Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Aspetti fisici e matematici della teoria della relatività ristretta

Relatore:
Prof. Paolo Albano

Presentata da:
Luca Morelli

Anno Accademico 2022/2023

Notazioni utilizzate

$\frac{d}{dt}$ è la derivata totale rispetto al tempo, indicata anche con un punto.
$\frac{\partial}{\partial x}$ è la derivata parziale rispetto a x.

Notazioni di grandezze tridimensionali

Il simbolo freccia denota un vettore $\vec{a} \in \mathbb{R}^3$ le cui cartesiane sono indicate con $v_x,\ v_y\ , v_z$.
$\vec{x}$ e $\vec{v}$ sono vettori posizione e velocità.
q_i e $\dot{q}_i$ sono le coordinate e le velocità generalizzate.
$\vec{p}$ e E sono la quantità di moto o impulso e l'energia.
Il campo elettrico e magnetico sono indicati con $\vec{E}$ e $\vec{B}$.
Si è deciso di fare uso delle unità del Sistema Internazionale così che campo elettrico e magnetico dipendono dalle costanti dielettriche e magnetiche del vuoto ϵ_0 e μ_0.
$\vec{a} \cdot \vec{b}$ indica il prodotto scalare euclideo.
$\vec{\nabla}$ è l'operatore $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$.

Notazioni di grandezze quadridimensionali

Con A^μ si indica in 4-vettore (quadrivettore) controvariante di $\mathbb{R}^4$ mentre con A_μ se ne intende uno covariante.
Le componenti di A^μ sono date dal variare di μ tra 0 e 1, anche indicate con $A^\mu = (A^0, \vec{A})$.
Due indici ripetuti sottintendono una sommatoria $A^\mu B^\mu = \sum_\mu A^\mu B^\mu$.
Matrice metrica $g_{\mu\nu}$ con segnatura $(+, -, -, -)$.
$|A^\mu|$ è la norma Minkowsky data da: $A^\mu A_\mu = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2$.

Abstract:

Riassunto breve del documento da creare dove si sintetizzano tutti i punti trattati in maniera tale da rendere più immediata la selezione del documento da parte di un sventurato lettore.

Indice

Introduzione

Capitolo 1

Fondamenti di relatività ristretta

1.1 La matematica dei sistemi di riferimento inerziali

I sistemi di riferimento costituiscono le fondamenta dello studio dei moti dei corpi sulle quali costruire leggi fisiche, è quindi opportuno descrivere matematicamente tali oggetti e le loro proprietà per poi addentrarsi nella teoria della relatività.

1.1.1 I sistemi di riferimento inerziali

Come ogni ambito della fisica la descrizione matematica dei sistemi di riferimento necessita di una serie di osservazioni sperimentali da utilizzare assiomaticamente.

Il primo fatto sperimentale che viene assunto è che lo spazio sia tridimensionale, isotropo e omogeneo, ossia che non esistano rispettivamente direzioni e punti privilegiati ma che siano tutti equivalenti. Inoltre si assume che lo spazio rispetti la geometria euclidea mentre il tempo sia ad una sola dimensione ed anch'esso isotropo ed omogeneo.
Sulla base di queste assunzioni si può quindi scegliere un punto dello spazio-tempo come vettore nullo di uno spazio vettoriale $\mathbb{R}\times\mathbb{R}^3$, quindi da questo vengono scelte tre direzioni spaziali arbitrarie lungo le quali sono orientati tre assi cartesiani, identificabili con la base di $\mathbb{R}^3$. Inoltre dall'istante corrispondente al punto spazio-temporale precedentemente scelto si inizia a misurare il tempo. Si osservi che tali scelte sono arbitrarie poiché spazio e tempo sono assunti essere isotropi ed omogenei.

Per poter formulare delle leggi che descrivano la realtà risulta necessario assumere che esistano una serie di sistemi di riferimento, detti inerziali, in cui tali leggi sono valide indipendentemente dal sistema in cui sono descritte. Questi sistemi sono sperimentalmente caratterizzati dalla proprietà di essere reciprocamente in moto rettilineo uniforme, ossia che il moto dell'origine di un sistema visto in un qualsiasi altro sistema di riferimento inerziale deve essere nella forma:

$$\vec{x}(t) = \vec{x}_0 + \vec{v}t \qquad \vec{x}_0, \vec{v} \in \mathbb{R}^3,\ t \in \mathbb{R}, \tag{1.1.1}$$

dove $\vec{x}$ è rappresnta un punto nello spazio, t un istante di tempo e $\vec{v}$ la velocità reciproca dei due sistemi di riferimento che risulta identicamente costante.
Una volta identificati questi particolari sistemi si può enunciare il principio di relatività: *in ogni sistema di riferimento inerziale tutte le leggi della fisica sono identiche.*

1.1.2 Trasformazioni di sistemi di riferimento inerziali

Si vuole ora studiare quali siano le più generiche trasformazioni che consentono di ottenere la descrizione di un moto in un sistema di riferimento inerziale K' conoscendo la descrizione di tale moto in un primo riferimento inerziale K. In questo modo la generica trasformazione da K a K' sarà un'applicazione $f : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R} \times \mathbb{R}^3$ invertibile; quest'ultima propreità è necessaria poiché come deve essere possibile passare da K a K' deve essere possibile fare il contrario.
Per poter caratterizzare le proprietà di tale applicazione è ora necessario tradurre matematicamente la richiesta imposta dal principio di relatività. Si osservi che per tale principio e per la definizione di sistema inerziale è necessario che in seguito ad una trasformazione tutti i sistemi inerziali restino tali, in altre parole è necessario che ogni retta spaziotemporale venga trasformata in un'altra retta spazitemporale, poiché rettilineo ed uniforme è il moto caratteristico di questi sistemi di riferimento. Questa condizione soddisfa le ipotesi di un teorema[3] di geometria che consente di identificare la famiglia di queste trasformazioni.

Teorema 1. *Sia $f : \mathbb{R}^n \to \mathbb{R}^n$ $(n > 1)$ una funzione biettiva che trasforma rette in altre rette. Allora f è una trasformazione affine.*

Per procedere alla dimostrazione di questo teorema è utile iniziare dimostrandone un caso particolare, ossia il caso $n = 2$. Per questo caso particolare è però necessario dimostrare in primo luogo un risultato preliminare:

Proposizione 1. *Sia $\varphi : \mathbb{R} \to \mathbb{R}$ un automorfismo di campo, ossia tale che se presi $x, y \in \mathbb{R}$ valgano $\varphi(x + y) = \varphi(x) + \varphi(y)$ e $\varphi(xy) = \varphi(x)\varphi(y)$. Allora $\varphi = id$.*

Dimostrazione. Si osservi che $\varphi(1) = 1$ infatti se $x \neq 0$ allora $\varphi(x) = \varphi(1 \cdot x) = \varphi(1)\varphi(x)$. Inoltre $\varphi(0) = 0$, infatti $\forall x \in \mathbb{R}$ vale $\varphi(0 \cdot x) = \varphi(0)\varphi(x) = \varphi(0)$, questa equazione implica $\varphi(0)(\varphi(x) - 1) = 0$ che risulta soddisfatta per ogni x solo se $\varphi(0)$ è identicamente nullo. Queste proprietà garantiscono che $\varphi(-1) = -1$, infatti $\varphi(1 - 1) = \varphi(1) + \varphi(-1) = \varphi(0) = 0$.
Si prenda ora $n \in \mathbb{N}$, n è esprimibile come somma ripetuta n volte dell'unità, in questo modo si ottiene che:

$$\varphi(n) = \varphi(1 + 1 + ... + 1) = n\varphi(1) = n.$$

Dalle proprietà precedentemente descritte è immediato che $\varphi(n) = n$ valga per ogni numero intero poiché $\varphi(-n) = -1\varphi(n) = -n$.
Si consideri adesso $q \in \mathbb{Q}$ allora $\varphi(q) = q$. Infatti, se $x, y \in \mathbb{R}$ tali che $xy = 1$, per quanto già detto, $\varphi(xy) = \varphi(x)\varphi(y) = 1$ e quindi $\varphi(\frac{1}{x}) = \frac{1}{\varphi(x)}$. Poiché ogni numero razionale è esprimibile come rapporto di numeri interi si ha che pure questi sono conservati da φ.
Siano $x, y \in \mathbb{R}$ con $x < y$, allora $\exists z \in \mathbb{R}$ tale che $z \neq 0$ e $y - x = z^2$. Utilizzando le proprietà supposte per ipotesi si ha quindi che:

$$\varphi(y) - \varphi(x) = \varphi(y - x) = \varphi(z^2) = \varphi(z)^2 > 0.$$

Questo implica che φ conservi l'ordinamento di $\mathbb{R}$.
Sia ora $x \in \mathbb{R}$, poiché $\varphi(x) \in \mathbb{R}$ e $\mathbb{Q}$ è denso in $\mathbb{R}$ allora se $x \neq \varphi(x)$ deve esistere un $q \in \mathbb{Q}$ compreso tra $\varphi(x)$ e x. Si supponga $\varphi(x) > x$, in maniera del tutto analoga si può procedere supponendo $\varphi(x) < x$, allora $\varphi(x) > q > x$, poiché φ conserva l'ordinamento si ha $\varphi(q) > \varphi(x)$, al contempo $q = \varphi(q)$. Da questo ragionamento si deduce che $\varphi(x) > q > \varphi(x)$, ossia che $q = \varphi(q) = \varphi(x)$. Questo risultato è però in contraddizione con la proprietà di densità per cui si conclude che $\varphi = \text{id}$. □

Teorema 2. *Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ una funzione biettiva che trasforma rette in altre rette. Allora f è una trasformazione affine.*

Dimostrazione. Sia $A : \mathbb{R}^2 \to \mathbb{R}^2$ una trasformazione affine invertibile tale che $(A \circ f)(0,0) = (0,0)$, $(A \circ f)(1,0) = (1,0)$ e $(A \circ f)(0,1) = (0,1)$. Si osservi che anche $A \circ f$ trasforma rette in rette poiché la trasformazione affine può agire su una retta solamente ruotandola e traslandola. Si dirà $(A \circ f)(x) = g(x)$.

Si considerino i punti x, y e $(0,0)$ non giacenti sulla stessa retta nel piano $\mathbb{R}^2$. Si prendano due rette passanti rispettivamente per l'origine e x e l'origine e y e le rispettive rette parallele passanti per y nel primo caso e x nel secondo, come mostrato in Figura 1.1, queste ultime due rette si intersecheranno nel punto $x+y$, formando quindi un parallelogrammo.
Per ipotesi queste quattro rette verranno mappate in altre quattro rette da g e poiché

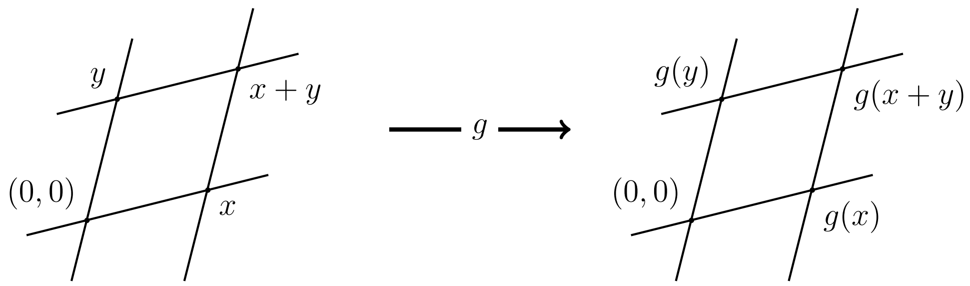

Figura 1.1: Rappresentazione grafica di come agisce la funzione g nel piano

questa funzione è biettiva le immagini di rette parallele saranno a loro volta parallele, altrimenti nel punto di incidenza si perderebbe l'iniettività, mentre i punti di incidenza potranno essere mappati solamente in altri punti di incidenza poiché questi appartenendo al dominio di due rette ognuno dovranno, per biettività, appartenere ad entrambe le immagini delle due rette nel codominio. Questo sta a significare che le quattro rette così ottenute formeranno un novo parallelogrammo di vertici $g(x)$, $g(y)$, $g(x+y)$ e $(0,0)$, poiché per costruzione $g(0,0) = (0,0)$. La regola del parallelogramma garantisce quindi che $g(x) + g(y) = g(x+y)$.
Se invece i punti x, y e $(0,0)$ giacciono sulla stessa retta è sufficiente prendere un punto $z \in \mathbb{R}^2$ che non giaccia sulla medesima retta dei punti precedenti. Si può quindi utilizzare quanto dimostrato nel caso precedente per ottenere:

$$g(x+y+z) = g(x+y) + g(z) = g(x) + g(y+z) = g(x) + g(y) + g(z)$$

da cui segue che $g(x+y) = g(x) + g(y)$ anche in questo caso.

Poiché per ipotesi g conserva le rette del piano e per costruzione mappa l'origine nell'origine e i punti $(0,1)$ e $(1,0)$ in se stessi allora g deve trasformare gli assi x e y in se stessi. Si possono quindi considerare due applicazioni $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ tali che $g(x,y) = (\alpha(x), \beta(y))$ con $x, y \in \mathbb{R}$. Si osservi che per quanto dimostrato fino ad ora $f(1,1) = f(1,0) + f(0,1) = (1,0) + (0,1) = (1,1)$, analogamente a come si è detto per gli assi x e y anche la bisettrice del primo quadrante è quindi mappata in se stessa per cui per ogni $x \in \mathbb{R}$ $\alpha(x) = \beta(x)$. Si proseguirà quindi studiando solo una di queste due funzioni poiché quanto si dirà per una è valido anche per l'altra.

Siano $a, b \in \mathbb{R}$ e si consideri una retta passante per l'origine.

$$L = \{(x, y) \in \mathbb{R}^2 \text{ tale che } y = ax\ \}$$

Il punto $(1, a) \in L$ e poiché $g(1, a) = (1, \alpha(a))$ si ottiene che L è mappata in una nuova retta passante per l'origine con coefficiente angolare $\alpha(a)$.

$$g(L) = \{(x, y) \in \mathbb{R}^2 \text{ tale che } y = \alpha(a)x\}$$

Pure $(b, ab) \in L$ e quindi $(\alpha(b), \alpha(ab)) \in g(L)$, ricordando però che $g(L)$ è una retta passante per l'origine con coefficiente angolare $\alpha(a)$ si deduce che dovrà sussistere la seguente relazione: $\alpha(ab) = \alpha(a)\alpha(b)$.

Si è quindi dimostrato che per α e β valgono le seguenti relazioni per ogni $x, y \in \mathbb{R}$:

$$\begin{aligned} \alpha(x + y) &= \alpha(x) + \alpha(y) \qquad & \alpha(xy) &= \alpha(x)\alpha(y) \\ \beta(x + y) &= \beta(x) + \beta(y) \qquad & \beta(xy) &= \beta(x)\beta(y). \end{aligned}$$

Queste relazioni implicano che α e β siano due automorfismi di campo di $\mathbb{R}$ e per la proposizione inizialmente dimostrata risultano entrambi l'identità, di conseguenza pure $g = A \circ f = \text{id}$. Infine è sufficiente appllicare a g la trasformazione inversa di A così da avere $A^{-1} \circ A \circ f = A^{-1} = f$, per cui si è dimostrato che f deve essere affine. □

Si procederà ora generalizzando questo teorema con la dimostrazione del teorema 1.

