

Dimitri Colferai

Appunti di Relatività Speciale

Dipartimento di Fisica ed Astronomia dell'Università di Firenze

anno accademico 2018-2019

Revision: 93 : 58d14394e706

Date: 2019/04/15 15:49:08

Contents

1	I postulati della Fisica Newtoniana e della Relatività Speciale	3
1.1	Meccanica newtoniana	3
1.1.1	La posizione è un concetto relativo	3
1.1.2	Il movimento è un concetto relativo	3
1.1.3	Lo spazio euclideo	4
1.1.4	Trasformazioni tra sistemi di riferimento	6
1.1.5	Il gruppo euclideo	8
1.1.6	Il tempo	9
1.1.7	Orologi in punti diversi	10
1.1.8	Coordinate spazio-temporali	10
1.1.9	Il tempo assoluto	11
1.1.10	Lo spazio assoluto	11
1.1.11	Trasformazioni di coordinate spazio-temporali	12
1.1.12	Sistemi inerziali	12
1.1.13	Il gruppo di Galileo	13
1.1.14	Addizione delle velocità	14
1.1.15	Covarianza delle leggi della dinamica	14
1.1.16	I postulati della meccanica Newtoniana	15
1.2	Le equazioni dell'elettromagnetismo	15
1.2.1	Trasformazione dei campi tra SDR inerziali	16
1.2.2	Etere...?	18
1.2.3	L'esperimento di Michelson-Morley	19
1.2.4	Tentativi di salvare l'ipotesi dell'etere	21
1.2.5	Principio di relatività di Einstein	22
1.2.6	Postulati della relatività	22
2	Cinematica relativistica	25
2.1	Relatività della simultaneità	25
2.2	Invarianza delle distanze ortogonali al moto	27
2.3	Dilatazione dei tempi	28
2.4	Contrazione delle lunghezze	29
2.5	Trasformazioni di Lorentz	30
2.5.1	Deduzione delle trasformazioni di Lorentz	30
2.5.2	Limite non-relativistico	35
2.5.3	Trasformazioni di Lorentz proprie non-omogenee	35
2.5.4	Unicità della velocità invariante	35
2.5.5	La velocità della luce come velocità limite	36
2.5.6	Contrazione delle lunghezze	36
2.5.7	Dilatazione dei tempi	36

2.6	Il tempo proprio	37
2.7	Ordine temporale degli eventi: causalità	38
2.8	Composizione delle velocità	40
3	La struttura dello spazio-tempo	43
3.1	Invarianza dell'intervallo	43
3.1.1	Coordinate omogenee	44
3.2	Lo spazio di Minkowski	45
3.2.1	Quadri-vettori	45
3.2.2	Classificazione dei quadri-vettori	45
3.2.3	Tensore metrico	46
3.2.4	Diagrammi spazio-temporali per trasformazioni di Lorentz speciali	49
3.3	Trasformazioni generali tra SDR inerziali	49
3.3.1	Il gruppo di Poincaré	50
3.3.2	Struttura del gruppo di Poincaré	52
3.3.3	Il gruppo di Lorentz	52
3.3.4	Struttura del gruppo proprio di Lorentz	54
3.3.5	Invarianza delle leggi della fisica per trasformazioni di Poincaré	55
4	Quadritensori e calcolo quadritensoriale	57
4.1	Introduzione	57
4.2	Quadritensori	58
4.2.1	Quadriscalari	58
4.2.2	Quadrivettori	58
4.2.3	Tensori doppi	61
4.2.4	Pseudo-tensori	62
4.2.5	Tensore di Levi-Civita	62
4.3	Algebra tensoriale	63
4.3.1	Prodotto scalare	63
4.3.2	Contrazione tensoriale	63
4.3.3	Prodotto diretto	64
4.3.4	Leggi di trasformazione dei campi	64
5	Meccanica relativistica	67
5.1	Cinematica del punto materiale	67
5.1.1	Quadri-velocità	68
5.1.2	Quadri-accelerazione	69
5.1.3	Quadri-impulso	69
5.1.4	Energia relativistica	71
5.1.5	Trasformazione di Lorentz speciale per energia ed impulso	73
5.1.6	Particelle di massa nulla	74
5.1.7	Difetto di massa nei nuclei: equivalenza tra massa ed energia di legame	74
5.2	Sistemi a molte particelle	75
5.2.1	Decadimenti: equivalenza tra massa ed energia cinetica	76
5.2.2	Sistema del centro di massa	79
5.2.3	Urti tra particelle	80
5.3	Dinamica del punto materiale	81
5.3.1	Quadri-forza	82
5.3.2	Forza di Newton	82
5.3.3	Teorema delle forze vive	84

