

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

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

1	Fondamenti di relatività speciale	3
1.1	La matematica dei sistemi di riferimento inerziali	3
1.1.1	I sistemi di riferimento inerziali	3
1.1.2	Trasformazioni di sistemi di riferimento inerziali	4
1.2	Fatti di Fisica Classica	7
1.2.1	La meccanica e le trasformazioni di Galileo	8
1.2.2	L'elettromagnetismo e le equazioni di Maxwell	10
1.3	L'articolo del 1905	11
1.3.1	La non invarianza delle Equazioni di Maxwell	11
1.3.2	I postulati di Einstein	13
1.4	Le trasformazioni di Lorentz	15
1.4.1	Derivazione	15
1.4.2	Contrazione delle lunghezze e dilatazione dei tempi	19
1.4.3	Trasformazione delle velocità	20
2	Meccanica relativistica	22
3	L'elettromagnetismo nella teoria della relatività	23
A	Sulla linearità delle trasformazioni di Lorentz	24
A.1	Prerequisiti matematici	24
A.2	La dimostrazione	25

Introduzione

Capitolo 1

Fondamenti di relatività speciale

1.1 La matematica dei sistemi di riferimento inerziali

Prima di poter descrivere la teoria della relatività ristretta e l'iter che portò alla sua fondazione è necessario definire e studiare i sistemi di riferimento. Di seguito si mostrerà quale struttura matematica è necessaria per descrivere tali sistemi e le loro proprietà. Infine si studierà le loro trasformazioni. Tutta la trattazione è finalizzata ad essere il più generale possibile così da essere valida sia classicamente parlando che nel campo della relatività ristretta.

1.1.1 I sistemi di riferimento inerziali

Trattando di materia strettamente aderente alla realtà, 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:

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

dove \vec{x} è utilizzata per rappresentare un punto nello spazio, t un istante di tempo e \vec{v} la velocità reciproca dei due sistemi di riferimento e questa risulta identicamente costante.

Una volta identificati questi particolari sistemi di riferimento 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 $f : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R} \times \mathbb{R}^3$ invertibile; quest'ultima proprietà 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 venga trasformata in un'altra retta, 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 $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ($n > 1$) una funzione biettiva che trasforma rette in altre rette. Allora f è una trasformazione affine.*

Per procedere nella 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 $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ un automorfismo di campo, ossia tale che se presi $x, y \in \mathbb{R}$ valgono $\varphi(x + y) = \varphi(x) + \varphi(y)$ 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:

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

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 $\varphi = \text{id}$. \square

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

Dimostrazione. Sia $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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

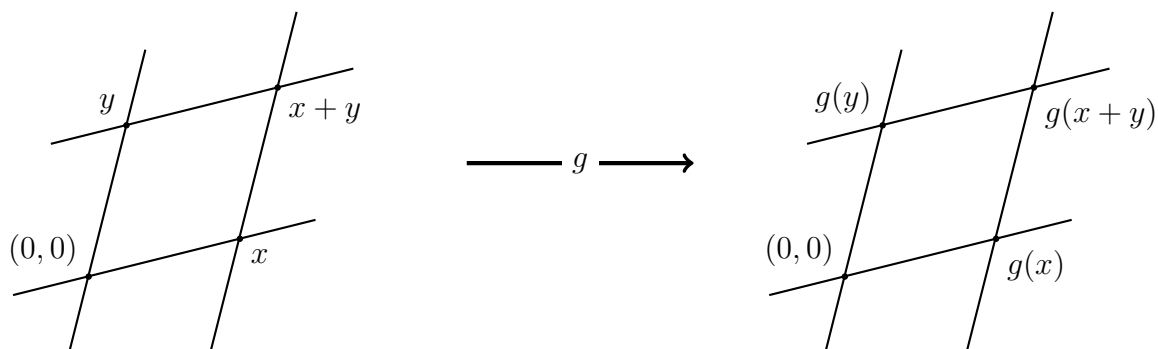


Figura 1.1: Rappresentazione grafica di come agisce la funzione g nel piano

questa funzione è biettiva le immagini di rette parallele saranno a loro volta parallele, altrimenti nel punto di incidenza delle rette 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 $\alpha, \beta : \mathbb{R} \rightarrow \mathbb{R}$ 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 $x, y \in \mathbb{R}$:

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

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 e di conseguenza pure $g = A \circ f = \text{id}$. In fine è sufficiente applicare 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 necessariamente una trasformazione affine. \square

Si procederà ora con la generalizzazione di questo teorema e quindi dimostrando il teorema 1.