Dimostrazione teorema 1 ($n > 2$). Si consideri la traslazione $T : \mathbb{R}^n \to \mathbb{R}^n$ tale che $(T \circ f)(0) = 0$, si dirà $(T \circ f)(x) = g(x)$.
Siano $x, y \in \mathbb{R}^n$ e π un piano tale che $x, y \in \pi$. Si prendano due rette $r_1, r_2 \in \pi$ e incidenti in O, e un punto $P \in \pi$. Se si considerano altre due rette r_3 e r_4 passanti per P e tali che r_3 sia parallela a r_1 e r_4 lo sia per r_2, così che r_3 intersecherà r_2 in un punto P_2 e analogamente r_4 intersecherà r_1 in un punto P_1, è possibile costruire un parallelogrammo con vertice P e lati giacenti sulle rette considerate. Poiché g è biettiva

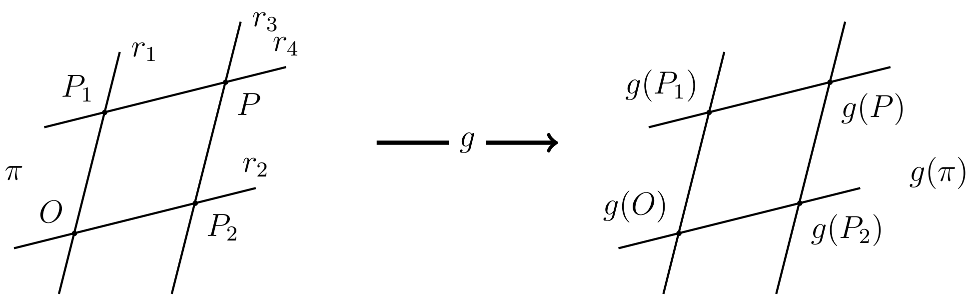

Figura 1.2: Rappresentazione grafica di come agisce la funzione g nel piano

e trasforma rette in rette allora, come si è già visto per dimostrare il teorema 2, questo parallelogrammo deve essere trasformato in un nuovo parallelogrammo con i lati giacenti sulle immagini delle precedenti rette e con i vertici dati dalle immagini dei vertici del precedente parallelogrammo. Detto π' il piano contente le immagini di r_1 e r_2 allora per costruzione del parallelogrammo si avrà che $g(P) \in \pi'$ e poiché $P \in \pi$ è arbitrario si deduce che ogni punto di π è mappato in π' ossia g trasforma piani in altri piani.

Si consideri ora un'applicazione lineare invertibile L tale che $(L \circ g)(\pi) = \pi$, in questo modo $L \circ g\big|_{\pi} : \mathbb{R}^2 \to \mathbb{R}^2$ è biettiva e trasforma rette in rette, per il teorema 2 allora questa è un'applicazione affine, nella fattispecie poiché $g(0) = 0$ per costruzione e L è lineare allora $(L \circ g)(0) = 0$, per cui $L \circ g$ è lineare. Risulta a questo punto sufficiente far uso dell'inversa di L per ottenere:

$$g(\alpha x+\beta y) = L^{-1}((L\circ g)(\alpha x+\beta y)) = L^{-1}(\alpha(L\circ g)(x)+\beta(L\circ g)(y)) = \alpha g(x)+\beta g(y) \quad \alpha, \beta \in \mathbb{R},$$

poiché la scelta di x e y fatta prima di determinare il piano π è arbitraria segue che g è lineare e considerando $f = T^{-1} \circ g$ allora si ha che f deve essere affine. □

Così facendo si conclude che ogni trasformazione di un sistema di riferimento inerziale in un altro debba essere necessariamente una trasformazione affine di qualche sorta. Ulteriori osservazioni sperimentali consentiranno di identificare famiglie più ristrette di queste trasformazioni.

1.2 Fatti di Fisica Classica

Quanto scoperto dai fisici fino alla fine dell'ottocento fu sviluppato sulle idee di Galilei e Newton nella loro formulazione classica della meccanica. Infatti, utilizzando i risultati di questa branca della fisica, scienziati come Coulomb, Ampère e Faraday furono in grado di descrivere fenomeni noti da svariati secoli dando vita alla teoria dell'elettromagnetismo. Poiché la nascita della teoria della relatività speciale è strettamente connessa a queste due branche si procederà illustrando brevemente i loro fondamenti.

1.2.1 La meccanica e le trasformazioni di Galileo

Alle osservazioni fatte nella sezione 1.1 vanno aggiunte ulteriore assunzioni di carattere sperimentale che consentono di delineare gli schemi della meccanica classica.

Oltre a quanto assunto nella sezione 1.1, ossia che lo spazio sia tridimensionale, isotropo, omogeneo e che rispetti la geometria euclidea mentre il tempo sia ad una sola dimensione, così da poter definire cosa sia un sistema di riferimento inerziale e il principio di relatività[1], si prende come assioma che le distanze spaziali e temporali siano assolute, ossia che ogni osservatore concordi sulla misura di queste.

Inoltre si assume che, in un riferimento inerziale, le posizioni e le velocità dei punti di un sistema ad un tempo iniziale ne determinino in maniera univoca l'evoluzione $\vec{x}(t) \in \mathbb{R}^3$ secondo la legge:

$$m\ddot{\vec{x}}(t) = \vec{F}(t, \vec{x}, \dot{\vec{x}}) \tag{1.2.1}$$

dove m è detta massa inerziale e $\vec{F}$ è una funzione caratteristica del sistema detta forza.

Si vuole quindi identificare quali applicazioni $\varphi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}$ consentono di cambiare sistema di riferimento inerziale, ossia quali trasformazioni non variano le leggi della natura. Queste applicazioni si chiamano trasformazioni di Galileo e, come si è dimostrato nella sezione 1.1, per soddisfare il principio di relatività devono essere applicazioni affini.

[1] In meccanica classica è anche detto principio di relatività galileiano

Una generica trasformazione di Galileo è quindi data dalla composizione di tre famiglie di applicazioni:

- una generica traslazione spazio temporale dell'origine, dedotta dalla proprietà di omogeneità dello spazio e del tempo:
$$\varphi_{\vec{r},s}(t,\vec{x}) = (t+s, \vec{x}+\vec{r}) \tag{1.2.2}$$
- una generica rotazione degli assi spaziali, dovuta alla proprietà di isotropia dello spazio:
$$\varphi_G(t,\vec{x}) = (t, G\vec{x}) \qquad G \in M_{3\times 3}(\mathbb{R}) : G^{-1} = G^t \tag{1.2.3}$$
- una traslazione di moto rettilineo uniforme, ammissibile grazie alle proprietà di muoversi di tale moto dei sistemi di riferimento inerziali:
$$\varphi_{\vec{v}}(t,\vec{x}) = (t, \vec{x}+\vec{v}t) \tag{1.2.4}$$

Quest'ultima tipologia di trasformazione è quella che viene comunemente studiata per caratterizzare le trasformazioni di sistemi inerziali. Si consideri quindi un sistema K, con coordinate $(t,\vec{x})$ e un sistema K', con coordinate $(t',\vec{x'})$, in moto a velocità $\vec{V}$ rispetto a K, si scriverà allora la (1.2.4) come:

$$\vec{x'} = \vec{x} - \vec{V}t \qquad t' = t. \tag{1.2.5}$$

Per determinare l'invarianza di un legge fisica rispetto alle trasformazioni (1.2.5) può essere necessario studiare come si trasformino gli operatori di differenziazione.
Se si vuole derivare una $f(\vec{x},t)$, dalla regola di Leibniz, si ha:

$$\frac{\partial}{\partial t'} = \frac{\partial t}{\partial t'}\frac{\partial}{\partial t} + \sum_{i=1}^{3}\frac{\partial x_i}{\partial t}\frac{\partial t}{\partial t'}\frac{\partial}{\partial x_i} \qquad\qquad \frac{\partial}{\partial x'_i} = \frac{\partial t}{\partial x'_i}\frac{\partial}{\partial t} + \sum_{i=1}^{3}\frac{\partial x_j}{\partial x'_i}\frac{\partial}{\partial x_j}$$

dove $\frac{\partial}{\partial t}$ è la derivata parziale rispetto al tempo nel sistema K, $\frac{\partial}{\partial t'}$ è la derivata parziale rispetto al tempo nel sistema K', $\frac{\partial}{\partial x_i}$ è la derivata parziale rispetto alla coordinata i-esima del sistema K e $\frac{\partial}{\partial x'_i}$ è la derivata parziale rispetto alla coordinata i-esima del sistema K'.
Osservando dalla (1.2.5) che $\frac{\partial t}{\partial t'} = 1$, che $\frac{\partial t}{\partial x'_i} = 0$, che $\frac{\partial x_i}{\partial t} = V_i$ e che $\frac{\partial x_j}{\partial x'_i} = \delta_{ij}$, dove δ_{ij} è una delta di Kronecker, si ottengono le trasformazioni degli operatori di differenziazione desiderate:

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \vec{V}\cdot\vec{\nabla} \qquad\qquad \frac{\partial}{\partial x'_i} = \frac{\partial}{\partial x_i} \tag{1.2.6}$$

dove $\vec{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$.
Dalla (1.2.6) si deduce che tutti gli operatori che comprendono solamente derivate rispetto alle coordinate spaziali sono lasciati inalterati dalle trasformazioni di Galileo. In particolare si ha che $\vec{\nabla}' = (\frac{\partial}{\partial x'}, \frac{\partial}{\partial y'}, \frac{\partial}{\partial z'}) = \vec{\nabla}$.

Note queste trasformazioni è possibile osservare che il vettore accelerazione $\ddot{\vec{x}}$ risulta galileo invariante[1], infatti considerando una traiettoria $\vec{x}(t)$:

$$\frac{d^2}{dt'^2}(\vec{x'}(t)) = \frac{d^2}{dt^2}(\vec{x}(t) - \vec{V}t) = \ddot{\vec{x}}(t) + \frac{d}{dt}(\vec{V}) = \ddot{\vec{x}}(t)$$

[1]Con galileo invariante si intende che un oggetto fisico è invariante sotto le trasformazioni di galileo.

questo fatto, assieme al principio di relatività galileiano applicato alla legge di Newton (1.2.1), impone che le forze esercitate su di un punto e misurate in due sistemi inerziali differenti debbano essere le medesime.

Analogamente a quanto appena fatto si possono ricavare le trasformazioni per la velocità di un punto in moto con traiettoria $\vec{x}(t)$:

$$\frac{d}{dt'}(\vec{x}'(t)) = \frac{d}{dt}(\vec{x}(t) - \vec{V}t) = \dot{\vec{x}}(t) + \vec{V}. \tag{1.2.7}$$

Per cui si ottiene dalle trasformazioni di Galileo che le velocità si compongono per somma algebrica.

1.2.2 L'elettromagnetismo e le equazioni di Maxwell

Per interpretare i fenomeni elettromagnetici, anche in questo caso, è necessario introdurre una serie di osservazioni sperimentali: in primo luogo esiste una proprietà della materia detta carica elettrica, che non dipende dal sistema di riferimento in cui è misurata e che consente ai corpi di interagire con due campi vettoriali: il campo elettrico $\vec{E}$ e il campo magnetico $\vec{B}$.
Un corpo puntiforme di carica q interagendo con questi subisce un forza $\vec{F}$ data da:

$$\vec{F} = q(\vec{E}(\vec{x}, t) + \vec{v} \wedge \vec{B}(\vec{x}, t)) \tag{1.2.8}$$

dove $\vec{v}$ è il vettore velocità di tale corpo.
In secondo luogo gli esperimenti mostrarono che questi due campi rispettano una serie di equazioni dette equazioni di Maxwell:

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \frac{\rho}{\epsilon_0} & \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \wedge \vec{E} &= -\frac{\partial \vec{B}}{\partial t} & \vec{\nabla} \wedge \vec{B} &= \mu_0 \vec{J} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} \end{aligned} \tag{1.2.9}$$

dove ρ è la densità di carica volumetrica, $\vec{J}$ è la densità di corrente superficiale e ϵ_0, μ_0 sono due costanti del vuoto.
Dalle equazioni di Maxwell (1.2.9) segue che le cariche sono sorgenti del campo elettrico mentre le correnti lo sono per il campo Magnetico. Per esempio è possibile mostrare che una carica puntiforme è sorgente di campo elettrico e tramite la forza di Lorentz (1.2.8) è possibile derivare la legge di Coulomb.
Si consideri una carica puntiforme Q posta nell'origine di un sistema di riferimento e si prenda una sfera $\mathcal{S}$ di raggio R contenete tale carica. Integrando la prima equazione di Maxwell e facendo uso del teorema di Gauss si ottiene:

$$\int_{\mathcal{S}} \vec{\nabla} \cdot \vec{E}\; d^3x = \int_{\mathcal{S}} \frac{\rho}{\epsilon_0}\; d^3x = \frac{Q}{\epsilon_0} = \oint_{\partial\mathcal{S}} \vec{E} \cdot \hat{n}\; d\Sigma$$

dove si è indicato con $\hat{n}$ il versore normale alla superficie $\mathcal{S}$ nel punto in cui si valuta l'integrando.
Si osservi che essendo lo spazio isotropo e la carica puntiforme allora qualsiasi rotazione degli assi coordinati mantiene immutato il sistema, ne segue che il Campo Elettrico

generato dalla carica deve essere radiale e costante su superfici sferiche centrate nell'origine. Infatti un'ipotetica seconda carica fissa a distanza R dall'origine del sistema di riferimento deve percepire sempre la medesima forza, indipendentemente dalla rotazione effettuata, che risulta connessa al campo elettrico per mezzo della (1.2.8), inoltre sempre dalle Equazioni di Maxwell, supponendo assenza di correnti si ha che il rotore di $\vec{E}$ risulta nullo. Grazie a queste deduzioni l'integrale di superficie si riduce a:

$$\oint_{\partial S} \vec{E} \cdot \hat{n}\, d\Sigma = 4\pi R^2 |\vec{E}| = \frac{Q}{\epsilon_0}.$$

Tenendo conto che $\vec{E}$ è radiale, come si è dedotto, si ha la legge di Coulomb

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{Q}{R^2}\hat{r} \quad \Rightarrow \quad \vec{F} = q\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{qQ}{R^2}\hat{r} \tag{1.2.10}$$

dove si è indicato con $\hat{r}$ il versore radiale in coordinate sferiche.