5.3.4	Forza e accelerazione	85
5.3.5	Moto in campo di forza omogeneo e costante (campo elettrico)	86
6	Elettrodinamica	89
6.1	Equazioni dell'elettromagnetismo classico	89
6.2	Densità di carica e corrente	90
6.2.1	La quadri-corrente per un sistema di cariche puntiformi	90
6.2.2	Conservazione della carica elettrica	92
6.3	Forma covariante a vista delle equazioni di Maxwell	92
6.4	Forma covariante a vista della forza di Lorentz	94
6.5	Le proprietà di trasformazione dei campi	95
6.5.1	Trasformazione per traslazioni	95
6.5.2	Trasformazione per rotazioni	96
6.5.3	Trasformazione per spinte	96
6.5.4	Trasformazione per inversione degli assi	97
6.6	Invarianti del campo elettromagnetico	98
6.7	Il quadri-potenziale	99
6.8	Invarianza di gauge	100
6.9	Equazioni di campo del quadri-potenziale	100
A	Calcolo tensoriale tri-dimensionale	103
A.1	Vettori e tensore metrico	103
A.2	Tensore di Levi-Civita	104
A.2.1	Prodotto vettoriale	104
A.2.2	Prodotto misto	105
A.2.3	Contrazioni del tensore di Levi-Civita	105
A.3	Tensori doppi antisimmetrici	107
B	Unità di energia	109
C	Annichilazione $e^+e^- \rightarrow \gamma\gamma$	111
D	Sistema di unità razionalizzate	115

Introduzione

Cos'è la Teoria della Relatività

La Teoria della Relatività descrive le relazioni che esistono tra entità fisiche misurate da differenti osservatori che siano in moto uno rispetto all'altro (moto relativo). La Teoria della Relatività deve il suo nome in quanto rifiuta il concetto di moto assoluto, ossia non ammette l'esistenza di un SDR in quiete assoluta, e riconosce che solo il moto relativo ha significato fisico.

La Teoria della Relatività è stata sviluppata in due fasi.

- Nella prima fase, chiamata "Relatività Speciale" (o "Ristretta"), le leggi della fisica sono formulate in sistemi di riferimento inerziali, e quindi tra osservatori in moto rettilineo uniforme uno rispetto all'altro. Da notare che questa teoria non include i fenomeni gravitazionali.
- Nella seconda fase ("Relatività Generale") la teoria viene generalizzata a tutti i tipi di osservatori, cioè le leggi della fisica sono formulate per osservatori qualsiasi, i quali si possono muovere in modo arbitrario (ruotante, accelerato, ecc.). Questa teoria include anche la gravitazione, anzi, ne costituisce la nuova formulazione.

Entrambe le teorie sono state elaborate da Albert Einstein (1905 e 1916), ma va detto che nel caso della Relatività Speciale molti risultati e conclusioni erano stati già enunciati da altri fisici, in particolare da Lorentz e soprattutto da Poincaré.

Relatività Speciale

- Ha rivoluzionato i concetti di spazio e di tempo, inquadrandoli in un'unica realtà (lo spazio-tempo) nella quale l'estensione spaziale dei corpi e lo scorrere del tempo dipendono e quindi variano tra osservatori in moto relativo. Essenzialmente, lo spazio ed il tempo sono come le componenti x ed y di un vettore, che assumono valori diversi a seconda del sistema di riferimento dell'osservatore.

- Su questa scia anche la quantità di moto, l'energia e la massa sono visti come aspetti diversi della stessa entità ($E = mc^2$).

Relatività Generale

- Ha ulteriormente rivoluzionato i concetti di spazio e di tempo;
- Fornisce un'interpretazione nuova dei fenomeni gravitazionali.

Lo spazio-tempo è curvo; l'apparente attrazione dei corpi è una conseguenza del loro moto libero nello spazio-tempo curvo — come delle formiche che partono parallelamente dall'equatore verso il polo nord o sud, si avvicinano sempre più velocemente fino ad incontrarsi; a sua volta la curvatura dello spazio tempo è generata dalla presenza dei corpi (dalla loro massa-energia), nonché dalla curvatura stessa — come dei corpi posati su un telo elastico lo deformano, ed altri corpi che si muovano sul telo vengono deviati da queste deformazioni.

In questo corso ci occuperemo esclusivamente della Teoria della Relatività Speciale.

Chapter 1

I postulati della Fisica Newtoniana e della Relatività Speciale

Com'è nata la Teoria della Relatività

La Teoria della Relatività (Speciale, come sempre d'ora in poi) è stata il compimento di una lunga serie di tentativi che avevano lo scopo di mettere insieme in un quadro coerente i due grandi blocchi su cui era basata la fisica nella seconda metà dell' '800:

- la Meccanica di Newton, con la sua legge di gravitazione;
- la Teoria dell'Elettromagnetismo, formulata da Maxwell.

Il problema di fondo è questo: mentre la meccanica di Newton è valida in ogni sistema di riferimento inerziale, la teoria dell'elettromagnetismo sembra valida solamente in un particolare sistema di riferimento (e negli altri sistemi di riferimento in quiete rispetto ad esso, ma non in moto rispetto ad esso).

