial_\mu \left[\epsilon\omega^\mu_\nu \mathcal{L} x^\nu - \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi} \epsilon\omega^\alpha_\beta \partial_\alpha \Phi x^\beta \right] = 0$$

da cui:

$$\epsilon\omega^\alpha_\nu x^\nu \left[\delta_{\alpha\mu}\mathcal{L} - \frac{\partial \mathcal{L}}{\partial\, \partial_\mu \Phi} \partial_\alpha \Phi \right] = 0$$

ovvero:

$$\omega^\alpha_\nu \partial_\mu (x^\nu T^\mu_\alpha) = 0$$

che grazie alla proprietà di antisimmetria si può scrivere:

$$\omega_{\alpha\nu} \partial_\mu \frac{1}{2} [x^\nu T^{\mu\alpha} - x^\alpha T^{\mu\nu}] = 0$$

dunque in analogia al caso precedente si può scrivere $\partial_\mu m^{\mu\nu\alpha} = 0$, con 6 correnti conservate:

$$M^{\nu\alpha} = -\frac{i}{c} \int m^{\nu 4 \alpha} d^3x$$

ed esplicitando le quantità:

$$M^{\nu\alpha} \propto \int [x^\nu T^{4\alpha} - x^\alpha T^{4\nu}] d^3x$$

che si presenta come una densità di momento angolare ($T^{4\alpha}$ è la densità di impulso). Le componenti temporali rappresentano analoghe cariche conservate: sono in effetti i generatori delle rotazioni, ovvero delle trasformazioni di Lorentz (*boosts*). Ora:

$$\partial_\mu m^{\mu\nu\alpha} = \delta^\nu_\mu T^{\mu\alpha} - \delta^\alpha_\mu T^{\mu\nu} = 0 \underbrace{\Rightarrow}_{somma\ su\ \mu} T^{\nu\alpha} = T^{\alpha\nu}$$

dunque il tensore $T^{\mu\nu}$ deve essere simmetrico.

Consideriamo ora il caso del campo elettromagnetico, la cui densità di Lagrangiana è data da:

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu}$$

Imponiamo l'invarianza per trasformazioni di Lorentz e ricordiamo la definizione del tensore elettromagnetico data in (3.9):

$$A'^\mu(x') = \Lambda^\mu_\nu A^\nu(\Lambda x)$$

Per trasformazioni infinitesime risulta:

$$
\begin{aligned}
\delta A_\mu(x) &= A'_\mu(x) - A_\mu(x) = (\delta_{\mu\nu} + \epsilon\omega_{\mu\nu})A_\nu(\Lambda^{-1}x) - A_\mu(x) = \\
&= (\delta_{\mu\nu} + \epsilon\omega_{\mu\nu})\left[A_\nu(x) - \epsilon\frac{\partial A_\nu}{\partial x_\rho}\omega_{\rho\sigma}x_\sigma\right] = \epsilon\left[\omega_{\mu\nu}A_\nu - \frac{\partial A_\nu}{\partial x_\rho}\omega_{\rho\sigma}x_\sigma\right] = \\
&= -\epsilon\omega_{\rho\sigma}[(\partial_\mu x_\sigma)A_\rho + \partial_\rho A_\mu x_\sigma] = -\epsilon\omega_{\rho\sigma}[(\partial_\mu x_\sigma A_\rho) - x_\sigma\partial_\mu A_\rho + \partial_\rho A_\mu x_\sigma] = \\
&= \epsilon\omega_{\rho\sigma}[-(\partial_\mu x_\sigma A_\rho) + x_\sigma(\partial_\mu A_\rho - \partial_\rho A_\mu)] = \epsilon\omega_{\rho\sigma}[-(\partial_\mu x_\sigma A_\rho) + x_\sigma F_{\mu\rho}] \qquad (5.2)
\end{aligned}
$$

Ora, ricordiamo questo risultato e calcoliamo la quantità conservata:

$$\delta\mathcal{L} = \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial\,\partial_\mu A_\nu}\delta A_\nu\right]$$

ma $\delta\mathcal{L} = -\epsilon\partial_\mu(\omega_{\mu\sigma}x_\sigma\mathcal{L})$, da cui sostituendo nell'equazione appena scritta:

$$\epsilon\partial_\mu\left[\omega_{\mu\sigma}x_\sigma\mathcal{L}\frac{\partial\mathcal{L}}{\partial\,\partial_\mu A_\nu}\delta A_\nu\right] = 0$$

e ricordando il risultato (5.2):

$$
\begin{aligned}
&\epsilon\partial_\mu\left\{-\frac{1}{16\pi}\omega_{\mu\sigma}x_\sigma F^2 - \frac{1}{4\pi}F_{\mu\nu}\omega_{\rho\sigma}[-\partial_\nu(x_\sigma A_\rho) + x_\sigma F_{\nu\rho}]\right\} = \\
&= \frac{\epsilon}{4\pi}\partial_\mu\left\{-\frac{1}{4}\omega_{\mu\sigma}x_\sigma F^2 + F_{\mu\nu}\omega_{\rho\sigma}\partial_\nu(x_\sigma A_\rho) - F_{\mu\nu}\omega_{\rho\sigma}x_\sigma F_{\nu\rho}\right\}
\end{aligned}
$$

siccome lungo un moto risulta $\partial_\mu F_{\mu\nu} = 0$:

$$\frac{\epsilon}{4\pi}\left\{F_{\mu\nu}\omega_{\rho\sigma}\partial_\mu\partial_\nu(x_\sigma A_\rho) - \partial_\mu\omega_{\rho\sigma}\left[F_{\mu\nu}F_{\nu\rho}x_\sigma - \delta_{\rho\mu}\frac{1}{4}F^2x_\sigma\right]\right\} = 0$$

il primo termine è antisimmetrico, pertanto nella somma degli indici si annulla. Ne risulta quindi il tensore energia-impulso nella forma simmetrica:

$$
\begin{aligned}
\frac{\epsilon}{4\pi}[\partial_\mu(\omega_{\rho\sigma}x_\sigma T_{\mu\rho})] = 0 \rightarrow \\
\frac{1}{2}\partial_\mu[\omega_{\rho\sigma}(x_\sigma T_{\mu\rho} - x_\rho T_{\mu\sigma})] = 0 \rightarrow \\
\partial_\mu[\omega_{\rho\sigma}m_{\sigma\mu\rho}] = 0
\end{aligned}
$$

Il tensore $T_{\mu\nu}^{\text{campo}}$ associato al campo genera quindi delle cariche conservate (del campo) $M_{\mu\nu}^\rho$. Il tensore energia–impulso $T_{\mu\nu}^{\text{part.}}$ associato alle particelle richiede l'introduzione di una densità di massa:

$$\mu = \sum_i m_i\delta(\vec{r}_i - \vec{r}'_i)$$

siccome inoltre $p_i = m_i c u_i$ con $u^\alpha = dx^\alpha/ds$:

$$\pi_\alpha = \mu c u^\alpha$$

e la relazione con il tensore è $p_\alpha = -\frac{i}{c}T_{4\alpha}$, per cui:

$$\mu c u^\alpha = -\frac{i}{c}T_{4\alpha} \Rightarrow T_{4\alpha} = -\frac{\mu c^2}{i}u^\alpha \Rightarrow T_{4\alpha} = i\mu c^2 u^\alpha$$

La densità di massa si può interpretare come la quarta componente di un quadrivettore $T^{4\alpha} = \mu\frac{dx^\alpha}{dt}cu^\alpha$, ovvero:

$$T^{\mu\alpha} = \mu c\frac{dx^\mu}{dt}\frac{dx^\alpha}{ds}$$

che riscritta nella forma:

$$T^{\mu\alpha} = \mu c \frac{dx^\mu}{dt} \frac{dx^\alpha}{dt} \frac{ds}{dt}$$

ne mostra chiaramente la simmetria. Si vede inoltre che risulta:

$$\partial_\mu \left[T^{\mu\nu}_{\text{campo}} + T^{\mu\nu}_{\text{part.}} \right] = 0$$

cioé la somma dei tensori del campo e delle particelle è costante.

Capitolo 6

Cinematica Relativistica

Daremo qui alcuni cenni di cinematica relativistica per mostrare come si procede in linea generale.

6.1 Generalità

Ricordiamo le equazioni base per trattare la cinematica relativistica:

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-\beta^2}} \qquad E = \frac{mc^2}{\sqrt{1-\beta^2}}$$

da cui, evidentemente:

$$\vec{p} = \frac{E\vec{v}}{c^2} \qquad p^\alpha = \left(\vec{p}, i\frac{E}{c}\right)$$

e se si considera per semplicità un moto lungo x le leggi di trasformazione hanno la forma:

$$\begin{cases} p'_x = \dfrac{1}{\sqrt{1-\beta^2}}\left[p_x - \dfrac{E}{c^2}v_x\right] \\ p'_y = p_y \\ p'_z = p_z \\ E' = \dfrac{1}{\sqrt{1-\beta^2}}[E - p_x v_x] \end{cases}$$

Sia $f(\vec{p})dp_x dp_y dp_z$ una funzione di distribuzione che dà il numero di particelle con impulso compreso fra p e $p + dp$. Ora, $dp_x dp_y dp_z$ è l'elemento di ipersuperficie perpendicolare a p_4, infatti dalla definizione di ipersuperficie:

$$dS^\alpha = \frac{1}{6}\epsilon^{\alpha\beta\gamma\delta} dp_\beta dp_\gamma dp_\delta$$

quindi dS_α risulta parallelo a p_4. Di conseguenza la quantità

$$\frac{dp_x dp_y dp_z}{E} = \text{invariante}$$

essendo il rapporto dei moduli. Più in dettaglio:

$$d^4p = dp_1 dp_2 dp_3 dp_4$$

e siccome deve valere la $p_\alpha p^\alpha = m^2c^2 \Rightarrow p_4^2 - \vec{p}^2 = m^2c^2$, ovvero:

$$E^2 = m^2c^4 + c^2p^2$$

si vede che le componenti non sono in realtà indipendenti. Si deve allora imporre questa relazione fra le variabili e per fare questo si deve far comparire una $\delta(m^2c^2 + p_\alpha^2)$ in un integrale integrato su p_4:

$$dp_x dp_y dp_z \qquad\qquad \int \delta(p_4^2 + \vec{p}^2 + m^2c^2)dp_4$$

Dalla definizione di δ: $\delta[f(x)] = \sum_i \frac{\delta(x-a_i)}{|f'(a_i)|}$ segue:

$$\begin{aligned}\int \delta(p_4^2 + \vec{p}^2 + m^2c^2)dp_4 &= \int \frac{1}{2p_4}\delta\left(p_4 - i\sqrt{m^2c^2 - \vec{p}^2}\right)dp_4 = \\ &= \int \frac{c}{2iE}\delta\left(\frac{iE}{c} - i\sqrt{m^2c^2 - \vec{p}^2}\right)\frac{idE}{c} = \\ &= \frac{c}{2i}\int \delta\left(E - \sqrt{m^2c^2 - \vec{p}^2}\right)\frac{dE}{E}\end{aligned}$$

allora:

$$\int \delta(p^2 + m^2c^2)d^4p = \frac{2c}{i}\int \delta\left(E - \sqrt{m^2c^2 - p^2}\right)\frac{dE}{E}dp_x dp_y dp_z = d^4p$$

e dunque $\frac{dp_x dp_y dp_z}{E}$ è invariante perché d^4p è invariante. Ne consegue allora che:

$$f(\vec{p})E\frac{dp_x dp_y dp_z}{E}$$

è invariante. Segue immediatamente che:

$$f(\vec{p}')E' = f(\vec{p})E$$

e perciò:

$$f(\vec{p}') = f(\vec{p})\frac{E}{E'}$$

Inoltre, se consideriamo che $d^3p = p^2dp\,d\Omega$ ed essendo $E^2 = m^2c^2 + \vec{p}^2$ e $EdE = p\,dp$, segue:

$$\frac{pEdEd\Omega}{E} = \frac{d^3p}{E} \Rightarrow \frac{d^3p}{E} = p\,dEd\Omega$$

che per particella a massa nulla diventa:

$$\frac{d^3p}{E} = \frac{E}{c}dEd\Omega$$

6.2 Decadimento di una particella

Consideriamo una particella di massa M che decade in due particelle più leggere. Nel sistema, che indicheremo con l'accento, del centro di massa S' le due particelle (identificate dai pedici 1 e 2) si allontanano a 180° l'una dall'altra.[1] La conservazione dell'energia impone che debba essere $Mc^2 = E_1' + E_2'$.([2]) Siccome vale $E_i'^2 = m_i^2c^4 + \vec{p}_i'^2c^2$, ne consegue che $E_i' > m_ic^2$. Pertanto:

$$\begin{aligned}M^2c^4 &= \left(m_1^2c^4 + \vec{p}_1'^2c^2\right) + \left(m_2^2c^4 + \vec{p}_2'^2c^2\right) + 2\sqrt{\left(m_1^2c^4 + \vec{p}_1'^2c^2\right)\left(m_2^2c^4 + \vec{p}_2'^2c^2\right)} \\ &= \left(m_1^2 + m_2^2\right)c^4 + \left(\vec{p}_1'^2 + \vec{p}_2'^2\right)c^2 + 2\sqrt{\left(m_1^2c^4 + \vec{p}_1'^2c^2\right)\left(m_2^2c^4 + \vec{p}_2'^2c^2\right)}\end{aligned}$$

[1] Ciò vale evidentemente per la conservazione dell'impulso: nullo all'inizio, dopo il decadimento deve essere nulla la somma delle componenti lungo i tre assi del sistema di riferimento, da cui l'angolo di 180°. Supponendo infatti che l'asse x sia orientato secondo il moto della particella 1, ne segue che $\vec{p}_1 = (p, 0, 0)$. Ma allora deve essere $\vec{p}_2 = (-p, 0, 0)$.

[2] Il primo membro rappresenta l'energia *prima* del decadimento nel sistema del centro di massa, in cui quindi non c'è moto. Il secondo membro rappresenta le due particelle che si allontanano dopo il decadimento.

Il decadimento può esserci solo se la massa della particella iniziale è superiore alla somma delle masse dei prodotti di decadimento. Per dimostrarlo, supponiamo per comodità di calcolo che sia $m_1 = m_2 = m$, per cui $\vec{p}_1 = \vec{p}_2 = \vec{p}$, si ricava allora:

$$M^2c^4 = 2m^2c^4 + 2p^2c^2 + \sqrt{2m^2c^4 + 2p^2c^2}$$
$$M^2 = 2m^2 + 2\frac{p^2}{c^2} + \sqrt{2\frac{m^2}{c^4} + 2\frac{p^2}{c^6}} > 2m^2$$

Nel centro di massa risulta quindi $\vec{p}_1'^2 = -\vec{p}_2'^2 \Rightarrow p_1'^2 = p_2'^2$. Ne consegue che:

$$E_1'^2 - m_1^2c^4 = E_2'^2 - m_2^2c^4$$

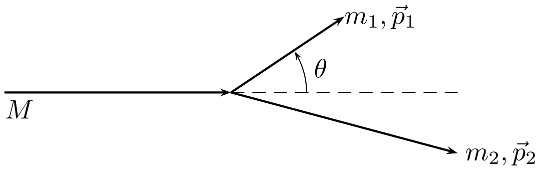
e dalla conservazione dell'energia:

$$\begin{cases} E_2' = Mc^2 - E_1' \\ \cancel{E_1'^2} - m_1^2c^4 = M^2c^2 - 2M^2c^2E_1' + \cancel{E_1'^2} - m_2^2c^4 \end{cases}$$

da cui si ricava:

$$\begin{cases} E_1' = c^2\dfrac{M^2 + m_1^2 - m_2^2}{2M} \\ E_2' = Mc^2 - E_1' \end{cases}$$

Naturalmente, bisogna calcolare gli angoli a cui vengono emesse le particelle nel sistema S del laboratorio.



$$\vec{p}_1 + \vec{p}_2 = \vec{P} \qquad E_1 + E_2 = E = \frac{Mc^2}{\sqrt{1-\beta^2}}$$

Essendo $\vec{p} = \frac{E}{c^2}\vec{v}$:

$$\vec{v} = \frac{\vec{p}c^2}{E} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2}{\sqrt{1-\beta^2}}\frac{c^2\sqrt{1-\beta^2}}{Mc^2} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2}{M}$$

quindi la velocità del centro di massa coincide con quella classica.

$$E' = \frac{1}{\sqrt{1-\beta^2}}(E - vp_1\cos\theta) \Rightarrow \cos\theta = \frac{E - E'\sqrt{1-\beta^2}}{vp}$$

ma $p^2 = \frac{E^2}{c^2} - m^2c^2$, quindi:

$$\cos\theta = \frac{E - E'\sqrt{1-\beta^2}}{v\sqrt{\frac{E^2}{c^2} - m^2c^2}}$$

e quadrando:

$$\left(1-\beta^2\right)E'^2 - 2\sqrt{1-\beta^2}EE' + E^2 - \left[v^2\left(\frac{E^2}{c^2} - m^2c^2\right)\right]\cos^2\theta = 0$$

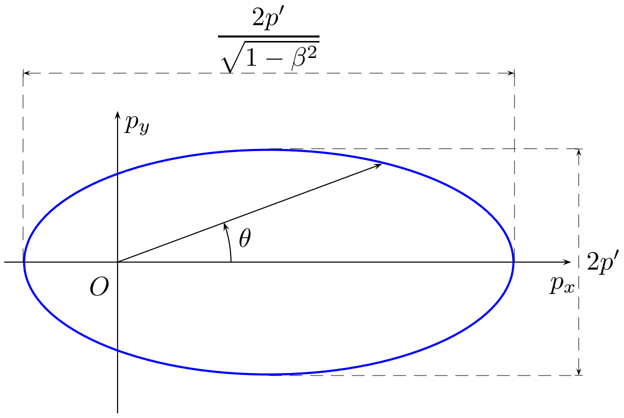
che stabilisce una relazione $E(\cos\theta)$. Ora:

$$E = \frac{1}{\sqrt{1-\beta^2}}(E' + vp'\cos\theta)$$
$$p_y' = p_y = p'\sin\theta$$
$$p_x = \frac{\frac{E'v}{c^2} + p'\cos\theta}{\sqrt{1-\beta^2}} \to p'\cos\theta = p_x\sqrt{1-\beta^2} - \frac{E'v}{c^2}$$

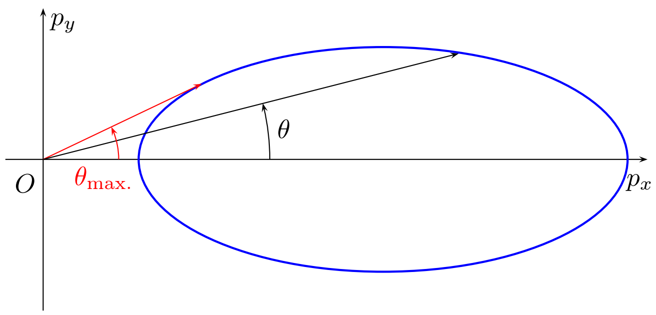
e quadrando e sommando $p' \cos\theta$ e $p' \sin\theta$:

$$p'^2 = \left[p_x \sqrt{1-\beta^2} - \frac{E^2}{c^2} \right]^2 + p_y^2$$

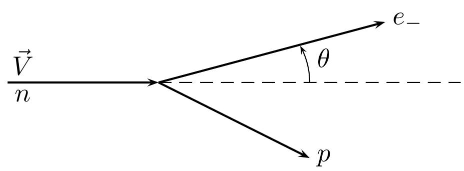
Questa equazione rappresenta una ellisse nel piano $p_x - p_y$ i cui semiassi sono p' e $\frac{p'}{\sqrt{1-\beta^2}}$ e traslata di $\frac{E^2}{c^2}\frac{1}{\sqrt{1-\beta^2}}$:



Siccome si tratta della somma di due termini che danno p'^2, l'origine O può essere interno o esterno all'ellisse, a seconda che risulti $p' \lessgtr \frac{E^2}{c^2\sqrt{1-\beta^2}}$. In particolare, se l'origine O risulta esterno all'ellisse, allora esistono due soluzioni per l'impulso ed esiste un angolo limite $\theta_{\text{max.}}$ entro cui è osservabile la particella:



Rendiamo ora l'esempio più specifico e reale, considerando il caso del decadimento del neutrone:[3]



$$n \longrightarrow p + e^-$$

Nel sistema di riferimento del neutrone, il decadimento è completamente isotropo, ovvero non ha direzioni privilegiate. Si consideri il piano passante per la traiettoria del neutrone e dell'elettrone, utilizzando la tipica convenzione di indicare con il sistema accentato il sistema del laboratorio, si può scrivere allora:

$$p'^e = \left(i\frac{E'}{c}, p' \cos\theta', p' \sin\theta', 0 \right)$$

[3]Omettiamo volutamente in questa trattazione la presenza dell'antineutrino elettronico $\bar{\nu}_e$. D'altra parte, l'osservazione di un risultato sperimentale *diverso* da quello calcolato in questo modo fu alla base dell'ipotesi di una particella fino ad allora sconosciuta che portasse via parte della quantità di moto.

Applicando la trasformazione:

$$\begin{cases} p_x & = \dfrac{p'_x + \frac{V}{c^2}E'}{\sqrt{1-\beta^2}} = \dfrac{p'\cos\theta + \frac{V}{c^2}E'}{\sqrt{1-\beta^2}} \\ p'_y & = p_y = p'\sin\theta \end{cases}$$

da cui segue:

$$p'\cos\theta = p_x\sqrt{1-\beta^2} - \frac{V}{c^2}E'$$

che quadrate e sommate fornisce:

$$p'^2 = p_y^2 + \left(p_x\sqrt{1-\beta^2} - \frac{V}{c^2}E'\right)^2$$

che costituisce, come abbiamo visto, un'ellisse nel piano $p_x - p_y$ centrato in $\frac{VE'}{\sqrt{1-\beta^2 c^2}}$. Questo ellisse può includere, essere tangente o essere all'esterno dell'origine, e questo a seconda del valore (ponendo $p_y = 0$):

$$p_x\sqrt{1-\beta^2} - \frac{V}{c^2}E' = -p' \Rightarrow \boxed{\bar{p}_x = \frac{\frac{V}{c^2}E' - p'}{\sqrt{1-\beta^2}}}$$

ovvero a seconda della condizione:

$$V - \frac{c^2p'}{E'} \lesseqgtr 0 \qquad \left(\frac{c^2p'}{E'} \text{ è la velocità dell'elettrone}\right)$$

nel caso l'origine sia all'esterno dell'ellisse, esiste un valore massimo all'impulso osservabile per l'elettrone.

Il decadimento β e i neutrini

Nel caso del decadimento β reale, il neutrone si può considerare inizialmente fermo perché contenuto nell'atomo e così il protone dopo il decadimento. L'unica particella a muoversi è quindi l'elettrone.

La conservazione dell'energia si scrive quindi come:

$$m_N c^2 = E_p + E_e \qquad \begin{cases} E_p = \sqrt{m_p^2c^4 + p_p^2c^2} \\ E_e = \sqrt{m_e^2c^4 + p_e^2c^2} \end{cases}$$

e trascurando il riculo del protone nel nucleo a causa della sua massa e del fatto che è legato:

$$m_N c^2 \simeq m_p c^2 + E_e + \sqrt{m_e^2c^4 + p_e^2c^2}$$

In queste condizioni, l'unica variabile è l'impulso dell'elettrone che deve quindi risultare univocamente determinato:

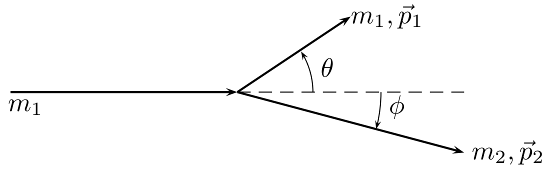
$$p_e = \sqrt{c^2(m_n - m_p) - m_e}$$

Sperimentalmente, invece, lo spettro di energia dell'elettrone era continuo, partendo da 0 fino ad arrivare circa 5 volte e mezza la masse dell'elettrone e passando per un massimo. Per risolvere il problema, Wolfgang Pauli (insieme con Enrico Fermi) propose che il decadimento non producesse due corpi ma tre e che il terzo portasse via una parte di energia variabile. Questa particella fu battezzata da Pauli "neutrone", per essere poi rinominata da Fermi con il termine ancora oggi adottato: *neutrino.*

La proposta di questa particella fu fatta per la prima volta nella famosa lettera di Pauli alla riunione del 4 dicembre 1930 della sezione di Tubinga della Società Tedesca di Fisica, che inizia con il famoso "Cari colleghi radioattivi..." e continua dicendo: "[..]*come vi spiegherá piú dettagliatamente il portatore di questa missiva, che vi prego di ascoltare attentamente, a causa dell'enigma del momento angolare dei nuclei N-14 e Li-6 e dello spettro continuo nel decadimento β ho escogitato un estremo rimedio per salvare la regola di composizione dei momenti angolari e la legge di conservazione dell'energia. È possibile che esistano nei nuclei particelle elettricamente neutre, che desidero chiamare neutroni, che hanno spin 1/2 e obbediscono al* *principio di esclusione e che inoltre differiscono dai quanti di luce nel fatto che non viaggiano con la velocità della luce. La massa dei neutroni dovrebbe essere dello stesso ordine di grandezza della massa degli elettroni e in ogni caso non superiore a 1% della massa del protone. Lo spettro continuo nel decadimento β sarebbe allora comprensibile assumendo che nel decadimento β vengano simultaneamente emessi un neutrone e un elettrone in modo tale che la somma delle energie del neutrone e dell'elettrone sia costante.*[...]"

6.3 Urto elastico su una particella in quiete

Supponiamo ora una particella che incida su un altra in quiete in modo elastico e supponiamo le masse uguali: $m_1 = m_2 = m$.



Nel sistema S del laboratorio valgono le relazioni iniziali:

$$\begin{cases} p_1 = \dfrac{mu_1}{\sqrt{1-\beta^2}} \\ p_2 = 0 \end{cases} \qquad \begin{cases} E_1 = \dfrac{mc^2}{\sqrt{1-\beta^2}} \\ E_2 = mc^2 \end{cases}$$

Nel sistema del centro di massa vale invece:

$$V = \frac{Pc^2}{E} = \frac{mu_1c^2}{\sqrt{1-\beta^2}} \frac{1}{mc^2\left(1+\frac{1}{\sqrt{1-\beta^2}}\right)} \Rightarrow V = \frac{u_1}{1+\sqrt{1+\beta^2}}$$

Siccome vale:

$$u' = \frac{u-v}{1-\frac{uv}{c^2}}$$

nel sistema S':

$$\begin{cases} u_1' = \dfrac{u_1 - \frac{u_1}{1+\sqrt{1-\beta_1^2}}}{1-\frac{u_1}{c}\frac{u_1}{1+\sqrt{1-\beta_1^2}}} \\ u_2' = \dfrac{u_2 - v}{1-\frac{u_2 v}{c^2}} = -V \end{cases}$$

ovvero:

$$\frac{u_1 - \frac{u_1}{1+\sqrt{1-\beta_1^2}}}{1-\frac{u_1}{c^2}\frac{u_1}{1+\sqrt{1-\beta_1^2}}} = \frac{\frac{u_1\sqrt{1-\beta_1^2}}{1+\sqrt{1-\beta_1^2}}}{1-\beta_1^2\frac{1}{1+\sqrt{1-\beta_1^2}}} = u_1\frac{\sqrt{1-\beta_1^2}}{1+\sqrt{1-\beta_1^2}-\beta_1^2} = \frac{u_1}{1+\sqrt{1-\beta_1^2}} \equiv V$$

e quindi il fattore β è uguale per entrambe le particelle:

$$S' = \begin{cases} p_1' = -p_2' \\ E_1' = E_2' = \dfrac{mc^2}{\sqrt{1-\beta^2}} \end{cases}$$

dopo l'urto, $p_1 = -p_2 \Rightarrow E_1' = E_2' = \frac{1}{2}E'$, con $E^2 = m^2c^4 + p^2c^2$. D'altra parte:

$$p_1'^2 = \frac{E_1'^2}{c} - m^2c^2 = \frac{m^2c^2}{\sqrt{1-\beta^2}} - m^2c^2 = \frac{mV^2}{\sqrt{1-\beta^2}} \Rightarrow p_1' = -p_2'$$

Troviamo ora finalmente la relazione sugli angoli di diffusione. Nel sistema del laboratorio:

$$\tan\theta = \frac{p_{1y}}{p_{1x}} \qquad \tan\phi = \frac{p_{2y}}{p_{2x}} \qquad \tan\theta\tan\phi = -\frac{p_{1y}p_{2y}}{p_{1x}p_{2x}}$$

e trasformando le relazioni nel sistema S' del centro di massa:

$$p_{1y}p_{2y} = p'_{1y}p'_{2y} = -P'^2_{1y} \qquad\qquad p_{1x}p_{2x} = \frac{1}{\sqrt{1-\beta^2}}P'^2_y$$

che deriva dalla $-p^2_{1x} + \frac{V^2}{c^2}E'^2_1 \equiv -p^2_{1x} + P'^2 = P'^2_{1y}$. Segue allora:

$$\tan\theta\tan\phi = 1-\beta^2 = 1-\frac{V^2}{c^2} = 1-\frac{u^2}{c^2(1+\sqrt{1-\beta^2})}$$

ovvero in definitiva:

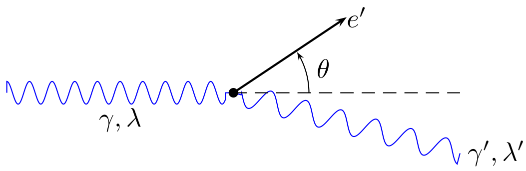
$$\boxed{\tan\theta\tan\phi = \frac{2}{1+\gamma}}$$

6.4 Diffusione Compton ed effetto Compton inverso

Si parla di diffusione Compton (o effetto Compton) quando un fotone di lunghezza d'onda λ¸ subisce una diffusione elastica da parte di una particella ed in seguito a questa la sua lunghezza d'onda varia diventando λ'.