Infine è importante per trattare la teoria della relatività osservare che è possibile ottenere, calcolando il rotore di ambo i membri delle ultime due equazioni di Maxwell

$$\vec{\nabla} \wedge \vec{\nabla} \wedge \vec{E} = -\frac{\partial}{\partial t}\vec{\nabla} \wedge \vec{B} \qquad \vec{\nabla} \wedge \vec{\nabla} \wedge \vec{B} = \mu_0\epsilon_0\frac{\partial}{\partial t}\vec{\nabla} \wedge \vec{E}$$

e supponendo assenza di cariche, per cui $\rho = 0$ e $\vec{J} = 0$, due equazioni che descrivono onde di campo elettrico e magnetico nel vuoto:

$$\vec{\nabla}^2\vec{E} = \mu_0\epsilon_0\frac{\partial^2\vec{E}}{\partial t^2} \qquad \vec{\nabla}^2\vec{B} = \mu_0\epsilon_0\frac{\partial^2\vec{B}}{\partial t^2} \tag{1.2.11}$$

dove si è indicato l'operatore laplaciano con la notazione $\vec{\nabla}^2 = (\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}, \frac{\partial^2}{\partial z^2})$.
Queste onde si propagano con una velocità $\frac{1}{\sqrt{\mu_0\epsilon_0}} = 299792458\ \frac{m}{s}$, che corrisponde con precisione ai valori sperimentalmente misurati della velocità della luce. Maxwell suppose allora che questa fosse quindi da intendere come un fenomeno elettromagnetico e successivi esperimenti, come quelli di Hertz, confermarono tale ipotesi.
Sulla base della teoria ondulatoria classica è però necessario identificare un mezzo nel quale queste onde possano propagarsi e rispetto al quale la loro velocità di propagazione debba essere intesa. Per questo motivo alla fine dell'ottocento venne ipotizzata l'esistenza di tale mezzo detto etere luminifero.

1.3 L'articolo del 1905

La Meccanica newtoniana e l'elettromagnetismo di Maxwell si rivelarono a gli occhi dei fisici dell'ottocento incompatibili fra loro, poiché le equazioni di Maxwell non risultarono invarianti per le trasformazioni di Galileo. Proprio per questo motivo i fisici dell'epoca dovettero rivalutare i principi alla base delle leggi della natura fino a quando l'incompatibilità trovò una soluzione nel 1905 con la relatività ristretta di Einstein. Si ripercorreranno ora i passi che Einstein stesso indicò nel suo articolo del 1905[1].

1.3.1 La non invarianza delle Equazioni di Maxwell

Si considerino due sistemi di riferimento K e K', inerziali e reciprocamente in moto in maniera tale che in K' l'origine del sistema K risulti in moto a velocità costante $\vec{V}$, allora in ogni sistema si misureranno rispettivamente $\vec{E}$, $\vec{B}$ e $\vec{E'}$, $\vec{B'}$.
Il principio di relatività galileiana impone che questi due campi, nei loro sistemi di riferimento, rispettino le equazioni di Maxwell (1.2.9). Inoltre, siccome la forza di Lorentz (1.2.8) deve essere la medesima in tutti i sistemi di riferimento inerziali, considerando una carica q in moto con velocità $\vec{v}$ in K e $\vec{V}+\vec{v}$ in K', essendo q invariante deve valere:

$$\begin{aligned} \vec{F'} = \vec{F} \quad &\Rightarrow \quad \vec{E'} + \vec{v} \wedge \vec{B'} = \vec{E} + (\vec{V} + \vec{v}) \wedge \vec{B} \\ &\Rightarrow \quad \vec{E'} + \vec{v} \wedge (\vec{B'} - \vec{B}) = \vec{E} + \vec{V} \wedge \vec{B'} \end{aligned}$$

Così facendo si possono ottenere delle relazioni tra $\vec{E}$, $\vec{B}$ e $\vec{E'}$, $\vec{B'}$ che diventano le trasformazioni dei capi elettrici e magnetici tra sistemi inerziali. Queste però possono dipendere esclusivamente dalle proprietà dei due sistemi di riferimento considerati e nella fattispecie dalla loro velocità reciproca, per questo motivo il termine contenente $\vec{v}$, ossia la velocità della carica nel sistema K, deve annullarsi, per cui le trasformazioni risultano:

$$\begin{cases} \vec{E'}(t,\vec{x'}) = \vec{E}(t,\vec{x'}) + \vec{V} \wedge \vec{B}(t,\vec{x}) \\ \vec{B'}(t,\vec{x'}) = \vec{B}(t,\vec{x}) \end{cases} \tag{1.3.1}$$

Bisogna ora studiare come si trasformino le grandezze generatici dei campi ρ e $\vec{J}$. Se si considera una volume ΔV, in cui è presente una carica Δq, allora la densità di carica è definita come

$$\rho = \lim_{\Delta V \to 0} \frac{\Delta q}{\Delta V}$$

e siccome le lunghezze sono assunte essere assolute allora devono esserlo pure i volumi ed essendo la carica non dipendente dal sistema di riferimento si conclude che pure la densità di carica non lo può essere. Per quanto riguarda la densità di corrente superficiale, definita come $\vec{J} = \rho\vec{v}$, è sufficiente applicare le trasformazioni delle velocità tra due sistemi in moto reciproco a velocità $\vec{V}$ per ottenere:

$$\vec{J'} = \vec{J} - \rho\vec{V} \tag{1.3.2}$$

I risultati appena ottenuti consentono di determinare l'invarianza delle equazioni di Maxwell per le Trasformazioni di Galileo.

Studiando la prima equazione di Maxwell in K' e considerandola valida in K, se si trasformano E' in E e analogamente per gli operatori di differenziazione secondo la (1.2.6), si ottiene una quantità che è nulla se questa equazione è valida in anche K' e così facendo è possibile verificare se le equazioni di Maxwell sono valide in ogni sistema di riferimento inerziale, come richiesto dal principio di relatività.

$$\begin{aligned} \vec{\nabla}' \cdot \vec{E'} - \frac{\rho}{\epsilon_0} &= \vec{\nabla} \cdot (\vec{E} + \vec{V} \wedge \vec{B}) - \frac{\rho}{\epsilon_0} \\ &= \left(\vec{\nabla} \cdot \vec{E} - \frac{\rho}{\epsilon_0} \right) - \vec{V} \cdot (\vec{\nabla} \wedge \vec{B}) = -\vec{V} \cdot (\vec{\nabla} \wedge \vec{B}) \end{aligned}$$

Il primo termine tra parentesi è identicamente nullo poiché valgono le equazioni di Maxwell in K mentre l'ultimo termine non è sempre nullo, nella fattispecie in presenza di campi elettrici variabili nel tempo, questo implica che quindi la prima equazione di Maxwell non sia invariante per le trasformazioni di Galileo.

Se si studia la seconda con lo stesso procedimento si scopre che questa invece è invariante:

$$\vec{\nabla}' \cdot \vec{B}' = \vec{\nabla} \cdot \vec{B} = 0$$

Analogamente per la terza equazione:

$$\begin{aligned}
\vec{\nabla}' \wedge \vec{E}' + \frac{\partial \vec{B}'}{\partial t'} &= \vec{\nabla} \wedge \vec{E}' + \frac{\partial \vec{B}'}{\partial t} + (\vec{V} \cdot \vec{\nabla})\vec{B}' \\
&= \vec{\nabla} \wedge (\vec{E} + \vec{V} \wedge \vec{B}) + \frac{\partial \vec{B}}{\partial t} + (\vec{V} \cdot \vec{\nabla})\vec{B} \\
&= \left(\vec{\nabla} \wedge \vec{E} + \frac{\partial \vec{B}}{\partial t}\right) + \vec{\nabla} \wedge \vec{V} \wedge \vec{B} + (\vec{V} \cdot \vec{\nabla})\vec{B} = 0
\end{aligned}$$

infatti il termine tra parentesi è identicamente nullo poiché in K vale la terza equazione di Maxwell e gli addendi restanti si annullano se sviluppati tramite le regole di differenziazione ricordando che $\vec{V}$ è costante:

$$\vec{\nabla} \wedge \vec{V} \wedge \vec{B} + (\vec{V} \cdot \vec{\nabla})\vec{B} = -(\vec{V} \cdot \vec{\nabla})\vec{B} + (\vec{V} \cdot \vec{\nabla})\vec{B} = 0$$

L'ultima equazione di Maxwell è invece non invariante infatti sempre con il medesimo procedimento si ottiene:

$$\begin{aligned}
&\vec{\nabla}' \wedge \vec{B}' - \mu_0\epsilon_0 \frac{\partial \vec{E}'}{\partial t'} - \mu_0 \vec{J}' = \vec{\nabla} \wedge \vec{B}' - \mu_0\epsilon_0 \frac{\partial \vec{E}'}{\partial t} - \mu_0\epsilon_0(\vec{V} \cdot \vec{\nabla})\vec{E}' - \mu_0 \vec{J}' \\
&= \vec{\nabla} \wedge \vec{B} - \mu_0\epsilon_0 \frac{\partial \vec{E}}{\partial t} - \mu_0\epsilon_0 \vec{V} \wedge \frac{\partial \vec{B}}{\partial t} - \mu_0\epsilon_0(\vec{V} \cdot \vec{\nabla})(\vec{E} + \vec{V} \wedge \vec{B}) - \mu_0 \vec{J} + \mu_0 \vec{V}\rho \\
&= \left(\vec{\nabla} \wedge \vec{B} - \mu_0\epsilon_0 \frac{\partial \vec{E}}{\partial t} - \mu_0 \vec{J}\right) - \mu_0\epsilon_0 \vec{V} \wedge \frac{\partial \vec{B}}{\partial t} - \mu_0\epsilon_0(\vec{V} \cdot \vec{\nabla})(\vec{E} + \vec{V} \wedge \vec{B}) + \mu_0 \vec{V}\rho \\
&= -\mu_0\epsilon_0 \vec{V} \wedge (-\vec{\nabla} \wedge \vec{E}) - \mu_0\epsilon_0(\vec{V} \cdot \vec{\nabla})(\vec{E} + \vec{V} \wedge \vec{B}) + \mu_0 \vec{V}\rho \\
&= -\mu_0\epsilon_0(\vec{V} \cdot \vec{\nabla})(\vec{V} \wedge \vec{B}) + \mu_0 \vec{V}\rho
\end{aligned}$$

che non è nullo in generale con le assunzioni fin qui fatte.

Si è quindi giunti alla conclusione che la teoria di Maxwell non è conciliabile con la meccanica di Newton e viceversa.

1.3.2 I postulati di Einstein

Per giungere ad una formulazione coerente della dinamica dei corpi carichi Einstein, nel suo articolo del 1905[1], propose di modificare gli assunti alla base della meccanica classica per prediligere un modello coerente con l'elettromagnetismo di Maxwell.
Infatti all'epoca erano noti alcuni risultati sperimentali che giocavano a sfavore della concezione classica della meccanica, primo fra tutti l'esperimento di Michelson e Morley che avrebbe dovuto consentire di misurare la velocità della Terra rispetto all'etere luminifero, detta vento d'etere. L'esperimento ebbe un esito inaspettato, infatti non

fu possibile misurare alcun vento d'etere portando i fisici a tre possibili spiegazioni: o l'Etere si muove assieme alla Terra, o l'apparato sperimentale si contrae lungo la direzione del moto terrestre oppure non esiste alcun Etere e la luce si propaga alla medesima velocità in ogni direzione e per ogni osservatore.

Einstein pose quindi a fondamento della sua teoria due postulati:

- Il principio di relatività: basato sull'assunzione che esistano una serie di sistemi di riferimento detti inerziali, reciprocamente in moto rettilineo uniforme, in cui le leggi della fisica sono identicamente valide.

- Il principio di costanza della velocità della luce: il quale asserisce che la luce nello spazio vuoto si propaghi sempre con modulo della velocità determinato ed identico per ogni osservatore inerziale, che si indicherà con c.

Il primo si rifà al principio di relatività galileiano mentre il secondo è una diretta conseguenza dell'esperimento di Michelson e Morley il cui risultato viene spiegato senza la necessità dell'introduzione dell'etere e di un sistema di riferimento privilegiato. Inoltre si continua ad intendere spazio e tempo come due enti omogenei e isotropi.

La primissima conseguenza dell'assunzione di questi due postulati è la non validità delle trasformazioni di Galileo. Infatti secondo queste un moto di velocità $\vec{v}$ in un sistema K, se osservato in un sistema K', nel quale K si muove a velocità $\vec{V}$, risulterà in un moto a velocità $\vec{v}+\vec{V}$. Il secondo postulato però richiede che se tale moto è di un fascio di luce questo debba risultare sia in K che in K' in un moto con modulo della velocità pari a c, in totale disaccordo con le trasformazioni di Galileo.

Inoltre nel suo articolo Einstein stesso propose, dopo aver enunciato i postulati, un'esperimento mentale che consente di mostrare come questi siano in diretto conflitto con le assunzioni classiche dell'assolutezza del tempo e delle lunghezze. Si considerino due orologi reciprocamente a riposo posizionati in due punti detti A e B. Einstein osservò che ogni orologio è in grado di misurare intervalli temporali solamente per eventi che avvengono nello stesso punto in cui ognuno è posizionato, questo poiché diventa necessario tener conto della velocità della luce che quindi propagandosi genera dei ritardi nella percezione degli eventi lontani.
Il secondo postulato consente però di sincronizzare gli orologi così che sia possibile confrontare i tempi misurati in A e in B. In primo luogo si ipotizzi di far partire all'istante $t_A = 0$, misurato dal primo orologio, un fascio di luce che viaggia da A e giunge in B quando l'orologio posizionato in tale punto segna un tempo t_B. In B il fascio è riflesso e fa ritorno in A quando il relativo orologio segna un tempo t_A. Poiché per il principio di costanza della velocità della luce il fascio luminoso deve propagarsi in ogni direzione con la stessa velocità e la distanza tra i due orologi è fissata e costante allora il tempo impiegato dalla luce per andare da A a B e vice versa deve essere il medesimo. Si conclude quindi che i due orologi sono sincronizzati solamente se vale:

$$2t_B = t_A \tag{1.3.3}$$

Chiaramente se l'orologio nel punto A si è rivelato sincrono con quello nel punto B, tramite la procedura appena descritta, è chiaro che è altrettanto vero che quello posto in B è sincrono con quello posto in A e che se si considera un orologio posto in un terzo

punto C che risulta sincrono con quello posto in B allora questo terzo orologio è sincrono con quello nel punto A.
Così facendo è possibile assegnare un immaginario orologio ad ogni punto dello spazio in maniera tale che siano tutti sincroni tra loro e sia possibile determinare quando due eventi lontani fra loro avvengono nello stesso istante in un determinato sistema di riferimento inerziale.
Fatta propria questa osservazione è possibile procedere analizzando l'esperimento mentale: si consideri un regolo di lunghezza l in un sistema di riferimento ad esso solidale detto K. Si considerino anche due orologi sincroni posti nelle due estremità del regolo dette A e B. In un sistema K' si osserva il regolo in moto a velocità v lungo la direzione in cui la parte lunga del regolo poggia. Sia detto A' il punto del sistema K' in cui si osserva l'emissione del fascio di luce dalla prima estremità del regolo, B' il punto, sempre in K', in cui si osserva la riflessione del fascio nel secondo estremo del regolo ed infine C' il punto in K' in cui si osserva il fascio fare ritorno alla prima estremità. In ognuno dei tre punti è presente un orologio sincronizzato con gli altri del sistema K' in maniera da osservare l'emissione del fascio ad un tempo $t'_{A'} = 0$. Se $t'_{B'}$ è l'istante in cui si osserva la riflessione in B', $t'_{C'}$ è l'istante in cui il fascio fa ritorno al primo estremo del regolo e l' è la lunghezza del regolo misurata nel sistema K' allora è possibile determinare in funzione di questi tempi le distanze tra i punti A', B' e C', infatti queste dipenderanno in parte dalla distanza percorsa del regolo e in parte dalla sua lunghezza:

$$\begin{aligned}\Delta x'_{A'B'} &= x'_{B'} - x'_{A'} = v(t'_{B'} - t'_{A'}) + l' = vt'_{B'} + l' \\ \Delta x'_{B'C'} &= x'_{B'} - x'_{C'} = l' - v(t'_{C'} - t'_{B'})\end{aligned}$$

Queste due distanze sono quelle percorse dal fascio luminoso rispettivamente in tempi $t'_{B'}$ e $t'_{C'} - t'_{B'}$, per cui si ottiene:

$$\Delta x'_{A'B'} = ct'_{B'} = vt'_{B'} + l' \quad \Rightarrow \quad t'_{B'} = \frac{l'}{c-v}$$
$$\Delta x'_{B'C'} = c(t'_{C'} - t'_{B'}) = l' - v(t'_{C'} - t'_{B'}) \quad \Rightarrow \quad t'_{C'} - t'_{B'} = \frac{l'}{c+v}$$

Questa semplice osservazione consente di concludere che i postulati enunciati da Einstein non ammettono la possibilità di assumere tempi assoluti: infatti l'osservatore in K' osserva che il fascio di luce impiega più tempo per giungere al secondo estremo di quanto ne trascorra tra la riflessione e il suo ritorno al primo estremo, mentre invece l'osservatore in K, solidale con il regolo, osserva tempi identici per questi due tragitti. Una volta sviluppata matematicamente questa nuova teoria della relatività si osserverà che in maniera analoga pure le lunghezze non possono più essere considerate assolute.