Cerchiamo di capire meglio questo aspetto riepilogando brevemente i fondamenti della meccanica di Newton, confrontandoli poi con le equazioni dell'elettromagnetismo.

1.1 Meccanica newtoniana

1.1.1 La posizione è un concetto relativo

Il problema centrale della meccanica è stabilire il movimento dei corpi, ossia determinare la loro posizione in funzione del tempo. Operativamente, ciò che si può misurare come posizione è la distanza tra coppie di punti (per esempio mediante un regolo rigido scelto come unità di lunghezza), oppure l'angolo compreso tra coppie di rette concorrenti. Quindi, la posizione di un oggetto A può essere definita solamente rispetto ad un determinato insieme di punti di riferimento S , comunemente chiamato Sistema di Riferimento (SDR), in modo tale che le misure di distanza tra A ed un certo numero di punti noti di S identifichi in modo univoco il luogo in cui A si trova.

1.1.2 Il movimento è un concetto relativo

Allo stesso modo, quando si parla di movimento, quello che si osserva è il cambiamento della posizione di un corpo relativamente ad un altro corpo: non c'è modo di osservare direttamente il moto rispetto ad un incorporato ed assoluto sistema di riferimento. Per esempio: quando sono su una nave, o in treno, o in ascensore, o sull'astronave (senza vibrazioni, con i finestrini abbassati), riesco a percepire se sono fermo o in movimento? Riesco a determinare a quale velocità sto andando senza interagire con l'esterno?

Pertanto, per descrivere quantitativamente la posizione dei corpi ed il loro moto occorre sempre specificare un sistema di riferimento. Tuttavia, la natura non ci dà un SDR privilegiato rispetto al quale riferire le nostre misure di posizione. Piuttosto, abbiamo imparato dalla meccanica che ogni osservatore sceglie il SDR che più gli è conveniente.

Rivediamo ora come rappresentare matematicamente la struttura dello spazio 3D (in modo consistente con le misure di distanza tra punti), in modo da poter trattare allo stesso modo SDR diversi.

1.1.3 Lo spazio euclideo

Le misure di posizione possono essere eseguite in modo statico, costruendo molti regoli identici al regolo campione, e trasportandoli in punti diversi. Questa procedura sembra non presentare problemi, purché i regoli non siano accorciati né deformati dopo il trasporto, e purché le regole della geometria rimangano inalterate nelle varie regioni dello spazio. Pensate ad un foglio di carta a quadretti, in cui ogni lato di un quadretto (il regolo) è identico agli altri e forma angoli retti con i lati adiacenti. L'esperienza ci incoraggia a ritenere valide queste ipotesi, che matematicamente si formalizzano descrivendo lo spazio delle posizioni come uno **spazio affine euclideo a 3 dimensioni**.

Osserviamo che uno spazio affine euclideo 2D non sarebbe adatto per descrivere, ad esempio, il mondo di esseri (amebe) 2D che vivono sulla superficie di una sfera: ponendo 4 regoli uguali ad angolo retto, non chiuderebbero un quadrilatero.

Lo spazio affine è una struttura matematica strettamente collegata a quella di spazio vettoriale. Intuitivamente, uno spazio affine assomiglia ad uno spazio vettoriale, con la differenza che tra i suoi punti non ce n'è uno (l'origine) "centrale" e "privilegiato" rispetto agli altri. In effetti lo spazio affine tridimensionale è stato introdotto per poter trattare lo spazio della fisica classica senza essere legati a sistemi di riferimento con una origine prefissata, ma invece per essere in grado di formulare leggi fisiche valide in riferimenti diversi.

Uno **spazio affine** (possiamo, anzi, dobbiamo immaginarci lo spazio che ci circonda) è costituito di punti. Questi punti non possono essere subito associati a vettori, per esempio non ha senso sommare 2 punti:

$$\text{Pisa} + \text{Firenze} = ?$$

I vettori ce li immaginiamo come frecce con un capo ed una coda. Due punti individuano un vettore.

Uno spazio affine \mathcal{M} è un insieme di elementi che chiamiamo punti dotato di una funzione

$$\psi : \mathcal{M} \times \mathcal{M} \rightarrow \mathcal{V} \quad (1.1)$$

a valori in uno spazio vettoriale \mathcal{V} che esprime, dati due punti A e B di \mathcal{M} , il "vettore applicato da A a B ". Questo vettore può essere indicato con $\psi(A, B)$ oppure \underline{AB} oppure $B - A$. La funzione ψ deve soddisfare i seguenti requisiti:

- (1) per ogni punto A fissato, la mappa che associa a B il vettore \underline{AB} è una biiezione da \mathcal{M} in \mathcal{V} ;
- (2) per ogni $A, B, C \in \mathcal{M}$ vale $\underline{AB} + \underline{BC} = \underline{AC}$.