Si parla di effetto Compton inverso, quando una particella subisce una diffusione elastica da parte di un fotone.[4]

6.4.1 Effetto Compton



La conservazione dell'energia si scrive come:

$$\begin{cases} p^e_\mu + p^\gamma_\mu & = p'^e_\mu + p'^\gamma_\mu \\ \left(p^e_\mu + p^\gamma_\mu - p'^\gamma_\mu\right)^2 & = -m^2c^2 \end{cases}$$

e sviluppando:[5]

$$\cancel{-m^2c^2} + 2p^e_\mu p^\gamma_\mu - 2p^e_\mu p'^\gamma_\mu - 2p^\gamma_\mu p^\gamma_\mu = \cancel{-m^2c^2}$$

con $p^e_\mu \equiv (imc, 0)$, $p^\gamma_\mu \equiv (i\frac{h}{\lambda}, \frac{h}{\lambda}\hat{k})$ e $p'^\gamma_\mu \equiv (i\frac{h}{\lambda'}, \frac{h}{\lambda'}\hat{k}')$. Sostituendo nella relazione si ricava:

$$-mc\frac{h}{\lambda} + mc\frac{h}{\lambda'} + \frac{h^2}{\lambda\lambda'} - \frac{h^2}{\lambda\lambda'}\hat{k}\cdot\hat{k}' = 0 \rightarrow$$
$$-mc(\lambda'-\lambda) + h(1-\cos\theta) = 0$$

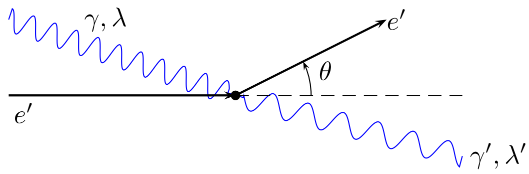
ovvero, in definitiva:

$$\boxed{\lambda' - \lambda = \frac{h}{mc}(1-\cos\theta)}$$

[4] È il caso, ad esempio, di un particella che incide su un gas di fotoni, come in un corpo nero.

[5] Si tenga presente che $\left(p^\gamma_\mu\right)^2 = 0$ e $\left(p'^\gamma_\mu\right)^2 = 0$

6.4.2 Effetto Compton inverso



Procedendo in maniera analogo all'effetto Compton:

$$\left(p^e_\mu + p^\gamma_\mu - p'^\gamma_\mu\right)^2 = -m^2c^2$$

solo che questa volta tutte le componenti sono diverse da zero:

$$\begin{aligned} \cancel{-m^2c^2} + \left(p^\gamma_\mu - p'^\gamma\mu\right) + 2p^e_\mu\left(p^\gamma_\mu - p'^\gamma\mu\right) = \cancel{-m^2c^2} \to \\ -\cancel{2}p^\gamma_\mu p'\gamma_\mu = \cancel{2}p^e_\mu\left(p^\gamma_\mu - p'^\gamma\mu\right) \to \\ -\frac{E^\gamma E'^\gamma}{c^2} + \frac{E^\gamma E'^\gamma}{c^2}\hat{k}\cdot\hat{k}' = -\frac{E^e}{c^2}\left(E^\gamma - E'^\gamma\right) + \cancel{\vec{p}}'\left(E^\gamma\hat{k} - E'^\gamma\hat{k}'\right) \end{aligned}$$

Consideriamo ora un angolo di 180°, quando si ha il massimo trasferimento di energia:

$$-2\frac{E^\gamma E'^\gamma}{c^2} = -\frac{E^2}{c^2}\left(E^\gamma - E'^\gamma\right) - \frac{p}{c}\left(E^\gamma + E'^\gamma\right) \Rightarrow E'^\gamma = \frac{E^\gamma(cp+E)}{2E^\gamma + E - pc}$$

con $pc = \sqrt{E^2 - m^2c^4}$. Se l'energia in gioco è molto superiore alla massa di riposo dell'elettrone $E^2 \gg m^2c^4$, si può sviluppare la radice in serie:

$$E\sqrt{1 - \frac{m^2c^4}{E^2}} \simeq E\left(1 - 1\frac{1}{2}\frac{m^2c^4}{E^2}\right)$$

che sostituita permette di ricavare:

$$\begin{aligned} E'^\gamma &= \frac{E^\gamma\left[E\left(1 - 1\frac{1}{2}\frac{m^2c^4}{E^2}\right) + E\right]}{2E^\gamma + E - E\left(1 - 1\frac{1}{2}\frac{m^2c^4}{E^2}\right)} = \frac{E^\gamma\left[E - \frac{1}{2}\frac{m^2c^4}{E} + E\right]}{2E^\gamma + \cancel{E} - \cancel{E} + \frac{1}{2}\frac{m^2c^4}{E}} = \\ &= \frac{\cancel{2E^\gamma}\left[E - \frac{1}{4}\frac{m^2c^4}{E}\right]}{\cancel{2E^\gamma}\left[1 + \frac{1}{4}\frac{m^2c^4}{EE^\gamma}\right]} \end{aligned}$$

ovvero:

$$\boxed{E'^\gamma = \frac{E - \frac{1}{4}\frac{m^2c^4}{E}}{1 + \frac{1}{4}\frac{m^2c^4}{EE^\gamma}}}$$

ed essendo tutti gli altri fattori piccoli, in caso di massimo trasferimento di energia si ha $E'^\gamma \simeq E$, ovvero un urto quasi elastico.

Parte II

Meccanica Quantistica Relativistica

Capitolo 7

Unificazione della Meccanica Quantistica con la Relatività Ristretta

7.1 Impostazione del problema

Nella meccanica quantistica non relativistica ad ogni osservabile è associato un operatore hermitiano. Valgono il principio di corrispondenza e di sovrapposizione (che deriva dalla linearità dell'equazione):

$$\begin{cases} \hat{x} & \to x \\ \hat{p} & \to \dfrac{\hbar}{i}\dfrac{\partial}{\partial x} \end{cases} \qquad\qquad \psi = \sum_n a_n \psi_n$$

e la media di un operatore $\hat{\Omega}$ è data da:

$$\langle\hat{\Omega}\rangle = \int \psi^* \hat{\Omega} \psi d^3x$$

L'interpretazione probabilistica della meccanica quantistica ci fornisce inoltre un'equazione di continuità per le probabilità:

$$\psi^*\psi = \varrho \qquad \frac{\partial}{\partial t}(\psi^*\psi) = i\hbar\vec{\nabla}\cdot\left[(\vec{\nabla}\psi^*)\psi - \psi^*\vec{\nabla}\psi\right] = \vec{\jmath}$$

dove la densità ϱ è definita positiva.

La conseguenza fondamentale della meccanica quantistica è la relazione di indeterminazione $\Delta x \Delta p \geq \hbar/2$. A questo, la relatività aggiunge "solo" la condizione $v < c$ e la relazione $E^2 = m^2c^4 + p^2c^2$. Ciò ha tuttavia grosse implicazioni sullo scarto indotto da p (e quindi da q) su E:

$$\Delta E = \Delta\left(\sqrt{m^2c^4 - p^2c^2}\right)$$

lo scarto indotto su E da p è esprimibile anche come $\Delta E = \frac{\partial E}{\partial p}\Delta p$, per cui:

$$\Delta E = \frac{c^2 p}{\sqrt{m^2c^4 - p^2c^2}}\Delta p = \frac{c^2p}{E}\Delta p$$

ricordando che $p = \frac{Ev}{c^2}$ ed utilizzando il principio di indeterminazione:

$$\Delta E = \frac{c^2p}{E}\Delta p = v\Delta p \to \frac{\delta q \Delta E}{v} \simeq \hbar \to \Delta q \Delta E \simeq v\hbar$$

e $v\hbar$ è una costante che al massimo vale $c\hbar$, per cui se $\delta q \to 0$, $\Delta E \to \infty$. Ma come abbiamo visto l'energia contribuisce alla massa (massa e energia sono interscambiabili: nota a pag.28), per cui all'aumentare di ΔE ci si deve aspettare la creazione di nuove particelle. *Una trattazione unipartitcellare non è quindi più valida, il numero di particelle deve essere una variabile dinamica e questo* ***non*** *è come trattare un sistema a più particelle.*

Vediamo cosa accade se cerchiamo direttamente una funzione d'onda che soddisfi la relazione relativistica fra energia e impulso. Il principio di corrispondenza permette di associare gli operatori:

$$\hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial \vec{x}} \qquad\qquad E = i\hbar\frac{\partial}{\partial t}$$

Le componenti di energia e impulso sono un quadrivettore, quindi:

$$p^{\mu} = \left(\vec{p}, \frac{iE}{c}\right) \quad p_4 = \frac{iE}{c} \rightarrow \hat{p}_4 = -\frac{\hbar}{c}\frac{\partial}{\partial t} = \hbar\left(\frac{1}{i \cdot i}\frac{\partial}{\partial t}\right) = \frac{\hbar}{i}\partial_4$$

l'operatore quantistico relativistico sarà quindi:

$$\boxed{\hat{p} = \frac{\hbar}{i}\partial_{\mu}}$$

Per applicare la relazione relativistica così com'è dovremmo considerare la radice di un operatore, il che comporta lo sviluppo in una serie infinita. Tuttavia, ricalcando la relazione $E^2 = m^2c^4 + p^2c^2$, consideriamone il quadrato:[1]

$$-\hbar^2\frac{\partial^2}{\partial t^2}\psi = \left[m^2c^4 - \hbar^2c^2\nabla^2\right]\psi$$

che introducendo l'operatore dalambertiano $\Box$ e scritta anche per la ψ^* fornisce le due equazioni (*equazione di Klein-Gordon*):

$$\begin{cases} [\Box - m^2c^2/\hbar^2]\psi = 0 \\ [\Box - m^2c^2/\hbar^2]\psi^* = 0 \end{cases}$$

Seguendo quindi la procedura standard per ricavare l'equazione di continuità, moltiplicando la prima per ψ^* e la seconda per ψ e sottraiamo:

$$[\psi^*\partial_\mu\partial_\mu\psi - (\partial_\mu\partial_\mu\psi^*)\psi] = 0 \Rightarrow \partial_\mu[\psi^*\partial_\mu\psi - (\partial_\mu\psi^*)\psi] = 0$$

si vede quindi che esiste sempre una corrente conservata, ma esplicitando la quarta componente, che è legata alla densità di proabilità:

$$ic\varrho = [\psi^*\partial_4\psi - (\partial_4\psi^*)\psi]$$

ci si rende conto che quest'ultima *non è definita positiva.*

7.2 L'equazione di Dirac

Tale problema condusse P.A.M. Dirac nel 1928 a formulare la sua equazione, che fra l'altro porta alla predizione dell'esistenza delle antiparticelle.[2]

L'equazione che Dirac cercava doveva soddisfare i seguenti requisiti:

1. Deve essere lineare (per conservare il principio di sovrapposizione proprio dell'equazione di Schrödinger);
2. Deve rispettare la condizione relativistica fra E e p, comprendere le derivate di spazio e tempo sullo stesso piano (per avere una formulazione relativistica);
3. Deve contenere solo le derivate prime (per non incorrere nel problema dell'equazione di Klein-Gordon);

[1] Ciò porterà tuttavia ad introdurre dei valori negativi nelle soluzioni.
[2] La situazione in cui il numero delle particelle è una variabile è detto *Seconda Quantizzazione*.

4. Deve tuttavia soddisfare l'equazione di Klein-Gordon $[\Box - m^2c^2/\hbar^2]\psi = 0$ (per contenere in sé la relazione relativistica (3.4));

5. Deve essere covariante a vista (per conservare l'invarianza per trasformazioni)

Impostiamo quindi una relazione lineare con le derivate prime, imponendo in tal modo i requisiti (1) e (3):

$$E = [\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta mc]\psi \Rightarrow [p_4 - \alpha_1 p_1 - \alpha_2 p_2 - \alpha_3 p_3 - \beta mc]\psi = 0 \tag{7.1}$$

I coefficienti α_i e β sono calcolati imponendo le altre condizioni. Moltiplichiamo infatti la (7.1) per $[p_4 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta mc]$ per imporre le condizioni (2) e (4). Si ottiene allora:

$$\left[p_4^2 - p_4 \sum_i \alpha_i p_i - p_4 \beta mc + \sum_i \alpha_i p_i \cdot p_4 - \sum_i \sum_j \frac{\alpha_i\alpha_j + \alpha_j\alpha_i}{2} p_i p_j + \right.$$
$$\left. - \sum_i \alpha_i \beta mc p_i + \beta mc p_4 - \beta mc \sum_i \alpha_i p_i - \beta^2 m^2 c^2 \right] \psi = 0$$

e tenendo presente che α_i e β commutano con p (sono in effetti dei coefficienti), si ricava:

$$\left[p_4^2 - \sum_i \sum_j \frac{\alpha_i\alpha_j + \alpha_j\alpha_i}{2} p_i p_j - \sum_i (\alpha_i\beta + \beta\alpha_i) mc p_i - \beta^2 m^2 c^2 \right] \psi = 0$$

Siamo giunti quindi ad un'equazione del II° ordine, che deve corrispondere alla (3.4) per la condizione (4). Esplicitando la componente p_4:

$$\left[-\frac{E^2}{c^2} + \sum_i \sum_j \frac{\alpha_i\alpha_j + \alpha_j\alpha_i}{2} p_i p_j + \sum_i (\alpha_i\beta + \beta\alpha_i) mc p_i + \beta^2 m^2 c^2 \right] \psi = 0$$

Per ottenere la (3.4) a partire da questa equazione, dobbiamo ritrovare il termine p^2 e i termini misti in mcp_i devono essere nulli. I coefficienti devono quindi sottostare alle condizioni:

$$\frac{\alpha_i\alpha_j + \alpha_j\alpha_i}{2} = \delta_{ij} \tag{7.2}$$
$$\alpha_i\beta + \beta\alpha_i = 0 \tag{7.3}$$
$$\beta^2 = \mathbb{I} \tag{7.4}$$

ed in questo caso la (7.1) diventa:

$$p_4\psi = \left[\sum_i \alpha_i p_i + \beta mc \right] \psi$$

Ricerchiamo ora dettagliatamente i coefficienti da inserire nell'equazione. I coefficienti α_i soddisfano la proprietà di *anticommutazione* (7.2) e quelle delle matrici. Questo implica immediatamente che ψ non può essere uno scalare, ma un vettore sul quale queste matrici possono operare. La proprietà (7.3) ci permette di dedurre:

$$\alpha_i\beta = -\beta\alpha_i \qquad \Rightarrow \qquad \det\alpha_i \cdot \det\beta = (-1)^N \det\beta \cdot \det\alpha_i$$

in N dimensioni. Ma il determinante è un numero, quindi **N deve essere pari**. Essendo anche il determinante diverso da 0, possiamo prendere le matrici inverse e sempre dalla (7.3) ricavare le relazioni:

$$\alpha_i^{-1}\alpha_i\beta = -\alpha_i^{-1}\beta\alpha_i \quad \Rightarrow \quad \beta = -\alpha_i^{-1}\beta\alpha_i$$

e prendendone le tracce:[3]

$$\mathrm{Tr}\,\beta = -\,\mathrm{Tr}\,\beta \quad \Rightarrow \quad \mathrm{Tr}\,\beta = 0$$

moltiplicando invece per β^{-1} si ricava:

$$\beta^{-1}\alpha_i\beta = -\beta^{-1}\beta\alpha_i \quad \Rightarrow \quad \alpha_i = -\beta^{-1}\alpha_i\beta \quad \Rightarrow \quad \mathrm{Tr}\,\alpha_i = \mathrm{Tr}(-\alpha_i) \quad \Rightarrow \quad \mathrm{Tr}\,\alpha_i = 0$$

le 3 matrici α_i e la β sono quindi matrici non unitarie a traccia nulla di dimensione pari. La matrice β, per via della traccia nulla e della condizione (7.4), deve pertanto avere la forma:

$$\beta = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}$$

La dimensione minima che soddisfa tutte queste richieste è 4,([4]) si noti tuttavia che questo numero non ha alcuna relazione con le dimensioni dello spazio-tempo. Notando che in uno spazio 2×2 una base matriciale è costituita dalla matrice identità e dalle tre matrici di Pauli σ_k:

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

e notando inoltre che le matrici di Pauli possono essere utilizzate a livello più generale come elementi di matrici a blocchi, si può senz'altro scegliere come base:[5]

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}$$

Questa scelta costituisce la cosiddetta *rappresentazione standard* o *rappresentazione di Dirac.*[6]

L'equazione in forma operatoriale assume questo aspetto:

$$i\frac{\partial}{\partial t}\psi = \frac{c}{i}\sum_k \alpha_k\partial_k\psi + \frac{mc^2}{\hbar}\beta\psi \tag{7.5}$$

la cui complessa coniugata è:[7]

$$-i\frac{\partial}{\partial t}\psi^\dagger = -\frac{c}{i}\sum_k \partial_k\psi^\dagger\alpha_k + \frac{mc^2}{\hbar}\psi^\dagger\beta$$

Seguendo il procedimento standard per trovare l'equazione di continuità, possiamo ora moltiplicare la prima per $\psi^\dagger$ a sinistra e la seconda per ψ a destra. Sottraendo poi la seconda dalla prima si ottiene:

$$i\left[\psi^\dagger\frac{\partial}{\partial t}\psi + \frac{\partial}{\partial t}\psi^\dagger\cdot\psi\right] = \frac{c}{i}\sum_k\left(\psi^\dagger\alpha_k\partial_k\psi + \partial_k\psi^\dagger\alpha_k\psi\right)$$

[3] In effetti vale il teorema $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)$.

[4] Non è possibile soddisfare queste condizioni in uno spazio a dimensione 2, in quanto una base è composta dalle tre matrici di Pauli più la matrice unitaria, a causa della particolare forma della terza matrice σ_3

[5] Ovvero, esplicitamente:

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad \alpha_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

[6] Altre due rappresentazione possibili sono quella di Weyl e quella di Majorana.

[7] Siccome si ha a che fare con vettori, con la notazione $\psi^\dagger$ si indica l'operazione di coniugazione complessa degli elementi più quella di inversione righe/colonne. In pratica, se $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ allora $\psi^\dagger = (\psi_1^*, \psi_2^*)$.

ovvero:

$$i\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) = \frac{c}{i}\sum_k \partial_k\left(\psi^\dagger\alpha_k\psi\right)$$

È possibile quindi definire una densità di probabilità ed una densità di corrente di probabilità rispettivamente come:

$$\begin{aligned} \varrho &= \psi^\dagger\psi \\ j_k &= c\psi^\dagger\alpha_k\psi \end{aligned}$$

ed ottenere infine l'equazione di continuità nella solita forma, dove il termine $c\vec{\alpha}$ risulta legato alla velocità ed in cui la densità di probabilità è effettivamente definita positiva:

$$\frac{\partial\varrho}{\partial t} + \vec{\nabla}\cdot\vec{j} = 0$$

Resta da soddisfare infine la condizione (5). L'equazione (7.5) è in effetti già covariante *in nuce* e si esplicita semplicemente moltiplicando a sinistra per β:

$$\frac{i\beta}{c}\frac{\partial}{\partial t}\psi + i\beta\alpha_k\partial_k\psi - \frac{mc}{\hbar}\psi = 0$$

Introduciamo ora la notazione standard di Dirac, facendo le posizioni (*Matrici di Dirac*):

$$i\beta \equiv \gamma_4 \qquad \beta\alpha_i \equiv \gamma_i$$

Questo permette di riscrivere l'equazione in forma più compatta:

$$\begin{aligned} i\gamma_4\partial_4\psi + i\vec{\gamma}\cdot\vec{\nabla}\psi - \frac{mc}{\hbar}\psi = 0 \rightarrow \\ -\gamma_4\frac{\hbar}{i}\partial_4\psi - \frac{\hbar}{i}\vec{\gamma}\cdot\vec{\nabla}\psi - mc\psi = 0 \rightarrow \\ \left(\gamma_4 p_4 + \vec{\gamma}\cdot\vec{p} + mc\right)\psi = 0 \end{aligned}$$

La covarianza a vista è ormai già evidente. Riscrivendo l'equazione in termini di quadrivettori si trova la forma definitiva per l'equazione di Dirac:

$$\boxed{(\gamma_\mu\cdot\hat{p}_\mu + mc)\psi = 0}$$

Si può introdurre anche la notazione $\hat{\not{P}} \equiv \gamma_\mu\cdot\hat{p}_\mu$,[8] che permette di scrivere l'equazione di Dirac nella forma eccezionalmente compatta:

$$\boxed{(\hat{\not{P}} + mc)\psi = 0}$$

Evidenziamo ora alcune proprietà delle matrici γ_μ. In primo luogo:

$$\left.\begin{aligned} \gamma_4^\dagger = -i\beta^\dagger = -\gamma_4 \\ \gamma_i^\dagger = \alpha_1^\dagger\beta^\dagger = \alpha_i\beta = -\beta\alpha_i = -\gamma_i \end{aligned}\right\} \Rightarrow \gamma_\mu^\dagger = -\gamma_\mu$$

Dalla 7.2 si ricava invece:

$$\alpha_i\beta^2\alpha_j + \alpha_j\beta^2\alpha_i = 2\delta_{ij} \Rightarrow -\beta\alpha_i\beta\alpha_j - \beta\alpha_j\beta\alpha_i = 2\delta ij \Rightarrow \gamma_i\gamma_j + \gamma_j\gamma_i = -2\delta_{ij}$$

ed in particolare, se $i = j$, $\gamma_i^2 = -\mathbb{I}$. Inoltre, per la quarta componente:

$$\alpha_i\beta + \beta\alpha_i = 0 \Rightarrow -i\gamma_i\gamma_4 - i\gamma_4\gamma_i = 0$$

e dunque in definitiva vale l'importante relazione:

$$\boxed{\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = -2\delta_{\mu\nu}}$$

[8] Questa notazione si usa per snellire le formule in cui compare il prodotto della matrice γ_μ, quindi in generale $\hat{\not{A}} \equiv \gamma_\mu\cdot\hat{A}^\mu$.

Capitolo 8

Sviluppi sull'Equazione di Dirac

8.1 Formalismo dei Bispinori. Interazione elettromagnetica

Iniziamo per semplicità con il considerare una particella a impulso nullo. L'equazione in forma operatoriale diventa quindi:

$$i\hbar\frac{\partial}{\partial t}\psi = \beta mc^2\psi \quad \Rightarrow \quad \frac{\partial}{\partial t}\psi = -\frac{i\beta mc^2}{\hbar}\psi$$

che può essere integrata componente per componente. In quattro dimensioni – come è il caso presente – una base di vettori è fornita banalmente da:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad ; \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad ; \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad ; \quad e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

ricordando che $\beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}$, si ritrovano per soluzione i ben noti esponenziali:

$$\psi_1(t) = e^{-\frac{imc^2}{\hbar}t}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad ; \qquad \psi_2(t) = e^{-\frac{imc^2}{\hbar}t}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

ai quali risultano ora affiancate altre due soluzioni che devono essere prese in considerazione:

$$\psi_3(t) = e^{\frac{imc^2}{\hbar}t}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad ; \qquad \psi_4(t) = e^{\frac{imc^2}{\hbar}t}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Le prime due soluzioni rappresentano l'evoluzione temporale di un'onda a due componenti: sono le funzioni rappresentanti lo *spin semintero*. Le seconde soluzioni sono invece la descrizione di particelle ad energia negativa, che la teoria interpreta come *antiparticelle*. Questa semplificazione – quella di trascurare il termine dell'impulso – corrisponde in realtà al limite non relativistico, in cui il termine predominante è proprio βmc.

Introduciamo ora nella equazione di Dirac l'interazione elettromagnetica, operando la sostituzione $p^\mu \to p^\mu + eA/c$. L'equazione in forma operatoriale si trasforma allora nella seguente:

$$\frac{i\hbar}{c}\frac{\partial}{\partial t}\psi = (\vec{\alpha}\cdot\vec{p} + \beta mc)\psi \qquad \Rightarrow \qquad \frac{i\hbar}{c}\frac{\partial}{\partial t}\psi = (\vec{\alpha}\cdot\vec{p} - \frac{e}{c}\vec{\alpha}\cdot\vec{A} + e\Phi + \beta mc)\psi$$

Ne segue che l'Hamiltoniana di questo sistema si può scrivere nella forma $H = H_0 + H'$, dove H' rappresenta l'Hamiltoniana di interazione con il campo elettromagnetico:

$$H' = -\frac{e}{c}\vec{\alpha} \cdot \vec{A} + e\Phi$$

e dove $\vec{\alpha}$, come già detto sopra, assume il ruolo della velocità. Questo permette di riscrivere il teorema di Ehrenfest come:

$$\frac{d\vec{r}}{dt} = \langle \frac{\hbar}{i}[H, \vec{r}] \rangle = \langle c\vec{\alpha} \rangle$$

Le conclusioni precedenti sulle soluzioni dell'equazione di Dirac ci autorizzano a scrivere la funzione d'onda ψ in termini di un bispinore:[1]

$$\psi = \begin{bmatrix} \varphi' \\ \chi' \end{bmatrix}$$

Scriviamo quindi l'equazione di Dirac comprendente l'interazione elettromagnetica nel formalismo dei bispinori:

$$i\hbar\frac{\partial}{\partial t}\begin{bmatrix} \varphi' \\ \chi' \end{bmatrix} = c\vec{\sigma} \cdot \vec{\pi}\begin{bmatrix} \chi' \\ \varphi' \end{bmatrix} + e\Phi\begin{bmatrix} \varphi' \\ \chi' \end{bmatrix} + mc^2\begin{bmatrix} \varphi' \\ -\chi' \end{bmatrix}$$

dove $\vec{\pi} = \vec{p} - e/c\vec{A}$ e si è tenuto presente che:

$$\begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}\begin{bmatrix} \varphi' \\ \chi' \end{bmatrix} = \begin{bmatrix} \sigma_k\chi' \\ \sigma_k\varphi' \end{bmatrix} = \sigma_k\begin{bmatrix} \chi' \\ \varphi' \end{bmatrix}$$

Nel limite non relativistico il termine predominante è ovviamente $mc^2\begin{bmatrix} \varphi' \\ -\chi' \end{bmatrix}$. Ricerchiamo ora le soluzioni per la particella libera nella forma:

$$\begin{bmatrix} \varphi' \\ \chi' \end{bmatrix} = e^{-i\frac{mc^2}{\hbar}t}\begin{bmatrix} \varphi \\ \chi \end{bmatrix}$$

e sostituendo questa forma nell'equazione ai bispinori scritta sopra:

$$\left\{i\hbar\frac{\partial}{\partial t}\begin{bmatrix} \varphi \\ \chi \end{bmatrix} + mc^2\begin{bmatrix} \varphi \\ \chi \end{bmatrix}\right\}e^{-i\frac{mc^2}{\hbar}t} = \left\{c\vec{\sigma} \cdot \vec{\pi}\begin{bmatrix} \chi \\ \varphi \end{bmatrix} + e\Phi\begin{bmatrix} \varphi \\ \chi \end{bmatrix} + mc^2\begin{bmatrix} \varphi \\ -\chi \end{bmatrix}\right\}e^{-i\frac{mc^2}{\hbar}t}$$

ovvero, eliminando il fattore esponenziale in comune ai due e mai nullo:

$$i\hbar\frac{\partial}{\partial t}\begin{bmatrix} \varphi \\ \chi \end{bmatrix} = c\vec{\sigma} \cdot \vec{\pi}\begin{bmatrix} \chi \\ \varphi \end{bmatrix} + e\Phi\begin{bmatrix} \varphi \\ \chi \end{bmatrix} - 2mc^2\begin{bmatrix} 0 \\ \chi \end{bmatrix} \tag{8.1}$$

Nel limite non relativistico predominano chiaramente i termini in c e c^2. Separando quindi l'equazione per le componenti alte e basse, nel limite non relativistico otteniamo:

$$\begin{cases} c\vec{\sigma} \cdot \vec{\pi}\chi & = 0 \\ c\vec{\sigma} \cdot \vec{\pi}\varphi - 2mc^2\chi & = 0 \end{cases}$$

Le componenti alte restituiscono quindi in questo caso ("classico") $\chi = 0$ e questo è corretto: a energie sufficientemente basse non ci sono componenti legate alle antiparticelle. Le componenti basse ci permettono invece di giungere alla relazione:

$$\chi = \frac{\vec{\sigma} \cdot \vec{\pi}}{2mc}\varphi$$

Questa relazione permette di identificare le componenti basse come le componenti "piccole" in rapporto alle componenti alte φ. Queste componenti sono infatti ridotte di un fattore $v/c \ll 1$ nel limite

[1] Ovvero in termini di un particolare vettore che può essere pensato composto di due parti, in cui ogni parte contiene le due componenti dello spin.

non relativistico rispetto alle φ. Sostituendo la forma delle χ nelle componenti alte dell'equazione (8.1) si ricava:

$$i\hbar\frac{\partial}{\partial t}\varphi = \frac{(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi})}{2mc}\varphi + e\Phi\varphi$$

siccome vale:[2]

$$(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi}) = \left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 - \frac{e\hbar}{c}\vec{\sigma}\cdot(\vec{\nabla}\times\vec{A})$$

l'equazione per le componenti alte (particelle) nel caso non relativistico diventa quindi (*Equazione di Pauli per particelle con spin*):

$$\boxed{i\hbar\frac{\partial}{\partial t}\varphi = \left[\frac{(\vec{p}-\frac{e}{c}\vec{A})^2}{2m} - \frac{e\hbar}{2m}\vec{\sigma}\cdot\vec{B} + e\Phi\right]\varphi}$$

Ritrovare questa equazione significa che l'equazione di Dirac è effettivamente un buon punto di partenza per costruire una teoria quantistica relativistica. Infatti, le due componenti alte sono sufficienti per i due gradi di libertà associato allo spin 1/2 di una particella libera, ritrovando anche il corretto momento magnetico dell'elettrone corrispondente al rapporto giromagnetico g=2. Infatti, tenendo presente che $\vec{A} = 1/2\vec{B}\times\vec{r}$ e prendendo solo i termini al primo ordine nel campo magnetico $\vec{B}$:

$$\frac{1}{2m}\left[\vec{p} - \frac{e}{2c}\vec{B}\times\vec{r}\right]^2 \simeq \frac{1}{2m}\left[p^2 - \frac{e}{c}\vec{p}\cdot(\vec{B}\times\vec{r})\right]$$

dalle proprietà del prodotto misto si sa che: $\vec{p}\cdot(\vec{B}\times\vec{r}) = (\vec{p}\times\vec{B})\cdot\vec{r} = (\vec{r}\times\vec{p})\cdot\vec{B} = \vec{L}\cdot\vec{B}$, per cui denotando con $\vec{S}\equiv 1/2\hbar\vec{\sigma}$ l'operatore di spin l'equazione di Pauli assume la forma:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left[\frac{p^2}{2m} - \frac{e}{2mc}(\vec{L}+2\vec{S})\cdot\vec{B}\right]\varphi$$

dove il coefficiente di interazione dello spin con il campo magnetico è appunto g=2.