1.4 Le trasformazioni di Lorentz

Come si è già visto, i postulati di Einstein non risultano compatibili con le trasformazioni di Galileo, per questo motivo il primo passo compiuto da Einstein stesso fu quello di identificare matematicamente quali siano le nuove trasformazioni dei sistemi di riferimento inerziali che derivano direttamente dai nuovi postulati.
In questa sezione si procederà a ricavare le trasformazioni di Lorentz seguendo i ragionamenti di Fock[2] e di Landau[4], per poi discuterne le implicazioni.

1.4.1 Derivazione

Si vogliono trovare le trasformazioni che, rispettando i due postulati di Einstein, trasformano i vettori dello spazio-tempo $\mathbb{R} \times \mathbb{R}^3$ misurati in un sistema di riferimento inerziale K in quelli misurati in un secondo sistema di riferimento inerziale K'. Il principio di relatività impone che tutte le leggi della fisica siano valide sia in K che in K' e in modo particolare il principio di costanza della velocità della luce. Si può richiedere questa condizione nel seguente modo: sia emesso rispetto a K un segnale luminoso all'istante t_0 e nel punto $\vec{r}_0$, allora in un istante t_1 si osserverà il segnale in $\vec{r}_1$ tale che rispettando il secondo postulato, la propagazione avvenga a velocità con modulo costante e pari a c, così che:

$$|\vec{r}_1 - \vec{r}_0|^2 = (x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2 = c^2(t_1 - t_0)^2 \tag{1.4.1}$$

e se in K' si osserva lo stesso fascio di luce emesso all'istante t'_0 nel punto $\vec{r'}_0$ analogamente in un istante t'_1 il segnale giungerà in $\vec{r'}_1$ e per il principio di costanza della velocità della luce anche in questo sistema di riferimento deve valere la seguente espressione.

$$|\vec{r'}_1 - \vec{r'}_0|^2 = (x'_1 - x'_0)^2 + (y'_1 - y'_0)^2 + (z'_1 - z'_0)^2 = c^2(t'_1 - t'_0)^2 \tag{1.4.2}$$

Il luogo dei punti in $\mathbb{R} \times \mathbb{R}^3$ che soddisfa queste relazioni è detto cono di luce ed è la rappresentazione spaziotemporale del moto luminare. I postulati di Einstein impongono che la trasformazione che si sta cercando trasformi sempre punti del cono di luce in K in altri punti del cono di luce in K'.

Come si è visto nella sezione 1.3.2 gli istanti temporali in cui si verifica un preciso evento non sono più assoluti e diversi osservatori di diversi sistemi di riferimento potrebbero non concordare sulla loro misura, per questo motivo è importate iniziare a ragionare utilizzando vettori appartenenti a $\mathbb{R}^4$, detti quadrivettori[1]. Le relazioni (1.4.1) e (1.4.2) suggeriscono di utilizzare come quadrivettore $r = (ct, x, y, z)$ poiché così facendo è possibile definire la norma quadra di un quadrivettore tramite il prodotto righe per colonne con una matrice g detta matrice metrica

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \Rightarrow \quad |r|^2 = (ct, x, y, z)\ g \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = c^2t^2 - x^2 - y^2 - z^2$$

e in tal modo si descriverà un quadrivettore rappresentante una traiettoria nello spazio-tempo di un fascio di luce indicando che la sua norma deve essere nulla.
Un secondo modo equivalente consiste nel considerare il vettore $r = (ict, x, y, z)$, dove $i^2 = -1$, il che consente di poter utilizzare il regolare prodotto scalare euclideo per definire la norma di un quadrivettore ottenendo la stessa espressione della precedente convenzione. Per passare da una convenzione all'altra è sufficiente far uso di un cambio di base P che mappi i vettori della base canonica in una nuova base $\{e'_0 = ie_0,\ e'_1 = e_1,\ e'_2 = e_2,\ e'_3 = e_3\}$.

$$P = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.4.3}$$

[1] Poiché con l'uso dei quadrivettori la coordinata temporale è considerata al pari di quelle spaziali si dirà da ora in poi che questi appartengono a $\mathbb{R}^4$ e non a $\mathbb{R} \times \mathbb{R}^3$, seppur queste due notazioni rappresentino lo stesso spazio vettoriale.

La trasformazione che si vuole trovare è quindi una trasformazione $f : \mathbb{R}^4 \to \mathbb{R}^4$, invertibile e che trasforma quadrivettori con norma nulla in altri quadrivettori con norma nulla. L'invertibilità è necessaria affinché sia possibile trasformare sia da un sistema di riferimento K ad uno K' sia viceversa. Questa trasformazione deve essere una trasformazione affine affinché il principio di relatività si soddisfatto, come si è dimostrato nella sezione 1.1; un'ulteriore dimostrazione, valida solamente per la relatività speciale, è fornita nell'appendice A.

Si osservi che la proprietà di mantenere la norma nulla, che è equivalente alla richiesta di soddisfare i postulati di Einstein, implica che $|f(r)|^2 = \lambda |r|^2$, con λ costante, questo fatto algebrico è dimostrato nel lemma 2 dell'appendice A. Si considerino 3 sistemi di riferimento K, K_1 e K_2, tali per cui in K l'origine di K_1 risulti in moto a velocità $\vec{v}_1$ e l'origine di K_2 risulti in moto a velocità $\vec{v}_2$. Sia r un quadrivettore in K e r_1 e r_2 il medesimo quadrivettore rispettivamente in K_1 e K_2, è possibile considerare tre trasformazioni f_1, f_2 f_{12}, tali che $f_1(r_1) = f_2(r_2) = r$ e $f_{12}(r_1) = r_2$, per cui si avrà:

$$|r_1|^2 = \lambda_1 |r|^2 \quad |r_2|^2 = \lambda_2 |r|^2 \quad |r_1|^2 = \lambda_{12} |r_2|^2 \quad \Rightarrow |r_1|^2 = \frac{\lambda_1}{\lambda_2} |r_2|^2 = \lambda_{12} |r_2|^2.$$

Ogni costante λ può dipendere esclusivamente dalla trasformazione considerata, ossia dalle velocità reciproche dei sistemi di riferimento tra cui questa agisce, infatti non è possibile ammettere una dipendenza da un qualche quadrivettore r siccome questa violerebbe la proprietà di omogeneità dello spazio-tempo. Analogamente non è possibile supporre che λ dipenda dalla direzione o dal verso delle velocità reciproche altrimenti sarebbe possibile definire una direzione preferenziale violando l'isotropia dello spazio-tempo. Si conclude che deve quindi valere la seguente relazione:

$$\frac{\lambda(|\vec{v}_1|)}{\lambda(|\vec{v}_2|)} = \lambda(|\vec{v}_{12}|). \tag{1.4.4}$$

Infine si osservi che la norma della velocità reciproca $|\vec{v}_{12}|$ dipenderà dalle direzioni in cui sono orientate le velocità $\vec{v}_1$ e $\vec{v}_2$, infatti se le due sono identiche la velocità reciproca dovrà essere nulla mentre se sono in norma uguale ma in direzioni differenti sarà possibile ottenere solamente velocità non nulle. Nella relazione (1.4.4), appena ottenuta, $\lambda(|\vec{v}_{12}|)$ è quindi dipendente dalle direzioni di $\vec{v}_1$ e $\vec{v}_2$ mentre il rapporto $\frac{\lambda(|\vec{v}_1|)}{\lambda(|\vec{v}_2|)}$ permane di valore fissato al variare dell'angolo tra $\vec{v}_1$ e $\vec{v}_2$. Si conclude quindi che λ deve assumere un valore costante indipendentemente dalla trasformazione considerata, inoltre dalla (1.4.4) è immediato concludere che tale costante è proprio pari a 1.

La trasformazione f deve quindi anche conservare le norme dei vettori r dello spazio tempo. Se si fa uso della convenzione per cui $r = (ict, x, y, z)$ si ottiene che il luogo dei punti con norma fissata R è dato da una 4-sfera

$$r^2 = (ict)^2 + x^2 + y^2 + z^2 = R^2 \tag{1.4.5}$$

e questo insieme di punti permane una 4-sfera solo per rotazioni e traslazioni, per cui l'applicazione lineare della trasformazione affine che si sta cercando deve essere una rotazione. Questa generica rotazione può essere decomposta in rotazioni nei piani xy,

xz, yz, xt, yt e yt: le rotazioni nei primi tre piani corrispondono alle classiche rotazioni spaziali mentre sono le ultime tre quelle più peculiari.
Si consideri ora una rotazione nel piano xt, le altre due rotazioni sono analizzabili in maniera del tutto analoga, questa può essere facilmente espressa nella forma:

$$\begin{pmatrix} ict' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \qquad \Rightarrow \qquad \begin{cases} t' = t\cos\theta + i\frac{x}{c}\sin\theta \\ x' = ict\sin\theta + x\cos\theta \\ y' = y \\ z' = z \end{cases}$$

dove t', x', y', z' sono le coordinate misurate in K' e t, x, y, z sono quelle misurate in K. Siccome si sta studiando la trasformazione tra due sistemi inerziali e quindi in moto rettilineo uniforme l'uno rispetto l'altro solamente nel piano xt, è naturale porre l'origine di K' in moto a velocità costante V lungo l'asse x, coincidente con x' siccome non si sono effettuate rotazioni spaziali. Se si considera la trasformazione del punto che ha come immagine l'origine di K' si ottiene:

$$\begin{cases} t' = t\cos\theta + i\frac{x}{c}\sin\theta \\ 0 = ict\sin\theta + x\cos\theta \\ 0 = y \\ 0 = z \end{cases} \qquad \Rightarrow \qquad V = \frac{x}{t} = -ic\tan\theta \qquad \Rightarrow \qquad \begin{cases} \cos\theta = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}} \\ \sin\theta = \frac{i\frac{V}{c}}{\sqrt{1-\frac{V^2}{c^2}}} \end{cases} .$$

Se si definisce ora $\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$ la trasformazione diventa:

$$\begin{pmatrix} ict' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -i\frac{V}{c}\gamma & 0 & 0 \\ i\frac{V}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \qquad \Leftrightarrow \qquad \begin{cases} t' = (t - \frac{V}{c^2}x)\gamma \\ x' = (x - Vt)\gamma \\ y' = y \\ z' = z \end{cases} .$$

Siccome la convenzione $r = (ict, x, y, z)$ non è quella comunemente utilizzata ma è stata adoperata solamente per evidenziare il carattere geometrico della trasformazione, è necessario ricondursi alla convenzione consueta dove $r = (ct, x, y, z)$ e come si è già detto è necessario effettuare un cambio di base tramite le matrici della (1.4.3):

$$P^{-1} \begin{pmatrix} \gamma & -i\frac{V}{c}\gamma & 0 & 0 \\ i\frac{V}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} P = \begin{pmatrix} \gamma & -\frac{V}{c}\gamma & 0 & 0 \\ -\frac{V}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

Questa è convenzionalmente riconosciuta come la trasformazione di Lorentz[1] Λ da K a K' con K' in moto a velocità v lungo l'asse x rispetto a K:

$$\Lambda = \begin{pmatrix} \gamma & -\frac{V}{c}\gamma & 0 & 0 \\ -\frac{V}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{cases} t' = (t - \frac{V}{c^2}x)\gamma \\ x' = (x - Vt)\gamma \\ y' = y \\ z' = z \end{cases} \qquad \gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}. \tag{1.4.6}$$

[1]Una trasformazione di Lorentz di questo tipo è più propriamente detta Boost di Lorentz.

Come si è già detto la trasformazione inversa di Λ, che consente di trasformare i vettori di K' in quelli di K, è ricavabile invertendo la matrice appena ottenuta.

Analoghe considerazioni possono essere fare per sistemi in moto in direzioni differenti a qui corrisponderanno trasformazioni descritte da rotazioni di piani differenti. Inoltre è possibile comporre tutte le trasformazioni di Lorentz tramite il prodotto righe per colonne con altre trasformazioni di Lorentz o rotazioni spaziali così da ottenere arbitrarie trasformazioni dei sistemi di riferimento inerziali.
Si osservi che in tutte queste trasformazioni appare un fattore che si è chiamato γ, caratteristico della trasformazione che si sta considerando. Questo fattore è dipendente esclusivamente dal modulo della velocità reciproca dei due sistemi di riferimento e assume un comportamento caratteristico e fondamentale per la teoria della relatività: in generale $\gamma \geq 1$ ed è pari a 1 solo per $V = 0$ mentre tende ad ∞ per V che tendono a $\pm c$.
Per $|V| > c$ il fattore $\gamma \in \mathbb{C}$, in questo modo ipotetici sistemi di riferimento in moto a velocità superluminare darebbero origine a trasformazioni che includono coordinate complesse e quindi di senso non fisico, questo suggerisce, come si osserverà più avanti, che c costituisca la velocità limite del moto.
Per piccoli valori di V rispetto a c, ossia nel limite in cui c è infinitamente grande, detto limite classico poiché coincide con la descrizione classica secondo cui la luce si propaga istantaneamente:

$$\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}} = 1 + \frac{1}{2}\frac{V^2}{c^2} + \frac{3}{8}\frac{V^4}{c^4} + o\left(\frac{V^4}{c^4}\right). \tag{1.4.7}$$

Così facendo, considerando $c \to \infty$, le trasformazioni di Lorentz divengono quelle di Galileo, per questo motivo la meccanica classica risulta solamente una prima approssimazione della corretta descrizione delle realtà.