Quest'ultima relazione esprime la nostra idea fisica che sommare due vettori significa effettuare due spostamenti: da A a B e poi da B a C . È facile dimostrare che, per ogni $A, B \in \mathcal{M}$, valgono $\underline{AA} = \underline{0}$ e $\underline{AB} = -\underline{BA}$.

Esercizio: Mostrare che $\underline{AA} = \underline{0}$ e $\underline{AB} = -\underline{BA} \quad \forall A, B \in \mathcal{M}$.

Dato un punto A in \mathcal{M} ed un vettore $\underline{v} \in \mathcal{V}$, possiamo definire il punto $A + \underline{v} \in \mathcal{M}$ come l'unico [per (1)] punto B tale che $\underline{AB} = \underline{v}$. Quindi, sommando un vettore \underline{v} ad un punto otteniamo un nuovo punto spostato rispetto a quello di partenza di \underline{v} .

L'aggettivo **euclideo** si riferisce al fatto che in questo spazio vettoriale \mathcal{V} è definito un prodotto scalare

simmetrico, bilineare e definito positivo

$$\cdot : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$$

$$\underline{u} \cdot \underline{v} \mapsto \underline{u} \cdot \underline{v} \quad (1.2)$$

$$\underline{u} \cdot \underline{v} = \underline{v} \cdot \underline{u} \quad (1.3)$$

$$\underline{u} \cdot (a_1 \underline{v}_1 + a_2 \underline{v}_2) = a_1 (\underline{u} \cdot \underline{v}_1) + a_2 (\underline{u} \cdot \underline{v}_2) \quad \forall a_1, a_2 \in \mathbb{R} \quad (1.4)$$

$$\underline{u} \cdot \underline{u} > 0 \quad \forall \underline{u} \neq \underline{0}. \quad (1.5)$$

Tramite il prodotto scalare possiamo definire i concetti di:

- *lunghezza* di un vettore: $|\underline{v}| \equiv \sqrt{\underline{v} \cdot \underline{v}}$, da cui la *distanza* tra due punti A e B definita da $|\underline{AB}|$;
- *angolo* ϕ tra due vettori: $\underline{u} \cdot \underline{v} = |\underline{u}| |\underline{v}| \cos(\phi)$, da cui l'angolo ϕ tra tre punti A, B, C tale che $\underline{AB} \cdot \underline{BC} = |\underline{AB}| |\underline{BC}| \cos(\phi)$. In particolare due vettori si dicono ortogonali quando il loro prodotto scalare è nullo.

La dimensione dello spazio vettoriale \mathcal{V} definisce la dimensione dello spazio affine \mathcal{M} . Lo spazio fisico ha **3 dimensioni** in quanto ogni base di \mathcal{V} è costituita da 3 vettori linearmente indipendenti.

Lo spazio affine euclideo non è nient'altro che l'ambiente in cui descrivere la ben nota geometria euclidea.

- ⊗ Alle rette della geometria corrispondono i sottospazi affini 1D dati da $\{A + \lambda \underline{v} : \lambda \in \mathbb{R}\}$ in cui $A \in \mathcal{M}$ è un punto (appartenente alla retta) e $\underline{v} \in \mathcal{V}$ è un vettore non nullo che ne individua la direzione.
- ⊗ Ai piani della geometria corrispondono i sottospazi affini 2D dati da $\{A + \lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2 : \lambda_i \in \mathbb{R}\}$ in cui $A \in \mathcal{M}$ è un punto (appartenente al piano) e $\underline{v}_1, \underline{v}_2 \in \mathcal{V}$ sono due vettori non paralleli.

Per poter fare dei conti, abbiamo bisogno di tradurre punti e vettori in grandezze numeriche, perché solo con i numeri possiamo dare relazioni quantitative tra grandezze, e quindi, fare dei calcoli. Questo si fa introducendo i **sistemi di riferimento** (SDR).

Un SDR è individuato da un punto $O \in \mathcal{M}$, detto origine, e da tre versori¹ ortogonali $(\underline{e}_1, \underline{e}_2, \underline{e}_3) \in \mathcal{V}$:

$$\underline{e}_i \cdot \underline{e}_j = \delta_{ij}, \quad \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad (1.6)$$

Ogni punto $A \in \mathcal{M}$ è quindi messo in corrispondenza biunivoca con la terna di numeri reali

$$(x, y, z) \equiv (x^1, x^2, x^3) \equiv \underline{x} \in \mathbb{R}^3$$

dette coordinate cartesiane, tale che

$$\underline{OA} = \underline{e}_1 x^1 + \underline{e}_2 x^2 + \underline{e}_3 x^3. \quad (1.7)$$

Quindi, dato un SDR, lo spazio affine dei punti dello spazio è messo in corrispondenza con \mathbb{R}^3 . Per ragioni che saranno spiegate in seguito, nella teoria relativistica è conveniente usare indici in alto per le coordinate ed indici in basso per i vettori di base. Tanto vale cominciare ad usare fin d'ora questa notazione.