8.2 Invarianza per trasformazioni

Si riconsideri ora l'equazione di Dirac:

$$[\gamma_\mu\cdot p_\mu + mc]\psi = 0 \qquad \text{ovvero} \qquad \left[\frac{\hbar}{i}\gamma_\mu\partial_\mu + mc\right]\psi = 0$$

La richiesta di covarianza per trasformazioni di Lorentz implica la ricerca di un operatore tale che $\psi(x)\to\psi'(x')$ in modo da soddisfare ancora la stessa equazione. Di conseguenza, questo equivale a cercare una matrice che soddisfi la trasformazione $\psi'(x') = S(\Lambda)\psi(x)$, e quindi che soddisfi anche le:

$$\psi'(x') = S(\Lambda)\psi(\Lambda^{-1}x')$$
$$\psi(x) = S^{-1}(\Lambda)\psi'(x')$$
$$\psi(x) = S(\Lambda^{-1})\psi'(x)$$

[2] Vale la relazione per il prodotto scalare: $(\vec{\sigma}\cdot\vec{a})(\vec{\sigma}\cdot\vec{b}) \equiv \vec{a}\cdot\vec{b} + i\sigma\cdot(\vec{a}\times\vec{b})$ e $\vec{\pi} = \vec{p} - e/c\vec{A}$. Il prodotto vettoriale esplicito fornisce:

$$i\epsilon_{ijk}\sigma_k\left(p - \frac{e}{c}A\right)_i\left(p - \frac{e}{c}A\right)_j = i\epsilon_{ijk}\sigma_k\left[p_ip_j + \frac{e^2}{c^2}A_iA_j - \frac{e}{c}A_ip_j - \frac{e}{c}p_iA_j\right]$$

I primi due termini scompaiono (si tratta in effetti di un prodotto vettore per sé stesso), per gli altri due termini vale ricordando che $\vec{p}$ e $\vec{A}$ sono operatori:

$$\epsilon_{ijk}\left(p_iA_j + A_ip_j\right) = \epsilon_{ijk}\left(A_jp_i + [p_i, A_j] + A_ip_j\right)$$

il termine $\epsilon_{ijk}(A_jp_i + A_ip_j)$ è nullo in quanto la parentesi è simmetrica in i e j e ϵ_{ijk} è antisimmetrico, resta quindi il termine $\epsilon_{ijk}[p_i, A_j]$ e esplicitando gli operatori:

$$-i\hbar\epsilon_{ijk}\left(\partial_iA_j - A_j\partial_i\right) \text{ ovvero } -i\hbar\vec{\nabla}\times\vec{A}$$

e cioé, in definitiva, che rispetti:

$$S^{-1}(\Lambda) = S(\Lambda^{-1})$$

Riscriviamo quindi l'equazione di Dirac in termini di ψ' e x':

$$\left[\frac{\hbar}{i}\tilde{\gamma}_\mu \partial'_\mu + mc\right]\psi'(x') = 0$$

dove abbiamo fatto la posizione:

$$\tilde{\gamma}_\mu = u^\dagger \gamma_\mu u \qquad u^\dagger u = \mathbb{I}$$

L'equazione di Dirac può essere espressa in questi termini:

$$\frac{\hbar}{i}\gamma_\mu \partial_\mu S^{-1}(\Lambda)\psi' + mcS^{-1}(\Lambda)\psi' = 0$$

che moltiplicata a sinistra per $S(\Lambda)$ fornisce:

$$\frac{\hbar}{i}S(\Lambda)\gamma_\mu \partial_\mu S^{-1}(\Lambda)\psi' + mc\psi' = 0$$

ricordando che vale la relazione:

$$\frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\mu}$$

l'equazione di sopra diventa:

$$\frac{\hbar}{i}S(\Lambda)\gamma_\mu S^{-1}(\Lambda)\Lambda_{\mu\nu}\partial'_\nu \psi' + mc\psi' = 0$$

la trasformazione $S(\Lambda)$ rimane quindi invariata l'equazione di Dirac se risulta verificata la relazione:

$$S(\Lambda)\gamma_\mu S^{-1}(\Lambda)\Lambda_{\mu\nu} = \gamma_\nu$$

siccome la matrice $\Lambda_{\mu\nu}$ commuta con $S(\Lambda)$ in quanto agiscono su spazi diversi, mandiamo Λ in Λ^{-1} e moltiplichiamo poi per Λ:

$$S^{-1}(\Lambda)\gamma_\mu S(\Lambda) = \Lambda_{\mu\nu}\gamma_\nu \tag{8.2}$$

Passiamo ora a considerare trasformazioni infinitesime:

$$\Lambda_{\mu\nu} = \delta_{\mu\nu} + \epsilon\omega_{\mu\nu}$$
$$S(\Lambda) = \mathbb{I} - \frac{i}{4}\epsilon\sigma_{\mu\nu}\omega_{\mu\nu}$$

dove le $\sigma_{\mu\nu}$ sono 4 matrici antisimmetriche. Sostituiamo queste nella relazione (8.2) e ricordando che gli indici muti possono essere rinominati:

$$\begin{aligned}
(\delta_{\mu\nu} + \epsilon\omega_{\mu\nu})\gamma_\nu &= (\mathbb{I} - \frac{i}{4}\epsilon\sigma_{\alpha\beta}\omega_{\alpha\beta})\gamma_\mu(\mathbb{I} - \frac{i}{4}\epsilon\sigma_{\lambda\rho}\omega_{\lambda\rho}) \rightarrow \\
\epsilon\omega_{\mu\nu}\gamma_\nu &= \frac{i}{4}\epsilon(\sigma_{\alpha\beta}\omega_{\alpha\beta}\gamma_\mu - \gamma_\mu\sigma_{\alpha\beta}\omega_{\alpha\beta}) \rightarrow \\
\epsilon\omega_{\mu\nu}\gamma_\nu &= \frac{i}{4}\epsilon\omega_{\alpha\beta}(\sigma_{\alpha\beta}\gamma_\mu - \gamma_\mu\sigma_{\alpha\beta}) \rightarrow \\
\frac{\omega_{\alpha\beta}}{2}(\delta_{\alpha\mu}\gamma_\beta - \delta_{\beta_\mu}\gamma_\alpha) &= \frac{i}{4}\epsilon\omega_{\alpha\beta}(\sigma_{\alpha\beta}\gamma_\mu - \gamma_\mu\sigma_{\alpha\beta}) \rightarrow \\
2(\delta_{\alpha\mu}\gamma_\beta - \delta_{\beta_\mu}\gamma_\alpha) &= i(\sigma_{\alpha\beta}\gamma_\mu - \gamma_\mu\sigma_{\alpha\beta})
\end{aligned}$$

che permette di ricavare la relazione per le matrici $\sigma_{\alpha\beta}$:

$$\sigma_{\alpha\beta} = \frac{i}{2}[\gamma_\alpha, \gamma_\beta]$$

e quindi le rotazioni infinitesime sono espresse da:

$$S = \mathbb{I} - \frac{1}{8}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)\epsilon\omega_{\mu\nu} = \mathbb{I} - \frac{1}{4}\gamma_\mu\gamma_\nu\epsilon\omega_{\mu\nu}$$

Consideriamo ora la parte spaziale del generatore delle rotazioni:

$$\vec{\gamma} = \beta\vec{\alpha} = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \cdot \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}$$

ora, ricordando anche che vale $\sigma_i\sigma_j = i\epsilon_{ijk}\sigma_k$:

$$\gamma_i\gamma_j = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} = i\epsilon_{ijk} \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \equiv i\epsilon_{ijk}\Sigma_k$$

dove con Σ_k abbiamo indicato $\begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}$, una sorta di matrice $\vec{\sigma}$ in 4 dimensioni. La parte spaziale risulta in definitiva:

$$S = \mathbb{I} + \frac{i}{4}\epsilon_{ijk}\Sigma_k\epsilon\omega_{ij}$$

che è una matrice simile a quelle di rotazione della ψ per particelle con spin. La parte temporale fornisce:

$$\gamma_i\gamma_4 = i\begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} = i\begin{pmatrix} 0 & -\vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} = -i\begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \equiv -i\vec{\alpha}$$

ma $c\vec{\alpha}$ risulta legata alla velocità, quindi le componenti temporali delle rotazioni infinitesime corrispondono alle trasformazioni di Lorentz propriamente dette.

Consideriamo ora l'equazione di Dirac scritta per le componenti $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$. L'equazione si scrive in questo caso:

$$-\frac{\hbar}{c}\frac{\partial}{\partial t}\psi^\dagger + \frac{\hbar}{i}\vec{\nabla}\psi^\dagger\alpha - \psi^\dagger\beta mc = 0$$

siccome $\beta^2 = \mathbb{I}$, l'equazione di sopra si può riscrivere come:

$$\hbar\partial_4\psi^\dagger\beta^2 + \frac{\hbar}{i}\psi^\dagger\beta^2\alpha - \psi^\dagger\beta mc = 0$$
$$\frac{\hbar}{i}\partial_4\psi^\dagger\beta\gamma_4 + \frac{\hbar}{i}\vec{\nabla}\psi^\dagger\beta\vec{\gamma} - \psi^\dagger\beta mc = 0$$

dove la quantità $\psi^\dagger\beta$ rappresenta in sostanza la $\psi^\dagger$ con le ultime due componenti cambiate di segno. Definendo $\bar{\psi} \equiv \psi^\dagger\beta$, l'equazione assume la forma:

$$\bar{\psi}(\gamma_\mu p_\mu - mc) = 0$$

Ne consegue che quello che deve avere senso fisico è la quantità $\bar{\psi}$ e non la semplice $\psi^\dagger$. Infatti, come abbiamo appena visto, è la quantità $\bar{\psi}$ sulla quale l'equazione di Dirac conserva la sua forma, ovvero risulta covariante. La densità di corrente per la $\bar{\psi}$ si può ricavare a partire dall'equazione scritta per la ψ, dove definendo $\varrho \equiv \psi^\dagger\psi$ e $\vec{\jmath} \equiv c\psi^\dagger\vec{\alpha}\psi$:

$$\frac{\partial}{\partial t}\varrho = c\vec{\nabla}\left[\psi^\dagger\vec{\alpha}\psi\right] = 0$$

Ma la densità di corrente e di probabilità possono essere scritte anche come:

$$\vec{\jmath} = c\psi^\dagger\beta^2\vec{\alpha}\psi = c\bar{\psi}\vec{\gamma}\psi$$
$$\varrho = \psi^\dagger\beta^2\psi = \bar{\psi}\beta\psi = -i\bar{\psi}\gamma_4\psi$$

ne consegue quindi che il quadrivettore $j_\mu = c\bar{\psi}\gamma_\mu\psi$ soddisfa anch'esso l'equazione di continuità:

$$\partial_\mu j_\mu = 0$$

Siamo ora pronti per vedere come trasforma la quantità $\bar{\psi}$ e quindi il quadrivettore corrente $j_\mu = c\bar{\psi}\gamma_\mu\psi$:

$$\psi'(x') = S(\Lambda)\psi(x) \qquad \psi'^\dagger(x') = \psi^\dagger(x)S^\dagger(x)$$

ma:

$$\psi'^\dagger(x')\beta \equiv \bar{\psi}'(x') = \psi^\dagger\beta^2 S^\dagger(\Lambda)\beta \equiv \bar{\psi}(x)\beta S^\dagger(\Lambda)\beta$$

siccome è dimostrabile che $\beta S^\dagger(\Lambda)\beta = S^{-1}(\Lambda)$, la precedente è riscrivibile nella forma:

$$\bar{\psi}'(x') = \bar{\psi}(x)S^{-1}(\Lambda)$$

Il termine della corrente di probabilità assume allora la forma:

$$\bar{\psi}\gamma_\mu\psi = \bar{\psi}S^{-1}(\Lambda)\gamma_\mu S(\Lambda)\psi$$

Ma $S^{-1}(\Lambda)\gamma_\mu S(\Lambda)$ è la relazione (8.2) che definisce $S(\Lambda)$ e quindi il termine appena scritto equivale a $\bar{\psi}\Lambda_{\mu\nu}\gamma_\nu\psi$. La legge di trasformazione diventa pertanto:

$$\bar{\psi}(x)\gamma_\mu\psi(x) = \bar{\psi}(x)\Lambda_{\mu\nu}\gamma_\nu\psi(x) = \Lambda_{\mu\nu}\bar{\psi}(x)\gamma_\nu\psi(x)$$

ed in definitiva:

$$\vec{j}'_\mu = \Lambda_{\mu\nu}\vec{j}_\nu$$

ovvero trasforma esattamente come un quadrivettore.

L'insieme delle matrici combinazioni di γ_μ, $\bar{\psi}$ e ψ aventi proprietà definite di trasformazione (di Lorentz) sono 16, si indicano complessivamente con Γ e sono denominate *covarianti bilineari*. Quelle riportate nella tabella seguente sono le uniche possibili, in quanto tutte le altre combinazioni di matrici γ_μ si riducono a queste sfruttando le proprietà di queste matrici.

Simbolo	Combinazione	Applicazione sulla funzione d'onda	Tipo di Oggetto
Γ	$\mathbb{I}$	$\bar{\psi}\psi$	Scalare
Γ_μ	γ_μ	$\bar{\psi}\gamma_\mu\psi$	Quadrivettore
$\Gamma_{\mu\nu}$	$i[\gamma_\mu, \gamma_\nu]$	$\bar{\psi}\Gamma_{\mu\nu}\psi$	Tensore Antisimmetrico di rango 2
γ_5	$\gamma_1\gamma_2\gamma_3\gamma_4$	$\bar{\psi}\gamma_5\psi$	Pseudoscalare
$\gamma_5\gamma_\mu$	$\gamma_5\gamma_\mu$	$\bar{\psi}\gamma_5\gamma_\mu\psi$	Pseudovettore

8.3 Ancora sui Bispinori. Soluzioni a massa nulla

Torniamo ora a considerare la ψ in termini di bispinori $\begin{bmatrix} \varphi \\ \chi \end{bmatrix}$.

Abbiamo visto che l'equazione in questo caso fornisce per le due componenti (eq.(8.1)):

$$\begin{cases} i\dfrac{\hbar}{c}\dfrac{\partial}{\partial t}\varphi & = \vec{\sigma}\cdot\vec{\pi}\chi + mc\varphi \\ i\dfrac{\hbar}{c}\dfrac{\partial}{\partial t}\chi & = \vec{\sigma}\cdot\vec{\pi}\varphi + mc\chi \end{cases}$$

Cerchiamo ora di disaccoppiare le due equazioni con una opportuna trasformazione unitaria: $\psi' \mapsto u\psi$ e $\gamma'_\mu \to u^\dagger\gamma_\mu u$ in modo che questa lasci l'equazione invariata. Utilizzando le matrici:

$$\beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \qquad \rho = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}$$

e definendo la trasformazione come $u \equiv \frac{1}{\sqrt{2}}(\beta+\rho)$, allora:

$$uu^\dagger = \frac{1}{\sqrt{2}}(\beta+\rho)\frac{1}{\sqrt{2}}(\beta+\rho) = \frac{1}{2}(\beta^2+\beta\rho+\rho\beta+\rho^2) = \mathbb{I}+\frac{1}{2}(\beta\rho+\rho\beta)$$

Essendo:

$$\beta\rho = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \cdot \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}$$
$$\rho\beta = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}$$

risulta effettivamente $uu^\dagger = \mathbb{I}$ e quindi u è unitaria. Applichiamola quindi alla ψ:

$$\psi' = u\psi = \frac{1}{\sqrt{2}}(\beta+\rho)\begin{bmatrix} \varphi \\ \chi \end{bmatrix}$$

e denominando $\psi' \equiv \begin{bmatrix} \xi \\ \eta \end{bmatrix}$, si ricava:

$$\begin{cases} \xi &= \frac{1}{\sqrt{2}}(\varphi+\chi) \\ \eta &= \frac{1}{\sqrt{2}}(\varphi-\chi) \end{cases}$$

Operando ora con la trasformazione u anche su α, β:

$$\alpha_i' = u\alpha_i u^\dagger = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}$$
$$\beta = u\beta u^\dagger = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}$$

e l'equazione diventa:

$$i\frac{\hbar}{c}\frac{\partial}{\partial t}\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \frac{\hbar}{i}\begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \cdot \vec{p}\begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} mc \begin{bmatrix} \xi \\ \eta \end{bmatrix}$$

ovvero, per le singole equazioni:

$$i\frac{\hbar}{c}\frac{\partial}{\partial t}\xi = \vec{\sigma}\cdot\vec{p}\,\xi + mc\eta$$

$$i\frac{\hbar}{c}\frac{\partial}{\partial t}\eta = -\vec{\sigma}\cdot\vec{p}\,\eta + mc\xi$$

che si disaccoppiano completamente per particelle a massa zero.[3] Occorre fare attenzione al fatto che questa equazione non vale per particella a massa zero come i fotoni, perché questi ultimi sono a spin 1, mentre l'equazione di Dirac vale solo per particelle a spin semintero $\pm 1/2$.

8.4 Teoria di Dirac. Antiparticelle

Ricerchiamo ora esplicitamente delle soluzioni all'equazione di Dirac per la particella libera. Per fare ciò, scriviamola in forma operatoriale:

$$i\frac{\hbar}{c}\frac{\partial}{\partial t}\psi = (\vec{\alpha}\cdot\vec{p} + \beta mc)\psi$$

[3] All'epoca del corso, nel 1992-93, si era tentati di interpretare queste soluzioni come neutrini, ma è stato finalmente dimostrato che queste particelle hanno una massa, che se estremamente piccola. Quindi, anche per i neutrini queste due equazioni non si disaccoppiano perfettamente, ma "quasi".

e ricerchiamo delle soluzioni in termini della trasformata di Fourier:

$$\psi(\vec{r},t) = \int \psi_k e^{i\vec{k}\cdot\vec{r}} d^3k \qquad \psi_k = \begin{bmatrix} \varphi_k \\ \chi_k \end{bmatrix}$$

le equazioni per i bispinori, che erano:

$$\begin{cases} i\dfrac{\hbar}{c}\dfrac{\partial}{\partial t}\varphi_k & = \vec{\sigma}\cdot\vec{k}\chi_k + mc\varphi_k \\ i\dfrac{\hbar}{c}\dfrac{\partial}{\partial t}\chi_k & = \vec{\sigma}\cdot\vec{k}\varphi_k + mc\chi_k \end{cases}$$

diventano, nel caso di soluzioni nella forma di onde piane $\psi_k = \psi_0(k)e^{-i\omega t}$:

$$\begin{cases} \left(\dfrac{\hbar}{c}\omega - mc\right)\varphi_k - \hbar\vec{\sigma}\cdot\vec{k}\chi_k & = 0 \\ -\hbar\vec{\sigma}\cdot\vec{k}\varphi_k + \left(\dfrac{\hbar}{c}\omega + mc\right)\chi_k & = 0 \end{cases}$$

Ora, questo è un sistema di equazioni nelle due componenti del bispinore e per ammettere soluzioni diverse dalla soluzione banale il suo determinante deve essere nullo. Pertanto, la condizione per avere delle soluzioni è la seguente:

$$\left(\frac{\hbar^2}{c^2}\omega^2 - m^2c^2\right) - \hbar^2(\vec{\sigma}\cdot\vec{k})(\vec{\sigma}\cdot\vec{k}) = 0$$

Ricordando ora che fra le proprietà delle matrici σ c'è la $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}$, si ricava:

$$(\sigma_i k_i)(\sigma_j k_j) = \frac{1}{2}(\sigma_i\sigma_j + \sigma_j\sigma_i)k_i k_j = k^2$$

e quindi per la condizione sul determinante:

$$\frac{\hbar^2}{c^2}\omega^2 = m^2c^2 + k^2$$

che, ricordando che $E = \hbar\omega$ e $p = \hbar k$, si riduce alla ben nota relazione relativistica fra energia ed impulso (3.4). In particolare deve valere:

$$\omega = \pm\frac{1}{\hbar}\sqrt{m^2c^4 + \hbar^2k^2c^2} \equiv \pm\epsilon$$

l'energia può quindi risultare negativa.

La ψ risulta somma di funzioni a frequenza positiva e a frequenza negativa: $\psi = \psi^{(+)} + \psi^{(-)}$. La più bassa frequenza positiva ammessa è mc^2, mentre la più alta negativa è $-mc^2$. Le trasformazioni di Lorentz, inoltre, non permettono di passare da soluzioni a frequenza positiva a soluzioni a frequenza negativa.

Se si lavora per comodità in *unità naturali*, dove $\hbar = c = 1$, le due soluzioni assumono la forma:

$$\psi^{(+)}(\vec{r},t) = \int \psi_0^{(+)}(\vec{k})e^{i(\vec{k}\cdot\vec{r}-\epsilon t)}d\vec{k}$$
$$\psi^{(-)}(\vec{r},t) = \int \psi_0^{(-)}(\vec{k})e^{i(\vec{k}\cdot\vec{r}+\epsilon t)}d\vec{k}$$

La $\psi^{(+)}(\vec{r},t)$ rappresenta soluzioni che si propagano "in avanti" nel tempo, mentre la $\psi^{(-)}(\vec{r},t)$ rappresenta onde che si propagano "all'indietro". Questo è reso evidente mandando $\vec{k} \to -\vec{k}$, in quanto la soluzione negativa diventa:

$$\psi^{(-)}(\vec{r},t) = \int \psi_0^{(-)}(\vec{k})e^{-i(\vec{k}\cdot\vec{r}-\epsilon t)}d\vec{k}$$

Nell'equazione operatoriale $\hat{E}\psi = \hbar\omega\psi$, gli autovalori dell'energia sono dati proprio da $\omega = \pm\sqrt{m^2 + k^2}$, e dunque $\psi^{(+)}(\vec{r}, t)$ e $\psi^{(-)}(\vec{r}, t)$ sono due autofunzioni relativi ad autovalori distinti e pertanto ortogonali fra loro.

In unità naturali, il sistema di equazioni ai bispinori si scrive:

$$\begin{cases} (\omega - m)\varphi_k - \vec{\sigma} \cdot \vec{k}\chi_k &= 0 \\ \vec{\sigma} \cdot \vec{k}\varphi_k + (\omega - m)\chi_k &= 0 \end{cases}$$

che per le autofunzioni fornisce:

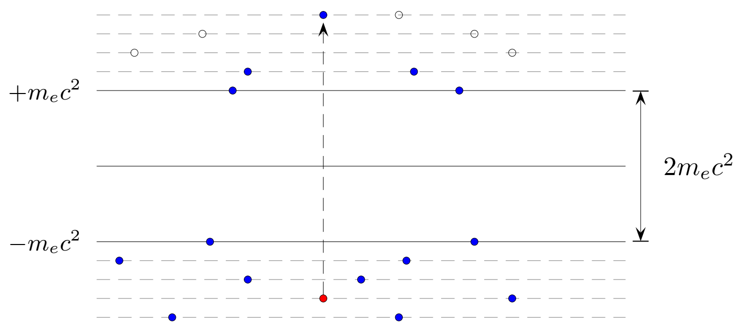
$$\chi_k^{(+)} = \frac{\vec{\sigma} \cdot \vec{k}}{(\epsilon - m)}\varphi_k^{(+)} \qquad\qquad \varphi_k^{(-)} = -\frac{\vec{\sigma} \cdot \vec{k}}{(\epsilon + m)}\chi_k^{(-)}$$

Queste relazioni evidenziano effettivamente che $\psi^{(+)}(\vec{r}, t)$ e $\psi^{(-)}(\vec{r}, t)$ sono ortogonali e quindi soddisfano la relazione:

$$\int \psi^{(+)}(\vec{r}, t)\psi^{(-)}(\vec{r}, t)d\vec{r} = 0$$

L'esistenza di energie negative con un limite superiore di $-mc^2$ ma senza limite inferiore rende impossibile la stazionarietà di qualunque sistema, in quanto in seguito a scambi di energia una particella potrebbe portarsi su livelli di energia via via sempre più bassi.

Nella teoria di Dirac, la stabilità della materia viene spiegata ipotizzando che tutti gli stati a energia negativa compatibili con il principio di Pauli siano occupati (*mare di elettroni*).



Questa teoria ha diverse conseguenze. Quando si cede al sistema un'energia pari ad almeno $2m_ec^2$, può accadere che un elettrone in un livello di energia negativa acquisti sufficiente energia e passi in uno stato a energia positiva, creando una buca. In questo caso si osserverà un elettrone (carica $-|e|$) di energia positiva $+E$ più una buca nel mare di elettroni a energia negativa. L'osservatore vedrà quindi anche l'assenza di un elettrone di carica $-|e|$ e energia $-E$, che sarà interpretata come la presenza di una particella di carica $+|e|$ e energia $+E$, ovvero un *positrone*.

In maniera analoga una buca si comporta come una trappola per qualunque elettrone a energia positiva e conduce quindi alla "annichilazione" dell'elettrone con la buca, con emissione di radiazione di energia pari ad almeno $2m_ec^2$.

La teoria di Dirac prevede quindi la presenza di *coppie di particelle con carica opposta* (*antiparticelle*) e la funzione d'onda deve tenere conto anche della possibilità di creare e distruggere particelle.

Le soluzioni relative all'autovalore negativo $-\epsilon$ sono quindi in realtà associate ad una frequenza negativa e hanno un'energia positiva, si tratta quindi delle antiparticelle con una carica positiva:[4] in

[4] L'equazione di Dirac descrive evidentemente gli elettroni.

altri termini, nella teoria di Dirac l'equazione ai bispinori rappresenta coppie identiche di particelle a carica opposta.

Dalla teoria delle buche di Dirac emerge quindi una fondamentale simmetria della natura: ad ogni particella deve corrispondere una antiparticella, l'esistenza dell'una implica quella dell'altra. Per esprimere formalmente questa simmetria occorre cercare un operatore che permetta di passare da una soluzione all'altra (l'operatore *coniugazione di carica* $\hat{C}$).

Consideriamo l'equazione di Dirac con il termine di interazione elettromagnetica:

$$\left[\gamma_\mu \left(p_\mu - \frac{e}{c}A_\mu\right) + mc\right]\psi = 0 \tag{8.3}$$

L'analoga equazione per il positrone avrà la carica e con il segno invertito:

$$\left[\gamma_\mu \left(p_\mu + \frac{e}{c}A_\mu\right) + mc\right]\psi^C = 0 \tag{8.4}$$

Nella teoria di Dirac il segno della carica e non ha in realtà alcun ruolo. La simmetria richiede che l'assenza di una soluzione a energia negativa dell'equazione dei positroni corrisponda a un elettrone di energia positiva. Esiste quindi una corrispondenza uno-a-uno fra le soluzioni dell'equazione di Dirac (8.3) e quella per i positroni (8.4). Per stabilire questa corrispondenza è necessario invertire il segno relativo dell'operatore impulso p_μ e campo elettromagnetico A_μ. Per fare questo, si prende innanzitutto la complessa coniugata della (8.3):

$$\left[\gamma_\mu \left(p_\mu - \frac{e}{c}A_\mu\right) + mc\right]^*\psi^* = \left[\left(-p_\mu + \frac{e}{c}A_\mu\right)\gamma_\mu^* + mc\right]\psi^* = 0$$

avendo ricordato che $p_\mu = -i\hbar\partial_\mu$. Se è possibile trovare una matrice non singolare C tale che:

$$(C\gamma_4)\gamma_\mu^*(C\gamma_4)^{-1} = -\gamma_\mu$$

allora l'equazione di Dirac assume la forma cercata per i positroni:

$$\left[\gamma_\mu \left(p_\mu + \frac{e}{c}A_\mu\right) + mc\right](C\gamma_4\psi^*) = 0$$

e in questo caso la funzione d'onda dei positroni ha la forma:

$$\psi^C = C\gamma_4\psi^* \equiv C\bar{\psi}^T$$

La matrice dell'operatore coniugazione di carica $\hat{C}$ può essere costruita esplicitamente. Siccome vale la proprietà $\gamma_4\gamma_\mu^*\gamma_4 = \gamma_\mu^T$, deve risultare:

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu \qquad \Rightarrow \qquad C^{-1}\gamma_\mu C = -\gamma_\mu^T$$

In questa rappresentazione $\hat{C}$ deve quindi commutare con γ_1 e γ_3 e anticommutare con γ_2 e γ_3. Una scelta ragionevole è:

$$\hat{C} = i\gamma_2\gamma_4 = -\hat{C}^{-1} = -\hat{C}^\dagger = -\hat{C}^T$$

È sufficiente riuscire a costruire l'operatore $\hat{C}$ in una qualunque rappresentazione, in quanto la trasformazione unitaria che passa da una rappresentazione all'altra una volta applicata alla matrice C fornisce la forma dell'operatore di coniugazione di carica in una qualunque altra rappresentazione. L'operartore $\hat{C}$ permette quidni di costuire esplicitamente la funzione d'onda del positrone.

Vediamo come agisce più in dettaglio la coniugazione di carica $\psi^C = C\hat{\psi}^T = i\gamma_2\psi^*$ su un elettrone libero di energia negativa e spin giù:

$$\psi_4(t) = e^{\frac{imc^2}{\hbar}t}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

L'operazione coniugazione fornisce:

$$i\gamma_2\psi_4^* = i\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{-\frac{imc^2}{\hbar}t} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{\frac{-imc^2}{\hbar}t} = \psi_1(t)$$

e quindi l'assenza di un elettrone di energia negativa con spin giù corrisponde alla presenza di un positrone a energia negativa e spin su. Nel caso di particella libera, quindi non interagente con un campo elettromagnetico, non c'è differenza fra elettroni e positroni e la coniugazione di carica trasforma quindi una soluzione in un'altra soluzione.