1.4.2 Contrazione delle lunghezze e dilatazione dei tempi

Come si è già visto il concetto di tempo assoluto non è conciliabile con i postulati di Einstein, questo fatto è ora deducibile dalle Trasformazioni di Lorentz (1.4.6).
Si consideri in un sistema di riferimento inerziale K il moto rettilineo ed uniforme a velocità V di un corpo che emette un segnale luminoso ogni Δt. Considerando un secondo sistema di riferimento inerziale K' solidale a tale corpo e tale da osservarlo nell'origine degli assi spaziali, allora le trasformazioni di Lorentz risulteranno:

$$\begin{cases} t = (t' + \frac{V}{c^2}x')\gamma \\ x = (x' + Vt')\gamma \\ y = y' \\ z = z' \end{cases} \iff \begin{cases} t' = (t - \frac{V}{c^2}x)\gamma \\ x' = (x - Vt)\gamma \\ y' = y \\ z' = z \end{cases} \qquad \gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}.$$

Si possono quindi correlare i tempi di emissione misurati in K con quelli in K' considerando la trasformazione di Lorentz rispetto al corpo posto nell'origine del sistema K' (r'=(ct',0,0,0)):

$$\begin{cases} t = t'\gamma \\ x = Vt'\gamma \\ y = 0 \\ z = 0 \end{cases} \Rightarrow \Delta\tau = t'_1 - t'_0 = \frac{(t_1 - t_0)}{\gamma} = \frac{\Delta t}{\gamma} = \Delta t\sqrt{1-\frac{V^2}{c^2}}. \tag{1.4.8}$$

L'intervallo di tempo $\Delta\tau$, misurato nel sistema K' solidale al corpo, è detto tempo proprio di tale corpo e considerando che γ assume valori maggiori di 1 per ogni $V \neq 0$ risulta l'intervallo di tempo tra due eventi più breve misurabile tra tutti gli intervalli Δt misurati in ogni sistema di riferimento inerziale.

Si consideri adesso il corpo come un parallelepipedo esteso di dimensioni l_1 l_2 l_3 rispettivamente lungo gli assi $x,\ y,\ z$ del sistema K e λ_1 λ_2 λ_3 rispettivamente lungo gli assi $x',\ y',\ z'$ del sistema K'. Operativamente le dimensioni dell'oggetto sono misurate ponendo dei regoli diretti lungo gli assi di ogni sistema di riferimento e non in moto in essi, quindi ad un determinato istante si misurano contemporaneamente le posizioni delle estremità del corpo utilizzando i regoli. Si osservi che è necessario effettuare misure contemporanee della posizione di ogni coppia di estremità per assicurarsi di non misurare anche una componente dovuta allo spostamento del corpo avvenuto tra le misure non contemporanee. Si effettui la misura come è appena stata descritta nell'istante $t = 0$, si ottene, facendo uso delle trasformazioni di Lorentz:

$$\begin{cases} 0 = (t' + \frac{V}{c^2}\lambda_1)\gamma \\ l_1 = (\lambda_1 + Vt')\gamma \\ l_2 = \lambda_2 \\ l_3 = \lambda_3 \end{cases} \Rightarrow \begin{cases} t' = -\frac{V}{c^2}\lambda_1 \\ l_1 = (1 - \frac{V^2}{c^2})\lambda_1\gamma \\ l_2 = \lambda_2 \\ l_3 = \lambda_3 \end{cases} \Rightarrow \begin{cases} \lambda_1 = l_1\gamma = \frac{l_1}{\sqrt{1-\frac{V^2}{c^2}}} \\ \lambda_2 = l_2 \\ \lambda_3 = l_3 \end{cases} . \tag{1.4.9}$$

Si osserva quindi che anche le lunghezze non possono più essere considerate assolute, nella fattispecie la lunghezza di un corpo misurata nella direzione del suo moto risulta contratta rispetto alla medesima lunghezza misurata nel sistema K' solidale al corpo, detta lunghezza propria. Inoltre, analogamente a come si è osservato con il tempo proprio, si ha che $\gamma > 1$ per ogni $V \neq 0$ da cui si deduce che la lunghezza propria è la lunghezza maggiore tra tutte quelle misurabili in differenti sistemi di riferimento inerziali.

Per concludere è possibile constatare che il fenomeno della contrazione delle lunghezze influenza anche il volume di un corpo, nel caso del parallelepipedo sopra considerato il volume $\mathcal{V}_0$ misurato in K' diminuirà dello stesso fattore della lunghezza λ_1 se misurato in K:

$$V = l_1\ l_2\ l_3 = \frac{\lambda_1\ \lambda_2\ \lambda_3}{\gamma} = \frac{V_0}{\gamma}. \tag{1.4.10}$$

Quest'ultima osservazione mette in luce un'ulteriore differenza tra la teoria che si sta descrivendo e la teoria classica, infatti poiché il volume di un copro può mutare in base al suo moto nella teoria della relatività non è possibile fare uso del così detto corpo rigido.

1.4.3 Trasformazione delle velocità

Si è già osservato che le trasformazioni di Galileo non risultano compatibili con i postulati di Einstein in quanto ammettono che un osservatore potrebbe misurare la velocità della luce con un valore differente da c. Dalle trasformazioni di Lorentz è possibile ricavare quali sono le effettive trasformazioni delle velocità nella teoria della relatività.

Per prima cosa è necessario studiare come si trasformi un generico moto $(ct, \vec{r}(t))$ quando si cambia sistema di riferimento da K a K', tale per cui nel primo il secondo si muove a velocità V lungo l'asse x, si avrà allora la trasformazione:

$$\begin{cases} t = (t' + \frac{V}{c^2} r'_x(t'))\gamma \\ r_x = (r'_x(t') + Vt')\gamma \\ r_y = r'_y(t') \\ r_z = r'_z(t') \end{cases} \iff \begin{cases} t' = (t - \frac{V}{c^2} r_x(t))\gamma \\ r'_x = (r_x(t) - Vt)\gamma \\ r'_y = r_y(t) \\ r'_z = r_z(t) \end{cases} \qquad \gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

Se si calcolano le componenti della velocità in K' facendo uso della regola di Leibniz e delle trasformazioni di $(ct, \vec{r}(t))$ si ottiene:

$$\begin{aligned} v'_{x'} &= \frac{dr'_{x'}}{dt'} = \gamma \frac{dt}{dt'} \frac{d}{dt}(r_x(t) - Vt) = \gamma^2 \left(1 + \frac{Vv'_{x'}}{c^c}\right)(v_x - V) \\ v'_{y'} &= \frac{dr'_{y'}}{dt'} = \gamma \frac{dt}{dt'} \frac{d}{dt}(r_y(t)) = \gamma^2 \left(1 + \frac{Vv'_{x'}}{c^c}\right) v_y \\ v'_{z'} &= \frac{dr'_{z'}}{dt'} = \gamma \frac{dt}{dt'} \frac{d}{dt}(r_z(t)) = \gamma^2 \left(1 + \frac{Vv'_{x'}}{c^c}\right) v_z \\ \frac{dt}{dt'} &= \gamma \left(1 + \frac{Vv'_{x'}}{c^2}\right). \end{aligned}$$

Risolvendo algebricamente per le componenti $v'_{x'}$, $v'_{y'}$ e $v'_{z'}$ si ottengono le trasformazioni:

$$\begin{cases} v'_{x'} = \frac{v_x - V}{1 - \frac{Vv_x}{c^2}} \\ v'_{y'} = \frac{v_y}{\gamma(1 - \frac{Vv_x}{c^2})} \\ v'_{z'} = \frac{v_z}{\gamma(1 - \frac{Vv_x}{c^2})} \end{cases} \qquad \gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}. \tag{1.4.11}$$

Si noti che queste trasformazioni per le velocità rispettano il principio di costanza della velocità della luce: se si considera il moto di un raggio di luce lungo l'asse x di K si ha[1]

$$v' = \frac{c - V}{1 - \frac{Vc}{c^2}} = \frac{c - V}{c - V} c = c.$$

Infine si consideri un corpo in moto $\vec{r}(t)$ in K e si derivi il quadrivettore posizione rispetto al tempo proprio del corpo τ:

$$u = \frac{d}{d\tau}(ct, \vec{r}(t)) = \left(c\frac{dt}{d\tau}, \vec{v}\frac{dt}{d\tau}\right) = (c, \vec{v})\gamma_0 \qquad \gamma_0 = \frac{1}{\sqrt{1 - \frac{|\vec{v}|^2}{c^2}}} \tag{1.4.12}$$

dove $\frac{dt}{d\tau}$ è stato calcolato dalla (1.4.8).
Siccome le trasfromazioni di Lorentz sono applicazioni lineari e non dipendono da τ è possibile trasformare il quadrivettore posizione prima di derivarlo così che sia possibile trasformare u da K a K' tramite la matrice Λ (1.4.6), in tal caso si ottiene:

$$(c, v'_x, v'_y, v'_z)\gamma'_0 = \Lambda \begin{pmatrix} c \\ v_x \\ v_y \\ v_z \end{pmatrix} \gamma_0 = \left(\left(c - \frac{Vv_x}{c}\right)\gamma, (v_x - V)\gamma, v_y, v_z\right)\gamma_0$$

[1] Questo caso in realtà è generale poiché è sufficiente comporre la trasformazione di Lorentz con una rotazione degli assi spaziali perché ci si riconduca a questo caso.

dove γ dipende dalla velocità reciproca dei sistemi di rifermento mentre γ_0 e γ_0' dipendono dalla velocità del corpo in K e K'. Dalla prima componente trasformata si ottiene che:

$$\frac{\gamma_0'}{\gamma_0} = \left(1 - \frac{V v_x}{c^2}\right)\gamma \tag{1.4.13}$$

che se sostituita nelle altre componenti spaziali consente di ottenere le trasformazioni delle velocità (1.4.11), queste infatti non sono altro che le trasformazioni di Lorentz del quadrivettore u, chiamato quadrivelocità. Si osservi che il modulo quadro di u è costante:

$$|u|^2 = \frac{c^2 - |\vec{v}|^2}{1 - \frac{|\vec{v}|^2}{c^2}} = c^2.$$

1.5 Quadrivettori e spazio-tempo

Nella sezione 1.4.1 è stato introdotto il concetto di 4-vettore (quadrivettore) tramite il 4-vettore posizione, successivamente si è osservato che è possibile costruire altri 4-vettori, come la 4-velocità, che si trasformano anch'essi con le trasformazioni di Lorentz. Volendo essere più precisi si chiameranno 4-vettori tutti i vettori di $\mathbb{R}^4$ che si trasformano come il 4-vettore posizione.

La struttura non più euclidea dello spazio-tempo richiede una maggior attenzione nella trattazione dei 4-vettori: è quindi necessario distinguere 4-vettori controvarianti, indicati con $A^\mu = (A^0, A^1, A^2, A^3)$, e covarianti, indicati con $A_\mu = (A_0, A_1, A_2, A_3)$. I primi si trasformano con le trasformazioni di Lorentz Λ^μ_ν (1.4.6), mentre i secondi con la trasformazione inversa e trasposta, così che utilizzando la convenzione di Einstein[1] tali trasformazioni si scrivono:

$$A'^\mu = \Lambda^\mu_\nu A^\nu \qquad\qquad A'_\mu = (\Lambda^{-1})^\nu_\mu A_\nu. \tag{1.5.1}$$

Utilizzando 4-vettori covarianti e controvarianti è possibile ottenere una quantità scalare invariante per le trasformazioni di Lorentz:

$$A'^\mu B'_\mu = \Lambda^\mu_\alpha A^\alpha (\Lambda^{-1})^\beta_\mu B_\beta = A^\alpha \delta^\beta_\alpha B_\beta = A^\alpha B_\alpha$$

dove δ^β_α è la delta di Kronecker data dal prodotto righe per colonne della matrice della trasformazione di Lorentz con la sua inversa. La quantità appena ottenuta è il prodotto scalare dei 4-vettori A e B, questo è esprimibile, come si è visto nella sezione 1.4.1, attraverso l'uso della matrice metrica $g_{\mu\nu}$, questa osservazione consente di trovare una relazione tra componenti controvarianti e covarianti. Fissato un 4-vettore A_μ allora $\forall B^\nu$:

$$B^\mu A_\mu = B^\mu g_{\mu\nu} A^\nu \quad \Rightarrow \quad A_\mu = g_{\mu\nu} A^\nu \;\Leftrightarrow\; A^0 = A_0,\; A^i = -A_i \;\; i \in \{1, 2, 3\}. \tag{1.5.2}$$

Si consideri ora la 4-vettore posizione x_μ, come si è già visto il luogo dei punti nei quali $x^\mu x_\mu = 0$ rappresenta moti che avvengono alla velocità della luce ed è detto cono di luce. In analogia con il 4-vettore posizione ed il cono di luce (Fig. 1.3) è possibile classificare un qualsiasi 4-vettore in base al suo modulo quadro:

[1] La convenzione di Einstein è una notazione abbreviata dell'operazione di sommatoria: ogni coppia di indici ripetuti corrisponde ad una sommatoria sottintesa su tale indice. Un esempio può essere: $\sum_k A_{i,k} B^{j,k} = A_{i,k} B^{j,k}$

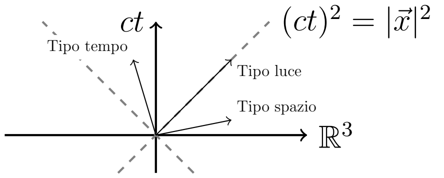

Figura 1.3: Cono di luce nel piano

$$A^\mu A_\mu = (A^0)^2 - |\vec{A}|^2 \begin{cases} > 0 & \text{Tipo tempo} \\ = 0 & \text{Tipo luce} \\ < 0 & \text{Tipo spazio} \end{cases}$$

Poiché come è già stato anticipato c rappresenta un limite per le velocità allora si può enunciare il seguente principio: *dall'origine di un sistema di riferimento l'informazione non può raggiungere 4-vettori posizione di tipo spazio, ossia al di fuori del cono di luce.*

Il concetto di 4-vettore può essere esteso ad oggetti a più indici che si trasformano con le trasformazioni di Lorentz detti 4-tensori, come la matrice metrica o la delta di Kronecker. Anche le componenti di questi oggetti possono essere covarianti o controvarianti, in base a come queste si trasformino, si fa quindi uso della medesima notazioni con indici bassi e alti tenendo conto che però un 4-tensore può avere anche indici misti.