Notate che per parlare di base ortonormale, abbiamo bisogno del prodotto scalare per stabilire quando due vettori sono ortogonali, e qual è la loro lunghezza. Il prodotto scalare di due vettori è dato da

$$\underline{u} \cdot \underline{v} = \sum_{i,j=1}^3 (\underline{e}_i u^i) \cdot (\underline{e}_j v^j) = \sum_{i,j} u^i v^j (\underline{e}_i \cdot \underline{e}_j) = \sum_{i,j} u^i v^j \delta_{ij} = \sum_i u^i v^i = u^1 v^1 + u^2 v^2 + u^3 v^3, \quad (1.8)$$

ove u^i e v^i sono le componenti di \underline{u} e \underline{v} .

Le componenti del generico vettore \underline{AB} sono date dalle differenze delle coordinate di B e quelle di A :

$$\underline{AB} = \underline{AO} + \underline{OB} = \underline{OB} - \underline{OA} = \sum_{i=1}^3 \underline{e}_i (x_B^i - x_A^i). \quad (1.9)$$

¹Un versore è un vettore di lunghezza uguale ad 1. Tale lunghezza corrisponde a quella del regolo campione o, equivalentemente, alla scelta dell'unità di misura.

La distanza tra A e B misurata in \mathcal{M} con i regoli, corrisponde alla distanza calcolata con il teorema di pitagora a partire dalle coordinate: ponendo $\underline{u} = \underline{v} = \underline{AB}$ nell'eq (1.8) troviamo

$$|\underline{AB}|^2 = (x_B^1 - x_A^1)^2 + (x_B^2 - x_A^2)^2 + (x_B^3 - x_A^3)^2 . \quad (1.10)$$

Le coordinate si rivelano particolarmente utili come strumento di calcolo per svariate grandezze fisiche. Pensiamo ad esempio al calcolo della distanza tra due punti separati da una montagna o da un'oceano: provarci con i regoli è poco pratico, se non impossibile. Oppure, dati n punti A_1, \dots, A_n in posizione nota, se volessimo conoscere la distanza tra un nuovo punto B e gli n punti noti, invece di effettuare n misure di distanza, basterebbe effettuarne 3 per determinare le coordinate di B , e procedere successivamente al calcolo delle n distanze $|\underline{A_k B}|$.

Diciamo infine che le coordinate di \mathbb{R}^3 servono per *rappresentare* lo spazio \mathcal{M} , così come una carta geografica della Toscana, con latitudini e longitudini, rappresenta (ma non è) la Toscana.

1.1.4 Trasformazioni tra sistemi di riferimento

Dato che i SDR possono essere scelti arbitrariamente, è importante stabilire la relazione che intercorre tra le coordinate che ciascun SDR attribuisce ai punti dello spazio. Consideriamo dunque due SDR $S = (O, \underline{e}_1, \underline{e}_2, \underline{e}_3)$ ed $S' = (O', \underline{e}'_1, \underline{e}'_2, \underline{e}'_3)$. Possiamo esprimere gli elementi di S' rispetto a quelli di S specificando la posizione dell'origine O' e l'orientazione degli assi \underline{e}'_i :

$$\underline{OO'} = \sum_{i=1}^3 \underline{e}_i q^i \quad (1.11a)$$

$$\underline{e}'_j = \sum_{i=1}^3 \underline{e}_i U^i_j , \quad (j = 1, 2, 3) . \quad (1.11b)$$

La seconda relazione esprime il fatto che ogni vettore di base di S' si può scrivere come combinazione lineare dei vettori di base di S . $U \equiv (U^i_j)_{j=1,2,3}^{i=1,2,3}$ è una matrice reale 3×3 , detta matrice di cambio di base da S ad S' . Con la convenzione di associare l'indice a sinistra all'*indice riga* e quello a destra con l'*indice colonna*, abbiamo che le 3 colonne di U non sono altro che le componenti dei 3 vettori \underline{e}'_1 , \underline{e}'_2 e \underline{e}'_3 nella base di S .

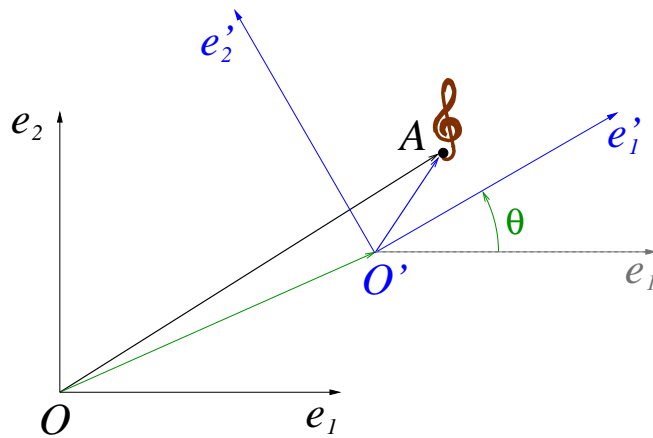


Figure 1.1: Sistemi di riferimento nello spazio affine: S (nero) ed S' (blu).