Dimostrazione teorema 1. Si consideri la traslazione $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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

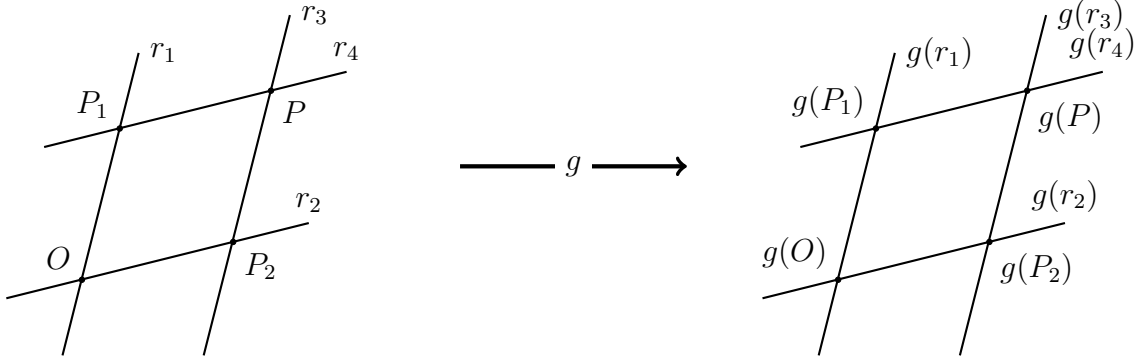


Figura 1.2: Rappresentazione grafica di come agisce la funzione g nel piano

con vertice P e lati giacenti sulle rette considerate. Poiché g è biettiva 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 rette e con i vertici dati dalle immagini dei vertici del precedente parallelogrammo. Detto π' il piano contenente le immagini di r_1 e r_2 allora per costruzione del parallelogrammo si avrà che $g(P) \in \pi'$ e poiché la scelta di $P \in \pi$ è arbitraria 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 $L \circ g|_{\pi} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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:

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

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. \square

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

Nella seconda metà del diciannovesimo secolo la fisica era costituita fondamentalmente dalla meccanica, dalla termodinamica e dall'elettromagnetismo.

La meccanica, fondata da Newton e Galileo, si occupava di studiare il moto dei corpi e fungeva da modello per tutta la fisica.

La termodinamica, grazie a molteplici risultati sperimentali, era riuscita a descrivere molti fenomeni tramite i così detti principi della termodinamica, la cui origine però non era ancora del tutto chiara.

L'elettromagnetismo invece, in seguito a numerosi esperimenti, aveva trovato una completa descrizione nelle equazioni di Maxwell.

Poiché la nascita della teoria della relatività speciale è strettamente connessa sia alla meccanica, sia all'elettromagnetismo, si procederà illustrando brevemente i loro fondamenti.

1.2.1 La meccanica e le trasformazioni di Galileo

Come fondamento della meccanica classica vi sono un serie di osservazioni di carattere sperimentale, queste possono essere utilizzate come assiomi da cui dedurre le leggi del moto dei corpi.

Il primo fatto sperimentale che viene assunto è che lo spazio sia tridimensionale, isotropo, omogeneo e che rispetti la geometria euclidea mentre il tempo sia ad una sola dimensione. Queste assunzioni, come si è visto nella sezione 1.1, consentono di definire cosa sia un sistema di riferimento. Inoltre si prende come assioma che le distanze spaziali e temporali siano assolute, ossia che ogni osservatore concordi sulla misura di queste.

Il secondo fatto sperimentale prende il nome di principio di relatività¹ e, basandosi sulla nozione di sistema di riferimento inerziale (sezione 1.1), afferma che: *in ogni sistema di riferimento inerziale tutte le leggi della fisica sono identiche.*

Infine 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:

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

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

Si vuole quindi identificare quali applicazioni $\varphi : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$ 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. 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:

$$\varphi_{\vec{r},s}(t, \vec{x}) = (t + s, \vec{x} + \vec{r}) \quad (1.2.2)$$

- una generica rotazione degli assi spaziali, dovuta alla proprietà di isotropia dello spazio:

$$\varphi_G(t, \vec{x}) = (t, G\vec{x}) \quad G \in M_{3 \times 3}(\mathbb{R}) : G^{-1} = G^t \quad (1.2.3)$$

- una traslazione di moto rettilineo uniforme, ammissibile grazie alle proprietà di muoversi di tale moto dei sistemi di riferimento inerziali:

$$\varphi_{\vec{v}}(t, \vec{x}) = (t, \vec{x} + \vec{v}t) \quad (1.2.4)$$

¹In meccanica classica è anche detto principio di relatività galileiano

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:

$$\vec{x}' = \vec{x} - \vec{V}t \quad t' = t \quad (1.2.5)$$

Si noti che questa trasformazione non muta la forma del vettore accelerazione $\ddot{\vec{x}}$, infatti considerando una traiettoria $\vec{x}(t)$:

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

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 $\vec{x}(t)$:

$$\frac{d}{dt'}(\vec{x}'(t)) = \frac{d}{dt}(\vec{x}(t) - \vec{V}t) = \dot{\vec{x}}(t) + \vec{V} \quad (1.2.6)$$

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

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:

$$\begin{aligned} \frac{\partial}{\partial t'} &= \frac{\partial t}{\partial t'} \frac{\partial}{\partial t} + \sum_{i=1}^3 \frac{\partial x_i}{\partial t'} \frac{\partial t}{\partial t'} \frac{\partial}{\partial x_i} \\ \frac{\partial}{\partial x'_i} &= \frac{\partial t}{\partial x'_i} \frac{\partial}{\partial t} + \sum_{j=1}^3 \frac{\partial x_j}{\partial x'_i} \frac{\partial}{\partial x_j} \end{aligned}$$

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:

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \vec{V} \cdot \vec{\nabla} \quad \frac{\partial}{\partial x'_i} = \frac{\partial}{\partial x_i} \quad (1.2.7)$$

dove $\vec{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$.

Dalla (1.2.7) 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 $\vec{\nabla}' = (\frac{\partial}{\partial x'}, \frac{\partial}{\partial y'}, \frac{\partial}{\partial z'}) = \vec{\nabla}$.

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 una forza \vec{F} data da:

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

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:

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

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 contenente 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 in:

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

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

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

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

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

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:

$$\vec{\nabla}^2 \vec{E} = \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} \quad \vec{\nabla}^2 \vec{B} = \mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} \quad (1.2.11)$$

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:

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

Così facendo si possono ottenere delle relazioni tra \vec{E} , \vec{B} e \vec{E}' , \vec{B}' che diventano le trasformazioni dei campi 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:

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

Bisogna ora studiare come si trasformino le grandezze generatrici 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:

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

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:

$$\vec{J}' = \vec{J} - \rho \vec{V} \quad (1.3.2)$$

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 (1.2.7), 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} , osservato 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:

$$2t_B = t_A \quad (1.3.3)$$

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:

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

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:

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

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.

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

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 quadrivettori². 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

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

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 $\{e'_0 = ie_0, e'_1 = e_1, e'_2 = e_2, e'_3 = e_3\}$.

$$P = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad P^{-1} = \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.4.3)$$

La trasformazione che si vuole trovare è quindi una trasformazione $f : \mathbb{R}^4 \rightarrow \mathbb{R}^4$, 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 sia 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 misurato in K e r_1 e

²Poiché 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.

r_2 il medesimo quadrivettore misurato 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 $f_{12}(r_1) = r_2$ e per quanto detto si avrà:

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

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:

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

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

$$r^2 = (ict)^2 + x^2 + y^2 + z^2 = R^2 \quad (1.4.5)$$

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 zt : le rotazioni nei primi tre piani corrispondono chiaramente 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:

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

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:

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

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

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

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 Lorentz³ Λ da K a K' con K' in moto a velocità v lungo l'asse x rispetto a K :

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

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 $\pm c$.

³Una trasformazione di Lorentz di questo tipo è più propriamente detta Boost di Lorentz.

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:

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

Così facendo, considerando $c \rightarrow \infty$, 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:

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

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' ($r'=(ct',0,0,0)$):

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

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:

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

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 :

$$V = l_1 l_2 l_3 = \frac{\lambda_1 \lambda_2 \lambda_3}{\gamma} = \frac{\mathcal{V}_0}{\gamma} \quad (1.4.10)$$

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' , per cui nel secondo il primo si muove a velocità V lungo l'asse x' , si avrà allora la trasformazione:

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

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:

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

Con un po' di algebra si ottengono le trasformazioni delle velocità:

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

Si noti che queste trasformazioni per le velocità rispettano il principio di costanza della velocità della luce: si consideri il moto di un raggio di luce lungo l'asse x di K .⁴

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

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 :

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

dove $\frac{dt}{d\tau}$ è stato calcolato dalla (1.4.8).

Se si calcola la trasformazione di u da K a K' si ottiene:

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

e dalla prima componente si ottiene che:

$$\frac{\gamma'_0}{\gamma_0} = (1 - \frac{Vv_x}{c^2})\gamma \quad (1.4.13)$$

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

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

Infine si può iterare questo procedimento riderivando u rispetto al tempo proprio per ottenere la quadriaccelerazione.

$$w = \frac{du}{d\tau} \quad (1.4.14)$$

⁴Questo 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.