Infine si osservi che definendo un campo scalare $\Phi(t, x, y, z)$ e calcolandone il 4-gradiente $\partial_\mu \Phi = (\frac{1}{c}\frac{\partial \Phi}{\partial t}, \frac{\partial \Phi}{\partial x^1}, \frac{\partial \Phi}{\partial x^2}, \frac{\partial \Phi}{\partial x^3})$ questo risulta essere un 4-vettore in coordinate covarianti, infatti secondo la regola di Leibniz:

$$\partial'_\mu \Phi = \partial_\nu \Phi \, \frac{\partial x^\nu}{\partial x'^\mu} = \partial_\nu \Phi \, (\Lambda^{-1})^\nu_\mu.$$

Analogamente è possibile calcolare la 4-divergenza di un 4-vettore, questa però risulta uno scalare infatti:

$$\partial'_\mu A'^\mu = \frac{\partial A^\delta}{\partial x^\nu}\frac{\partial x^\nu}{\partial x'^\mu}\Lambda^\mu_\delta = \frac{\partial A^\delta}{\partial x^\nu}(\Lambda^{-1})^\nu_\mu \Lambda^\mu_\delta = \partial_\nu A^\nu = \frac{1}{c}\frac{\partial A^0}{\partial t} + \vec{\nabla} \cdot \vec{A}.$$

Considerazioni di questo tipo possono essere fatte per altri operatori di differenziazione che possono anche agire su 4-tensori risultanti quindi in altri 4-tensori o in scalari, questo poiché sia gli operatori di differenziazione, tramite la regola di Leibniz, sia i 4-tensori, per definizione, si trasformano con Λ[1]. La notazione ad indici alti e bassi diventa quindi utile per evidenziare come si possano costruire quantità tensoriali o scalari invarianti sfruttando quanto appena detto: infatti per esempio la notazione $\partial_\mu A^\mu$ ricorda il prodotto scalare, che come si è visto da luogo ad un invariante di Lorentz come avviene anche per la 4-divergenza.

1.6 Trasformazioni del campo elettromagnetico

Nella sezione 1.3.1 si è visto come le equazioni di Maxwell (sezione 1.2.2) non risultino Galileo invarianti, si procederà ora a derivare le trasformazioni relativistiche del campo elettrico e magnetico in maniera analoga a come fece Einstein nel suo articolo del 1905[1].

[1] Questa osservazione è valida solamente in relatività speciale dove la trasformazione delle coordinate è lineare, in relatività generale infatti non è più vero che la derivata di un 4-tensore è ancora un 4-tensore o che la 4-divergenza è uno scalare.

Per prima cosa è opportuno ricavare le trasformazioni degli operatori di differenziazione considerando le trasformazioni di Lorentz (1.4.6) tra due sistemi K e K' in moto reciproco a velocità V. Usando la regola di Leibniz si ottiene:

$$\frac{\partial}{\partial x'} = \gamma\frac{\partial}{\partial x} - \gamma\frac{V}{c^2}\frac{\partial}{\partial t}, \quad \frac{\partial}{\partial y'} = \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial z'} = \frac{\partial}{\partial z}, \quad \frac{\partial}{\partial t'} = \gamma\frac{\partial}{\partial x} - \gamma\frac{V}{c^2}\frac{\partial}{\partial x}. \tag{1.6.1}$$

Siccome la relatività presuppone che la teoria di Maxwell sia corretta, è naturale partire proprio da questa, infatti si considereranno due delle quattro equazioni di Maxwell:

$$\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \qquad \vec{\nabla} \cdot \vec{B} = 0. \tag{1.6.2}$$

Se si studia la prima trasformando gli operatori di derivazione, componente per componente, secondo le (1.6.1), si può dedurre come debba trasformarsi il campo elettromagnetico.

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{\partial B_x}{\partial t} \qquad \qquad \frac{\partial E_z}{\partial y'} - \frac{\partial E_y}{\partial z'} = -\left(\frac{\partial B_x}{\partial t'} - v\frac{\partial B_x}{\partial x'}\right)\gamma \tag{1.6.3}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial z} = -\frac{\partial B_y}{\partial t} \qquad \Rightarrow \qquad \frac{\partial E_x}{\partial z'} - \frac{\partial E_z}{\partial x'} + \gamma\frac{V}{c^2}\frac{\partial E_z}{\partial t'} = -\left(\frac{\partial B_y}{\partial t'} - v\frac{\partial B_y}{\partial x'}\right)\gamma \tag{1.6.4}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -\frac{\partial B_z}{\partial t} \qquad \qquad \frac{\partial E_y}{\partial x'} - \gamma\frac{V}{c^2}\frac{\partial E_y}{\partial t'} - \frac{\partial E_x}{\partial y'} = -\left(\frac{\partial B_z}{\partial t'} - v\frac{\partial B_z}{\partial x'}\right)\gamma. \tag{1.6.5}$$

Infatti siccome le equazioni di Maxwell devono valere sia in K che in K' è possibile identificare quali termini devono corrispondere alle coordinate di $\vec{E}'$ e $\vec{B}'$ nelle equazioni trasformate affinché queste si riducano a nuove equazioni di Maxwell:

$$\frac{\partial E_x}{\partial z'} - \frac{\partial}{\partial x'}\left[\gamma\left(E_z + vB_y\right)\right] = -\frac{\partial}{\partial t'}\left[\gamma\left(B_y + \frac{v}{c^2}E_z\right)\right],$$
$$\frac{\partial}{\partial x'}\left[\gamma\left(E_y - vB_z\right)\right] - \frac{\partial E_x}{\partial y'} = -\frac{\partial}{\partial t'}\left[\gamma\left(B_z - \frac{v}{c^2}E_y\right)\right].$$

Così facendo si ottengono le trasformazioni di tutte le componenti tranne che per B_x:

$$E'_{x'} = E_x, \qquad E'_{y'} = (E_y - vB_z)\gamma, \qquad E'_{z'} = (E_z + vB_y)\gamma,$$
$$B'_{y'} = (B_y + \frac{V}{c^2}E_z)\gamma, \qquad B'_{z'} = (B_z - \frac{V}{c^2}E_y)\gamma.$$

Per conoscere come si trasforma B_x si sostituiscano le espressioni di E_y e E_z, ricavabili dalle trasformazioni appena ottenute, nella (1.6.3) così che utilizzando la (1.6.2) e il procedimento precedentemente adoperato si ottenga:

$$\frac{\partial E'_z}{\partial y'} - \frac{\partial E'_y}{\partial z'} = v(\vec{\nabla}' \cdot \vec{B}') - \frac{\partial B_x}{\partial t'} = -\frac{\partial B_x}{\partial t'} \qquad \Rightarrow \qquad B'_{x'} = B_x.$$

Queste trasformazioni, essendo ricavate dall'uso combinato della relatività e dell'elettromagnetismo di Maxwell, sono quindi coerenti con entrambe.

Capitolo 2

Meccanica relativistica

2.1 Il principio di minima azione

La meccanica classica, brevemente illustrata nel capitolo 1.2.1, descrive il moto di sistemi di corpi tramite l'equazione di Newton, si può dimostrare che questa però è del tutto equivalente ad un principio più generale noto come *principio di minima azione*.
Questo principio assume che per ogni sistema meccanico sia possibile costruire una funzione $\mathcal{L}(q_1...q_n, \dot{q}_1...\dot{q}_n, t)$ detta *lagrangiana*, dove $q_1...q_n$ sono le coordinate generalizzate e $\dot{q}_1...\dot{q}_n$ sono le velocità generalizzate date dagli n gradi di libertà del sistema, che è uso comune indicare solamente con q e $\dot{q}$.
In questo modo il principio di minima azione afferma che: se si suppone che ad un istante t_1 iniziale e ad un istante finale t_2 il sistema si trovi a coordinate fissate l'evoluzione del sistema avverrà secondo un moto $(q(t), \dot{q}(t))$ tale che il funzionale

$$\mathcal{S}[q(t)] = \int_{t_1}^{t_2} \mathcal{L}(q(t), \dot{q}(t), t)\ dt, \tag{2.1.1}$$

detto *azione*, assuma un valore stazionario, nella fattispecie si può dimostrare che per intervalli sufficientemente piccoli assume il valore minimo.

Si vogliono ora ottenere le equazioni che consentono, dato questo principio, di ottenere le equazioni del moto: si consideri una funzione $q(t) = (q_1(t), ..., q_n(t))$ e una seconda funzione arbitraria $\delta q(t) = (\delta q_1(t), ..., \delta q_n(t))$, detta variazione di $q(t)$, tale che $\delta q(t_1) = \delta q(t_2) = 0$ e che assuma valori piccoli nel suo dominio $[t_1, t_2]$. In questo modo si può costruire una seconda funzione $q(t) + \delta q(t)$ che soddisfa ancora le ipotesi per cui devono essere fissate le coordinate agli istanti t_1 e t_2.
La variazione δS, data dalla variazione δq, deve risultare nulla affinché $q(t)$ sia la funzione che rende stazionaria l'azione, così che il principio di minima azione sia equivalente a richiedere che

$$\int_{t_1}^{t_2} \mathcal{L}(q+\delta q, \dot{q}+\delta\dot{q}, t)\ dt - \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t)\ dt = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left[\frac{\partial \mathcal{L}}{\partial q_i}\delta q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\delta \dot{q}_i\right] dt = 0, \tag{2.1.2}$$

nella quale si è espressa esplicitamente la variazione dell'azione rispetto alla variazione δq e alla sua derivata prima $\delta\dot{q}$ tramite uno sviluppo in serie al prim'ordine con le derivate parziali della lagrangiana $\frac{\partial \mathcal{L}}{\partial q_i}$, $\frac{\partial \mathcal{L}}{\partial \dot{q}_i}$.

Integrando per parti gli addendi contenenti $\delta\dot{q}_i = \frac{d}{dt}\delta q_i$ e ricordando che $\delta q(t)$ si annulla in t_1 e in t_2, si ottiene

$$\int_{t_1}^{t_2}\sum_{i=1}^{n}\left[\frac{\partial\mathcal{L}}{\partial q_i}\delta q_i - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\delta q_i\right]\,dt + \sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\delta\dot{q}_i\bigg|_{t_1}^{t_2} = \int_{t_1}^{t_2}\sum_{i=1}^{n}\left[\frac{\partial\mathcal{L}}{\partial q_i} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right]\delta q_i\;dt = 0.$$

Poiché questo integrale deve annullarsi per ogni δq che rispetti le ipotesi fin qui fatte è necessario che si annulli identicamente l'integrando, nella fattispecie ogni addendo singolarmente siccome ognuno di questi è moltiplicato per una componente di $\delta q(t)$ che è una funzione arbitraria.
Si ottengono quindi le *equazioni di Eulero Lagrange* che consentono di determinare $q(t)$:

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i} - \frac{\partial\mathcal{L}}{\partial q_i} = 0 \qquad \forall i \in 1, 2, ..., n. \tag{2.1.3}$$

Date le coordinate generalizzate $q(t) = (q_1(t), ..., q_n(t))$ si definisco i momenti generalizzati associati a tali coordinate

$$p_{q_i} = \frac{\partial\mathcal{L}}{\partial\dot{q}_i}. \tag{2.1.4}$$

Se la lagrangiana non dipende esplicitamente da una coordinata $q_i(t)$ allora si dice che questa è una coordinata ciclica e dalle equazioni di Eulero Lagrange segue che sulla curva del moto il momento associato a tale coordinata resta costante nel tempo.
Oltre ai momenti è utile definire l'energia meccanica del sistema

$$E = \sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\dot{q}_i - \mathcal{L}, \tag{2.1.5}$$

che risulta una quantità conservata durante il moto se $\mathcal{L}$ non dipende esplicitamente dal tempo, infatti derivando rispetto al tempo la lagrangiana sulla curva del moto si ha:

$$\begin{aligned}\frac{d\mathcal{L}}{dt}\bigg|_{q(t)} &= \sum_i \frac{\partial\mathcal{L}}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\ddot{q}_i + \frac{\partial\mathcal{L}}{\partial t} = \sum_i\left(\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right)\dot{q}_i + \sum_i p_i\ddot{q}_i + \frac{\partial\mathcal{L}}{\partial t}\\ &= \sum_i \dot{p}_i\dot{q}_i + \sum_i p_i\ddot{q}_i + \frac{\partial\mathcal{L}}{\partial t} = \sum_i \frac{d}{dt}(p_i\dot{q}_i) + \frac{\partial\mathcal{L}}{\partial t} \qquad \Rightarrow \qquad \frac{dE}{dt}\bigg|_{q(t)} = -\frac{\partial\mathcal{L}}{\partial t}.\end{aligned}$$

Come si è già visto l'azione di un sistema è una quantità estremamente importante per la meccanica lagrangiana, questa però può essere intesa anche come funzione delle coordinate e del tempo. In questo caso a differenza del funzionale (2.1.1) si mantengono libere le coordinate e gli istanti finali, che divengono le variabili di dipendenza dell'azione, e si calcola l'integrale lungo la curva del moto, così da avere la funzione

$$S(\tilde{q}, t_2) = \int_{t_1}^{\tilde{q}, t_2}\mathcal{L}(q, \dot{q}, t)\;dt. \tag{2.1.6}$$

Questa è nota come *funzione principale di Hamilton* o funzione d'azione ed è indicata con la lettera S normale per distinguerla dal funzionale d'azione $\mathcal{S}$.
Si consideri ora una variazione delle coordinate finale $\tilde{q}_i \to \tilde{q}_i + \delta\tilde{q}_i$, poiché ogni coppia di coordinate iniziali e finali determina tramite il principio di minima azione la traiettoria del moto, a $\delta\tilde{q}$ corrisponde una variazione $\delta q(t)$ di tale curva che mantiene invariata solamente

le coordinate nell'istante iniziale. La variazione della funzione d'azione si calcola quindi come per il funzionale azione e si ottiene:

$$\delta S = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i \right] dt + \sum_{i=1}^{n} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i \bigg|_{t_1}^{t_2}.$$

Ricordando che questo integrale è calcolato su curve del moto si ha che l'integrando deve annullarsi, poiché devono essere soddisfatte le equazioni di Eulero Lagrange, mentre come si appena osservato $\delta q_i(t_1) = 0$ ma $\delta q_i(t_2) = \delta \tilde{q} \neq 0$, per cui si deduce che:

$$\delta S = \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i \bigg|_{t_2} \quad \Rightarrow \quad \frac{\partial S}{\partial \tilde{q}_i} = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = p_{q_i}.$$

Se ora si considera la derivata totale rispetto al tempo di $S(q, t)$, per il teorema fondamentale del calcolo integrale, si ha che questa è proprio la lagrangiana $\mathcal{L}$ e quindi :

$$\frac{dS}{dt} = \sum_{i=1}^{n} \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t} = \mathcal{L} \quad \Rightarrow \quad \frac{\partial S}{\partial t} = \mathcal{L} - \sum_{i=1}^{n} \frac{\partial S}{\partial q_i} \dot{q}_i = -E.$$

Si osservi che quanto appena descritto necessita solamente del principio di minima azione e di nessun'altra assunzione sperimentale. Per questo motivo questa trattazione della meccanica è utilizzabile anche nell'ambito della relatività dove sarà la lagrangiana a dover tenere conto di quanto descritto nel capitolo 1.