Se le basi $\{\underline{e}_i\}$ ed $\{\underline{e}'_j\}$ sono entrambe *orientate*, ossia ortonormali e destrorse, allora U è una matrice di

rotazione. Per esempio, se S' si ottiene ruotando S di un angolo θ rispetto all'asse z , abbiamo

$$j \rightarrow \quad (1.12)$$

$$U = \begin{matrix} i \\ \downarrow \end{matrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.13)$$

$$\begin{matrix} \uparrow & \uparrow & \uparrow \\ \underline{e}'_1 & \underline{e}'_2 & \underline{e}'_3 \end{matrix} \quad (1.14)$$

Più in generale, se richiediamo soltanto che le due basi siano ortonormali, abbiamo (U^T indica la matrice trasposta di U):

$$\begin{aligned} \delta_{ij} &\stackrel{(1.6)}{=} \underline{e}'_i \cdot \underline{e}'_j \stackrel{(1.11b)}{=} \sum_{m,n=1}^3 (\underline{e}_m U^m_i) \cdot (\underline{e}_n U^n_j) = \sum_{m,n} U^m_i U^n_j (\underline{e}_m \cdot \underline{e}_n) \stackrel{(1.6)}{=} \sum_{m,n} U^m_i U^n_j \delta_{mn} \\ &= \sum_m U^m_i U^m_j = \sum_m (U^T)_i^m U^m_j = (U^T U)_{ij} . \end{aligned} \quad (1.15)$$

Essendo $\delta_{ij} = \mathbb{I}_{ij}$ gli elementi della matrice identità, ne concludiamo che la matrice U deve essere una matrice *ortogonale*: $U^T U = \mathbb{I}$,² che si scrive anche $U \in O(3, \mathbb{R})$ gruppo ortogonale. Le matrici di rotazione che collegano SDR orientati sono matrici ortogonali con l'ulteriore condizione $\det U = 1$, e si scrive anche $U \in SO(3, \mathbb{R})$ gruppo speciale ortogonale.

Consideriamo ora un punto $A \in \mathcal{M}$, a cui S ed S' associano le coordinate \mathbf{x} ed \mathbf{x}' rispettivamente:

$$\underline{OA} = \sum_i \underline{e}_i x^i \quad (1.16)$$

$$\underline{O'A} = \sum_j \underline{e}'_j x'^j . \quad (1.17)$$

Usando le relazioni (1.11) abbiamo

$$\begin{aligned} \underline{OA} &= \underline{OO'} + \underline{O'A} = \sum_i \underline{e}_i q^i + \sum_j \underline{e}'_j x'^j \\ &= \sum_i \underline{e}_i q^i + \sum_j \sum_i \underline{e}_i U^i_j x'^j = \sum_i \underline{e}_i \left[\sum_j U^i_j x'^j + q^i \right] . \end{aligned} \quad (1.18)$$

Confrontando le espressioni (1.16) e (1.18), otteniamo la relazione tra le coordinate

$$x^i = \sum_j U^i_j x'^j + q^i \quad (1.19)$$

Raccogliendo le coordinate in colonne a 3 componenti

$$\mathbf{x} \equiv \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} , \quad \mathbf{q} \equiv \begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix} , \quad \text{ecc.} \quad (1.20)$$

possiamo scrivere la precedente relazione nella più comoda forma matriciale

$$\mathbf{x} = U \mathbf{x}' + \mathbf{q} . \quad (1.21)$$

Per scrivere le \mathbf{x}' in funzione delle \mathbf{x} , sottraiamo \mathbf{q} ad entrambi i membri e quindi moltiplichiamo per la matrice inversa di U :

$$\mathbf{x}' = U^{-1}(\mathbf{x} - \mathbf{q}) = R \mathbf{x} + \mathbf{b} , \quad (R \equiv U^{-1} = U^T , \quad \mathbf{b} \equiv -U^{-1} \mathbf{q}) \quad (1.22)$$

²La precedente relazione si capisce ricordando che le colonne della matrice U sono le componenti dei vettori di base \underline{e}'_j , quindi le righe di U^T sono di nuovo le componenti dei vettori \underline{e}'_j , e nel prodotto matriciale $U^T U$ tra la i -esima riga di U^T e la j -esima colonna di U altro non si sta facendo che il prodotto scalare $\underline{e}'_i \cdot \underline{e}'_j$.

in cui le a^j sono le coordinate di O rispetto ad S' . È importante constatare che il cambio di coordinate ha la *stessa forma* sia che si esprimano le \mathbf{x}' in termini delle \mathbf{x} , sia che si esprimano le \mathbf{x} in termini delle \mathbf{x}' ; la differenza tra le due formule consiste solo negli scambi $U \leftrightarrow R$ e $q \leftrightarrow a$. Essendo R la matrice inversa di una matrice ortogonale U , anche R è ortogonale: $R^T R = \mathbb{I}$.