Capitolo 9

Applicazioni al Campo Elettromagnetico

9.1 Riscrittura delle equazioni di Maxwell nello spazio k

Consideriamo le equazioni di Maxwell e lavoriamo in unità naturali:

$$\begin{cases} \vec{\nabla} \cdot \vec{H} &= 0 \\ \vec{\nabla} \times \vec{H} &= \dfrac{\partial \vec{E}}{\partial t} \end{cases} \qquad \begin{cases} \vec{\nabla} \cdot \vec{E} &= 0 \\ \vec{\nabla} \times \vec{E} &= -\dfrac{\partial \vec{H}}{\partial t} \end{cases}$$

Poiché non ha senso scriverle come un'equazione d'onda nello spazio delle x,[1] passiamo nello spazio delle k tramite la trasformata di Fourier:

$$\vec{E}(\vec{r}) = \int \vec{E}_k e^{i\vec{k}\cdot\vec{r}} d\vec{k}$$
$$\vec{H}(\vec{r}) = \int \vec{H}_k e^{i\vec{k}\cdot\vec{r}} d\vec{k}$$

che sostituite nelle equazioni di Maxwell forniscono:

$$\begin{cases} \vec{k} \cdot \vec{H}_k &= 0 \\ i\vec{k} \times \vec{H}_k &= \dot{\vec{E}}_k \end{cases} \qquad \begin{cases} \vec{k} \cdot \vec{E}_k &= 0 \\ -i\vec{k} \times \vec{E}_k &= \dot{\vec{H}}_k \end{cases}$$

Bisogna ora naturalmente garantire che $\vec{E} = \vec{E}^*$ e $\vec{H} = \vec{H}^*$.[2] Siccome vale:

$$\vec{E}^*(\vec{r}) = \int \vec{E}_k^* e^{-i\vec{k}\cdot\vec{r}} d\vec{k} = \int \vec{E}_{-k}^* e^{i\vec{k}\cdot\vec{r}} d\vec{k}$$

deve quindi risultare:

$$\begin{cases} \vec{E}_k^* = \vec{E}_{-k} \\ \vec{H}_k^* = \vec{H}_{-k} \end{cases}$$

che rappresentano sei condizioni.

Siccome:[3]

$$i\vec{k} \times \vec{H}_k = \dot{\vec{E}}_k \quad \to \quad i\vec{k} \times [\vec{k} \times \vec{H}_k] = \vec{k} \times \dot{\vec{E}}_k \quad \to \quad i(\vec{k} \cdot \vec{H}_k)\vec{k} - ik^2\vec{H}_k = \vec{k} \times \dot{\vec{E}}_k$$

[1] Il fotone è a massa nulla e viaggia sempre a velocità $v \equiv c$, di conseguenza lo spazio delle x, le posizioni, non ha senso per esso.

[2] Richiesta necessaria per garantire che i campi siano reali.

[3] Si sfrutta qui la relazione del doppio prodotto vettore $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$.

si deduce senz'altro che:[4]

$$\vec{H}_k = \frac{i}{k^2}\vec{k} \times \dot{\vec{E}}_k$$

Per svincolarsi ora dalla condizione di realtà si fa la posizione:

$$\begin{cases} \vec{E}_k = N(k)\left(\vec{f}_k + \vec{f}^*_{-k}\right) \\ \dot{\vec{E}}_k = -ikN(k)\left(\vec{f}_k + \vec{f}^*_{-k}\right) \end{cases}$$

dove $N(k)$ è una costante di normalizzazione che sarà calcolata fra breve. Questa relazione implica a sua volta:

$$\begin{cases} \vec{E}_k = \dot{\vec{E}}_{-k} \\ \vec{H}_k = \dot{\vec{H}}_{-k} \end{cases}$$

Si consideri ora l'equazione di Maxwell:

$$i\vec{k} \times \vec{E}_k = \dot{\vec{H}}_k$$

sotituendo in questa l'espressione per $\vec{H}_k$ trovata sopra si trova:

$$i\vec{k} \times \vec{E}_k = \frac{i}{k^2}\vec{k} \times \frac{\partial^2}{\partial t^2}\vec{E}_k \to i\vec{k}\left(\frac{1}{k^2}\frac{\partial^2}{\partial t^2}\vec{E}_k + \vec{E}_k\right) = 0 \to \left(\frac{\partial^2}{\partial t^2} + k^2\right)\vec{E}_k$$

ovvero formalmente:

$$\left(\frac{\partial}{\partial t} + ik\right)\left(\frac{\partial}{\partial t} - ik\right)\vec{E}_k = 0$$

Essendo ora:

$$\left(\frac{\partial}{\partial t} - ik\right)\vec{E}_k = \dot{\vec{H}}_k - ik\vec{H}_k = ikN(k)\left(\vec{f}_k - \vec{f}^*_{-k}\right) - ikN(k)\left(\vec{f}_k + \vec{f}^*_{-k}\right) = -2ikN(k)\vec{f}_k$$

si può sostituire questo risultato nella precedente ottenendo:

$$-\left(\frac{\partial}{\partial t} + ik\right)2ikN(k)\vec{f}_k = 0$$

ovvero:

$$\boxed{i\frac{\partial}{\partial t}\vec{f}_k = k\vec{f}_k}$$

Dalle rimanenti equazioni si ricava invece:

$$\vec{k} \cdot \vec{E}_k = 0 \to \vec{k} \cdot \vec{f}_k + \vec{k} \cdot \vec{f}^*_{-k} = 0$$

$$\vec{k} \cdot \left(\frac{\vec{k} \times \dot{\vec{E}}_k}{k^2}\right) = 0 \to \vec{k} \cdot \dot{\vec{E}} = 0 \to \vec{k} \cdot \vec{f}_k + \vec{k} \cdot \vec{f}^*_{-k} = 0$$

che implicano:

$$\begin{cases} \vec{k} \cdot \vec{f}_k = 0 \\ \vec{k} \cdot \vec{f}^*_{-k} = 0 \end{cases}$$

ovvero:

$$\boxed{\vec{k} \cdot \vec{f}_k = 0}$$

Combinando ora le due relazioni di sopra si ricava la seguente:

$$i\frac{\partial}{\partial t}\vec{f}_k = k\vec{f}_k - \frac{\vec{k}}{k}\left(\vec{k} \cdot \vec{f}_k\right) \to i\frac{\partial}{\partial t}\vec{f}_k = k\left[\vec{f}_k - \frac{\vec{k}}{k^2}\left(\vec{k} \cdot \vec{f}_k\right)\right]$$

[4] I vettori $\vec{k}$, $\vec{H}$ e $\vec{E}$ sono perpendicolari fra loro.

che scritta per componenti fornisce:

$$\boxed{i\frac{\partial}{\partial t}\left(\vec{f}_k\right)_i = \hat{W}_{ij}\left(\vec{f}_k\right)_j \qquad \hat{W}_{ij} \equiv k\left(\delta_{ij} - \frac{k_i \cdot k_j}{k^2}\right)}$$

Consideriamo ora il prodotto scalare di questa con le k_i:

$$i\frac{\partial}{\partial t}\vec{k}\cdot\vec{f}_k = k_i\hat{W}_{ij}\left(\vec{f}_k\right)_j$$

ma:

$$k_i\hat{W}_{ij} = k(\vec{k}\cdot\vec{f}_k - \vec{k}\cdot\vec{f}_k)$$

e dunque $\vec{k}\cdot\vec{f}_k$ è soluzione di $i\frac{\partial}{\partial t}\vec{k}\cdot\vec{f}_k = 0$. In altre parole, imposto il valore $\vec{k}\cdot\vec{f}_k$ all'inizio, questo si conserva nel tempo.

Le equazioni di Maxwell si riducono quindi in definitiva a:

$$\begin{cases} i\dfrac{\partial}{\partial t}\vec{k}\cdot\vec{f}_k = k_i\hat{W}_{ij}\left(\vec{f}_k\right)_j \\ \vec{k}\cdot\vec{f}_k = 0\Big|_0 \qquad \text{Condizione Iniziale} \\ \hat{W}_{ij} = k\left(\delta_{ij} - \dfrac{k_i\cdot k_j}{k^2}\right) \end{cases}$$

che di fatto è *un'equazione di tipo Schrödinger con $\hat{W}_{ij}$ come operatore.*

9.2 L'operatore $\hat{W}$ come energia. L'impulso. Il momento angolare

Dimostreremo ora che l'operatore $\hat{W}$ può essere effettivamente interpretato come operatore energia, in analogia all'equazione di Schrödinger.

Dalla sua definizione, il valore medio è dato da:

$$\langle\hat{W}\rangle = \int_{-\infty}^{+\infty} \vec{f}_k^*\hat{W}\vec{f}_k d\vec{k} = \int_{-\infty}^{+\infty} k\vec{f}_k^*\vec{f}_k d\vec{k}$$

Si noti incidentalmente che per particelle a massa nulla, siccome vale $E^2 - k^2 = 0$, la relazione risulta automaticamente dimostrata.

L'energia del campo elettromagnetico è data da:

$$\mathcal{E} = \frac{1}{2}\int_{-\infty}^{+\infty}\left(\vec{E}^2 + \vec{H}^2\right)d\vec{r} = \frac{1}{2}\int_{-\infty}^{+\infty}\left(\vec{E}_k\cdot\vec{E}_{k'} + \vec{H}_k\cdot\vec{H}_{k'}\right)e^{i(\vec{k}+\vec{k}')\cdot\vec{r}}d\vec{k}\,d\vec{k}'\,d\vec{r}$$

e siccome vale la relazione:

$$\int_{-\infty}^{+\infty} e^{i\vec{k}\cdot\vec{r}} = (2\pi)^2\delta(\vec{k})$$

si ricava per l'energia:

$$\mathcal{E} = \frac{(2\pi)^3}{2}\int_{-\infty}^{+\infty}\left(\vec{E}_k\cdot\vec{E}_{-k} + \vec{H}_k\cdot\vec{H}_{-k}\right)d\vec{k}$$

ma $\vec{H}_k = \frac{i}{k^2}\vec{k}\times\dot{\vec{E}}_k$ per cui risulta:[5]

$$\vec{H}_k\cdot\vec{H}_{-k} = \frac{1}{k^2}\dot{\vec{E}}_k\dot{\vec{E}}_{-k}$$

[5] Si è utilizzata la proprietà $(\vec{a}\times\vec{b})\cdot(\vec{c}\times\vec{d}) = (\vec{a}\cdot\vec{c})\cdot(\vec{b}\cdot\vec{d}) - (\vec{a}\cdot\vec{d})\cdot(\vec{b}\cdot\vec{c})$.

e si ottiene:

$$\mathcal{E} = 4\pi^3 \int_{-\infty}^{+\infty} \left(\vec{E}_k \cdot \vec{E}_{-k} + \frac{1}{k^2} \dot{\vec{E}}_k \dot{\vec{E}}_{-k} \right) d\vec{k}$$

In termini di $\vec{f}$ risulta:

$$\mathcal{E} = 4\pi^3 \int_{-\infty}^{+\infty} N^2(k) \left[\left(\vec{f}_k + \vec{f}^*_{-k} \right) \left(\vec{f}_{-k} + \vec{f}^*_k \right) - \left(\vec{f}_k - \vec{f}^*_{-k} \right) \left(\vec{f}_{-k} - \vec{f}^*_k \right) \right] d\vec{k}$$

ovvero:

$$\begin{aligned} \mathcal{E} = 4\pi^3 \int_{-\infty}^{+\infty} N^2(k) \Big[& \cancel{\vec{f}_k \cdot \vec{f}_{-k}} + \vec{f}_k \cdot \vec{f}^*_k + \vec{f}_{-k} \cdot \vec{f}^*_{-k} + \cancel{\vec{f}^*_{-k} \cdot \vec{f}^*_k} - \\ & \cancel{- \vec{f}_k \cdot \vec{f}_{-k}} + \vec{f}_k \cdot \vec{f}^*_k + \vec{f}^*_{-k} \vec{f}_{-k} - \cancel{\vec{f}^*_{-k} \cdot \vec{f}^*_k} \Big] d\vec{k} = \\ = 16\pi^3 \int_{-\infty}^{+\infty} & N^2(k) \vec{f}^*_k \cdot \vec{f}_k d\vec{k} \end{aligned}$$

e questa è esattamente la definizione di valor medio di $\hat{W}$ a patto di porre:

$$N(k) = \sqrt{\frac{k}{4\pi^3}}$$

e cioé l'operatore $\hat{W}$ rappresenta effettivamente l'energia se:

$$\begin{cases} \vec{E}_k & = \sqrt{\dfrac{k}{4\pi^3}} \left(\vec{f}_k + \vec{f}^*_{-k} \right) \\ \dot{\vec{E}}_k & = -\dfrac{i}{4\pi} \sqrt{\dfrac{k^3}{\pi}} \left(\vec{f}_k + \vec{f}^*_{-k} \right) \end{cases}$$

Sia ora una soluzione con $k = \omega$, cioé del tipo $f_k = f^0_k e^{-i\omega t}$. L'energia è quindi fissata da ω:

$$\langle \omega \rangle = \omega \int_{-\infty}^{+\infty} \vec{f}^*_k \cdot \vec{f}_k d\vec{k}$$

il che implica che le $\vec{f}_k$ siano normalizzate.

Riscriviamo ora l'impulso associato al campo. Il vettore di Poynting ha la forma:

$$\vec{p} = \int_{-\infty}^{+\infty} \left(\vec{E} \times \vec{H} \right) d\vec{r}$$

ovvero:

$$\vec{p} = \int_{-\infty}^{+\infty} \left[\vec{E}_k \times \vec{H}_{k'} \right] e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} d\vec{k}\, d\vec{k}'\, d\vec{r} = 2\pi^3 \int_{-\infty}^{+\infty} \left(\vec{E}_k \times \vec{H}_k \right) d\vec{k}$$

ma:

$$\vec{E}_k \times \vec{H}_k = \vec{E}_k \times \left(-\frac{i}{k^2} \vec{k} \times \dot{\vec{E}}_{-k} \right)$$

e l'ultimo termine per la regola sul prodotto triplo fornisce:

$$-\frac{i}{k^2} \left(\vec{E}_k \cdot \dot{\vec{E}}_{-k} \right) \vec{k}$$

da cui:

$$\vec{p} = -(i2\pi^3) \int_{-\infty}^{+\infty} \left(\vec{E}_k \cdot \dot{\vec{E}}_{-k} \right) \frac{\vec{k}}{k^2} d\vec{k}$$

ed esprimendo in questa equazione i campi in termini delle $\vec{f}_k$:

$$\vec{p} = -(i2\pi^3)\int_{-\infty}^{+\infty} N^2(k)\left(\vec{f}_k + \vec{f}^*_{-k}\right) ik \left(\vec{f}_{-k} - \vec{f}^*_k\right) d\vec{k} =$$
$$= \frac{(2\pi)^3}{16\pi^3}\int_{-\infty}^{+\infty} k^2 \frac{\vec{k}}{k^2}\left[\vec{f}_k \cdot \vec{f}_{-k} - \vec{f}_k \cdot \vec{f}^*_k + \vec{f}^*_{-k}\vec{f}_{-k} - \vec{f}^*_{-k}\vec{f}^*_k\right] d\vec{k}$$

ma $(\vec{f}_k \cdot \vec{f}_{-k})^* = (\vec{f}_k \cdot \vec{f}_{-k})$ in quanto sono reali. Si ottiene pertanto:

$$\vec{p} = \int_{-\infty}^{+\infty} \vec{k}\left(\vec{f}_k \cdot \vec{f}^*_k\right) d\vec{k}$$

e quindi $\vec{k}$ è proprio l'impulso associato al fotone. Nello spazio delle x:

$$\vec{F}(\vec{r}) = \int_{-\infty}^{+\infty} \vec{f}_k e^{i\vec{k}\cdot\vec{r}} d\vec{k}$$

ma in realtà la $\vec{F}(\vec{r})$ non è interpretabile dal punto di vista probabilistico. Infatti, per poter localizzare un fotone, questo deve necessariamente interagire con i campi $\vec{E}$ e $\vec{H}$:

$$\vec{E}_k = N(\vec{k})\left(\vec{f}_k + \vec{f}^*_{-k}\right) \qquad\qquad \vec{E}(\vec{r}) = \int_{-\infty}^{+\infty} \vec{E}_k e^{i\vec{k}\cdot\vec{r}} d\vec{k}$$

La presenza di una radice di k([6]) ci conduce infatti ad una $\sqrt[4]{\nabla^2}$ e questo significa che determinare la $\vec{F}(\vec{r})$ non è sufficiente per determinare i campi $\vec{E}(\vec{r})$ e $\vec{H}(\vec{r})$. In altri termini: il fotone non è localizzabile esattamente e addirittura per lunghezze d'onda inferiori a quelle del fotone non ha nemmeno senso il concetto di localizzazione.

Consideriamo ora il momento angolare associato al campo elettromagnetico.

$$\vec{M} = \int_{-\infty}^{+\infty} (\vec{r} \times \vec{p}) d\vec{r} = \int_{-\infty}^{+\infty} \vec{r} \times (\vec{E} \times \vec{H}) d\vec{r} = \int_{-\infty}^{+\infty} \vec{r} \times (\vec{E}_k \times \vec{H}_{k'}) e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} d\vec{k}\, d\vec{k}'\, d\vec{r}$$

in questo caso non si può estrarre la delta dall'integrale. Però, notando che:

$$\int_{-\infty}^{+\infty} \vec{r} e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} d\vec{r} = -i\vec{\nabla}_{k'}\int_{-\infty}^{+\infty} e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} d\vec{r} = -i(2\pi)^3\vec{\nabla}_{k'}\delta^3(\vec{k}+\vec{k}')$$

e sostituendo:

$$\vec{M} = -i(2\pi)^3\int_{-\infty}^{+\infty}\left[\vec{\nabla}_{k'}\delta^3(\vec{k}+\vec{k}')\right] \times (\vec{E}_k \times \vec{H}_{k'}) d\vec{k}\, d\vec{k}'$$

si può quindi integrare per parti e (trascurando i termini ad infinito che si annullano) ottenere:

$$\vec{M} = -i(2\pi)^3\int_{-\infty}^{+\infty} \delta^3(\vec{k}+\vec{k}')\vec{\nabla}_{k'} \times (\vec{E}_k \times \vec{H}_{k'}) d\vec{k}\, d\vec{k}'$$

che riscritta in termini delle $\vec{f}_k$:

$$\vec{M} = \int_{-\infty}^{+\infty}\left\{-i\left[\vec{k} \times \vec{\nabla}(\vec{f}^{*c}_k \cdot \vec{f}_k)\right] - i(\vec{f}^*_k \times \vec{f}_k)\right\} d\vec{k}$$

dove con la notazione $\vec{f}^{*c}_k$ si intende "considerato costante sotto l'operatore $\vec{\nabla}$".

Questa equazione rappresenta di fatto il valore medio dell'operatore momento angolare:

$$M_j = \int_{-\infty}^{+\infty}\left[-i\vec{f}^{*i}_k\left(\vec{k} \times \vec{\nabla}\right)_j \vec{f}^i_k - i\epsilon_{ilj}\vec{f}^{*i}_k \vec{f}^l_k\right] d\vec{k}$$

[6] Presente nel fattore $N(\vec{k})$

In questa equazione può essere facilmente riconosciuto un termine che rappresenta il noto momento angolare:

$$\boxed{\hat{M} = \vec{k} \times \vec{\nabla} \qquad \textit{Operatore Momento Angolare}}$$

più un altro termine di momento angolare $i\epsilon_{ilj}\vec{f}_k^{*i}\vec{f}_k^l$ che deve pertanto essere riconosciuto come il momento angolare associato allo *spin*, il cui operatore è quindi definito da:

$$\boxed{(s_i\vec{f}_k)_j = -(s_j\vec{f}_k)_i = -i\epsilon_{ilj}\vec{f}_k^l \qquad \textit{Operatore Spin}}$$

Formalmente quindi:

$$\boxed{\vec{M} = \int_{-\infty}^{+\infty} \vec{f}_k^* \left[-i\vec{k} \times \vec{\nabla} + \vec{s}\right] \vec{f}_k d\vec{k}}$$

Parte III

Teoria della Relatività Generale

Capitolo 10

Premesse Concettuali e Fenomenologiche

10.1 Il principio di equivalenza

Nell'ambito della meccanica classica, esiste – ed era già chiaro dai tempi di Newton e Galileo – una differenza concettuale fra la massa gravitazionale m_g, che misura la forza di interazione gravitazionale, e la cosiddetta massa inerziale m_i, che misura invece l'inerzia di un corpo ed interviene quindi nelle leggi del moto implicite nel II° principio della dinamica:

$$\vec{F}_{\text{grav.}} = -G\frac{M_g m_g}{r^2}\hat{r}$$
$$\vec{F} = m_i\vec{a}$$

queste relazioni implicano che l'accelerazione gravitazionale è indipendente dalla massa **solo** se risulta $m_g \propto m_i$. La determinazione del loro rapporto e l'eventuale coincidenza delle due è pertanto una questione di primaria importanza nell'ambito della fisica. Si iniziò a determinare questo numero con una certa precisione a partire dall'esperimento con la bilancia a torsione di Eötvös effettuato nel 1922. In quell'anno, Eötvös dimostrò l'equivalenza delle due masse con una precisione di $5 \cdot 10^{-9}$.([1])

L'esperienza di Eötvös

L'esperienza di Eötvös fu un famoso esperimento della fisica della fine del XIX secolo che misurò la correlazione tra massa inerziale e massa gravitazionale, dimostrandone l'equivalenza con una precisione fino ad allora impossibile da raggiungere. L'esperimento principale fu compiuto da Loránd Eötvös nel 1885, con vari miglioramenti nell'arco di tempo compreso tra il 1906 e il 1909. La squadra di Eötvös proseguì con una serie di esperimenti simili ma più precisi, con vari materiali e in vari luoghi della Terra, che dimostrarono tutti la stessa equivalenza. Questi esperimenti diedero un forte contributo alla moderna nozione del principio di equivalenza, codificato nella teoria della relatività generale.

L'apparato sperimentale originario di Eötvös consisteva in due masse agli estremi di un'asta rigida, appesa al suo centro tramite un sottile filo. Uno specchio applicato all'asta, o al filo, rifletteva la luce in un piccolo telescopio. Ogni piccola rotazione dell'asta avrebbe causato la deflessione del raggio di luce, rilevabile attraverso il telescopio.

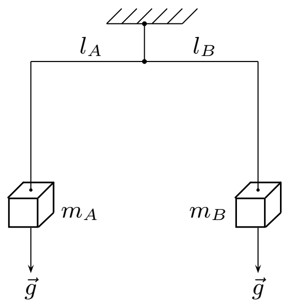
Le due principali forze agenti sul sistema sono la forza di gravità e la forza centrifuga dovuta alla rotazione terrestre. La prima è calcolata tramite la legge di gravitazione universale di Newton e dipende dalla massa gravitazionale; la seconda è ricavata da considerazioni cinematiche riguardanti i sistemi di riferimento non inerziali e dipende dalla massa inerziale. L'esperimento era predisposto in modo tale che se le due masse (inerziale e gravitazionale) dei corpi alle estremità della sbarra fossero state diverse, le due forze non si sarebbero cancellate esattamente e nel tempo l'asta sarebbe ruotata.

[1] Nel luglio 1999 si è raggiunta una precisione di $5 \cdot 10^{-13}$ sempre con metodo di bilancia a torsione e si conta di raggiungere una precisione di 10^{-17} con un metodo satellitare (Progetto STEP sul satellite Galileo Galilei).

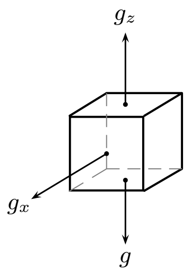
Gli esperimenti iniziali, attorno al 1885, confermarono la sostanziale equivalenza dei due tipi di massa, e Eötvös stesso migliorò l'esperimento per dimostrarlo con una migliore precisione. Nel 1889 usò lo stesso dispositivo con masse di diverso materiale per vedere se l'equivalenza dipendesse o no dal materiale utilizzato: non venne però misurato nessun cambiamento, entro una precisione di 1 su 20 milioni. Eötvös pubblicò i suoi risultati nel 1890.

L'anno successivo cominciò a lavorare ad una versione modificata dell'apparato sperimentale, da lui chiamata "variometro orizzontale". In questo nuova disposizione una delle due masse ad un capo dell'asta è sostenuta da un filo, invece di essere semplicemente applicata alla fine della sbarra. Questo permise di misurare la torsione in due dimensioni e quindi anche la componente locale orizzontale dell'accelerazione di gravità; la precisione totale dello strumento risultò ancora migliorata. Oggi ci si riferisce generalmente a questo apparato come *bilancia di Eötvös* ed esso viene correntemente usato nella ricerca di variazioni locali della densità della crosta terrestre.

La configurazione sperimentale (molto schematizzata) era la seguente:



Principio schematico della Bilancia di Eötvös



Dettaglio delle accelerazioni sulle masse

La componente di forza centrifuga g_x è separabile da quella della forza peso e sulla massa agiscono sia la forza peso (che coinvolge la massa gravitazionale $m^{(g)}$) che le forze apparenti (che coinvolgono la massa inerziale $m^{(i)}$). L'apparato sperimentale è tale che se queste due masse sono diverse appare una componente di torsione che fa ruotare il pendolo.

La condizione di equilibrio è data da:

$$l_A\left(m_A^{(g)}g - m_A^{(i)}g_z\right) = l_B\left(m_B^{(g)}g - m_B^{(i)}g_z\right)$$

mentre il momento torcente è dato da (ricavando l_B dalla relazione precedente):

$$T = l_A m_A^{(i)} g_x - l_B m_B^{(i)} g_x \qquad \Rightarrow \qquad T = l_A m_A^{(i)} g_x \left[1 - \frac{m_A^{(g)}g - m_A^{(i)}g_z}{m_B^{(g)}g - m_B^{(i)}g_z}\frac{m_B^{(i)}}{m_A^{(i)}}\right]$$

e sviluppando per $g_z \ll g$:

$$T = l_A m_A^{(i)} g_x \left[1 - \frac{m_B^{(i)}}{m_A^{(i)}}\frac{m_A^{(g)}}{m_B^{(g)}}\frac{1 - \frac{m_A^{(i)}}{m_A^{(g)}}\frac{g_z}{g}}{1 - \frac{m_A^{(i)}}{m_A^{(g)}}\frac{g_z}{g}}\right]$$

trascurando quindi i termini in g_z/g:

$$T = l_A m_A^{(i)} g_x \left[1 - \frac{m_A^{(g)}/m_A^{(i)}}{m_B^{(g)}/m_B^{(i)}}\right]$$

e cioé in definitiva:

$$T = \frac{m_B^{(i)}}{m_B^{(g)}} l_A m_A^{(i)} g_x \left[\frac{m_B^{(g)}}{m_B^{(i)}} - \frac{m_A^{(g)}}{m_A^{(i)}}\right]$$

ovvero: il momento torcente è proporzionale alla differenza dei rapporti fra massa gravitazionale e inerziale.

Questo discorso sottintende un punto fondamentale: se la massa gravitazionale coincide con la massa inerziale (eventualmente a meno di un fattore costante), è possibile allora eliminare la forza di gravità in un sistema di riferimento opportunamente accelerato, in altri termini un sistema di riferimento accelerato può assumere il ruolo di un sistema di riferimento inerziale soggetto a gravitazione. In termini più precisi:

Principio di Equivalenza Forte: *Un sistema di riferimento non inerziale in cui le forze apparenti cancellino le forze gravitazionali si comporta* localmente *come un sistema di riferimento inerziale in senso newtoniano.*[2]

La validità del principio di equivalenza forte implica la validità del:

Principio di Equivalenza Debole: *La massa gravitazionale di un corpo è numericamente uguale alla sua massa inerziale.*

Gli appellativi di forte e debole si giustificano dal momento che se vale il principio di equivalenza nella forma forte deve valere anche quello nella forma debole, mentre da un punto di vista logico l'implicazione non è invertibile: il principio di equivalenza debole non *implica* quello forte.[3] Anche se il principio in forma debole è stato sperimentalmente verificato entro limiti strettissimi, questa caratteristica fa sì che non sia possibile garantire lo stesso grado di certezza per la forma forte *che deve dunque essere considerata ancora come un postulato.*

Chiariamo immediatamente con un esempio perché nel principio di equivalenza forte abbiamo specificato *localmente.* Supponiamo un corpo esteso rigido in orbita intorno alla terra, la cui lunghezza lungo il raggio $\hat{r}$ sia $2l$. Se il corpo si mantiene in orbita, significa che le due componenti dell'accelerazione centrifuga e dell'attrazione gravitazionale devono essere uguali ed opposte:

$$a = \frac{GM}{r^2} - \frac{v^2}{r} = 0 \qquad\qquad v^2 = \frac{GM}{r}$$

Le due accelerazioni si bilanciano perfettamente però solo alla distanza r, distanza a cui si trova il centro di massa del corpo. Infatti, al suo estremo più lontano le accelerazioni valgono:

$$\begin{aligned} a = \frac{GM}{(r+l)^2} - \frac{v^2}{r+l} = \frac{GM}{(r+l)^2} - \frac{GM}{r(r+l)} &= GM\left[\frac{1}{(r+l)^2} - \frac{1}{r(r+l)}\right] = \\ &= \frac{GM}{r+l}\left[\frac{r-r-l}{r(r+l)}\right] = -\frac{GMl}{r(r+l)^2} \end{aligned}$$

esiste quindi un'accelerazione residua $a_r \propto \pm\frac{GMl}{r^3}$, dove il segno indica se predomina la forza di attrazione o quella centrifuga. All'estremità più lontana predomina la forza centrifuga, mentre su quella più vicina predomina quella di attrazione: l'effetto netto su un corpo esteso è quindi quello di "stirarlo" lungo l'asse di attrazione.[4]

Questo dimostra che un sistema di riferimento non inerziale, che fa comparire un campo di forze apparenti, può annullare effettivamente la forza di gravità solo a livello locale. Inoltre, l'uguaglianza dei due campi è *numerica* (*quantitativa*) e non certamente sostanziale: laddove i campi apparenti sono generati da un cambiamento di sistema di riferimento, il campo gravitazionale è generato da una distribuzione di massa e questo si riflette anche nel diverso comportamento *globale* che hanno i due tipi di campo.[5]

L'importanza del principio di equivalenza forte risiede in due grandi problemi della fisica:

---

[2] Ed è questa l'essenza dell'esperimento concettuale dell'ascensore di Einstein: "Se dentro un ascensore che sta precipitando verso il basso eseguo un qualsiasi esperimento di fisica, non potrò distinguerlo da un esperimento fatto in un riferimento inerziale come lo intendeva Newton".

[3] In termini formali: il principio nella forma forte è condizione sufficiente ma non necessaria per la validità della forma debole, mentre la forma debole è condizione necessaria ma non sufficiente per la validità della forma forte.

[4] La forza residua $\vec{F}_r = ma_r\hat{r}$ è quindi quella all'origine del cosiddetto *effetto marea* e prende pertanto il nome di *forza di marea.*

[5] I campi fittizi, come ad esempio la forza centrifuga, crescono illimitatamente a infinito o al massimo tendono ad un valore costante, mentre i campi fisici "reali" tendono a zero.