2.2 La lagrangiana relativistica

Si procederà in questa sezione ricavando la lagrangiana relativistica e quanto segue dalla forma di questa. Siccome una generica lagrangiana è data dalla somma di una parte di corpo libero $\mathcal{L}_{lib}$ con una seconda di interazione $\mathcal{L}_{int}$ è opportuno in primis dedurre la forma di $\mathcal{L}_{lib}$ studiando una particella libera.

2.2.1 La particella libera

Si consideri il moto di una particella nello spazio tempo, il principio di minima azione determina tale moto indipendentemente dalla scelta del sistema di riferimento in cui esso è descritto. Come si è già visto nella sezione 1.4.1 le trasformazioni di Lorentz conservano le lunghezze nello spazio-tempo, per questo motivo tale lunghezza per una curva del moto è un candidato ad essere l'azione relativistica. Una curva di questo tipo può essere parametrizzata dal tempo misurato nel sistema in cui si osserva il moto, così facendo la curva sarà data dal 4-vettore posizione $x^\mu(t) = (ct, \vec{x}(t))$. La lunghezza della curva, e la supposta azione, è quindi:

$$\mathcal{S}[x^\mu(t)] = \alpha \int_{t_1}^{t_2} |\dot{x}^\mu(t)| \, dt = \alpha c \int_{t_1}^{t_2} \sqrt{1 - \frac{|\vec{v}|^2}{c^2}} \, dt = \alpha c \int_{\tau_1}^{\tau_2} d\tau, \tag{2.2.1}$$

dove si è indicato $\vec{v} = \dot{\vec{x}}$ che corrisponde al vettore tridimensionale velocità e con τ il tempo proprio della particella (sezione 1.4.2), ossia il tempo misurato in un sistema di riferimento inerziale istantanemtne solidale al moto, ossia nel quale la velocità è nulla.

A questo punto si deve verificare se questa azione nel limite classico, ossia quando il rapporto $\frac{|\vec{v}|^2}{c^2}$ tende a 0, è riconducibile all'azione classica. Ricordando che la lagrangiana classica della particella libera è data da $\frac{m|\vec{v}|^2}{2}$:

$$\mathcal{L} = \alpha c\sqrt{1 - \frac{|\vec{v}|^2}{c^2}} \longrightarrow \alpha c - \frac{\alpha|\vec{v}|^2}{2c^2} \qquad \text{per } \frac{|\vec{v}|^2}{c^2} \to 0.$$

A meno di costanti moltiplicative e additive che non mutano la forma delle soluzioni delle equazioni di Eulero Lagrange, il limite classico ottenuto è quindi riconducibile alla lagrangiana classica se $\alpha = -mc$.
Questa evidenza quindi corrobora[1] l'ipotesi iniziale così che sia lecito affermare che la lagrangiana relativistica della particella libera sia:

$$\mathcal{L} = -mc^2\sqrt{1 - \frac{|\vec{v}|^2}{c^2}}. \tag{2.2.2}$$

2.2.2 Energia e momenti

Identificata la lagrangiana relativistica per la particella libera è ora possibile sfruttare tutti gli strumenti illustarti nella sezione 2.1 per studiare grandezze quali l'energia e i momenti di un sistema.

La prima quantità che si può studiare è il momento associato alle coordinate cartesiane, questo è il vettore $(\frac{\partial \mathcal{L}}{\partial v_x}, \frac{\partial \mathcal{L}}{\partial v_y}, \frac{\partial \mathcal{L}}{\partial v_z})$ noto come impulso $\vec{p}$: dalla (2.2.2) risulta

$$\vec{p} = m\vec{v}\gamma = \frac{m\vec{v}}{\sqrt{1 - \frac{|\vec{v}|^2}{c^2}}}. \tag{2.2.3}$$

Anche relativisticamente, come in meccanica classica, l'impulso di una particella libera risulta conservato, infatti tutte le coordinate della particella libera sono cicliche. Questo implica che il moto di una particella libera è rettilineo ed uniforme.
Si osservi che nel limite classico, in cui $|\vec{v}| << c$, l'impulso diviene la classica espressione $\vec{p} = m\vec{v}$, mentre per velocità tendenti in modulo a c l'impulso tende ad infinito.

Analogamente è possibile calcolare l'energia della particella libera $E = \vec{p}\cdot\vec{v} - \mathcal{L}$ che risulta una quantità conservata poiché non vi è dipendenza esplicita dal tempo nella lagrangiana. Calcolandola dalla (2.2.2) si ottiene:

$$E = mc^2\gamma = \frac{mc^2}{\sqrt{1 - \frac{|\vec{v}|^2}{c^2}}}. \tag{2.2.4}$$

In questo caso è opportuno osservare che per particelle libere a riposo l'energia non è nulla, come accade in meccanica classica, bensì assume il valore $E_0 = mc^2$ noto come energia a riposo della particella. Sottraendo all'energia totale del corpo il termine di energia a riposo si ottiene la quantità di energia esclusivamente dovuta al moto, ossia l'energia cinetica:

$$T = mc^2\gamma - mc^2 = mc^2(\gamma - 1) \tag{2.2.5}$$

[1] In realtà solamente un confronto tra i risultati teorici che si dedurranno e quelli sperimentali può giustificare l'ipotesi iniziale.

Anche in questo caso studiando il limite classico si ritrova l'energia cinetica classica mentre nel limite di velocità prossime a c si ottiene che l'energia del corpo tende ad infinito. Quest'ultima osservazione consente di affermare senza alcuna ombra di dubbio che c è la velocità limite di ogni moto, infatti un qualsiasi corpo dotato di massa necessita di un'energia infinita per poter essere accelerato fino a tale velocità.

Si consideri ora il 4-gradiente della funzione d'azione, siccome $S(q, t)$ è uno scalare Lorentz invariante, questo è un 4-vettore covariante (sezione 1.5). Come si è mostrato nella sezione 2.1, l'azione è legata ad energia e momenti dalle sue derivate parziali: $\vec{\nabla} S = \vec{p}$ e $\frac{\partial S}{\partial t} = -E$. Facendo uso di queste relazioni si ottiene che:

$$\frac{\partial S}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial S}{\partial t}, \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z}\right) = \left(-\frac{E}{c}, \vec{p}\right),$$

si definisce l'analogo 4-vettore controvariante e cambiato di segno 4-impulso

$$p^\mu = -\frac{\partial S}{\partial x_\mu} = \left(\frac{E}{c}, \vec{p}\right). \tag{2.2.6}$$

Poiché energia ed impulso costituiscono un 4-vettore è opportuno, analogamente a quanto accade per la 4-velocità, determinare come queste due quantità si trasformino con un cambio di riferimento. Considerando la trasformazione $p'^\mu = \Lambda^\mu_\nu p^\nu$ si ha:

$$p'_x = \left(p_x - \frac{V}{c^2}E\right)\gamma, \qquad p'_y = p_y, \qquad p'_z = p_z, \qquad E' = (E - Vp_x)\gamma. \tag{2.2.7}$$

Il formalismo quadrivettoriale garantisce che il modulo del 4-impulso sia Lorentz invariante, è quindi utile determinare tale quantità in un sistema di riferimento inerziale solidale al moto della particella, in questo sistema l'impulso e l'energia cinetica sono nulli per cui, detto p'^μ il 4-impulso in tale sistema, si ha

$$p'^\mu p'_\mu = \frac{E_0^2}{c^2} = m^2c^2.$$

Siccome tale modulo è lo stesso in ogni sistema di riferimento inerziale si ottene la relazione tra l'energia e l'impulso di una particella:

$$p^\mu p_\mu = \frac{E^2}{c^2} - |\vec{p}|^2 = m^2c^2. \tag{2.2.8}$$

Da quest'ultima relazione, se non si determinano i valori delle derivate parziali dell'azione, è possibile ottenere *l'equazione di Hamilton Jacobi* relativistica:

$$p^\mu p_\mu = \frac{\partial S}{\partial x^\mu}\frac{\partial S}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial S}{\partial t}\right)^2 - \left(\frac{\partial S}{\partial x}\right)^2 - \left(\frac{\partial S}{\partial y}\right)^2 - \left(\frac{\partial S}{\partial z}\right)^2 = m^2c^2. \tag{2.2.9}$$

2.3 Sistemi di più particelle

Si vuole adesso studiare un insieme di N particelle che formano un sistema isolato.
Per cominciare si consideri una sola particella delle N descritta del 4-vettore x^μ, se a tutto il sistema di fa compiere una rotazione infinitesima nello spazio-tempo questo 4-vettore

muterà di una quantità infinitesima. La variazione di questo 4-vettore può essere espressa in termini di una trasformazione lineare con coefficienti infinitesimi:

$$\delta x^\mu = x_\nu \delta\Omega^{\mu\nu}.$$

Una rotazione quadridimensionale non varia le lunghezze del 4-vettore posizione così, trascurando termini infinitesimi di ordine superiore al primo, si ha:

$$x^\mu x_\mu = (x^\mu + \delta x_\mu)(x_\mu + \delta x^\mu) = x^\mu x_\mu + 2x^\mu \delta x_\mu + \delta x_\mu \delta x_\mu \ \Rightarrow \ x^\mu x^\nu \delta\Omega_{\mu\nu} = 0.$$

Poiché x^μ non è stato scelto in maniera specifica quanto appena detto vale per ogni vettore in $\mathbb{R}^4$. Si osservi che siccome $x^\mu x^\nu$ forma un tensore simmetrico e la sua contrazione su $\delta\Omega_{\mu\nu}$ è nulla allora quest'ultimo deve esse un tensore antisimmetrico[1]:

$$\delta\Omega_{\mu\nu} = -\delta\Omega_{\nu\mu}. \tag{2.3.1}$$

Si è mostrato nella sezione 2.1 che una variazione delle coordinate spaziotemporali rispetto alle quali è calcolata la funzione d'azione genera una variazione di questa. Considerando le precedenti osservazioni per ogni particella del sistema in questo caso si ottiene:

$$\delta S = -\sum_{\text{Particelle}} p^\mu \delta x_\mu = -\delta\Omega_{\mu\nu} \sum_{\text{Particelle}} p^\mu x^\nu.$$ [2]

Se si scompone il tensore $p^\mu x^\nu$ nella somma di una parte simmetrica e una antisimmetrica, siccome questo è contratto su un tensore antisimmetrico, si ottiene che la sola parte antisimmetrica da un contributo non nullo:

$$\delta S = -\frac{\delta\Omega_{\mu\nu}}{2} \sum_{\text{Particelle}} (p^\mu x^\nu - x^\nu p^\mu).$$

Poiché lo spazio-tempo è isotropo qualsiasi rotazione di tutto il sistema, che è isolato, non ne genera una mutazione, per questo motivo se si considerano $\Omega_{\mu\nu}$ come alcune coordinate generalizzate queste non possono apparire in $\mathcal{L}$, il che le rende coordinate cicliche. Ne segue che il loro momento associato $\frac{\partial S}{\partial\Omega_{\mu\nu}}$ deve conservarsi.
Definendo 4-tensore *momento angolare* la quantità:

$$M^{\mu\nu} = \sum_{\text{Particelle}} (p^\mu x^\nu - x^\nu p^\mu) \tag{2.3.2}$$

questa deve essere conservata durante l'evoluzione del sistema. Si osservi che questo tensore è composto in alcune componenti dalla definizione classica di momento angolare $\vec{M}$, nella fattispecie:

$$M_x = M^{23},\ M_y = -M^{13},\ M_z = M^{12}.$$

Le componenti $M^{01},\ M^{02},\ M^{03}$ formano un secondo vettore pari a $\sum(t\vec{p} - \frac{E\vec{x}}{c^2})$. Siccome deve conservarsi il 4-tensore momento angolare anche questi due vettori dovranno conservarsi. Nella fattispecie, siccome anche l'energia del sistema deve conservarsi, il

[1] minidimostrazione.

[2] In questa formula, e da qui in avanti, sono omessi gli indici che indicano la particella di appartenenza per evitare confusioni con gli indici delle componenti, ogni sommatoria sottintende che questa avviene su ogni particella del sistema.

secondo vettore diviso per l'energia totale del sistema deve essere ancora un vettore costante:

$$\frac{\sum E\vec{x}}{\sum E} - t\frac{c^2\sum \vec{p}}{\sum E} = \text{costante}. \tag{2.3.3}$$

Si definiscono rispettivamente i vettori

$$\vec{x}_{CM} = \frac{\sum E\vec{x}}{\sum E} \qquad \vec{v}_{CM} = \frac{c^2\sum \vec{p}}{\sum E} \tag{2.3.4}$$

centro di massa relativistico e velocità del centro di massa relativistica. Dalla (2.3.3) si ha che il vettore $\vec{x}_{CM}$ consente di descrivere il moto traslazionale dell'intero sistema come di un solo punto che si muove alla velocità $\vec{v}_{CM}$.
Qual'ora le velocità dei punti del sistema fossero piccole rispetto a c l'energia a riposo di ogni particella risulterebbe molto maggiore di quella cinetica e la formula delle coordinate del centro di massa relativistico tenderebbero a quella classica $\frac{\sum m\vec{x}}{\sum m}$.

Capitolo 3

L'elettromagnetismo nella teoria della relatività

Appendice A

Lemmi e teoremi per le trasformazioni di Lorentz

Nelle seguenti pagine sono riportati alcuni risultati utili per la derivazione delle trasformazioni di Lorentz tra cui una dimostrazione dell'affinità delle trasformazioni di Lorentz riportata da Fock nel suo manuale[2].

A.1 Risultati generali

Lemma 1. *Siano* $v, u \in \mathbb{R}^n$, $<,>$ *sia il prodotto scalare euclideo e siano definite* $f_u : \mathbb{R}^n \to \mathbb{R}$, *tale che* $f_u(v) = < u, v >$, *e* $q : \mathbb{R}^n \to \mathbb{R}$ *forma quadratica tale che* $q(v) = < v, Av >$, *con* $A \in M_{n\times n}(\mathbb{R})$ *simmetrica. Se, fissato* $\bar{v}$ *tale che* $f_u(\bar{v}) = 0$ *implica che* $q(\bar{v}) = 0$,*allora* $\exists \psi \in \mathbb{R}^n$ *tale che* $q(v) = f_u(v) < \psi, v > \forall v \in \mathbb{R}^n$.