Osserviamo che la trasformazione di coordinate (1.22) $\mathbf{x} \rightarrow \mathbf{x}'$ è in qualche modo inversa alla trasformazione (1.11) tra sistemi di riferimento $S \rightarrow S'$:

- Se l'origine viene traslata di \mathbf{q} , le coordinate sono traslate di $-\mathbf{q}$;
- Se gli assi vengono ruotati di U , le coordinate sono ruotate con U^{-1} .

Quanto detto si può visualizzare confrontando la rappresentazione tramite coordinate che S ed S' danno di uno stesso sistema, come in fig. 1.2.

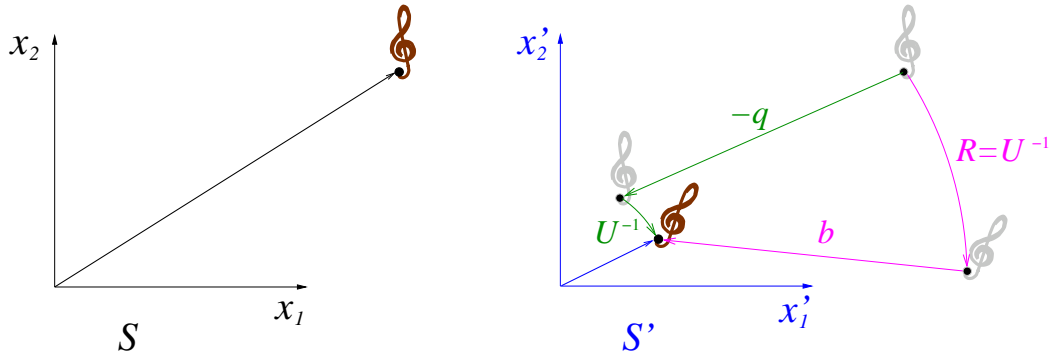


Figure 1.2: Rappresentazione di uno stesso oggetto tramite due sistemi di coordinate diversi: S (nero), S' (blu).

Questo modo di intendere la trasformazione (1.22) come coordinate in SDR diversi dello stesso oggetto (il punto A in questo caso) viene chiamato *punto di vista passivo*.

Un altro modo fisicamente accettabile di interpretare la relazione (1.22) è quella di pensare \mathbf{x} ed \mathbf{x}' come coordinate di due oggetti diversi A ed A' nello stesso SDR S . In altre parole, mantenendo immutato il sistema di riferimento, A' si ottiene traslando e ruotando l'oggetto A dalla posizione di coordinate \mathbf{x} a quella di coordinate \mathbf{x}' . Questo modo di leggere la trasformazione (1.22) è detto *punto di vista attivo*, in quanto il sistema fisico A è attivamente trasformato in un nuovo sistema A' .

Riassumendo: la trasformazione di coordinate $\mathbf{x}' = R\mathbf{x} + \mathbf{b}$ si può intendere in due modi diversi:

p.v. attivo: dato un unico SDR S , le \mathbf{x}' sono coordinate di un sistema fisico A' ottenuto ruotando di R e traslando di \mathbf{b} il sistema fisico A di coordinate \mathbf{x} ;

p.v. passivo: dato un unico sistema fisico A , le \mathbf{x} e le \mathbf{x}' sono le coordinate di A in due diversi SDR S ed S' , in cui S' si ottiene traslando S di $-\mathbf{b}$ e poi ruotandolo con R^{-1} attorno ad O .

Osservazione: il fatto che la legge di trasformazione delle coordinate sia “contraria” a quella dei vettori di base giustifica l’aver introdotto indici in posizioni diverse per le due quantità. In generale, le componenti delle grandezze che trasformano in modo analogo ai vettori di base verranno chiamate *co-varianti*, e saranno indicate con indici in basso; le componenti di grandezze che trasformano in modo inverso, come le coordinate spaziali, verranno chiamate *contro-varianti*, e denotate con indici in alto.

1.1.5 Il gruppo euclideo

Avendo assunto che lo spazio delle posizioni non presenti punti né direzioni privilegiate, ovvero che sia *omogeneo* ed *isotropo*, il metodo per calcolare le distanze a partire dalle coordinate deve essere lo stesso per tutti gli osservatori: questo metodo è il teorema di Pitagora ricordato nell’eq. (1.10). Infatti, siccome la formula (1.10) è stata ricavata per un generico osservatore S , essa dovrà valere anche per qualsiasi altro osservatore S' .