- la teoria della relatività ristretta permette di scrivere le equazioni della fisica in forma covariante[6] solo nei sistemi di riferimento inerziali. Una trasformazione fra sistemi in moto non uniforme comporta infatti l'introduzione di termini di forza (le cosiddette *forze apparenti*) che rendono le equazioni non covarianti nei sistemi di riferimento non inerziali. Occorre quindi trovare una formulazione che sia più generale rimuovendo lo status privilegiato di cui godono ancora i sistemi di riferimento inerziali,

- occorre rendere covariante anche la teoria della gravitazione, e siccome – se vale il principio di equivalenza forte – questa coincide localmente con un sistema di riferimento non inerziale, la teoria della relatività ristretta si rivela adeguata per il compito.

Il principio di equivalenza forte è quindi quello che permette di risolvere i due problemi con una sola teoria, denominata appunto *teoria della relatività generale*.

La base concettuale di questa teoria è quindi questa: *i campi di forze apparenti sono generati dal passaggio da un sistema inerziale a uno non inerziale ed è sempre possibile trovare una trasformazione di coordinate per passare dall'uno all'altro e viceversa. Siccome il campo gravitazionale localmente può essere assimilato ad un sistema non inerziale, è quindi sempre possibile trovare una trasformazione di coordinate che porti da un sistema soggetto a campo gravitazionale ad un sistema di riferimento inerziale in cui il campo gravitazionale è trascurabile a livello locale. In questo intorno infinitesimo dell'evento quadridimensionale è quindi valida la teoria della relatività ristretta.*

10.2 Descrizione non formale delle conseguenze del principio di equivalenza

Consideriamo una piattaforma rotante come esempio semplice di sistema non inerziale.

Una traiettoria rettilinea in questo sistema evidentemente non è più diritta.[7]

Il fatto che la traiettoria sia curvilinea in questo sistema di riferimento, essendo una relazione di natura puramente geometrica, deve valere anche per i raggi luminosi. Ma in base al principio di equivalenza, è legittimo dunque affermare che la luce deve curvare in presenza di campi gravitazionali. Ciò appare strano a prima vista, in quanto la luce è essenzialmente un campo elettromagnetico e per di più il fotone ha massa nulla.[8]

Nella teoria della Relatività Generale si assume come postulato il Principio di Equivalenza Forte. *Il problema di descrivere il campo gravitazionale diventa quindi quello di trovare le equazioni che descrivono il cambiamento di coordinate verso il sistema di riferimento in cui il campo gravitazionale si annulla punto per punto.*

Supponiamo ora due orologi, uno al centro e l'altro alla periferia della piattaforma rotante di cui sopra: siccome il secondo orologio è in moto rispetto al primo (e dunque subirà il fenomeno della dilatazione temporale), essi non possono essere sincronizzati anche se siamo nello stesso sistema di riferimento: *gli orologi segnano un tempo diverso a seconda della loro posizione nel campo gravitazionale.*

Supponiamo invece un regolo orientato secondo il raggio della solita piattaforma rotante: esso non subisce la contrazione di Lorentz. Un regolo disposto lungo la circonferenza – e la circonferenza stessa – invece subisce una contrazione nel senso della lunghezza. Misurando quindi il raggio e la circonferenza nello stesso sistema di riferimento si deduce che *il rapporto fra circonferenza e raggio non è* 2π.

[6] Ovvero, formalmente indipendente dal sistema di riferimento.

[7] È facilmente dimostrabile ricordando che leggi di trasformazione per questo sistema, supponendolo in rotazione intorno all'asse z, sono:

$$\begin{cases} x' &= x\cos\omega t + y\sin\omega t \\ y' &= -x\sin\omega t + y\cos\omega t \end{cases}$$

[8] Bisogna dire però che il rapporto $h\nu/c^2$ è ben definito: ed è questo fattore che in pratica fa risentire al fotone della gravità. Notare che curiosamente questo termine si può pensare derivato dall'aver imposto l'uguaglianza di energia a riposo relativistica e quantistica: $h\nu = mc^2$.

Quest'ultima osservazione ha una conseguenza importante: la geometria euclidea della fisica classica (e quella pseudoeuclidea della relatività speciale) non è necessariamente valida se si vuole tenere conto anche della gravitazione. La conclusione cui siamo forzati è pertanto che:

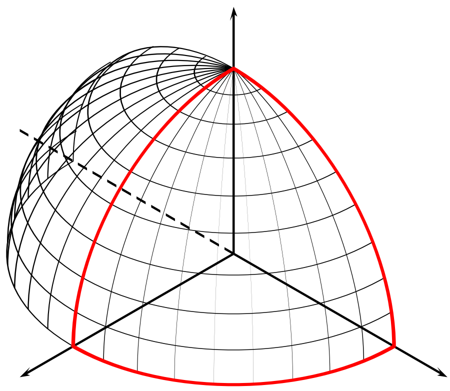
In presenza di gravitazione, lo spazio-tempo non è più una varietà euclidea.

Ovvero, detto in termini più comuni: la gravitazione deve modificare in qualche modo la struttura geometrica dello spazio-tempo.

La conseguenza più strettamente "geometrica" è che fallisce anche la nota relazione fra gli angoli interni di un triangolo – la loro somma non è più 180° – e quindi il 5° postulato di Euclide sul parallelismo delle rette, di cui questa è diretta conseguenza. Il fallimento della geometria euclidea implica che la descrizione dello spaziotempo in termini di una geometria piatta non è corretta. Occorre quindi descriverlo utilizzando la geometria di uno spazio curvo.

Dobbiamo qui porre particolare attenzione ad un aspetto concettuale importante. Nella relatività ristretta (e in fisica classica) vale il principio di inerzia: un corpo si muove di moto inerziale finché non interviene una forza che lo modifica. Ora, abbiamo trovato che le traiettorie dei raggi di luce possono curvare in presenza di campi gravitazionali (ovvero in sistemi non inerziali), perché linee rette si trasformano in questo caso in linee curve. È possibile spiegare questo in due modi: 1) La geometria dello spazio è euclidea (o pseudoeuclidea), ma intervengono delle forze che provocano l'incurvamento delle geodetiche; 2) la struttura dello spazio tempo non è in realtà piatta (euclidea). Questa è la strada intrapresa da Einstein: quella di considerare la geometria dello spazio-tempo come una geometria a curvatura variabile punto per punto (perché il campo gravitazionale varia punto per punto) e che si riduce alla geometria pseudoeuclidea localmente. In questo caso, i raggi di luce e le geodetiche in genere sono curvi non per la presenza di forze, ma a causa della curvatura stessa dello spaziotempo in cui si trovano "immersi". In un certo senso, è sempre la traiettoria di un moto inerziale come in fisica classica e in relatività ristretta, ma la curvatura dello spaziotempo dovuta al campo gravitazionale ce la fa apparire come curva.

Mostriamo ora come il fallimento delle relazioni di Euclide sui triangoli è la conseguenza di uno spazio a curvatura non nulla e introduciamo i necessari strumenti formali. Consideriamo a titolo di esempio una superficie sferica (figura seguente), su tale sfera si consideri un triangolo che ha i lati formati da due mezzi meridiani e dalla porzione di equatore delimitata da essi. È facile rendersi conto che la somma degli angoli interni è maggiore di π: già la somma degli angoli di base è π, essendo per definizione i meridiani ortogonali ai paralleli. In particolare, se si scelgono come lati due meridiani ortogonali fra loro la somma degli angoli interni vale $3/2\pi$.



Formalizzeremo ora questo importante concetto introducendo la definizione di *curvatura*. Supponiamo che la sfera abbia raggio R e che la superficie del triangolo sferico sia A. Allora si verifica la seguente relazione:[9]

$$\alpha + \beta + \gamma - \pi = \frac{A}{R^2}$$

Questa relazione può essere utilizzata come un indicatore della curvatura di uno spazio. A partire da essa definiamo quindi la *curvatura* k nel modo seguente:

$$\underbrace{k^2 \equiv \lim_{A \to 0} \frac{\alpha + \beta + \gamma - \pi}{A}}_{definizione} \qquad \left(\text{per una sfera: } \frac{\alpha + \beta + \gamma - \pi}{A} = \frac{1}{R^2} \right) \tag{10.1}$$

che in questo caso banale coincide proprio con $1/R$, l'inverso del raggio della sfera.[10] In generale la curvatura k è funzione del punto $k = k(x)$, in questo caso vale la relazione:

$$\alpha + \beta + \gamma - \pi = \int_{\triangle} k(x) d\sigma$$

dove con la notazione $\int_{\triangle} \cdots d\sigma$ si intende l'integrale di superficie esteso al triangolo sferico. Per superfici riconducibili senza deformazioni a piani (coni, cilindri, piani), la curvatura è nulla.

La relazione (10.1) definisce la cosiddetta *curvatura intrinseca*: il termine indica che questa curvatura è definibile tramite elementi dello stesso spazio senza alcun riferimento a spazi di dimensione superiore ed è pertanto una caratteristica intrinseca – appunto – dello spazio stesso.[11]

Ricordiamo ora che l'azione per un punto materiale è data da:

$$S = -mc^2 \int \sqrt{1 - \beta^2} dt \qquad \Rightarrow \qquad L = -mc^2 \sqrt{1 - \beta^2}$$

e che per $v \ll c$ si ritrova $L = -mc^2 + 1/2mv^2$. Prendiamo ora in considerazione anche il campo gravitazionale e ricerchiamo l'espressione dell'azione che fornisce nel limite classico la Lagrangiana:

$$L = -mc^2 + \frac{1}{2}mv^2 + \underbrace{\frac{GMm}{r}}_{\equiv -m\Phi}$$

questa Lagrangiana deve avere la forma:

$$L = -mc^2 \sqrt{1 + \frac{2\Phi}{c^2} - \beta^2} \simeq -mc^2 \left[1 + \frac{1}{2} \left(\frac{2\Phi}{c^2} - \beta^2 \right) \right] = -mc^2 + \frac{1}{2}mv^2 - m\Phi$$

$$S = \int L dt = -mc^2 \int \sqrt{1 + \frac{2\Phi}{c^2} - \beta^2}\, dt = -mc \int \sqrt{(c^2 + 2\Phi)dt^2 - dr^2}$$

[9] Per rendersene conto, nel caso particolare dell'esempio si ha che l'area del triangolo scelto è esattamente 1/8 di superficie sferica, quindi:

$$\alpha + \beta + \gamma - \pi = \frac{1}{8} \cdot 4\pi R^2 \cdot \frac{1}{R^2} = \frac{1}{2}\pi \quad \Rightarrow \quad \alpha + \beta + \gamma = \frac{3}{2}\pi$$

Inoltre, se il raggio della sfera tende a infinito, allora la somma degli angoli interni tende proprio a π.

[10] La ragione per definire la curvatura come $k = 1/R$ è per formalizzare il fatto che una circonferenza è sempre "meno curva" man mano che il raggio aumenta.

[11] Si distinguono due tipi di curvatura:

- *curvatura estrinseca*: è la curvatura posseduta dall'oggetto in relazione ad uno spazio piatto di dimensione superiore in cui è immerso e determinabile solo confrontando elementi dell'oggetto in relazione ad elementi dello spazio contenitore;
- *curvatura intrinseca*: è la curvatura determinabile utilizzando solo operazioni eseguite su elementi dell'oggetto medesimo

Un esempio di curvatura estrinseca è quella di una superficie cilindrica nello spazio tridimensionale: le linee tracciate sul cilindro sono curve se confrontate con le rette dello spazio in cui il cilindro è immerso. La geometria intrinseca del cilindro è invece piatta, in quanto su di essa valgono tutti gli assiomi del piano euclideo. Una sfera ha invece una curvatura intrinseca, determinabile rimanendo all'interno della superficie stessa: sulla Terra, un percorso che parte dal polo nord scendendo lungo un meridiano, ruota ad angolo retto lungo un parallelo e nuovamente ad angolo retto lungo un altro meridiano, ritorna al punto di partenza. Un percorso analogo eseguito su un piano non ripassa mai lungo lo stesso punto.

Il campo gravitazionale si presenta quindi come un coefficiente di dt^2, cioé si comporta come una sorta di "indice di rifrazione" per il tempo. Il quadrintervallo si scrive allora:

$$ds^2 = (c^2 + 2\Phi)dt^2 - dr^2 = g_{\mu\nu}dx^\mu dx^\nu \neq dx_\mu dx^\mu$$

e rispetto al caso senza campo gravitazionale è variato $g_{\mu\nu}$, che è in effetti l'oggetto che definisce come calcolare le distanze e che viene chiamato *metrica*.[12]

Vale la pena notare immediatamente che la relazione precedente implica una condizione in cui il dt si annulla:

$$c^2 + 2\Phi = 0 \to c^2 - \frac{2GM}{R_0} = 0 \Rightarrow \frac{2GM}{R_0 c^2} = 1$$

per una massa M esiste quindi una distanza R_0 alla quale $dt = 0$, ovvero oltre la quale alcun segnale può viaggiare (il quadrintervallo diventa infatti in questo caso di tipo spazio). Con una terminologia impropria possiamo dire che è come se la velocità della luce si riducesse a 0: è quello che si chiama *buco nero*.[13]

Riassumendo quanto detto finora, il principio di equivalenza esprime due fatti:

- la massa inerziale è uguale alla massa gravitazionale,
- La massa inerziale dipende quindi dalla distribuzione delle masse dell'universo.

Con tali premesse, anche l'accelerazione di un corpo varia con il sistema di riferimento.

Il moto rettilineo di un corpo in un sistema di riferimento inerziale visto da un qualsiasi altro sistema di riferimento viene trasformato in una curva. Tale curva prende il nome di *moto naturale*.

In un sistema di riferimento qualunque, inoltre, lo scorrere del tempo è diverso da punto a punto.

Dal momento che ogni sistema di riferimento si può considerare come un sistema inerziale immerso in un campo gravitazionale, si deduce che:

***La gravità influisce sul tempo*.**

[12] Vedremo che si tratta in realtà di un tensore e si parla dunque tecnicamente di *tensore metrico*.

[13] Generalmente per un corpo finito la distanza R_0 è calcolata dal centro di massa e cade all'interno del corpo stesso, per cui non si ha alcuna conseguenza fisica in quanto questa è la forma che il campo assume all'esterno del corpo. Diverso è il caso in cosmologia, dove possono esistere dei corpi contratti (stelle collassate) le cui dimensioni cadono al disotto di R_0.

Capitolo 11

Fondamenti Matematici

In questo capitolo riprenderemo alcune delle definizioni già date in §2.1 per approfondirle e generalizzarle in maniera da poter essere utilizzate in relatività generale.

11.1 Introduzione

Consideriamo ancora la nostra piattaforma rotante e valutiamo il quadrintervallo, che nel sistema del laboratorio è dato da $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$.

Le trasformazioni di coordinate nel sistema rotante intorno all'asse z sono date da:

$$\begin{cases} x & = x' \cos \omega t - y' \sin \omega t \\ y & = x' \sin \omega t + y' \cos \omega t \\ z & = z' \\ t & = t' \end{cases}$$

e quindi:

$$\begin{aligned} dx &= - \sin(\omega t)\omega dt \, x' + \cos(\omega t)dx' - \cos(\omega t)\omega dt \, y' - \sin(\omega t)dy' \\ dy &= \cos(\omega t)\omega dt \, x' + \sin(\omega t)dx' - \sin(\omega t)\omega dt \, y' - \cos(\omega t)dy' \end{aligned}$$

e cioé:

$$ds^2 = [c^2 - \omega^2(x' + y')^2]dt^2 - dx'^2 - dy'^2 - dz'^2 + 2\omega y'dx'dt - 2\omega x'dy'dt$$

questa espressione non può mai essere ridotta alla somma dei quadrati dei differenziali. Quindi in un sistema di riferimento non inerziale il ds^2 sarà una forma quadratica generale dei differenziali delle coordinate. Nel particolare caso esaminato, il tensore metrico $g_{\alpha\beta}$ ha le seguenti componenti:

$$g_{\alpha\beta} = \begin{pmatrix} 1 - \omega^2(x'^2 + y'^2) & 2\omega y' & -2\omega x' & 0 \\ 2\omega y' & -1 & 0 & 0 \\ -2\omega x' & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

e, come si vede, il passaggio ad un sistema di riferimento non inerziale ha modificato la forma della metrica.

In definitiva:

$$ds^2 = g_{\alpha\beta}dx'^{\alpha}dx'^{\beta}$$

Se consideriamo quindi il sistema rotante come un sistema inerziale + campo gravitazionale, in accordo al principio di equivalenza si ha che:

$$\boxed{\textbf{\textit{Il campo gravitazionale altera la metrica}}}$$

Le coordinate gaussiane sono delle coordinate atte a stabilire in maniera univoca i punti di una superficie.[1] Il principio di equivalenza sancisce inoltre che le leggi fisiche devono rimanere invariate per passaggi fra sistemi di coordinate. Il punto della teoria è quindi ricercare delle trasformazioni di coordinate tali che:

$$\underbrace{g_{\mu\nu}}_{\textit{Coordinate Gaussiane}} \quad \rightarrow \quad \underbrace{\eta_{\mu\nu}}_{\textit{Coordinate Piatte}}$$

tenendo conto, naturalmente, che questo tipo di trasformazione non è applicabile in presenza di campi gravitazionali che non possano essere ridotti.

11.2 Definizioni di base

Introduciamo ora gli elementi per una formalizzazione più forte. Diamo le seguenti definizioni:

Connessione: In un punto del cronotopo (x_0, x_1, x_2, x_3) con $x_0 = ct$ deve essere possibile una qualsiasi trasformazione $x \mapsto f(x)$ invertibile: cioé tale che $x'^{\alpha} = x'^{\alpha}(x^{\beta})$. Per invertibile, si intende che lo jacobiano della trasformazione è diverso da zero.

Scalare: per trasformazioni il valore deve restare lo stesso, cioé deve valere $\Phi'(x') = \Phi(x)$ e $\Phi(x(x')) = \Phi(x')$. In due punti vicini $\Phi(\bar{x}) - \Phi(x)$ deve essere invariante (e quindi uno scalare). Le differenze infinitesime sono date da $d\Phi = \frac{\partial \Phi}{\partial x^{\alpha}} dx^{\alpha}$.

Vettore Covariante: un insieme di quattro componenti che trasformano come Φ:

$$\frac{\partial \Phi}{\partial x'^{\alpha}} = \frac{\partial \Phi}{\partial x^{\beta}} \frac{\partial x^{\beta}}{\partial x'^{\alpha}}$$

Vettore Controvariante: un insieme di quattro componenti che trasformano come dx^{α}:

$$dx'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\beta}} dx^{\beta}$$

Le componenti covarianti si indicano con gli indici in basso, mentre le componenti controvarianti con gli indici in alto e si intende una somma sottintesa solo fra indici covarianti e controvarianti ripetuti.[2] Quindi, riassumendo, i vettori covarianti e controvarianti trasformano secondo le regole:

$$\begin{aligned} A'_{\mu} &= \frac{\partial x^{\beta}}{\partial x'^{\mu}} A_{\beta} \qquad \textit{Vettore Covariante} \\ B'^{\mu} &= \frac{\partial x'^{\mu}}{\partial x^{\beta}} B^{\beta} \qquad \textit{Vettore Controvariante} \end{aligned}$$

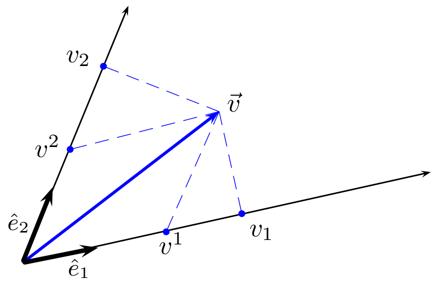
È possibile dare un'utile rappresentazione "visuale" delle componenti di un vettore. Consideriamo un sistema di riferimento non cartesiano, bidimensionale per semplicità, un vettore $\vec{v}$ in uno spazio $\mathbb{R}^2$ e siano $\hat{e}_i$ i versori degli assi. Le componenti covarianti e controvarianti sono definite come:

$$\begin{array}{lcll} \vec{v} & = & v^1 \hat{e}_1 + v^2 \hat{e}_2 & \text{Componenti Controvarianti } v^i \\ v_1 & = & \vec{v} \cdot \hat{e}_1 & \text{Componenti Covarianti } v_i \end{array}$$

Questa rappresentazione permette una semplice interpretazione geometrica: le componenti controvarianti rappresentano le *componenti* del vettore $\vec{v}$ nella base $\hat{e}_i$, mentre le componenti covarianti rappresentano la *proiezione* del vettore $\vec{v}$ sui versori $\hat{e}_i$ della base, come si vede dalla figura seguente. Nel caso di un sistema di riferimento ortogonale, come anticipato nella prima parte, le due rappresentazioni coincidono.

[1] Il sistema di meridiani e paralleli su una sfera sono di questo tipo.

[2] La convenzione di Einstein sugli indici ripetuti.



Questo ci permette di definire un vettore covariante come $v_\alpha = (v_1, v_2)$ e un vettore controvariante come $v^\alpha = (v^1, v^2)$, entrambi associati al vettore $\vec{v}$ dello spazio $\mathbb{R}^2$. La generalizzazione ad altri spazi e dimensioni è immediata.

Il prodotto scalare risulta evidentemente invariante: $A_\mu B^\mu = A'_\mu B'^\mu$.

Nello stesso modo si possono costruire i tensori:

$$\begin{aligned} T_{\mu\nu} &= \text{Tensore Covariante} \\ T^{\mu\nu} &= \text{Tensore Controvariante} \\ T^\nu_\mu &= \text{Tensore Misto} \end{aligned}$$

Può, naturalmente, essere definito un tensore nullo. Diamo la legge di contrazione degli indici: *se in un tensore esiste un indice covariante ed uno controvariante ripetuto, allora si sottintende una sommatoria sull'indice ed il risultato è ancora un tensore di rango diminuito di due*:

$$T^{\alpha\mu}_{\alpha\nu} = T^\mu_\nu$$

Un tensore numerico fondamentale (equivalente del δ di Kronecker) è il seguente:

$$\delta^\alpha_\beta = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$$

che trasformando come:

$$\delta'^\alpha_\beta = \frac{\partial x'^\beta}{\partial x^\mu} \frac{\partial x^\nu}{\partial x'^\alpha} \delta^\nu_\mu$$

risulta invariante.

La relazione $A_{\mu\alpha} A^{\alpha\nu} = \delta^\nu_\mu$ permette di definire l'inverso di un tensore.

Consideriamo ora un tensore di rango 2:

$$g'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}$$

Un tensore di questo tipo si può considerare come una matrice 4×4 di rango 4. In effetti possiamo definire a partire da questo tensore un determinante che trasforma come:

$$g' = \left| \frac{\partial x^\alpha}{\partial x'^\nu} \right|^2 g$$

ovvero tale che il determinante del trasformato è il prodotto dello Jacobiano della trasformazione per il determinante. Ne consegue anche:

$$\boxed{\sqrt{g'} = \left| \frac{\partial x^\alpha}{\partial x'^\nu} \right| \sqrt{g}}$$

Ora, si consideri uno scalare A. Si può definire per esso un integrale di volume, il cui elemento infinitesimo trasforma come:

$$\int_V A d^4x \qquad \rightarrow \qquad \int_V A \mathcal{J} d^4x' \int_V A \left| \frac{\partial x}{\partial x'} \right| d^4x'$$

Dove $\mathcal{J}$ indica lo Jacobiano della trasformazione. Ma questo è equivalente ad integrare una variabile che trasforma come:

$$A \to A' = A \left| \frac{\partial x}{\partial x'} \right|$$

uno scalare che trasforma in questo modo si chiama *densità scalare*. Ne consegue che *la radice di un determinante è una densità scalare*.

Ora, in analogia alle matrici, il determinante di un tensore di rango 2 può essere scritto in termini dei suoi minori:

$$g \equiv |g_{\mu\nu}| = g_{\mu\nu} M^{\mu\nu} \qquad M^{\mu\nu} = \text{ minore corrispondente di } g_{\mu\nu}$$

relazione che può essere riscritta nella forma:

$$g_{\mu\nu} M^{\alpha\nu} = \delta^\alpha_\mu g \to g_{\mu\nu} \frac{M^{\alpha\nu}}{g} = \delta^\alpha_\mu$$

da cui si deduce che $g^{\mu\nu} = \frac{M^{\mu\nu}}{g}$ è un tensore controvariante di rango 2.

Il tensore antisimmetrico:

$$T'_{1234} = \frac{\partial x^\alpha}{\partial x'^1} \frac{\partial x^\beta}{\partial x'^2} \frac{\partial x^\mu}{\partial x'^3} \frac{\partial x^\nu}{\partial x'^4} T_{\alpha\beta\mu\nu} = \left| \frac{\partial x^\alpha}{\partial x'_1} \right| T_{1234}$$

può essere pensato come densità scalare. In maniera analoga, da uno scalare si può costruire un tensore, ed in effetti si definisce il *Tensore di Levi-Civita* $\epsilon^{\mu\nu\alpha\beta}$.

Sia dato ora un tensore a due indici $\phi_{\mu\nu}$. La quantità:

$$\phi = \frac{1}{8} \epsilon^{\mu\nu\alpha\beta} \phi_{\mu\nu} \phi_{\alpha\beta} = \phi_{12}\phi_{34} + \phi_{23}\phi_{14} + \phi_{31}\phi_{24}$$

è una densità scalare. Analogamente, la quantità:

$$f^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} \phi_{\alpha\beta}$$

è chiamata *densità tensoriale*.

11.3 Estensione del concetto di derivata

Consideriamo ora la derivata di uno scalare:

$$\frac{\partial}{\partial x^\alpha} \frac{\partial \varphi}{\partial x^\beta} = \frac{\partial}{\partial x^\alpha} A_\beta \neq T_{\alpha\beta}$$

Questa non risulta essere un tensore. Infatti, se si cambia semplicemente sistema di riferimento:

$$\begin{aligned}
\frac{\partial}{\partial x'^\alpha} \frac{\partial \varphi}{\partial x'^\beta} &= \frac{\partial}{\partial x'^\alpha} \left[\frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial \varphi}{\partial x^\mu} \right] \\
&= \frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial}{\partial x'^\alpha} \frac{\partial \varphi}{\partial x^\mu} + \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta} \frac{\partial \varphi}{\partial x^\mu} \\
&= \frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial^2 \varphi}{\partial x^\mu \partial x^\nu} \frac{\partial x^\nu}{\partial x'^\alpha} + \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta} \frac{\partial \varphi}{\partial x^\mu} \\
&= \frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial x^\nu}{\partial x'^\alpha} \frac{\partial^2 \varphi}{\partial x^\mu \partial x^\nu} + \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta} \frac{\partial \varphi}{\partial x^\mu}
\end{aligned}$$

la derivata seconda è quindi composta da un termine che trasforma come un tensore (il primo) più un altro termine: il risultato quindi non è un tensore covariante a due indici.

Un caso analogo si presenta per un vettore:

$$\frac{\partial}{\partial x'^\alpha} A'_\beta = \frac{\partial}{\partial x'^\alpha} \left[\frac{\partial x^\mu}{\partial x'^\beta} A_\mu \right] = \frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial A_\mu}{\partial x^\nu} \frac{\partial x^\nu}{\partial x'^\alpha} + \frac{\partial}{\partial x'^\alpha} \frac{\partial x^\mu}{\partial x'^\beta} A_\mu$$

Questo significa che *se la derivata è nulla in un sistema, non si può garantire che sia nulla in tutti i sistemi.* I termini "estranei" si possono però eliminare. Infatti, introducendo per comodità la notazione:

$$\frac{\partial}{\partial x^\alpha}\varphi \equiv \varphi_{,\alpha} \qquad \frac{\partial}{\partial x^\alpha}A_\beta \equiv A_{\beta,\alpha}$$

risulta:

$$\varphi_{,\alpha,\beta} - \varphi_{,\beta,\alpha} = 0$$

Si può poi definire la quantità

$$T_{\mu\nu} = \frac{\partial}{\partial x^\mu}A_\nu - \frac{\partial}{\partial x^\nu}A_\mu$$

che è la differenza di due quantità che non sono tensori, ma essendo:

$$T'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\left[\frac{\partial}{\partial x^\alpha}A_\beta - \frac{\partial}{\partial x^\beta}A_\alpha\right] + \frac{\partial^2 x^\beta}{\partial x'^\mu \partial x'^\nu}A_\beta - \frac{\partial^2 x^\alpha}{\partial x'^\mu \partial x'^\nu}A_\alpha$$

gli ultimi due termini si cancellano e quindi $T_{\mu\nu}$ è effettivamente un tensore di rango 2 antisimmetrico (*rotore generalizzato*).

Se consideriamo ora la divergenza ciclica del tensore $T_{\mu\nu}$:

$$T_{\alpha\mu\nu} = \frac{\partial}{\partial x^\alpha}\left(\frac{\partial}{\partial x^\mu}A_\nu - \frac{\partial}{\partial x^\nu}A_\mu\right) + \text{Permutazioni Cicliche}$$

questo risulta essere ancora un tensore di rango 3. E infine, è ancora un tensore:

$$T_{\alpha\mu\beta\nu} = \frac{\partial}{\partial x^\alpha}T_{\mu\beta\nu} + \text{Permutazioni Cicliche}$$

11.4 La connessione affine

Riconsideriamo la relazione:

$$\frac{\partial}{\partial x'^\alpha}A'_\beta - \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta}A_\mu = \frac{\partial x^\mu}{\partial x'^\beta}\frac{\partial x^\nu}{\partial x'^\alpha}\frac{\partial A_\mu}{\partial x^\nu}$$

siccome vale senz'altro la relazione:

$$A_\mu = \frac{\partial x'^\nu}{\partial x^\mu}A'_\nu$$

la si può sostituire nella relazione precedente:

$$\frac{\partial}{\partial x'^\alpha}A'_\beta - \frac{\partial x'^\nu}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta}A'_\nu = \frac{\partial x^\mu}{\partial x'^\beta}\frac{\partial x^\nu}{\partial x'^\alpha}\frac{\partial A_\mu}{\partial x^\nu}$$

Definendo ora la quantità (che *non è un tensore*):

$$\Gamma'^\nu_{\alpha\beta} \equiv \frac{\partial x'^\nu}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta}$$

la derivata si può riscrivere in questi termini:

$$\frac{\partial}{\partial x'^\alpha}A'_\beta - \Gamma'^\nu_{\alpha\beta}A'_\nu = \frac{\partial x^\mu}{\partial x'^\beta}\frac{\partial x^\nu}{\partial x'^\alpha}\frac{\partial A_\mu}{\partial x^\nu}$$

Ora, se supponiamo nulla la derivata del vettore in x, in x' essa sarà in genere data da $\Gamma'^\nu_{\alpha\beta}A'_\nu$. Questo è dovuto in ultima analisi al fatto che in relatività generale le trasformazioni di coordinate non sono più lineari e quindi le nuove coordinate non sono più funzione lineare delle vecchie. Se lo fossero, il termine $\frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta}$ sarebbe nullo e la normale derivata sarebbe effettivamente un tensore di rango 2. Vale qui la pena di notare che le trasformazioni di Lorentz sono trasformazioni lineari, quindi in relatività speciale la derivata è già covariante: l'usuale operazione di derivazione è infatti invariante per variazioni di sistema di riferimento. Poiché invece le trasformazioni di coordinate considerate in relatività generale non sono lineari, l'usuale operazione di derivazione non è più covariante. E questo rappresenta un problema serio.