Dimostrazione. Si osservi che è possibile decomporre v in una componente ortogonale ad u ed una parallela:

$$v = v^{\parallel} + v^{\perp} = \frac{u\ < u, v >}{\|u\|^2} + v - \frac{u\ < u, v >}{\|u\|^2} + \tag{A.1.1}$$

si esprima quindi $q(v)$ secondo questa decomposizione:

$$\begin{aligned} q(v) &= < v^{\parallel} + v^{\perp}, A(v^{\parallel} + v^{\perp}) > \\ &= < v^{\parallel}, Av^{\parallel} > +2 < v^{\perp}, Av^{\parallel} > + < v^{\perp}, Av^{\perp} > \end{aligned}$$

il termine $< v^{\perp}, Av^{\perp} >$ è per definizione $q(v^{\perp})$ e poichè per costruzione di $v^{\perp}$ si ha $f_u(v^{\perp}) = 0$, per le ipotesi $q(v^{\perp})$deve annullarsi per cui:

$$q(v) =< v^{\parallel}, Av^{\parallel} > +2 < v^{\perp}, Av^{\parallel} >$$

sostituiendo le espressioni complete di $v^{\perp}$ e $v^{\parallel}$, dalla (A.1.1), si ha:

$$\begin{aligned} q(v) &=< u, Au > \left(\frac{< u, v >}{\|u\|^2}\right)^2 + 2 < v^{\perp}, Au > \frac{< u, v >}{\|u\|^2} \\ &= < u, Au > \left(\frac{< u, v >}{\|u\|^2}\right)^2 + 2 < v, Au > \frac{< u, v >}{\|u\|^2} - 2 < u, Au > \left(\frac{< u, v >}{\|u\|^2}\right)^2 \end{aligned}$$

Infine se si raccoglie un termine $< u, v >$ e si sommano i termini uguali si ottiene un'espressione per $q(v)$ composta da un termine lineare rispetto a v moltiplicato per il prodotto scalare $< u, v >$:

$$q(v) =< u, v > \left[2\frac{< v, Au >}{\|u\|^2} - \frac{< u, Au >< u, v >}{\|u\|^4} \right] \qquad \forall v \in \mathbb{R}^4 \tag{A.1.2}$$

□

Lemma 2. *Sia $v \in \mathbb{R}^4$, per comodità si scriverà $v = (v_0, v_1, v_2, v_3) = (v_0, \vec{v})$, tale che $\vec{v} \neq 0$. Sia quindi $q : \mathbb{R}^4 \to \mathbb{R}$ una forma quadratica tale che, se fissato $\bar{v}$ con $\bar{v}_0^2 - |\vec{\bar{v}}|^2 = 0$ implica $q(\bar{v}) = 0$, allora $\exists \lambda$ tale che $q(v) = \lambda(v_0^2 - |\vec{v}|^2)$ $\forall v \in \mathbb{R}^4$.*

Dimostrazione. Si osservi che $q(v)$ può essere scomposto, seguendo le regole del prodotto righe per colonne $q(v) = v^t Av$ (dove $A \in M_{3x3}(\mathbb{R})$ simmetrica), nella seguente somma:

$$q(v) = a_{00}v_0^2 + 2v_0(a_{01}v_1 + a_{02}v_2 + a_{03}v_3) + \hat{q}(\vec{v})$$

dove $\hat{q}$ è ancora una forma quadratica in $\mathbb{R}^3$ che agisce su $\vec{v}$.
Si consideri ora un vettore $\bar{v} \in \mathbb{R}^4$ e tale che considerando v allora $\bar{v} = (-|\vec{v}|, \vec{v})$; per costruzione$(\bar{v}_0)^2 - |\vec{\bar{v}}|^2 = |\vec{v}|^2 - |\vec{v}|^2 = 0$, per cui per le ipotesi assunte si ha $q(\bar{v}) = 0$.
Essendo $q(\bar{v}) = 0$ si può scrivere:

$$\begin{aligned} q(v) &= q(v) - q(\bar{v}) \\ &= a_{00}(v_0^2 - |\vec{v}|^2) + 2(v_0 - |\vec{v}|)(a_{01}v_1 + a_{02}v_2 + a_{03}v_3) + \hat{q}(\vec{v}) - \hat{q}(\vec{\bar{v}}) \\ &= a_{00}(v_0^2 - |\vec{v}|^2) + 2(v_0 - |\vec{v}|)(a_{01}v_1 + a_{02}v_2 + a_{03}v_3) \end{aligned} \tag{A.1.3}$$

inoltre calcolando la (A.1.3) per $\bar{v}$ si ha:

$$q(\bar{v}) = -2|\vec{v}|(a_{01}v_1 + a_{02}v_2 + a_{03}v_3) = 0$$

Per ipotesi $|\vec{v}| \neq 0$ per cui $a_{01} = a_{02} = a_{03} = 0$ che dimostra il lemma poiché implica:

$$q(v) = a_{00}(v_0^2 - |\vec{v}|^2) \quad \forall v \in \mathbb{R}^4 \tag{A.1.4}$$

□

A.2 Una dimostrazione della linearità delle trasformazioni di Lorentz

Siano $x, x' \in \mathbb{R}^4$ due vettori dello spazio-tempo misurati in due sistemi di riferimento inerziali K e K', si vogliono determinare le proprietà dell'applicazione $f : \mathbb{R}^4 \to \mathbb{R}^4$ che trasforma i vettori misurati in K negli stessi vettori misurati in K' e viceversa considerando validi il principio di Relatività e il principio di costanza della velocità della luce.

Per prima cosa si supporà che f sia sufficentemente regolare, almeno C^2, ed è necessario che sia invertibile poichè deve poter trasformare sia da K a K', sia viceversa: questa condizione è equivalente a richiedere che det $J_f(x) \neq 0$ $\forall x \in \mathbb{R}^4$.

Si considerino ora $\xi, \gamma \in \mathbb{R}^4$, rispettivamente detti punto iniziale e velocità della parametrizzazione, che consentono di parametrizzare un moto rettilineo uniforme nel seguente modo:

$$x = \xi + s\gamma \qquad s \in \mathbb{R}$$

infatti è possibile esplicitare la dipendenza di x_0 da s per ottenere la forma di un moto rettilineo uniforme, per questo motivo è necessario che $\gamma_0 \neq 0$.

$$s = \frac{x_0 - \xi_0}{\gamma_0} \qquad \Rightarrow \qquad x_i = \xi_i + \frac{x_0 - \xi_0}{\gamma_0}\gamma_i \qquad i = 1, 2, 3$$

Secondo il principio di relatività se si calcola $f(x) = x'$ è necessario che $x' = \xi' + s'\gamma'$ così che un moto rettilineo uniforme resti tale in ogni sistema di riferimento inerziale. Derivando $f(\xi + s\gamma)$ rispetto al parametro s' si ottiene:

$$\frac{dx'}{ds'} = \gamma' = J_f(\xi + s\gamma)\frac{ds}{ds'}\gamma$$

si consideri ora la i-esima componente di γ', dove $i = 0, 1, 2, 3$, e la si divida per γ_0:

$$\frac{\gamma'_i}{\gamma'_0} = \frac{\sum_{k=0}^{3} \partial_k f_i(\xi + s\gamma)\gamma_k}{\sum_{k=0}^{3} \partial_k f_0(\xi + s\gamma)\gamma_k}$$

dove $\partial_k = \frac{\partial}{\partial x_k}$. Da ora in poi si utilizzerà la convenzione degli indici ripetuti di Einstein per cui $\sum_k x_k y_k = x_k y_k$.
Il rapporto $\frac{\gamma'_i}{\gamma'_0}$ è costante da cui segue immediatamente che $\frac{d}{ds}\frac{\gamma'_i}{\gamma'_0} = 0$ e quindi:

$$\frac{d}{ds}\frac{\partial_k f_i(\xi + s\gamma)\gamma_k}{\partial_j f_0(\xi + s\gamma)\gamma_j}$$
$$= \frac{\left(\partial^2_{hn} f_i(\xi + s\gamma)\gamma_h\gamma_n\right)\left(\partial_k f_0(\xi + s\gamma)\gamma_k\right) - \left(\partial_k f_i(\xi + s\gamma)\gamma_k\right)\left(\partial^2_{hn} f_0(\xi + s\gamma)\gamma_h\gamma_n\right)}{\left(\partial_j f_0(\xi + s\gamma)\gamma_j\right)^2} = 0$$

che può annullarsi solo se si annulla il numeratore della frazione. Da questa osservazione e dal fatto che preso un $x \in \mathbb{R}^4$ esistono $\gamma \in \mathbb{R}^4, s \in \mathbb{R}$ tali che $x = \xi + s\gamma$ si ottiene:

$$\left(\partial^2_{hn} f_i(x)\gamma_h\gamma_n\right)\left(\partial_k f_0(x)\gamma_k\right) = \left(\partial_k f_i(x)\gamma_k\right)\left(\partial^2_{hn} f_0(x)\gamma_h\gamma_n\right) \qquad \forall x, \gamma \in \mathbb{R}^4 \quad i = 0, 1, 2, 3 \tag{A.2.1}$$

Siccome si è supposto inizialmente $\det J_f(x) \neq 0$ deve necessariamente esistere almeno un j tale che $\forall \gamma \in \mathbb{R}^4 \quad \partial_k f_j(x)\gamma_k \neq 0$, inoltre se si suppone di prendere $\bar{\gamma}$ tale che $\partial_k f_0(x)\bar{\gamma}_k = 0$ allora per la (A.2.1) necessariamente anche $\partial^2_{hn} f_0(x)\bar{\gamma}_h\bar{\gamma}_n = 0$.

Si applichi quindi il Lemma (1) il quale, in virtù del fatto che $\partial_k f_0(x)\bar{\gamma}_k = 0 \Rightarrow \partial^2_{hn} f_0(x)\bar{\gamma}_h\bar{\gamma}_n = 0$, consente di scrivere:

$$\partial^2_{hn} f_0(x)\gamma_h\gamma_n = < \psi(x), \gamma > \; \left(\partial_k f_0(x)\gamma_k\right) \qquad \forall x, \gamma \in \mathbb{R}^4 \quad i = 0, 1, 2, 3$$

che inserito nella (A.2.1) risulta in:

$$\left(\partial^2_{hn} f_i(x)\gamma_h\gamma_n\right)\left(\partial_k f_0(x)\gamma_k\right) = \left(\partial_k f_i(x)\gamma_k\right) \; < \psi(x), \gamma > \; \left(\partial_k f_0(x)\bar{\gamma}_k\right)$$

infine siccome $\gamma_0' \neq 0$ affinché sia x' sia un moto rettilineo uniforme, a maggior ragione la su espressione data dalla jacobiana di f soddisfa $(\partial_k f_0(x)\gamma_k \frac{ds}{ds'}) \neq 0$ ed è quindi possibile elidere da ambo i lati tale membro:

$$\partial^2_{hn} f_i(x)\gamma_h\gamma_n = (\partial_k f_i(x)\gamma_k) \ < \psi(x), \gamma > \qquad \forall x, \gamma \in \mathbb{R}^4 \quad i = 0, 1, 2, 3 \tag{A.2.2}$$

Si osservi che $\partial^2_{hn} f_i(x)$ è una matrice simmetrica per il teorema di Schwarz per cui dall'espressione appena ottenuta si può scrivere:

$$\partial^2_{hk} f_i(x) = \frac{1}{2}\left(\partial_k f_i(x)\psi_h(x) + \partial_h f_i(x)\psi_k(x)\right) \tag{A.2.3}$$

Si consideri ora x e x' parametrizzazioni del moto di un raggio di luce, per il principio di costanza della velocità della luce se in K x è parametrizzato sul cono di luce, lo stesso deve avvenire in K', il che è esprimibile matematicamente con le due condizioni:

$$\gamma_0^2 - (\gamma_1^2 + \gamma_2^2 + \gamma_3^2) = 0$$
$$(\gamma_0')^2 - [(\gamma_1')^2 + (\gamma_2')^2 + (\gamma_3')^2] = 0$$

e poiché si può esprimere ogni componente di γ' tramite la jacobiana di f e γ, come già si è fatto calcolando $\frac{\gamma_i'}{\gamma_0'}$:

$$(\partial_k f_0(x)\gamma_k)^2 - \left[\sum_{i=1}^{3} (\partial_k f_i(x)\gamma_k)^2\right] = 0$$

Il principio di costanza della velocità della luce consente di utilizzare il lemma (2) poiché l'annullarsi di una forma quadratica implica l'annullarsi dell'altra, da cui si ha che:

$$(\partial_k f_0(x)\gamma_k)^2 - \left[\sum_{i=1}^{3} (\partial_k f_i(x)\gamma_k)^2\right] = \lambda(x)\left[\gamma_0^2 - (\gamma_1^2 + \gamma_2^2 + \gamma_3^2)\right]$$

Sia ora g_{ij} una matrice tale per cui $g_{00} = 1$, $g_{ii} = -1$ per $i = 1, 2, 3$ e $g_{ij} = 0$ se $i \neq j$, così facendo:

$$\lambda(x)\left[\gamma_0^2 - (\gamma_1^2 + \gamma_2^2 + \gamma_3^2)\right] = \lambda(x) g_{kh}\gamma_k\gamma_h = g_{ij}(\partial_k f_i(x)\gamma_k)(\partial_h f_j(x)\gamma_h)$$

$$\Rightarrow \qquad \lambda(x) g_{hk} = g_{ij}(\partial_k f_i(x))(\partial_h f_j(x)) \tag{A.2.4}$$

Si derivi ora rispetto a x_s ad ambo i membri così da ottenere:

$$g_{hk}\partial_s\lambda = 2 g_{ij}\partial^2_{ks} f_i(x)\partial_h f_j(x)$$

sostituendo la derivata seconda parziale tramite la (A.2.3) questa espressione è semplificabile e si ottiene:

$$g_{hk}\partial_s\lambda = g_{ij}\left[\partial_k f_i(x)\psi_s + \partial_s f_i(x)\psi_k\right]\partial_h f_j(x) \tag{A.2.5}$$

infine applicando nuovamente la (A.2.4) si possono sostituire tutti i prodotti di derivate parziali:

$$g_{hk}\partial_s\lambda = g_{kh}\lambda\psi_s + g_{sh}\lambda\psi_k$$

Si considerino queste due differenti casistiche:

- Siano $k = h \neq s$ allora:

$$g_{hk}\partial_s\lambda = g_{kh}\lambda\psi_s \Rightarrow \partial_s\lambda = \psi_s\lambda$$

- Siano $k = h = s$ allora:

$$g_{hk}\partial_s\lambda = 2g_{kh}\lambda\psi_s \Rightarrow \partial_s\lambda = 2\psi_s\lambda$$

Queste due relazioni appena ottenute implicano però che $\lambda\psi_i = 0$ per $i = 0, 1, 2, 3$, λ però non può annullarsi altrimenti la trasformazione f mapperebbe ogni moto in un moto a velocità luminare, il che viola il Principio di Relatività.
Segue quindi che $\psi_i = 0$ per $i = 0, 1, 2, 3$, per cui dalla (A.2.3) risulta:

$$\partial^2_{kh} f_i(x) = 0 \qquad \forall i, k, h = 0, 1, 2, 3 \forall x \quad \in \mathbb{R}^4 \tag{A.2.6}$$

per cui $f(x)$ può essere esclusivamente una trasformazione affine.

Bibliografia

[1] A. Einstein. Sull'elettrodinamica dei corpi in movimento. *Annalen der Physik*, 17:891–921, 1905. Traduzione di S. Antoci.

[2] V. Fock. *The Theory of Space, Time and Gravitation*. Pergamon, 1964.

[3] J. Jeffers. Lost theorems of geometry. *The American Mathematical Monthly*, 107(9):800–812, 2000.

[4] Evgenij M. Lifšits Lev D. Landau. *Fisica teorica 2*. Editori Riuniti, 1976.