Capitolo 2

Meccanica relativistica

Capitolo 3

L'elettromagnetismo nella teoria della relatività

Appendice A

Sulla linearità delle trasformazioni di Lorentz

Le trasformazioni di Lorentz sono una famiglia di trasformazioni affini, ossia composte da una generica traslazione in quattro dimensioni e un'applicazione lineare. Questa loro proprietà è deducibile dai postulati di Einstein e solitamente, nella maggior parte delle trattazioni della teoria della relatività, è data come proprietà ovvia. Nelle seguenti pagine è presentata la dimostrazione di tale proprietà seguendo la dimostrazione riportata da Fock nel suo manuale [2].

A.1 Prerequisiti matematici

In primo luogo è necessario dimostrare due Lemmi che si rivelano utili anche per altri aspetti della teoria della relatività.

Lemma 1. *Siano $v, u \in \mathbb{R}^n$, $<, >$ sia il prodotto scalare euclideo e siano definite $f_u : \mathbb{R}^n \rightarrow \mathbb{R}$, tale che $f_u(v) = < u, v >$, e $q : \mathbb{R}^n \rightarrow \mathbb{R}$ 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:

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

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:

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

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

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

Infine se si raccoglie un termine $\langle u, v \rangle$ 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 $\langle u, v \rangle$:

$$q(v) = \langle u, v \rangle \left[2 \frac{\langle v, Au \rangle}{\|u\|^2} - \frac{\langle u, Au \rangle \langle u, v \rangle}{\|u\|^4} \right] \quad \forall v \in \mathbb{R}^4 \quad (\text{A.1.2})$$

□

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 $q : \mathbb{R}^4 \rightarrow \mathbb{R}$ una forma quadratica tale che, se fissato \bar{v} con $\bar{v}_0^2 - |\vec{v}|^2 = 0$ implica $q(\bar{v}) = 0$, allora $\exists \lambda$ tale che $q(v) = \lambda(v_0^2 - |\vec{v}|^2) \quad \forall v \in \mathbb{R}^4$.

Dimostrazione. Si osservi che $q(v)$ può essere scomposto, seguendo le regole del prodotto righe per colonne $q(v) = v^t A v$ (dove $A \in M_{3 \times 3}(\mathbb{R})$ 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 $(\bar{v}_0)^2 - |\vec{v}|^2 = |\vec{v}|^2 - |\vec{v}|^2 = 0$, per cui per le ipotesi assunte si ha $q(\bar{v}) = 0$.

Essendo $q(\bar{v}) = 0$ si può scrivere:

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

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:

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

□

A.2 La dimostrazione

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 $f : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ 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 supponrà che f sia sufficientemente 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:

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

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 $\gamma_0 \neq 0$.

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

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:

$$\begin{aligned} & \frac{d}{ds} \frac{\partial_k f_i(\xi + s\gamma) \gamma_k}{\partial_j f_0(\xi + s\gamma) \gamma_j} \\ &= \frac{(\partial_{hn}^2 f_i(\xi + s\gamma) \gamma_h \gamma_n) (\partial_k f_0(\xi + s\gamma) \gamma_k) - (\partial_k f_i(\xi + s\gamma) \gamma_k) (\partial_{hn}^2 f_0(\xi + s\gamma) \gamma_h \gamma_n)}{(\partial_j f_0(\xi + s\gamma) \gamma_j)^2} = 0 \end{aligned}$$

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:

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

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 $\partial_{hn}^2 f_0(x) \bar{\gamma}_h \bar{\gamma}_n = 0$.

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

$$\partial_{hn}^2 f_0(x) \gamma_h \gamma_n = \langle \psi(x), \gamma \rangle (\partial_k f_0(x) \gamma_k) \quad \forall x, \gamma \in \mathbb{R}^4 \quad i = 0, 1, 2, 3$$

che inserito nella (A.2.1) risulta in:

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

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:

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

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

$$\partial_{hk}^2 f_i(x) = \frac{1}{2} (\partial_k f_i(x) \psi_h(x) + \partial_h f_i(x) \psi_k(x)) \quad (\text{A.2.3})$$

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:

$$\begin{aligned} \gamma_0^2 - (\gamma_1^2 + \gamma_2^2 + \gamma_3^2) &= 0 \\ (\gamma'_0)^2 - [(\gamma'_1)^2 + (\gamma'_2)^2 + (\gamma'_3)^2] &= 0 \end{aligned}$$

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:

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

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:

$$\begin{aligned} \lambda(x) [\gamma_0^2 - (\gamma_1^2 + \gamma_2^2 + \gamma_3^2)] &= \lambda(x) g_{kh} \gamma_k \gamma_h = g_{ij} (\partial_k f_i(x) \gamma_k) (\partial_h f_j(x) \gamma_h) \\ \Rightarrow \lambda(x) g_{hk} &= g_{ij} (\partial_k f_i(x)) (\partial_h f_j(x)) \end{aligned} \quad (\text{A.2.4})$$

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

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

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

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

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:

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

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.