Il problema si risolve definendo una nuova operazione di derivazione (*Derivata Covariante di un vettore covariante*) con la relazione:

$$D_\alpha A_\beta \equiv \left[\frac{\partial}{\partial x'^\alpha} A'_\beta - \Gamma'^\nu_{\alpha\beta} A'_\nu\right]$$

definizione necessaria per eliminare la classe privilegiata di sistemi di riferimento in cui $\Gamma'^\nu_{\alpha\beta} = 0$.[3] Infatti, mentre variando in maniera generica il sistema di riferimento la derivata potrà assumere valori diversi da zero, la quantità $\frac{\partial}{\partial x'^\alpha} A'_\beta - \Gamma'^\nu_{\alpha\beta} A'_\nu = 0$ sarà invece sempre nulla. D'altra parte in un sistema di riferimento generico si può sempre considerare la quantità (che definiamo *Connessione Affine* o *Affinità*):

$$\boxed{\Gamma'^\nu_{\alpha\beta} \equiv \frac{\partial x'^\nu}{\partial x^\mu} \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta} \qquad \text{Connessione Affine}}$$

in modo che la definizione di sopra abbia senso.[4]

Si giustifica quindi il fatto che la connessione affine non sia un tensore essa stessa: dovendo infatti trasformare in modo da ridurre una quantità non tensoriale ad una tensoriale, è evidente che non può lei stessa essere un tensore. Questo ha una conseguenza importante: potendo essere nulla in alcuni sistemi e diversa da zero in altri, allora *è possibile trovare una particolare trasformazione di coordinate in cui la connessione sia nulla: in questo particolare sistema di riferimento, la derivata covariante si riduce localmente alla derivata usuale, e la relatività generale si riduce a quella speciale.* Abbiamo quindi potuto dare un formulazione formale al principio di equivalenza.

La derivata covariante si indica con la notazione:

$$D_\alpha A_\beta \equiv A_{\beta;\alpha}$$

e per uno scalare ci si aspetta che valga $\varphi_{,\alpha} = \varphi_{;\alpha}$.

Si dimostra che continuano a valere le regole di derivazione usuali, di modo che la derivata covariante possa considerarsi effettivamente come una estensione del concetto di derivata "classica".

Consideriamo ora il vettore covariante A_α. Allora si ha $A^\alpha = \delta^\beta_\alpha A_\beta$, per cui la sua derivata covariante fornisce:

$$A_{\alpha;\mu} = \delta^\beta_{\alpha;\mu} A_\beta + \delta^\beta_\alpha A_{\beta;\mu} = \delta^\beta_{\alpha;\mu} A_\beta + A_{\alpha;\mu} \Rightarrow \delta^\beta_{\alpha;\mu} = 0$$

per cui per il prodotto scalare vale $(A^\alpha B_\alpha)_{,\mu} = (A^\alpha B_\alpha)_{;\mu}$. Da questa relazione si ricava:

$$\cancel{A_{\alpha,\mu} B^\alpha} + A_\alpha B^\alpha_{,\mu} = \left(A_{\alpha,\mu} - \Gamma^\nu_{\mu\alpha} A_\nu\right) B^\alpha + A_\alpha B^\alpha_{;\mu} = \cancel{A_{\alpha,\mu} B^\alpha} - \Gamma^\nu_{\mu\alpha} A_\nu B^\alpha + A_\alpha B^\alpha_{;\mu}$$

rinominando l'indice muto α in ν e viceversa:[5]

$$A_\alpha B^\alpha_{,\mu} = -\Gamma^\alpha_{\mu\nu} A_\alpha B^\nu + A_\alpha B^\alpha_{;\mu}$$

ovvero:

$$A_\alpha \left[B^\alpha_{;\mu} - B^\alpha_{,\mu} - \Gamma^\alpha_{\mu\nu} B^\nu\right] = 0 \qquad \forall A_\alpha$$

che ci permette di scrivere la *Derivata Covariante di un vettore controvariante* nella forma:

$$B^\alpha_{;\mu} = B^\alpha_{,\mu} + \Gamma^\alpha_{\nu\mu} B^\nu$$

definizione che può essere generalizzata ad un vettore qualsiasi.

Per un tensore, la generalizzazione è immediata. Per esempio per un tensore una volta covariante e due volte controvariante vale:

$$T^{\mu\nu}_{\alpha;\lambda} = T^{\mu\nu}_{\alpha,\lambda} + \Gamma^\mu_{\sigma\lambda} T^{\sigma\nu}_\alpha + \Gamma^\nu_{\lambda\sigma} T^{\mu\sigma}_\alpha - \Gamma^\sigma_{\lambda\alpha} T^{\mu\nu}_\sigma$$

[3] Ovvero i sistemi di riferimento inerziali!

[4] Per la connessione stessa vale:

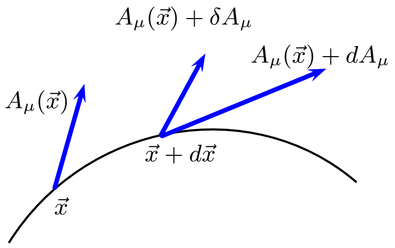
$$\Gamma'^\nu_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x_\rho}{\partial x'^\beta} \frac{\partial x'^\nu}{\partial x_\lambda} \Gamma^\lambda_{\mu\rho} + \frac{\partial x'^\nu}{\partial x^\mu} \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta}$$

[5] Si tratta di indice *muto* perché, essendo sottintesa una sommatoria, è irrilevante il modo in cui lo si chiama.

11.4.1 Interpretazione geometrica. Il trasporto parallelo.

La differenza tra derivata normale e derivata covariante è suscettibile di una interessante interpretazione geometrica. Come è noto, la derivata normale si effettua calcolando il rapporto incrementale del vettore, il che richiede implicitamente di effettuare la differenza di due vettori in due punti vicini.

In un sistema di coordinate curvilineo, questa operazione non è banale e deve essere opportunamente generalizzata. Consideriamo dunque un vettore $A_\mu(\vec{x})$ applicato nel punto $\vec{x}$. In un sistema generale di coordinate si deve fare distinzione fra il vettore applicato nel punto vicino $A_\mu(\vec{x}+d\vec{x}) = A_\mu(\vec{x}) + dA_\mu$ ed il vettore ottenuto trasportando parallelamente il vettore originario fino al punto $\vec{x}+d\vec{x}$: $A_\mu(\vec{x}) \to A_\mu(\vec{x}) + \delta A_\mu$.



Nel *trasporto parallelo* di un vettore in un sistema generico di coordinate le sue componenti cambiano, a differenza di ciò che accade in metrica piatta: per trasporto parallelo si intende infatti il trasporto del vettore effettuato *mantenendo costante l'angolo che esso forma con la superficie curva.* Per calcolare il rapporto incrementale si deve pertanto prima trasportare parallelamente il vettore nel punto $\vec{x}+d\vec{x}$ e poi calcolare la differenza dei due vettori applicati ora nello stesso punto $[A_\mu(\vec{x}) + dA_\mu - (A_\mu(\vec{x}) + \delta A_\mu)]$. In generale la variazione delle componenti del vettore è espressa da:[6]

$$\delta A_\mu = \Gamma^\rho_{\mu\nu}\,(A_\rho dx^\nu)$$

che sostituita nel rapporto incrementale ed effettuato il limite, fornisce proprio la derivata covariante.

Si impone quindi la seguente interpretazione: *la connessione affine è la proprietà di assegnare al campo come trasportare un vettore parallelamente, con l'assunzione che dipenda linearmente da A_ρ e da dx^ν, cioè dal vettore stesso e dallo spostamento del punto di applicazione.*

Ne consegue subito che in caso di coordinate cartesiane risulta $\Gamma^\rho_{\mu\nu} = 0$. Infatti, partendo dalla definizione di rapporto incrementale:

$$DA_\mu = \lim_{\Delta x \to 0}\left[\frac{dA_\mu - \delta A_\mu}{\Delta x}\right] = \left[\frac{\partial A_\mu}{\partial x^\nu} - \Gamma^\rho_{\mu\nu} A_\rho\right] dx^\nu$$

si vede che poiché la derivata in coordinate cartesiane coincide con la derivata classica, allora la connessione deve essere nulla.

Consideriamo ora un vettore $h_\mu(x)$ e costruiamo un campo vettoriale per trasporto parallelo di questo vettore h_μ: $\delta h_\mu = dh_\mu$. Questo implica allora che nel campo vettoriale così costruito valga ovunque $D_\alpha h_\mu(x) \equiv 0$.

Consideriamo ora una curva quadridimensionale ed un vettore tangente ad essa. Vale senz'altro:

$$DA_\mu = \left[\frac{\partial A_\mu}{\partial x^\nu} - \Gamma^\rho_{\mu\nu} A_\rho\right] dx^\nu$$

[6]Può essere ricavata integrando l'equazione della geodetica su una linea chiusa:

$$\frac{d}{ds}A_\mu - \Gamma^\nu_{\mu\lambda}\frac{dx^\lambda}{ds}A_\nu = 0 \to A_\mu = \oint \Gamma^\nu_{\mu\lambda} A_\nu dx^\lambda + A_\mu(p)$$

che permette di scrivere la variazione delle componenti del vettore lungo una curva chiusa come:

$$\delta A_\mu = \oint \Gamma^\nu_{\mu\lambda} A_\nu dx^\lambda$$

una curva quadridimensionale può essere individuata da un parametro τ (ad esempio, il tempo proprio), quindi:

$$DA_\mu = \frac{\partial A_\mu}{\partial \tau} d\tau - \Gamma'^{\rho}_{\mu\nu} A_\rho \frac{dx^\nu}{d\tau} d\tau$$

e quindi la derivata lungo la curva τ è data da:

$$\frac{DA_\mu}{d\tau} = \left[\frac{\partial A_\mu}{\partial \tau} - \Gamma'^{\rho}_{\mu\nu} A_\rho \frac{dx^\nu}{d\tau} \right]$$

dove $dx^\nu/d\tau$ risulta essere la quadrivelocità. Dunque, il vettore si sposta parallelamente alla curva τ quando la sua derivata covariante rispetto alla curva è nulla.

11.5 La metrica

Introduciamo ora il concetto di metrica.

Consideriamo il quadrintervallo $ds^2 = c^2 dt^2 - d\vec{r}^2$. Come abbiamo visto, in un sistema non inerziale questa relazione di può generalizzare nella forma:

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$$

dove il termine $g_{\alpha\beta}$ definisce come costruire la distanza nel sistema di riferimento ed è un tensore covariante a due indici. In particolare:

$$g'_{\alpha\beta} = \frac{dx^\mu}{dx'^\alpha} \frac{dx^\nu}{dx'^\beta} g_{\mu\nu}$$

Siccome poi si può considerare un tensore di rango 2 come una matrice 4×4, si può porre:

$$\det g' = \mathcal{J}^2 \det g$$

Il particolare valore di $g_{\alpha\beta}$ che riproduce il quadrintervallo ds^2 è:

$$g_{\alpha\beta} \equiv \eta_{\alpha\beta} = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & -1 \end{array} \right)$$

Questo è, appunto, il caso di coordinate cartesiane (pseudo-euclidee). Il determinante di questo tensore è sempre negativo, dunque in genere nelle notazioni si indica con $\sqrt{-g}$. Ponendo inoltre un tensore che soddisfi:

$$g_{\alpha\beta} g^{\beta\mu} = \delta^\mu_\alpha$$

esso permette di passare dalle componenti covarianti e controvarianti e viceversa, ovvero di innalzare gli indici come $A^\mu = g^{\mu\nu} A_\nu$ e abbassarli come $A_\mu = g_{\mu\nu} A^\nu$.

L'interpretazione geometrica di questo tensore è la seguente.

Si consideri una base $\vec{e}_i$. Allora $\vec{A} = A^i \vec{e}_i$ e $A_i = \vec{A} \cdot \vec{e}_i$, il prodotto scalare di due vettori risulta allora espresso da:

$$\vec{A} \cdot \vec{B} = A^i B^j \vec{e}_i \cdot \vec{e}_j$$

quindi, se il sistema di riferimento non è cartesiano ortogonale, risulta proprio:

$$g_{ij} = \vec{e}_i \cdot \vec{e}_j$$

Questa interpretazione giustifica per il tensore $g_{\alpha\beta}$ il nome di *Tensore Metrico*. In quanto simmetrico, esso ha 10 componenti.

Ora:

$$D_\lambda A_\mu = \frac{D}{Dx^\lambda} g_{\mu\nu} A^\nu = \left(\frac{D}{Dx^\lambda} g_{\mu\nu} \right) A^\nu + g_{\mu\nu} \frac{DA^\nu}{Dx^\lambda}$$

innalzando invece l'indice direttamente al termine di derivata:

$$D_\lambda A_\mu = g_{\mu\nu} D_\lambda A^\nu$$

da cui si ricava l'importante proprietà:

$$\boxed{g_{\mu\nu;\lambda} = 0 \qquad \textbf{Postulato Metrico}}$$

In sostanza, il postulato metrico appena scritto afferma che *un tensore a due indici la cui derivata covariante sia nulla è un tensore metrico*. Questa è una condizione sufficiente, ma non necessaria: esistono infatti dei tensori metrici la cui derivata covariante non è nulla. È il caso, ad esempio, dei tensori metrici di spazi con torsione.

Capitolo 12

Il Punto Materiale nel Campo Gravitazionale

12.1 Generalità

Si supponga inizialmente un campo uniforme e indichiamo le coordinate con ξ^α. Supponiamo le equazioni del moto del tipo:[1]

$$mc\frac{d^2\xi^\alpha}{d\tau^2} = 0 \qquad d\tau^2 = -\eta_{\alpha\beta}d\xi^\alpha d\xi^\beta$$

Sia $x^\alpha \equiv x^\alpha(\xi^\alpha)$ un riferimento in cui compare esplicitamente il campo gravitazionale. L'equazione di sopra si riscrive allora come:

$$\frac{d}{d\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\right) = \frac{d^2x^\mu}{d\tau^2}\frac{\partial\xi^\alpha}{\partial x^\mu} + \frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}\frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau}$$

che moltiplicata per $\frac{\partial x^\beta}{\partial\xi^\alpha}$ e sommata su α fornisce:

$$\frac{\partial x^\beta}{\partial\xi^\alpha}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \frac{\partial x^\beta}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}\frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau} = 0$$

ovvero:

$$\frac{d^2x^\beta}{d\tau^2} + \frac{\partial x^\beta}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}u^\nu u^\mu = 0$$

che ricordando la definizione di connessione può essere messa nella forma:

$$\frac{d^2x^\beta}{d\tau^2} + \Gamma^\beta_{\mu\nu}u^\nu u^\mu = 0$$

Se confrontiamo ora questa equazione con la (3.5):

$$mc\frac{du^\beta}{d\tau} = -mc\,\Gamma^\beta_{\mu\nu}u^\nu u^\mu$$

si deduce che il campo gravitazionale[2] è espresso da:

$$\boxed{-mc\,\Gamma^\beta_{\mu\nu}u^\nu u^\mu}$$

[1]Sulla base delle (3.5).
[2]Presente nell'equazione appena scritta perché si è utilizzato un sistema in cui compare esplicitamente.

cioé *la connessione è legata alla forza gravitazionale.*

Questa è l'equazione della geodetica, ovvero del moto di un punto materiale in un campo gravitazionale. È importante notare che l'equazione del moto è in pratica la derivata covariante della quadrivelocità, cioé l'equazione del moto ha la forma $\frac{Du^\beta}{D\tau}$. Quindi un punto materiale si muove in modo che il quadrivettore velocità rimanga parallelo a sé stesso: queste sono per definizione le cosiddette *geodetiche*.

Notiamo che siccome è possibile trovare una trasformazione generale di coordinate in cui la connessione sia nulla, ciò significa che è possibile trovare una particolare trasformazione di coordinate in cui il campo gravitazionale è localmente nullo.[3]

La connessione risulta quindi legata alla forza gravitazionale, cosa possiamo dire invece della metrica? Se la connessione è la proprietà di trasportare un vettore e la metrica definisce la distanza in un determinato sistema di riferimento, intuitivamente deve essere possibile stabilire una qualche relazione fra queste due quantità. Ricerchiamo dunque l'espressione che le lega. Sia:

$$d\tau^2 = -\eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}dx^\mu dx^\nu$$

Ora, una trasformazione di coordinate generiche applicata al tensore di rango due $\eta_{\alpha\beta}$ fornisce proprio:

$$g_{\mu\nu} = \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}$$

quindi:

$$d\tau^2 = -g_{\mu\nu}dx^\mu dx^\nu$$

Calcoliamo ora a partire dalle relazioni precedenti la derivata del tensore metrico rispetto a x_λ:

$$\frac{\partial}{\partial x^\lambda}g_{\mu\nu} = \frac{\partial^2\xi^\alpha}{\partial x^\lambda\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}\eta_{\alpha\beta} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\lambda\partial x^\nu}\eta_{\alpha\beta}$$

che ricordando la definizione di connessione $\Gamma^\rho_{\mu\nu} = \frac{\partial x^\rho}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu x^\nu}$ si può riscrivere come:

$$g_{\mu\nu,\lambda} = \Gamma^\rho_{\mu\lambda}g_{\rho\nu} + \Gamma^\rho_{\lambda\nu}g_{\rho\mu} \tag{12.1}$$

Naturalmente, il nome dell'indice non è importante, per cui nella formula appena scritta si possono scambiare gli indici $\mu \leftrightarrow \lambda$ e $\nu \leftrightarrow \lambda$ ottenendo due formule analoghe:

$$\begin{aligned} g_{\lambda\nu,\mu} &= \Gamma^\rho_{\lambda\mu}g_{\rho\nu} + \Gamma^\rho_{\mu\nu}g_{\rho\lambda} \qquad \mu \leftrightarrow \lambda \\ g_{\mu\lambda,\nu} &= \Gamma^\rho_{\mu\nu}g_{\rho\lambda} + \Gamma^\rho_{\nu\lambda}g_{\rho\mu} \qquad \nu \leftrightarrow \lambda \end{aligned}$$

sommando ora la prima di queste e sottraendo la seconda alla (12.1)[4] per separare la connessione e riscriverla in funzione della metrica:

$$g_{\mu\nu,\lambda} + g_{\lambda\nu,\mu} - g_{\mu\lambda,\nu} = \Gamma^\rho_{\mu\lambda}g_{\rho\nu} + \cancel{\Gamma^\rho_{\lambda\nu}g_{\rho\mu}} + \Gamma^\rho_{\lambda\mu}g_{\rho\nu} + \cancel{\Gamma^\rho_{\mu\nu}g_{\rho\lambda}} - \cancel{\Gamma^\rho_{\mu\nu}g_{\rho\lambda}} - \cancel{\Gamma^\rho_{\nu\lambda}g_{\rho\mu}}$$

tenendo presente che la metrica è simmetrica per scambio di indici e che se vale il teorema di Schwarz la connessione è a sua volta simmetrica per scambio degli indici inferiori, si ricava:

$$g_{\mu\nu,\lambda} + g_{\lambda\nu,\mu} - g_{\mu\lambda,\nu} = 2\Gamma^\rho_{\mu\lambda}g_{\rho\nu}$$

introducendo infine un tensore metrico controvariante tale da soddisfare la relazione $g_{\rho\nu}g^{\alpha\nu} = \delta^\alpha_\rho$ (*tensore inverso*), moltiplicando per esso l'ultima espressione e sommando sull'indice ρ, si ricava:

$$\boxed{\Gamma^\alpha_{\mu\lambda} = \frac{1}{2}g^{\alpha\nu}\left(g_{\mu\nu,\lambda} + g_{\lambda\nu,\mu} - g_{\mu\lambda,\nu}\right)} \tag{12.2}$$

[3] E cioé è possibile tramite una opportuna trasformazione di coordinate, non lineare perché deve annullare il termine di connessione, eliminare il campo gravitazionale dal sistema di riferimento. Di nuovo il *principio di Equivalenza.*

[4] Naturalmente, una qualunque combinazione che preveda due segni positivi ed un segno negativo andrebbe bene, l'unica differenza sarebbe il nome degli indici ottenuti nella formula finale, ma questo è ovviamente irrilevante.

È evidente che essa ha 40 componenti: gli indici μ e λ forniscono 10 componenti: 4 dalla diagonale e 6 fuori diagonale (il tensore metrico è simmetrico). L'indice α va da 0 a 3, fornendo quindi un totale di $10 \times 4 = 40$ componenti.

La connessione risulta quindi legata alle derivate prime della metrica, ne consegue che la metrica è legata al potenziale.

La connessione espressa in termini della metrica si indica con il simbolo:

$$\Gamma^{\alpha}_{\mu\lambda} = \left\{ \begin{array}{c} \alpha \\ \mu \ \lambda \end{array} \right\}$$

che prende il nome di *Simbolo di Christoffel di II ordine.*

12.2 L'azione del campo gravitazionale. Il limite classico

Ricerchiamo ora l'azione associata al campo gravitazionale.

L'azione del campo ha la forma generica:

$$S = -mc \int_1^2 ds \qquad\qquad -ds^2 = g_{\mu\nu} dx^{\mu} dx^{\nu}$$

da cui:

$$\delta ds^2 = -2ds\delta ds = \frac{\partial g_{\mu\nu}}{\partial x^{\alpha}} \delta x^{\alpha} dx^{\mu} dx^{\nu} + 2g_{\mu\nu} dx^{\mu} d\delta x^{\nu}$$

e siccome vale evidentemente:

$$\delta S = -mc \int_1^2 \delta ds$$

si può sostituire in questa l'espressione del δds:([5])

$$\begin{aligned}
\delta S &= mc \int_1^2 \left[\frac{1}{2} g_{\mu\nu,\alpha} \delta x^{\alpha} \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} + g_{\mu\nu} \frac{dx^{\mu}}{ds} \frac{d\delta x^{\nu}}{ds} \right] ds \\
&= mc \int_1^2 \left[\frac{1}{2} g_{\mu\nu,\alpha} u^{\mu} u^{\nu} \delta x^{\alpha} - \frac{d}{ds} \left(g_{\mu\nu} \frac{dx^{\mu}}{ds} \right) \delta x^{\nu} \right] ds \\
&= mc \int_1^2 \left[\frac{1}{2} g_{\mu\alpha,\nu} u^{\mu} u^{\alpha} - \frac{d}{ds} \left(g_{\mu\nu} \frac{dx^{\mu}}{ds} \right) \right] \delta x^{\nu} ds
\end{aligned}$$

Nel secondo passaggio si è integrato il secondo termine per parti e si è tenuto conto del fatto che δx^{ν} si annulla agli estremi.

Ora, imporre che la variazione dell'azione sia nulla $\forall \delta x^{\nu}$ e $\forall ds$ implica che:

$$\frac{1}{2} g_{\mu\alpha,\nu} u^{\mu} u^{\alpha} - \frac{d}{ds} \left(g_{\mu\nu} \frac{dx^{\mu}}{ds} \right) = 0$$

ovvero:

$$\begin{aligned}
\frac{1}{2} g_{\mu\alpha,\nu} u^{\mu} u^{\alpha} - \frac{d}{ds} \left(g_{\mu\nu} \frac{dx^{\mu}}{ds} \right) = 0 \rightarrow \\
\frac{1}{2} g_{\mu\alpha,\nu} u^{\mu} u^{\alpha} - g_{\mu\nu,\alpha} \frac{dx^{\alpha}}{ds} u^{\mu} - g_{\mu\nu} \frac{du^{\mu}}{ds} = 0
\end{aligned}$$

Notando che vale la relazione:

$$g_{\mu\nu,\alpha} u^{\alpha} u^{\mu} = -\frac{1}{2} u^{\alpha} u^{\mu} \left(g_{\mu\nu,\alpha} + g_{\alpha\nu,\mu} \right)$$

[5]Vale infatti:

$$g_{\mu\nu,\alpha} dx^{\mu} dx^{\nu} + 2g_{\mu\nu} dx^{\mu} \delta dx^{\nu} = -2ds\delta ds$$

da cui:

$$\delta ds = -\frac{1}{2} \left[g_{\mu\nu,\alpha} u^{\mu} dx^{\nu} + g_{\mu\nu} u^{\mu} d\delta x^{\nu} \right]$$

si può riscrivere la relazione precedente come:

$$-\frac{1}{2}\left[g_{\mu\nu,\alpha}+g_{\alpha\nu,\mu}-g_{\mu\alpha,\nu}\right]u^\alpha u^\mu - g_{\mu\nu}\frac{du^\mu}{ds}=0$$

moltiplicando ora per $g^{\lambda\nu}$ si ottiene proprio l'espressione della connessione in funzione della metrica:

$$-\frac{1}{2}g^{\lambda\nu}\left[g_{\mu\nu,\alpha}+g_{\alpha\nu,\mu}-g_{\mu\alpha,\nu}\right]u^\alpha u^\mu - g^{\lambda\nu}g_{\mu\nu}\frac{du^\mu}{ds}=0$$

ricordando infatti l'espressione (12.2), si ricava finalmente:

$$\Gamma^\lambda_{\mu\alpha}u^\mu u^\alpha+\frac{du^\lambda}{ds}=0$$

dunque *la metrica è effettivamente legata al potenziale.*

Consideriamo ora il limite classico. Per limite classico si intendono le seguenti approssimazioni:

- velocità piccole: $\frac{d\vec{x}}{d\tau}\ll\frac{dt}{d\tau}$ $\left(\frac{dx}{d\tau}\ll\frac{dt}{d\tau}\right)$,
- piccoli campi gravitazionali: $g_{\alpha\beta}=\eta_{\alpha\beta}+h_{\alpha\beta}$, $h_{\alpha\beta}\ll 1$ (*limite di campo debole*)

In tali approssimazioni il termine prevalente nell'equazione della geodetica è rappresentato da:

$$\frac{d^2x^\mu}{d\tau^2}+\Gamma^\mu_{00}\left(\frac{dt}{d\tau}\right)^2=0$$

e siccome vale (ancora dalla (12.2)):

$$\Gamma^\mu_{00}=\frac{1}{2}g^{\mu\lambda}\left(g_{\lambda 0,0}+g_{0\lambda,0}-g_{00,\lambda}\right)=-\frac{1}{2}\vec{\nabla}h_{00}$$

si può sostituire quest'ultima nell'equazione completa e separando la parte spaziale dalla parte temporale:

$$\begin{cases}\dfrac{d^2x^\mu}{d\tau^2}=\dfrac{1}{2}\vec{\nabla}h_{00}\left(\dfrac{dt}{d\tau}\right)^2 & \text{Parte Spaziale}\\[2ex] \dfrac{d^2t}{d\tau^2}=0 & \text{Parte Temporale}\end{cases}$$

La parte temporale ci dice in effetti che $\frac{dt}{d\tau}$ è costante, per cui è possibile dividere la parte spaziale per questa quantità, operazione che fornisce:[6]

$$\frac{d^2x^\mu}{dt^2}=\frac{1}{2}\vec{\nabla}h_{00}$$

ma l'equazione classica è data da:$\frac{d^2\vec{x}}{dt^2}=-\vec{\nabla}\Phi$, dunque imporre che nel limite non relativistico le geodetiche si riducano all'equazione classica implica che sussista la relazione:

$$h_{00}=-2\Phi$$

quindi nel limite classico $g_{\mu\nu}$ rispecchia il potenziale gravitazionale. Essendo infatti $g_{\mu\nu}=\eta_{\mu\nu}+h_{\mu\nu}$ e $\eta_{00}=-1$, risulta:

$$\boxed{g_{00}=-\left(1+2\Phi\right)}$$

Risulta inoltre che $\det(\eta_{\mu\nu})=\eta=-1$, e siccome vale la relazione:

$$\det g'=|\mathcal{J}|^2\det g$$

ne consegue che il determinante conserva lo stesso segno: nella particolare segnatura scelta, g conserva sempre il segno negativo. Più precisamente:

Il numero di autovalori positivi e negativi del tensore metrico $g_{\mu\nu}$ è conservato da trasformazioni generali di coordinate.

[6]Vale infatti:

$$\frac{d^2x^\mu}{d\tau^2}\left(\frac{d\tau}{dt}\right)^2=\frac{1}{2}\vec{\nabla}h_{00}\left(\frac{dt}{d\tau}\right)^2\left(\frac{d\tau}{dt}\right)^2$$

ed il quadrato dell'operatore derivata corrisponde alla doppia derivazione.

Capitolo 13

Secondo Intermezzo Matematico

13.1 La divergenza covariante

Definiamo, a partire dalla derivata covariante, una *divergenza covariante*. Supponiamo di avere un vettore controvariante: come nel caso classico, la divergenza si ottiene saturando gli indici della derivata.

$$D_\mu A^\mu = \partial_\mu A^\mu + \Gamma^\mu_{\mu\alpha} A^\alpha$$

Si vede che compare un termine aggiuntivo dovuto alla connessione $\Gamma^\mu_{\mu\alpha} A^\alpha$. Esplicitando questo termine:

$$\Gamma^\mu_{\mu\alpha} A^\alpha = \frac{1}{2} g^{\mu\sigma} (\partial_\alpha g_{\sigma\mu} + \partial_\mu g_{\alpha\sigma} - \partial_\sigma g_{\mu\alpha}) = \frac{1}{2} g^{\mu\sigma} \partial_\alpha g_{\sigma\mu}$$

infatti, μ e σ sono indici di somma simmetrici. Esplicitiamo meglio questo termine:

$$\frac{\partial}{\partial x^\alpha} g = \frac{\partial g}{\partial g_{\mu\sigma}} \frac{\partial g_{\mu\sigma}}{\partial x^\alpha} \qquad (g = g_{\mu\sigma} G^{\mu\sigma})$$

con $G^{\mu\sigma}$ minore corrispondente. Essendo:

$$g_{\mu\sigma} \frac{G^{\lambda\sigma}}{g} = \delta^\lambda_\mu \Rightarrow \frac{G^{\lambda\sigma}}{g} = g^{\lambda\sigma}$$

si ricava:

$$\frac{\partial}{\partial x^\alpha} g = g g^{\mu\sigma} \frac{\partial g_{\mu\sigma}}{\partial x^\alpha}$$

per cui:

$$\Gamma^\mu_{\mu\alpha} A^\alpha = \frac{1}{2} g^{\mu\sigma} \partial_\alpha g_{\sigma\mu} = \frac{1}{2g} \frac{\partial g}{\partial x^\alpha}$$

siccoma vale anche:

$$\frac{\partial}{\partial x^\alpha} \ln \sqrt{-g} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\alpha} \sqrt{-g} = \frac{1}{\sqrt{-g}} \frac{-1}{2\sqrt{-g}} \frac{\partial g}{\partial x^\alpha} = \frac{1}{2g} \frac{\partial g}{\partial x^\alpha}$$

ne consegue che il termine della connessione può essere scritto come:

$$\Gamma^\mu_{\mu\alpha} = \frac{\partial}{\partial x^\alpha} \ln \sqrt{-g}$$

La divergenza assume allora la forma defintiva:

$$D_\mu A^\mu = \partial_\mu A^\mu + \left(\frac{\partial}{\partial x^\alpha} \ln \sqrt{-g} \right) A^\alpha$$

ovvero:

$$\boxed{D_\mu A^\mu = \frac{1}{\sqrt{-g}}\,\partial_\mu\left(\sqrt{-g}A^\mu\right)}$$

L'integrale dell'espressione di sopra ci fornisce l'equivalente del teorema della divergenza:

$$\int D_\mu A^\mu \sqrt{-g} d^4x = \int \partial_\mu\left(\sqrt{-g}A^\mu\right) d^4x$$

Ricerchiamo ora le condizioni per cui la lunghezza di un vettore $|a^\mu|^2 = g_{\mu\nu}a^\mu a^\nu$ resta invariata per trasporto parallelo, ovvero:

$$g_{\mu\nu}(p)a^\mu(p)a^\nu(p) = g_{\mu\nu}(p')a^\mu(p')a^\nu(p')$$

dove:

$$a^\mu(p') = a^\mu(p) - \Gamma^\mu_{\alpha\beta}a^\alpha dx^\beta \qquad\qquad g_{\mu\nu}(p') = g_{\mu\nu}(p) + g_{\mu\nu,\lambda}dx^\lambda$$

Sostituendo queste relazioni:

$$g_{\mu\nu}a^\mu a^\nu = \left[g_{\mu\nu} + g_{\mu\nu,\lambda}dx^\lambda\right]\left[a^\mu - \Gamma^\mu_{\alpha\beta}a^\alpha dx^\beta\right]\left[a^\nu - \Gamma^\nu_{\alpha\beta}a^\alpha dx^\beta\right]$$

e dall'uguaglianza risulta la relazione (trascurando termini di ordine superiore in dx e rinominando successivamente gli indici muti):

$$\begin{aligned}
g_{\mu\nu,\lambda}a^\mu a^\nu dx^\lambda - g_{\mu\nu}\Gamma^\mu_{\alpha\beta}a^\alpha dx^\beta a^\nu - g_{\mu\nu}\Gamma^\nu_{\alpha\beta}a^\alpha dx^\beta a^\mu = 0 \rightarrow\\
\left[g_{\mu\nu,\lambda}a^\mu a^\nu dx^\lambda - g_{\alpha\nu}\Gamma^\alpha_{\mu\lambda}a^\mu a^\nu dx^\lambda - g_{\mu\alpha}\Gamma^\alpha_{\nu\lambda}a^\mu a^\nu dx^\lambda\right] = 0 \rightarrow\\
\left[g_{\mu\nu,\lambda}dx^\lambda - \Gamma^\alpha_{\mu\lambda}g_{\alpha\nu} - \Gamma^\alpha_{\nu\lambda}g_{\mu\alpha}\right]a^\mu a^\nu dx^\lambda = 0
\end{aligned}$$

ma questa è la derivata covariante del tensore metrico:

$$D_\lambda g_{\mu\nu} = 0$$

Quindi *il postulato metrico esprime il fatto che la lunghezza di un vettore non varia per trasporto parallelo.*

La condizione espressa dal postulato metrico è simmetrica in μ e ν, pertanto sono 40 equazioni in 4 dimensioni: $\frac{D^2(D+1)}{2}$. Dalla metrica si può pertanto ricavare la connessione simmetrica, che ha appunto 40 componenti.

Ora, solo gli integrali delle quantità scalari hanno senso, ma essi non sono invarianti:

$$\int_\Omega \varphi d\Omega = \int_\Omega \varphi\left|\frac{dx}{dx'}\right| d\Omega'$$

si è quindi ricondotti a considerare delle densità scalari definite da:

$$\varphi' = \varphi\left|\frac{dx}{dx'}\right|$$

In realtà:

$$g'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}g_{\mu\nu}$$

per cui:

$$\det g'_{\mu\nu} = \mathcal{J}^2 \det g_{\mu\nu} \qquad\qquad \sqrt{-g}\mathcal{J}\sqrt{-g}$$

e cioé il determinante risulta una densità scalare:

$$\int_\Omega \sqrt{-g}d\Omega = \int_\Omega \frac{\sqrt{-g}}{\mathcal{J}}\mathcal{J}d\Omega'$$

e l'elemento di volume invariante è dato da $\sqrt{-g}d\Omega$. In particolare quindi il valore di:

$$\int_{\Omega} D_{\mu}A^{\mu}\sqrt{-g}d\Omega$$

non dipende dal sistema di riferimento. Il teorema della divergenza prende quindi la forma:

$$\int_{\Omega} \frac{\partial}{\partial x^{\mu}}(\sqrt{-g}A^{\mu})d\Omega = \int_{\Sigma_{\mu}} \sqrt{-g}A^{\mu}d\Sigma_{\mu}$$

con $d\Sigma_{\mu}$ elemento di ipersuperficie.

Capitolo 14

Formulazione dell'Elettromagnetismo

14.1 L'elettromagnetismo in relatività generale

Si riconsideri l'elettromagnetismo classico.

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

Ricerchiamo la scrittura del tensore elettromagnetico $F_{\mu\nu}$ per trasformazioni generali di coordinate.

Consideriamo:

$$F_{\mu\nu} = D_\mu A_\nu - D_\nu A_\mu = \partial_\mu A_\nu - \Gamma^\alpha_{\mu\nu} A_\alpha - \partial_\nu A_\mu + \Gamma^\alpha_{\nu\mu} A_\alpha = \partial_\mu A_\nu - \partial_\nu A_\mu$$

La connessione infatti è simmetrica per scambio degli indici inferiori.

Le prime due equazioni di Maxwell senza sorgenti si scrivono:

$$\partial_\mu F_{\alpha\beta} + \partial_\alpha F_{\beta\mu} + \partial_\beta F_{\alpha\mu} = 0$$

Vediamo come trasforma questa equazione:

$$\begin{aligned}\frac{\partial}{\partial x'^\alpha} F'_{\mu\nu} = \frac{\partial}{\partial x'^\alpha}\left[\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} F_{\rho\sigma}\right] = \\ = \frac{\partial x^\lambda}{\partial x'^\alpha}\frac{\partial}{\partial x^\lambda} F_{\rho\sigma}\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} + \frac{\partial^2 x^\rho}{\partial x'^\alpha \partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} F_{\rho\sigma} + \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial^2 x^\sigma}{\partial x'^\alpha \partial x'^\nu} F_{\rho\sigma}\end{aligned}$$

Ora sommando fra loro i tre termini:

$$\frac{\partial x^\lambda}{\partial x'^\alpha}\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}\left[\frac{\partial}{\partial x^\lambda} F_{\rho\sigma} + \text{ permutazioni cicliche}\right] + \text{altri termini}$$

I termini delle derivate seconde hanno fattori che permutando gli indici si annullano (si ricordi che $F_{\rho\sigma}$ è antisimmetrico). Si ricava allora:

$$\epsilon_{\mu\alpha\beta\rho}\partial_\alpha F_{\beta\rho} = 0$$

e cioé le equazioni senza sorgenti si conservano.

Passiamo ora ad analizzare le equazioni con le sorgenti.

Per queste equazioni, è necessario introdurre innanzitutto un vettore corrente:

$$dq = \varrho dV$$

la carica deve, ovviamente, essere un invariante di Lorentz, quindi:

$$dq dx^\alpha = \varrho \frac{dx^\alpha}{dx^0} dV dx^0 = \frac{\varrho}{\sqrt{-g}}\frac{dx^\alpha}{dx^0}\sqrt{-g} d\Omega$$

con $d\Omega = dVdx^0$ elemento di volume. Quindi la corrente in forma invariante è:

$$j^\alpha = \frac{\varrho}{\sqrt{-g}} c \frac{dx^\alpha}{dx^0}$$

La seconda coppie di equazioni allora diventa:

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} j^\nu \qquad \rightarrow \qquad D_\mu F^{\mu\nu} = \frac{4\pi}{c} j^\nu$$

dove la corrente j^ν è data dall'espressione appena data sopra.

La densità lagrangiana del campo elettromagnetico diventa:

$$\mathcal{L} = -\frac{1}{4\pi} F_{\mu\nu} F^{\mu\nu} \qquad \rightarrow \qquad -\frac{1}{4\pi} F_{\mu\nu} F_{\rho\sigma} g^{\mu\rho} g^{\nu\sigma}$$

ed in maniera analoga:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi + V(\Phi) \qquad \rightarrow \qquad \mathcal{L} = g^{\mu\nu} \partial_\mu \Phi \partial_\nu \Phi + V(\Phi)$$

cioé nella densità di lagrangiana compare un termine (esplicito) di metrica. Questo termine è quello che in ultima analisi permette di accoppiare il campo elettromagnetico al campo gravitazionale e che permette, per esempio, di descrivere il fenomeno della luce deviata in un campo gravitazionale.

Consideriamo un sistema continuo, in cui abbiamo una densità di lagrangiana $\mathcal{L}(\Phi, \partial_\mu \Phi)$. L'azione è allora espressa da:

$$S = \int \mathcal{L} d^4 x$$

che in sistemi generali di coordinate si generalizza come:

$$S = \frac{1}{c} \int \mathcal{L} \sqrt{-g} d\Omega$$

Ovviamente, ci si attende che l'azione rimanga comunque invariata. Ricaviamo pertanto le equazioni del moto da questa forma generale:

$$\begin{aligned} \delta S\Big|_{\delta\Phi} &= \frac{1}{c} \int \left[\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial \Phi} \delta\Phi + \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\mu \Phi} \delta \partial_\mu \Phi \right] d\Omega = \\ &= \frac{1}{c} \int \left\{ \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial \Phi} \delta\Phi - \left[\partial_\mu \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\mu \Phi} \right] \delta\Phi + \partial_\mu \left[\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\mu \Phi} \delta\Phi \right] \right\} d\Omega \end{aligned}$$

l'ultimo termine nell'integrale si può trasformare – tramite il teorema della divergenza – in un integrale di superficie, dove Φ si annulla. L'equazione del moto diventa allora:

$$\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial \Phi} - \partial_\mu \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\mu \Phi} = 0$$

che è l'equazione di Eulero-Lagrange scritta per sistemi di coordinate arbitrari, ed in cui le derivate rispetto a $\partial_\mu \Phi$ coinvolgono $\sqrt{-g}$, mentre la derivata in Φ non coinvolge $\sqrt{-g}$. In metrica piatta, questa si riduce proprio all'equazione di Eulero-Lagrange classica.

Supponiamo ora una trasformazione generale di coordinate. La trasformazione indotta sui campi induce una variazione sull'azione, ma anche sulla metrica. Cioé si deve considerare anche un termine:

$$\delta S\Big|_{\delta g^{\mu\nu}}$$

Per trasformazioni infinitesime $x'_\mu = x_\mu + \xi_\mu$, la trasformazione indotta sulla metrica $\delta g^{\mu\nu}$ è data da:

$$\begin{aligned} \delta S\Big|_{\delta g^{\mu\nu}} &= \frac{1}{c} \int \left[\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial g^{\mu\nu}} \delta g^{\mu\nu} + \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\alpha g^{\mu\nu}} \delta \partial_\alpha g^{\mu\nu} \right] d\Omega = \\ &= \int \left[\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial g^{\mu\nu}} - \partial_\alpha \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\alpha g^{\mu\nu}} \right] \delta g^{\mu\nu} \end{aligned}$$

avendo, come prima, integrato per parti il II° membro e posto nulle le variazioni di $g^{\mu\nu}$ al bordo. In realtà, $g^{\mu\nu}$ ha 10 componenti e con questo procedimento se ne variano solo 4, pertanto la variazione non è completamente arbitraria.

L'annullamento del termine sotto integrale permette di definire una generalizzazione del *Tensore Energia-Impulso* $T_{\mu\nu}$:

$$\underbrace{\frac{1}{2}\sqrt{-g}T_{\mu\nu} \equiv -\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial g^{\mu\nu}} + \partial_\alpha \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\, \partial_\alpha g^{\mu\nu}}}_{\textit{definizione}} \tag{14.1}$$

Il calcolo del tensore energia-impulso della lagrangiana elettromagnetica sarà effettuato nel capitolo seguente.

Capitolo 15

La Geometria dello Spazio-Tempo

15.1 Variazioni della metrica. Equazioni di Killing

Si consideri una variazione infinitesima $x'^{\mu} = x^{\mu} + \xi^{\mu}$. Questa induce, come abbiamo visto, una variazione sulla metrica:

$$\delta S = \delta S\Big|_{\delta\Phi} + \delta S\Big|_{\delta g^{\mu\nu}}$$

in termini del tensore $T_{\mu\nu}$ definito nel capitolo precedente, la variazione indotta sulla metrica si può esprimere in questo modo:

$$\delta S = -\frac{1}{2c}\int \sqrt{-g}\; T_{\mu\nu}\delta g^{\mu\nu} d\Omega$$

e le variazioni delle componenti di $g^{\mu\nu}$ non sono tutte indipendenti. Per una variazione generica di queste, vale:

$$\begin{aligned} g'^{\mu\nu}(x') &= \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}} g^{\alpha\beta}(x) = \\ &= \left(\delta^{\mu}_{\alpha} + \frac{\partial \xi^{\mu}}{\partial x^{\alpha}}\right)\left(\delta^{\nu}_{\beta} + \frac{\partial \xi^{\nu}}{\partial x^{\beta}}\right) g^{\alpha\beta}(x) \end{aligned}$$

sviluppando il prodotto e fermandosi al I° ordine in ξ:

$$g'^{\mu\nu}(x') = g^{\mu\nu}(x) + \frac{\partial \xi^{\mu}}{\partial x^{\alpha}} g^{\alpha\nu} + \frac{\partial \xi^{\nu}}{\partial x^{\beta}} g^{\mu\beta} + O(2)$$

siccome vale:

$$\delta g^{\mu\nu}(x) = g'^{\mu\nu}(x) - g^{\mu\nu}(x)$$

segue:

$$g'^{\mu\nu}(x) + \frac{\partial g'^{\mu\nu}}{\partial x^{\alpha}}\xi^{\alpha} = g^{\mu\nu}(x) + \frac{\partial \xi^{\mu}}{\partial x^{\alpha}} g^{\alpha\nu} + \frac{\partial \xi^{\nu}}{\partial x^{\beta}} g^{\mu\beta}$$

se si fa ora l'assunzione che:[1]

$$\frac{\partial g'^{\mu\nu}}{\partial x^{\alpha}} \simeq \frac{\partial g^{\mu\nu}}{\partial x^{\alpha}}$$

si ricava:

$$\delta g^{\mu\nu}(x) = -\frac{\partial g^{\mu\nu}}{\partial x^{\alpha}}\xi^{\alpha} + \frac{\partial \xi^{\mu}}{\partial x^{\alpha}} g^{\alpha\nu} + \frac{\partial \xi^{\nu}}{\partial x^{\beta}} g^{\mu\beta} \tag{15.1}$$

ovvero che:

$$\delta g^{\mu\nu}(x) = \xi^{\mu;\nu} + \xi^{\nu;\mu}$$

[1] Questa assunzione è ragionevole. Intuitivamente, significa dire che la "rapidità" con cui varia la metrica passando ad un punto vicino non è sostanzialmente alterata dal cambiamento di coordinate. In parole ancora più semplici, si assume che se la metrica cambia poco passando dal punto x_1 al punto x_2 e molto passando al punto x_3, questo comportamento non è alterato da una trasformazione di coordinate.

Infatti:

$$\xi^{\mu;\nu} = g^{\nu\alpha}\partial_\alpha\xi^\mu + g^{\nu\alpha}\Gamma^\mu_{\alpha\beta}\xi^\beta$$
$$\xi^{\nu;\mu} = g^{\mu\alpha}\partial_\alpha\xi^\nu + g^{\nu\alpha}\Gamma^\nu_{\alpha\beta}\xi^\beta$$

che sommate fra di loro danno un termine di connessione:

$$\xi^{\mu;\nu} + \xi^{\nu;\mu} = g^{\nu\alpha}\partial_\alpha\xi^\mu + g^{\mu\alpha}\partial_\alpha\xi^\nu + g^{\nu\alpha}\Gamma^\mu_{\alpha\beta}\xi^\beta + g^{\mu\alpha}\Gamma^\nu_{\alpha\beta}\xi^\beta$$

Il primo ed il secondo termine a destra corrispondono proprio al secondo e terzo termine a destra della (15.1). Consideriamo quindi il termine di connessione $g^{\nu\alpha}\Gamma^\mu_{\alpha\beta}\xi^\beta + g^{\mu\alpha}\Gamma^\nu_{\alpha\beta}\xi^\beta$. Siccome vale la relazione (12.2) che lega la connessione alla metrica:

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2}g^{\mu\sigma}\left(\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta}\right)$$
$$\Gamma^\nu_{\alpha\beta} = \frac{1}{2}g^{\nu\sigma}\left(\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta}\right)$$

sostituendo nel termine di connessione si ricava:

$$g^{\nu\alpha}\Gamma^\mu_{\alpha\beta}\xi^\beta + g^{\mu\alpha}\Gamma^\nu_{\alpha\beta}\xi^\beta = \left[\frac{1}{2}g^{\nu\alpha}g^{\mu\sigma}\left(\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta}\right) + \frac{1}{2}g^{\mu\alpha}g^{\nu\sigma}\left(\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta}\right)\right]\xi^\beta$$

questa quantità è simmetrica negli indici muti α e σ, scambiando pertanto i due indici il termine si riduce a:

$$g^{\nu\alpha}\Gamma^\mu_{\alpha\beta}\xi^\beta + g^{\mu\alpha}\Gamma^\nu_{\alpha\beta}\xi^\beta = g^{\mu\alpha}g^{\nu\sigma}\partial_\beta g_{\alpha\sigma}\xi^\beta$$

Siccome ora vale $g^{\nu\sigma}g_{\alpha\sigma} = \delta^\nu_\sigma$, derivando si ricava la relazione:

$$\partial_\beta g^{\nu\sigma}g_{\sigma\alpha} = -g^{\nu\sigma}\partial_\beta g_{\alpha\sigma}$$

che permette di riscrivere il termine della connessione come:

$$g^{\mu\alpha}g^{\nu\sigma}\partial_\beta g_{\alpha\sigma}\xi^\beta = -g^{\mu\alpha}g_{\sigma\alpha}(\partial_\beta g^{\nu\sigma})\xi^\beta = -\delta^\mu_\sigma(\partial_\beta g^{\nu\sigma})\xi^\beta = -\partial_\beta g^{\mu\nu}\xi^\beta$$

e questo termine corrisponde al primo termine a destra della (15.1). Risultano pertanto dimostrate le cosiddette *Equazioni di Killing*:[2]

$$\boxed{\delta g^{\mu\nu} = \xi^{\mu;\nu} + \xi^{\nu;\mu}}$$

La variazione dell'azione assume allora la forma

$$\delta S = -\frac{1}{2c}\int \sqrt{-g}\; T_{\mu\nu}\delta(\xi^{\mu;\nu} + \xi^{\nu;\mu})d\Omega$$

Considerando ora che sia il tensore energia-impulso $T_{\mu\nu}$ che le equazioni di Killing sono simmetriche negli indici μ e ν, la variazione diventa:

$$\delta S = -\frac{1}{c}\int \sqrt{-g}\; T_{\mu\nu}\delta\xi^{\mu;\nu}d\Omega$$

[2]Vale la pena notare che le equazioni di Killing $\xi^{\mu;\nu} + \xi^{\nu;\mu} = 0$ definiscono le trasformazioni che lasciano inalterato il tensore metrico. In termini formali, dato un campo vettoriale X differenziabile, si dimostra che esiste sempre un sistema di coordinate locali, $(x^1, \ldots, x^n)$, definite nell'intorno di ogni punto, in cui una delle coordinate, per esempio x^1, è data dalle linee integrali di X. Nel caso in cui X soddisfi anche le equazioni di Killing nella regione in cui è definito il sistema di coordinate $(x^1, \ldots, x^n)$, allora la metrica espressa in tali coordinate ha coefficienti $g_{\mu\nu}$ che *non* dipendono da x^1.

che può ancora essere riscritta nella forma:

$$-\frac{1}{c}\int\sqrt{-g}\,[D^\nu(T_{\mu\nu}\xi^\mu)-(D^\nu T_{\mu\nu})\xi^\mu]\,d\Omega=$$
$$=-\frac{1}{c}\int\sqrt{-g}\frac{1}{\sqrt{-g}}\partial^\nu(\sqrt{-g}\,T_{\mu\nu}\xi^\mu)d\Omega+\frac{1}{c}\int\sqrt{-g}(D^\nu T_{\mu\nu})\xi^\mu d\Omega$$

Ora, l'integrale del termine $\partial^\nu(\sqrt{-g}\,T_{\mu\nu}\xi^\mu)$ può essere trasformato in un integrale di superficie, che si annulla in quanto si assume sempre che le variazioni ξ^μ si annullino ai bordi. Il δS risulta quindi essere:

$$\delta S=-\frac{1}{c}\int\sqrt{-g}(D^\nu T_{\mu\nu})\xi^\mu d\Omega$$

che si traduce nella condizione:

$$D^\nu T_{\mu\nu}=0$$

il che implica, come di solito, la presenza di cariche conservate.

15.2 Calcolo del Tensore Energia–Impulso

15.2.1 Campo scalare

Consideriamo ora la forma del tensore energia-impulso nel caso di un campo scalare. La densità di lagrangiana in questo caso è data da:

$$\mathcal{L}=\frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi+m^2\Phi^2 \qquad\rightarrow\qquad \mathcal{L}=\frac{1}{2}g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi+m^2\Phi^2$$

Il tensore energia impulso assume allora la forma:

$$T_{\mu\nu}=\frac{2}{\sqrt{-g}}\left[-\frac{\partial\sqrt{-g}}{\partial g^{\mu\nu}}\mathcal{L}-\sqrt{-g}\frac{\partial\mathcal{L}}{\partial g^{\mu\nu}}\right]$$

siccome vale:

$$\frac{\partial\sqrt{-g}}{\partial g^{\mu\nu}}=\frac{1}{2}\frac{-1}{\sqrt{-g}}\frac{\partial g}{\partial g^{\mu\nu}}=-\frac{1}{2\sqrt{-g}}g\,g_{\mu\nu}$$

insieme alle relazioni sui determinanti $\frac{G^{\nu\alpha}}{g}g^{\nu\alpha}$ e $g_{\mu\nu}G^{\mu\nu}=g$, si ricava:

$$T_{\mu\nu}=-\partial_\mu\Phi\partial_\nu\Phi-\frac{2}{\sqrt{-g}}\left(-\frac{1}{2\sqrt{-g}}g\,g_{\mu\nu}\right)\mathcal{L}$$

ovvero la forma:

$$T_{\mu\nu}=-\partial_\mu\Phi\partial_\nu\Phi+g_{\mu\nu}\mathcal{L}$$

che nel caso di metrica piatta ($g_{\mu\nu}\rightarrow\eta_{\mu\nu}$) si riduce alla nota relazione.

15.2.2 Campo elettromagnetico

Abbiamo visto che la densità di lagrangiana del campo elettromagnetico è espressa nella forma:

$$\mathcal{L}=-\frac{1}{16\pi}F_{\mu\nu}F_{\alpha\beta}g^{\mu\alpha}g^{\nu\beta}$$

Il tensore energia-impulso ha la forma:

$$T_{\mu\nu}=\frac{2}{\sqrt{-g}}\left[-\frac{\partial\sqrt{-g}}{\partial g^{\mu\nu}}\mathcal{L}-\sqrt{-g}\frac{\partial\mathcal{L}}{\partial g^{\mu\nu}}\right]$$

in questo caso vale:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial g^{\mu\nu}} &= -\frac{1}{16\pi}\left\{F_{\lambda\sigma}F_{\alpha\beta}\left[\left(\delta^{\lambda\mu}\delta^{\alpha\nu}+\delta^{\lambda\nu}\delta^{\alpha\mu}\right)g^{\sigma\beta}+\left(\delta^{\sigma\mu}\delta^{\beta\nu}+\delta^{\sigma\nu}\delta^{\beta\mu}\right)g^{\alpha\lambda}\right]\right\} = \\
&= -\frac{1}{16\pi}\left[F_{\mu\sigma}F_{\nu\beta}g^{\sigma\beta}+F_{\nu\sigma}F_{\mu\beta}g^{\sigma\beta}+F_{\lambda\mu}F_{\alpha\nu}g^{\alpha\lambda}+F_{\lambda\nu}F_{\alpha\mu}g^{\alpha\lambda}\right] = \\
&= -\frac{1}{16\pi}\left[F_\mu^\beta F_{\nu\beta}+F_\nu^\beta F_{\mu\beta}+F_\mu^\alpha F_{\alpha\nu}+F_\nu^\alpha F_{\alpha\mu}\right] = \\
&= -\frac{1}{8\pi}F_{\mu\alpha}F_\nu^\alpha
\end{aligned}
$$

(avendo rinominato β in α), ovvero in definitiva:

$$
T_{\mu\nu} = -g_{\mu\nu}\mathcal{L} + \frac{1}{4\pi}F_{\mu\alpha}F_\nu^\alpha
$$

che è l'espressione gauge-invariante già trovata e pertanto si ripropongono le stesse considerazioni sull'energia e sull'impulso.

15.3 Il tensore di curvatura

Si consideri ora una curva qualunque. Abbiamo già visto che un vettore per trasporto parallelo si trasforma secondo:

$$
\delta A_\mu = \Gamma_{\mu\nu}^\alpha A_\alpha dx^\nu \qquad\qquad dA_\mu = \frac{\partial A_\mu}{\partial x^\alpha}dx^\alpha
$$

Ricerchiamo quindi la condizione sotto la quale il trasporto del vettore non dipenda dal cammino particolare seguito, ovvero per cui risulti $dA_\mu = \delta A_\mu$.

In effetti, la variazione del vettore è data da (vedere nota a pag. 103):

$$
\Delta A_\mu = \oint \Gamma_{\mu\nu}^\alpha A_\alpha dx^\nu
$$

per cui la condizione equivale a dire che sussiste la relazione:

$$
\frac{\partial A_\mu}{\partial x^\nu} = -\Gamma_{\mu\nu}^\alpha A_\alpha
$$

Se tale relazione vale $\forall A_\alpha$ si dice che la connessione è integrabile. In particolare, per il vettore tangente ad una curva questa relazione vale sempre.

Le condizioni per cui la variazione del vettore non dipende dal percorso implicano che la connessione sia simmetrica:

$$
\begin{aligned}
\frac{\partial^2 A^\mu}{\partial x^\lambda \partial x^\nu} &= -\frac{\partial}{\partial x^\lambda}\Gamma_{\alpha\nu}^\mu A^\alpha \\
\frac{\partial^2 A^\mu}{\partial x^\nu \partial x^\lambda} &= -\frac{\partial}{\partial x^\nu}\Gamma_{\alpha\lambda}^\mu A^\alpha
\end{aligned}
$$

il che equivale a dire che sussiste la relazione:

$$
-\frac{\partial}{\partial x^\lambda}\left(\Gamma_{\alpha\nu}^\mu A^\alpha\right) + \frac{\partial}{\partial x^\nu}\left(\Gamma_{\alpha\lambda}^\mu A^\alpha\right) = 0
$$

che, esplicitando le derivate dei prodotti, diventa:

$$
-\frac{\partial}{\partial x^\lambda}\left(\Gamma_{\alpha\nu}^\mu\right)A^\alpha - \Gamma_{\alpha\nu}^\mu\frac{\partial}{\partial x^\lambda}A^\alpha + \frac{\partial}{\partial x^\nu}\left(\Gamma_{\alpha\lambda}^\mu\right)A^\alpha + \Gamma_{\alpha\lambda}^\mu\frac{\partial}{\partial x^\nu}A^\alpha = 0
$$

la quale, ricordando la relazione $\frac{\partial}{\partial x^\nu}A^\alpha = -\Gamma_{\alpha\lambda}^\mu A^\alpha$ e l'analoga ottenuta scambiando $\lambda \leftrightarrow \nu$, diventa ancora:

$$
\left[\left(-\partial_\lambda\Gamma_{\alpha\nu}^\mu + \partial_\nu\Gamma_{\alpha\lambda}^\mu\right)A^\alpha + \Gamma_{\alpha\nu}^\mu\Gamma_{\beta\lambda}^\alpha A^\beta - \Gamma_{\alpha\lambda}^\mu\Gamma_{\beta\nu}^\alpha A^\beta\right] = 0
$$

che rinominando l'indice muto β in α si può riscrivere finalmente come:

$$\left[-\partial_\lambda \Gamma^\mu_{\alpha\nu} + \partial_\nu \Gamma^\mu_{\alpha\lambda} + \Gamma^\mu_{\beta\nu}\Gamma^\beta_{\alpha\lambda} - \Gamma^\mu_{\beta\lambda}\Gamma^\beta_{\alpha\nu}\right] A^\alpha = 0$$

Definendo ora la quantità:

$$\boxed{R^\mu_{\alpha\lambda\nu} \equiv -\partial_\lambda \Gamma^\mu_{\alpha\nu} + \partial_\nu \Gamma^\mu_{\alpha\lambda} + \Gamma^\mu_{\beta\nu}\Gamma^\beta_{\alpha\lambda} - \Gamma^\mu_{\beta\lambda}\Gamma^\beta_{\alpha\nu}}$$

la condizione ricercata si esprime come:

$$R^\mu_{\alpha\lambda\nu} A^\alpha = 0 \qquad \forall A^\alpha$$

e dovendo valere, appunto, per ogni A^α, deve risultare:

$$\boxed{R^\mu_{\alpha\lambda\nu} = 0}$$

La quantità $R^\mu_{\alpha\lambda\nu} A^\alpha$ trasforma come un vettore a 4 indici, pertanto ha $4^4 = 256$ componenti; tuttavia relazioni di simmetria e antisimmetria le riducono a sole 20 indipendenti. Il tensore $R^\mu_{\alpha\lambda\nu}$ è estremamente importante e prende il nome di *Tensore di Curvatura*.

Dunque, *se il tensore di curvatura è nullo, la connessione è integrabile.*

Supponiamo quindi di essere in uno spazio a tensore di curvatura nullo. Sia $A^{(a)}_\mu$ un vettore, e scegliamo una condizione di ortonormalità:

$$|A^{(4)}_\mu A^{\mu(4)}| = -1 \qquad\qquad |A^{(i)}_\mu A^{\mu(i)}| = 1 \qquad \text{(i=1,2,3)}$$

allora, siccome vale $\partial_\nu A^{(a)}_\mu = \Gamma^\rho_{\nu\mu} A^{(a)}_\rho$, ed essendo la connessione simmetrica:

$$\partial_\nu A^{(a)}_\mu = \Gamma^\rho_{\nu\mu} A^{(a)}_\rho = \Gamma^\rho_{\mu\nu} A^{(a)}_\rho = \partial_\mu A^{(a)}_\nu$$

vale la relazione:

$$\partial_\nu A^{(a)}_\mu - \partial_\mu A^{(a)}_\nu = 0$$

ovvero, il rotore generalizzato di $A^{(a)}_\mu$ è nullo. Questo implica che il vettore si possa scrivere come il gradiente di una quantità scalare φ:

$$A^{(a)}_\mu = \partial_\mu \varphi$$

Supponiamo ora che $R^\mu_{\alpha\lambda\nu} \neq 0$. Il ΔA_μ rappresenta, come abbiamo visto, la differenza di due vettori trasportati parallelamente nello stesso punto. Valutiamo ora questa quantità.

$$\Delta A_\mu = \oint \Gamma^\alpha_{\mu\nu} A_\alpha dx^\nu$$

Il teorema di Stokes permette di passare da un integrale circuitale ad uno di superficie:

$$\oint A_\mu dx^\mu = \int_\Sigma \frac{\partial A_\mu}{\partial dx^\nu} dS^{\mu\nu} = \frac{1}{2}\int_\Sigma \left(\frac{\partial A_\mu}{\partial dx^\nu} - \frac{\partial A_\nu}{\partial dx^\mu}\right) dS^{\mu\nu}$$

che nel caso in esame diventa:

$$\Delta A_\mu = \frac{1}{2}\int_\Sigma \left[\frac{\partial}{\partial x^\beta}\left(\Gamma^\alpha_{\mu\nu} A_\alpha\right) - \frac{\partial}{\partial x^\nu}\left(\Gamma^\alpha_{\mu\beta} A_\alpha\right)\right] dS^{\beta\nu}$$

Ma *in* una superficie il trasporto parallelo non è ben definito, a meno di considerare superfici infinitesime (e trascurando infinitesimi di ordine superiore):

$$\begin{aligned}\Delta A_\mu &= \frac{1}{2}\left[\frac{\partial}{\partial x^\beta}\left(\Gamma^\alpha_{\mu\nu} A_\alpha\right) - \frac{\partial}{\partial x^\nu}\left(\Gamma^\alpha_{\mu\beta} A_\alpha\right)\right] \Delta S^{\beta\nu} = \\ &= \frac{1}{2}\left\{\left[\partial_\beta \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\mu\beta}\right] A_\alpha + \Gamma^\alpha_{\mu\nu}\partial_\beta A_\alpha - \Gamma^\alpha_{\mu\beta}\partial_\nu A_\alpha\right\} \Delta S^{\beta\nu}\end{aligned}$$

Per i termini nella derivata di A_α, essendo dS infinitesimo, si può prendere la derivata sul bordo ottenendo:

$$\boxed{\Delta A_\mu = \frac{1}{2} R^\alpha_{\mu\nu\beta} A_\alpha dS^{\nu\beta}}$$

che permette di concludere che *un vettore trasportato parallelamente a sé stesso non varia se e solo se il tensore di curvatura dello spazio considerato è nullo.*

In uno spazio quadridimensionale piatto il tensore di curvatura è evidentemente nullo. Infatti, in questo tipo di spazio è possibile scegliere un sistema di riferimento in cui la connessione si annulla e quindi il tensore di curvatura è nullo. Il carattere tensoriale di $R^\mu_{\alpha\lambda\nu}$ assicura quindi che questo tensore sia nullo in qualsiasi altro sistema di ccordinate.

È possibile anche dimostrare l'inverso: se il tensore di curvatura $R^\mu_{\alpha\lambda\nu}$ è nullo, allora lo spazio quadridimensionale è piatto. Infatti, in ogni spazio si può scegliere un sistema galileiano di coordinate in una regione infinitesima. Se vale $R^\mu_{\alpha\lambda\nu} = 0$, allora per quanto appena visto il trasporto parallelo è un'operazione univoca e trasportando questo sistema galileiano dalla regione infinitesima considerata in tutte le altre regioni si può costruire un sistema galileiano in tutto lo spazio, il che prova l'asserzione fatta.

L'annullarsi del tensore di curvatura $R^\mu_{\alpha\lambda\nu}$ è quindi condizione necessaria e sufficiente perché uno spazio quadridimensionale sia piatto.

È da notare, tuttavia, che anche se in determinato punto di uno spazio curvo è possibile scegliere localmente un sistema galileiano, il tensore di curvatura non si annulla nel punto dato, perché anche se le $\Gamma^\alpha_{\mu\nu}$ si annullano le loro derivate sono diverse da zero.

15.4 Proprietà del tensore di curvatura

Consideriamo ora alcune proprietà del tensore di curvatura $R^\mu_{\nu\alpha\beta}$. Iniziamo con le proprietà di simmetria e antisimmetria.

Abbassiamo l'indice controvariante tramite il tensore metrico:

$$g_{\mu\rho} R^\mu_{\nu\alpha\beta} = R_{\rho\nu\alpha\beta}$$

nel caso particolare[3] il tensore metrico coincida con la metrica piatta di Minkowski $g_{\mu\rho} = \eta_{\mu\rho}$:

$$R_{\rho\nu\alpha\beta} = \eta_{\rho\mu} \left[\partial_\alpha \Gamma^\mu_{\nu\beta} - \partial_\beta \Gamma^\mu_{\nu\alpha} \right]$$

essendo $\Gamma^\mu_{\nu\beta} = \frac{1}{2} g^{\mu\sigma} \left(\partial_\nu g_{\sigma\beta} + \partial_\beta g_{\nu\sigma} - \partial_\sigma g_{\beta\nu} \right)$, risulta che le derivate sono date da:

$$\begin{aligned}
\partial_\alpha \Gamma^\mu_{\nu\alpha} &= \frac{1}{2} \eta^{\mu\sigma} \left(\partial_\alpha \partial_\nu g_{\sigma\beta} + \partial_\alpha \partial_\beta g_{\nu\sigma} - \partial_\alpha \partial_\sigma g_{\beta\nu} \right) \\
\partial_\beta \Gamma^\mu_{\nu\alpha} &= \frac{1}{2} \eta^{\mu\sigma} \left(\partial_\beta \partial_\nu g_{\sigma\alpha} + \partial_\beta \partial_\alpha g_{\nu\sigma} - \partial_\beta \partial_\sigma g_{\alpha\nu} \right)
\end{aligned}$$

che sottratte fra di loro permettono di riscrivere il tensore di curvatura come:

$$\begin{aligned}
R_{\rho\nu\alpha\beta} &= \eta_{\rho\mu} \frac{1}{2} \eta^{\mu\sigma} \left(\partial_\alpha \partial_\nu g_{\sigma\beta} + \partial_\beta \partial_\sigma g_{\alpha\nu} - \partial_\alpha \partial_\sigma g_{\beta\nu} - \partial_\beta \partial_\nu g_{\sigma\alpha} \right) \\
&= \frac{1}{2} \left(\partial_\alpha \partial_\nu g_{\rho\beta} + \partial_\beta \partial_\rho g_{\alpha\nu} - \partial_\alpha \partial_\rho g_{\beta\nu} - \partial_\beta \partial_\nu g_{\rho\alpha} \right)
\end{aligned}$$

dove nell'ultimo passaggio si è effettuata la contrazione dell'indice μ e applicata la δ^σ_ρ risultante. Da questa relazione si deduce che per gli scambi:

- $\rho \leftrightarrow \nu$ $\quad R_{\rho\nu\alpha\beta}$ è antisimmetrico
- $\alpha \leftrightarrow \beta$ $\quad R_{\rho\nu\alpha\beta}$ è antisimmetrico

[3]Caso particolare scelto solo per comodità di calcolo, non ha naturalmente alcuna influenza sul risultato.

- $(\rho\nu) \leftrightarrow (\alpha\beta)$ $\quad R_{\rho\nu\alpha\beta}$ è simmetrico

una volta che sia stato abbassato l'indice controvariante, gli indici di questo tensore vanno quindi pensati a coppie.

Vale la relazione:

$$R^{\mu}_{\alpha\lambda\nu} + R^{\mu}_{\lambda\nu\alpha} + R^{\mu}_{\nu\alpha\lambda} = 0$$

Si consideri ora la derivata seconda $D_\alpha D_\beta A^\mu \equiv A^\mu_{;\alpha;\beta}$ e valutiamo il termine $D_\alpha D_\beta A^\mu - D_\beta D_\alpha A^\mu$. In un sistema localmente pseudoeuclideo vale $\Gamma(x) = 0$, ma in genere vale ancora $R^{\mu}_{\alpha\lambda\nu} \neq 0$. Ora $A^\mu_{;\alpha} = \partial_\alpha A^\mu + \Gamma^\mu_{\alpha\beta} A^\beta$, quindi:

$$\begin{aligned} D_\beta \left(D_\alpha A^\mu\right) &= D_\beta \left(\partial_\alpha A^\mu + \Gamma^\mu_{\alpha\beta} A^\beta\right) = \\ &= \partial_\beta \partial_\alpha A^\mu + \partial_\beta \Gamma^\mu_{\alpha\nu} A^\nu + \textit{altri termini in } \Gamma \end{aligned}$$

dove gli altri termini in Γ sono nulli.

Si mostra facilmente che il termine $D_\beta D_\alpha A^\mu$ ha la stessa forma con α e β scambiati, pertanto:

$$[D_\alpha, D_\beta] A^\mu = -R^\mu_{\alpha\beta\nu} A^\nu$$

Vale la cosiddetta *Identità di Bianchi*:

$$\boxed{R^\mu_{\nu\alpha\beta;\rho} + R^\mu_{\nu\beta\rho;\alpha} + R^\mu_{\nu\rho\alpha;\alpha} = 0}$$

Infatti ponendoci, ad esempio, in sistemi di riferimento in cui la connessione sia nulla, vale:

$$R^\mu_{\nu\alpha\beta;\rho} = \partial_\rho R^\mu_{\nu\alpha\beta} = -\partial_\rho \partial_\alpha \Gamma^\mu_{\nu\beta} + \partial_\rho \partial_\beta \Gamma^\mu_{\alpha\nu}$$

e analogamente per gli altri termini che, tenendo conto delle varie relazioni di simmetria/antisimmetria, si annullano.

Notiamo incidentalmente che nel tensore metrico espresso in termini della metrica compaiono le derivate seconde di quest'ultima.[4]

Per contrazione successiva degli indici del tensore metrico (eventualmente innalzando un indice covariante), si ottengono i seguenti importanti tensori. Saturando l'indice controvariante:

$$R^\mu_{\nu\alpha\mu} = R_{\nu\alpha} \qquad \textbf{Tensore di Ricci} \text{ (simmetrico)}$$

saturando il tensore di Ricci:

$$R^\nu_\nu = R \qquad \textbf{Curvatura Scalare}$$

La curvatura scalare per superfici bidimensionali coincide con la curvatura gaussiana.[5]

Applichiamo ora la contrazione successiva degli indici alla identità di Bianchi (notiamo preliminarmente che $R^\mu_{\nu\alpha\mu;\rho} \equiv g^\beta_\mu D_\rho R^\mu_{\nu\alpha\beta}$):

$$R^\mu_{\nu\alpha\mu;\rho} + R^\mu_{\nu\mu\rho;\alpha} + R^\mu_{\nu\rho\alpha;\mu} = 0 \quad \rightarrow \quad R_{\nu\alpha;\rho} + R_{\nu\rho;\alpha} + R^\mu_{\nu\rho\alpha;\mu} = 0$$

contraendo ancora gli indici α e ν:[6]

$$R^\alpha_{\alpha;\rho} - R^\alpha_{\rho;\alpha} + R^{\mu\alpha}_{\rho\alpha;\mu} = 0 \quad \rightarrow \quad R_{;\rho} - R^\alpha_{\rho;\alpha} - R^\mu_{\rho;\mu} = 0$$

quindi:

$$R_{;\rho} - 2R^\alpha_{\rho;\alpha} = 0 \quad \rightarrow \quad D_\alpha\left[\delta^\alpha_\rho R - 2R^\alpha_\rho\right] = 0$$

[4] Infatti il tensore di curvatura dipende dalle derivate prima della connessione, che è legata alle derivate prime del tensore metrico.

[5] In altri termini, coincide con il concetto classico di curvatura.

[6] Infatti vale $R^{\mu\alpha}_{\rho\alpha;\mu} = -R^{\mu\alpha}_{\alpha\rho;\mu} = -R^\mu_{\rho;\mu}$.

il che permette di definire l'importantissimo tensore:

$$\boxed{G^{\alpha}_{\rho} \equiv R^{\alpha}_{\rho} - \frac{1}{2}\delta^{\alpha}_{\rho} R \qquad \textbf{Tensore di Einstein}}$$

che risulta pertanto a derivata covariante nulla:

$$\boxed{D_\alpha G^{\alpha}_{\rho} = 0}$$

Scrivendo la curvatura scalare in termini del tensore metrico (e cioé utilizzando il tensore metrico per saturare gli indici):

$$R = g^{\mu\nu} g^{\lambda\rho} R_{\mu\nu\lambda\rho}$$

si ricava la relazione:[7]

$$R = \left[2g^{11}g^{22} - 2g^{12}g^{12}\right] R_{1212} = 2 \det g^{\mu\nu} R_{1212} = \frac{2}{g} R_{1212}$$

Ne consegue allora:

$$R_{\alpha\beta\mu\nu} = \left[g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\nu}g_{\beta\mu}\right] \frac{R_{1212}}{g}$$

ed essendo $r_{1212} = \frac{Rg}{2}$:

$$R_{\alpha\beta\mu\nu} = \frac{1}{2} R \left[g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\nu}g_{\beta\mu}\right]$$

e se si considera una sfera di raggio R, questa risulta proprio essere la curvatura gaussiana.

Vediamo finalmente l'espressione del tensore di curvatura in funzione della metrica.

Siccome vale la (12.2):

$$\Gamma^{\rho}_{\mu\alpha} = \frac{1}{2} g^{\rho\sigma} \left[g_{\sigma\alpha,\mu} + g_{\mu\sigma,\alpha} - g_{\mu\alpha,\sigma}\right]$$

si può ricavare da questa direttamente il termine nelle derivate della connessione. Ancora, dall'espressione della connessione (12.2):

$$\Gamma^{\rho}_{\mu\alpha} g_{\rho\beta} = \frac{1}{2} g_{\rho\beta} g^{\rho\sigma} \left[g_{\sigma\alpha,\mu} + g_{\mu\sigma,\alpha} - g_{\mu\alpha,\sigma}\right] = \frac{1}{2} \left[g_{\beta\alpha,\mu} + g_{\mu\beta,\alpha} - g_{\mu\alpha,\beta}\right]$$

mutando gli indici:

$$\Gamma^{\rho}_{\mu\beta} g_{\rho\alpha} = \frac{1}{2} \left[g_{\beta\alpha,\mu} + g_{\mu\alpha,\beta} - g_{\mu\beta,\alpha}\right]$$

e quindi, sommando:

$$\partial_\mu g_{\alpha\beta} = \Gamma^{\rho}_{\mu\alpha} g_{\rho\beta} + \Gamma^{\rho}_{\mu\beta} g_{\rho\alpha}$$

In definitiva il tensore di curvatura in termini del tensore metrico assume la forma:

$$R_{\lambda\mu\nu\alpha} = \frac{1}{2} \left[\frac{\partial^2 g_{\lambda\nu}}{\partial x^\alpha \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\alpha \partial x^\lambda} + \frac{\partial^2 g_{\mu\alpha}}{\partial x^\nu \partial x^\lambda} - \frac{\partial^2 g_{\lambda\alpha}}{\partial x^\mu \partial x^\nu}\right] + g_{\beta\sigma} \left[\Gamma^{\beta}_{\nu\lambda}\Gamma^{\sigma}_{\mu\alpha} - \Gamma^{\beta}_{\alpha\lambda}\Gamma^{\sigma}_{\mu\nu}\right]$$

e quindi compaiono le derivate seconde del tensore metrico in modo lineare e le derivate prime in modo quadratico. Questo permetterà di scrivere l'azione come uno scalare invariante (curvatura scalare).

[7] Si ricordi che vale $g_{\mu\nu} g^{\nu\alpha} = \delta^{\alpha}_{\mu}$ e $g \det g^{\nu\alpha} = 1 \Rightarrow \det g^{\nu\alpha} = \frac{1}{g}$.

15.5 Le equazioni del moto

Ricerchiamo ora finalmente l'equazione del moto. Nel caso classico, l'equazione deve ridursi alle:

$$\nabla^2\Phi = 4\pi G\varrho \qquad \Phi = -\int \frac{\varrho(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}' \qquad \varrho(\vec{x}) = \sum_i m_i \delta^{(3)}(\vec{x}-\vec{x}')$$

Si è già notato (in §12.2, pag.110) che $g_{00} = -c^2\left(1+\frac{2\Phi}{c^2}\right)$. Ora si richiede che:

$$\mathcal{L} = \mathcal{L}_0 + \frac{GmM}{r}$$

dove $\mathcal{L}_0$ è la lagrangiana libera. Inoltre, nel caso di metrica pseudoeuclidea ($g_{\alpha\beta} \equiv \eta_{\alpha\beta}$), deve essere riprodotta un'azione del tipo:

$$S = -mc\int dS = -mc^2\int_0^{t_1}\sqrt{1-\beta^2}dt \approx \int_0^{t_1}\left[-mc^2+\frac{1}{2}mv^2\right]dt$$

Ovviamente, nel caso risulti $g_{\alpha\beta} \simeq \eta_{\alpha\beta}$ l'azione sarà del tipo:

$$S \simeq -mc\int\sqrt{c^2\left(1+\frac{2\Phi}{c^2}\right)dt^2 - dr^2}$$

e nel caso generale:

$$S = -mc\int\sqrt{g_{\mu\nu}dx^\mu dx^\nu}$$

in cui compare un termine $\nabla^2 g_{00}$.

Per particelle puntiformi valeva la relazione:

$$p^\alpha = mcu^\alpha \equiv \left(\vec{p}, \frac{iE}{c}\right) \qquad \vec{p} = \frac{E\vec{v}}{c^2}$$

se invece si ha a che fare con densità, si dovrà prendere in considerazione una densità di impulso $\pi^\alpha = \mu c u^\alpha$. Il tensore energia-impulso per particelle puntiformi permette di scrivere:

$$T_{\alpha\beta} \Rightarrow \begin{cases} p_j &= \int_V T_{0j}d\vec{x} \\ p_0 &= \int_V T_{00}d\vec{x} \end{cases}$$

o analogamente:

$$p_j = \frac{i}{c}\int_{S^\alpha} T_{\alpha j}dS^\alpha \qquad p_0 = \frac{i}{c}\int_{S^\alpha} T_{\alpha 0}dS^\alpha$$

Ora, dalla densità di impulso (in metrica pseudoeuclidea) $T_{4\alpha} = -\frac{i}{c}\mu c u^\alpha$ e con un discorso analogo a quello fatto per la carica, ovvero:

$$de\cdot dx^\mu = \varrho d\vec{x}\frac{dx^\mu}{dt}dt = \varrho dV\frac{dx^\mu}{dt}$$

(dalla quale si ricava fra l'altro l'espressione della corrente di carica $j_\mu = \varrho\frac{dx^\mu}{dt} \equiv (\varrho\vec{v}, ic\varrho)$), se ne deduce che la densità di massa è la quarta componente di un quadrivettore:

$$\mu = \frac{1}{ic}j_0 \qquad j_\mu = \mu\frac{dx^\mu}{dt}$$

Si può quindi in genere definire il tensore:

$$T_{\beta\alpha} = c\mu\frac{dx^\beta}{dt}\frac{dx^\alpha}{ds} = \mu c\frac{dx^\beta}{ds}\frac{dx^\alpha}{ds}\frac{ds}{dt}$$

in seguito al quale l'equazione del moto assume la forma:

$$T_{00} = \mu c \frac{dct}{dt} \cdot \frac{d(ct)}{ct\sqrt{1-\beta^2}} \qquad (= \mu c^2 \text{ se } \beta \ll 1)$$

il che permette di ricavare la relazione:

$$c^2 \nabla^2 g^{00} = c^2 4\pi G T^{00}$$

L'azione richiesta deve quindi essere lineare nelle derivate seconde e simmetrica negli indici, ovvero: $\nabla^2 g^{\mu\nu} \propto T^{\mu\nu}$. Essa sarà pertanto del tipo:

$$S = \int G\sqrt{-g}\, dx^4 + S(campi\ in\ studio)$$

Dalla definizione del tensore energia-impulso (14.1), la lagrangiana del punto materiale sarà data da:

$$\mathcal{L} = -mc\frac{\sqrt{-g_{\mu\nu}dx^\mu dx^\nu}}{dt}$$

che, una volta integrata, restituisce effettivamente l'azione.

Ora, possiamo ricavare Φ dalla già ricordata relazione:

$$g_{00} = -\left(1 + \frac{2\Phi}{c^2}\right) \Rightarrow \Phi = -\frac{1}{2}c^2(1 + g_{00})$$

il che permette di scrivere:

$$-\frac{1}{2}c^2\nabla^2 g_{00} = 4\pi G\frac{T_{00}}{c^2}$$

e cioé:

$$\nabla^2 g_{00} = -\frac{8\pi G}{c^2}T_{00}$$

da cui l'equazione generale:

$$\nabla^2 g_{\mu\nu} = -\frac{8\pi G}{c^2}T_{\mu\nu} \tag{15.2}$$

L'azione sarà quindi l'integrale di una lagrangiana gravitazionale e di una lagrangiana di materia:

$$S = k\int \mathcal{L}_G\sqrt{-g}\, d^4x + \frac{1}{c}\int \mathcal{L}_M\sqrt{-g}\, d^4x$$

L'equazione (15.2) sopra scritta è alle derivate seconde nella metrica, è simmetrica negli indici ed il fatto che $T_{\mu\nu}$ sia a divergenza covariante nulla, implica che al primo membro debba esserci un tensore a divergenza covariante nulla.

Il termine dell'azione gravitazionale richiede che la lagrangiana gravitazionale sia uno scalare nelle derivate della metrica: un tale oggetto è stato incontrato in precedenza ed è la curvatura scalare, che fra l'altro comprende già le derivate seconde della metrica.

Poniamo quindi $\mathcal{L}_G = R$. Le derivate seconde della metrica vi compaiono in maniera lineare, si può pertanto integrare su un volume per parti ritornando alle derivate prime. Sia dunque:

$$S_G = k\int R\sqrt{-g}\, d^4x$$

e calcoliamo le variazioni dell'azione rispetto alla metrica $\delta g_{\mu\nu}$:

$$\begin{aligned} dS_G &= k\delta\left[\int g^{\alpha\beta}R_{\alpha\beta}\sqrt{-g}d\Omega\right] = \\ &= k\int_\Omega \left[\delta g^{\alpha\beta}R_{\alpha\beta}\sqrt{-g} + g^{\alpha\beta}\delta R_{\alpha\beta}\sqrt{-g} + g^{\alpha\beta}R_{\alpha\beta}\delta\sqrt{-g}\right] d\Omega \end{aligned}$$

Ora, siccome vale:

$$d\sqrt{g} = \frac{1}{2}\sqrt{g}\; g^{\mu\nu} dg_{\mu\nu} \qquad\qquad g^{\mu\nu} g_{\nu\alpha} = \delta^{\mu}_{\alpha}$$

il terzo membro $\delta\sqrt{-g}$ si può riscrivere come:

$$\delta\sqrt{-g} = \frac{\partial\sqrt{-g}}{\partial g_{\mu\nu}}\delta g_{\mu\nu} = \frac{1}{2}\sqrt{g}\; g^{\mu\nu}$$

mentre il termine $g^{\alpha\beta}\delta R_{\alpha\beta}\sqrt{-g}$ dà contributo nullo. Verifichiamo quest'ultima affermazione.

In un sistema galileiano, dove la connessione e le derivate prime della metrica sono nulle, vale:

$$g^{\mu\nu}\delta R_{\mu\nu} = g^{\mu\nu}\delta\left(\partial_\mu \Gamma^\alpha_{\alpha\nu} - \partial_\alpha \Gamma^\alpha_{\mu\nu}\right) = g^{\mu\nu}\left(\partial_\mu \delta\Gamma^\alpha_{\alpha\nu} - \partial_\alpha \delta\Gamma^\alpha_{\mu\nu}\right)$$

La variazione della connessione è costituita da un tensore a tre indici. Essendo la derivata del tensore metrico nulla si può portarla all'interno della derivata:

$$\partial_\mu\left(g^{\mu\nu}\delta\Gamma^\alpha_{\alpha\nu}\right) - \partial_\alpha\left(g^{\mu\nu}\delta\Gamma^\alpha_{\mu\nu}\right) = \partial_\mu\left[g^{\mu\nu}\delta\Gamma^\alpha_{\alpha\nu} - g^{\alpha\nu}\delta\Gamma^\alpha_{\alpha\nu}\right] \equiv \partial_\mu W^\mu$$

dove W^μ ha un solo indice controvariante libero. Dunque $g^{\mu\nu}\delta R_{\mu\nu} = \partial_\mu W^\mu$ e W^μ è un vettore. In un sistema di riferimento generale si avrebbe allora:

$$g^{\mu\nu}\delta R_{\mu\nu} = D_\mu W^\mu$$

con W^μ dato ancora da:

$$W^\mu = g^{\mu\nu}\delta\Gamma^\alpha_{\alpha\nu} - g^{\alpha\nu}\delta\Gamma^\alpha_{\alpha\nu}$$

D'altra parte:

$$W^\mu_{;\mu} = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}W^\mu)$$

per cui:

$$k\int_\Omega g^{\mu\nu}\delta R_{\mu\nu}\sqrt{-g}d\Omega = k\int_\Omega \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}W^\mu)\sqrt{-g}d\Omega$$

che, in virtù del teorema di Gauss, equivale al flusso di un vettore costituito da variazioni della connessione. Siccome le variazioni si annullano all'infinito, l'integrale risulta nullo.

Vale la pena di sottolineare che questo è anche l'unico pezzo che contiene le derivate seconde della metrica.

In definitiva, quindi la variazione dell'azione assume la forma finale:

$$dS_G = k\int_\Omega\left[-R\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu} + R_{\mu\nu}\sqrt{-g}\delta g^{\mu\nu}\right]d\Omega = k\int_\Omega\left[R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right]d\Omega$$

Per il campo in assenza di sorgenti (termine dell'azione dipendente dal campo gravitazionale) deve valere allora:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0$$

Per quanto riguarda invece il termine dell'azione totale associato alla materia:

$$S = k\int R\sqrt{-g}\; d^4x + \frac{1}{c}\int \mathcal{L}_M\sqrt{-g}\; d^4x$$

si ricava che la variazione del secondo termine è esprimibile in termini del tensore energia-impulso:

$$\delta S_M = \frac{1}{2c}k\int T_{\mu\nu}\delta g^{\mu\nu}\sqrt{-g}d\Omega$$

che permette di scrivere l'azione totale del campo gravitazionale più le sue sorgenti come:

$$\delta S = \int_\Omega \sqrt{-g}\left[k\left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right) - \frac{1}{2c}T_{\mu\nu}\right]\delta g^{\mu\nu}d\Omega$$

ed essendo le variazioni della metrica $\delta g^{\mu\nu}$ arbitrarie:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{1}{2kc}T_{\mu\nu}$$

Ora, il limite classico delle equazioni scritte sopra si ritrova ponendo $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Il termine $h_{\mu\nu}$ è quello che compare nel laplaciano ($\eta_{\mu\nu}$ è costante), ricavando:

$$k = \frac{c^3}{16\pi G}$$

che sostituite nelle precedenti permette di ricavare infine le *Equazioni di Einstein*:

$$\boxed{R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}} \tag{15.3}$$

Queste costituiscono un sistema di equazioni differenziali alle derivate parziali autoaccoppiate, cioé sono le equazioni del moto del campo gravitazionale in presenza di materia e contenenti in sé stesse il moto della materia e l'interazione con il campo da loro prodotte.[8]

L'aggiunta di un termine costante $\Lambda g_{\mu\nu}$ (*costante cosmologica*) al primo membro è possibile, ma altera la struttura dello spazio-tempo.

L'uguaglianza di due tensori simmetrici determina 10 equazioni. Sebbene le componenti di $g_{\mu\nu}$ siano 10, esse si riducono a 6 tramite l'identità di Bianchi, il che renderebbe indeterminato il tensore $g_{\mu\nu}$. Tuttavia, con un arbitrario passaggio di coordinate si possono occupare ancora 4 componenti: si può cioé cambiare il tensore $g_{\mu\nu}$ tramite una trasformazione di coordinate senza alterare la distribuzione di metrica.

≈≈ $\mathcal{FINE}$ ≈≈

[8] E questo costituisce la differenza con le equazioni del campo elettromagnetico.