In termini delle coordinate, possiamo esprimere questo concetto con queste parole: la *forma quadratica* (1.10) è invariante per trasformazioni di coordinate del tipo (1.22), con R matrice ortogonale. Possiamo verificare questa

affermazione direttamente. Per prima cosa usiamo la notazione matriciale per scrivere il prodotto scalare tramite le componenti cartesiane dei vettori in un qualsiasi SDR:

$$\underline{u} \cdot \underline{v} \stackrel{(1.8)}{=} \sum_i u^i v^i = \begin{pmatrix} u^1 & u^2 & u^3 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \underline{u}^T \underline{v}, \quad (1.23)$$

(a parole, il prodotto scalare $\underline{u} \cdot \underline{v}$ è dato dal prodotto della matrice riga \underline{u}^T per la matrice colonna \underline{v}). La distanza tra A e B calcolata in S' , secondo l'eq. (1.10) si può quindi scrivere

$$\begin{aligned} \sum_{i=1}^3 (x_B'^i - x_A'^i)^2 &= (\underline{x}_B' - \underline{x}_A')^T (\underline{x}_B' - \underline{x}_A') \stackrel{(1.22)}{=} (R\underline{x}_B + \underline{b} - R\underline{x}_A - \underline{b})^T (R\underline{x}_B + \underline{b} - R\underline{x}_A - \underline{b}) \\ &= (\underline{x}_B - \underline{x}_A)^T R^T R (\underline{x}_B - \underline{x}_A) = \sum_{i=1}^3 (x_B^i - x_A^i)^2 \end{aligned} \quad (1.24)$$

e, come ci aspettavamo, coincide con la distanza calcolata in S .

Ci possiamo chiedere se ci siano altre trasformazioni di coordinate che preservano la forma quadratica (1.10). Si può dimostrare che tutte e sole queste **trasformazioni euclidee** hanno la forma (1.22). L'insieme di tutte queste trasformazioni $\{(R, \underline{b}) : R \in O(3, \mathbb{R}), \underline{b} \in \mathbb{R}^3\}$ forma un *gruppo* rispetto all'operazione di “composizione di trasformazioni”:

- Se $\underline{x}' = R\underline{x} + \underline{b}$ e $\underline{x}'' = R'\underline{x}' + \underline{b}'$, allora il “prodotto” delle trasformazioni $\underline{x} \rightarrow \underline{x}' \rightarrow \underline{x}''$ dà

$$\underline{x}'' = R'(R\underline{x} + \underline{b}) + \underline{b}' = (R'R)\underline{x} + (R'\underline{b} + \underline{b}') \quad (1.25)$$

che è ancora del tipo (1.22), in quanto il prodotto di matrici ortogonali è ortogonale;

- L'elemento neutro, cioè la trasformazione identità, appartiene al gruppo: $\underline{x}' = \underline{x}$ si ha con $R = \mathbb{I}$ ed $\underline{b} = \underline{0}$;
- Ogni trasformazione $\underline{x}' = R\underline{x} + \underline{b}$ ammette una e una sola trasformazione inversa, data da $\underline{x} = R^{-1}\underline{x}' + (-R^{-1}\underline{b})$. Inoltre la matrice $R^{-1} = R^T$ è ortogonale, e quindi anche la trasformazione inversa è euclidea.

Il gruppo di tutte queste trasformazioni che preservano la forma quadratica (1.22) si chiama **gruppo Euclideo**. È facile dimostrare che le trasformazioni del gruppo euclideo lasciano invariato anche il prodotto scalare tra due vettori

$$\underline{AB} \cdot \underline{CD} = \sum_i (x_B'^i - x_A'^i)(x_D'^i - x_C'^i) = \sum_i (x_B^i - x_A^i)(x_D^i - x_C^i). \quad (1.26)$$

Quest'invarianza si può anche intendere dal punto di vista attivo: se i punti A e B di coordinate \underline{x}_A ed \underline{x}_B sono trasformati nei punti A' e B' di coordinate \underline{x}_A' ed \underline{x}_B' secondo la relazione (1.22) (sempre con R ortogonale), allora la distanza tra i due punti è preservata: $|\underline{AB}| = |\underline{A'B'}|$. Lo stesso dicasi per il prodotto scalare di una qualsiasi coppia di vettori.

1.1.6 Il tempo

Per specificare quantitativamente il moto di un corpo, dobbiamo attribuire allo scorrere del tempo un significato fisico quantitativo. Cos'è il tempo? Filosoficamente è una domanda molto impegnativa. Forse un bambino risponderebbe che il tempo è ciò che si misura con l'orologio. Questa è effettivamente la risposta corretta per i fisici. Come funziona l'orologio? L'orologio misura l'intervallo di tempo che intercorre tra due eventi contando quante volte si ripete un determinato intervallo di tempo scelto come unità di misura, per esempio il periodo di oscillazione di un fenomeno ciclico che si assume per definizione costante. Anche la “posizione temporale” quindi si misura in modo relativo, cioè rispetto ad un istante di tempo di riferimento.

1.1.7 Orologi in punti diversi

Supponiamo ora di voler misurare, nel SDR S , il vettore velocità (media) di un corpo H lungo un certo tratto del suo cammino. Sia quindi A il punto, fisso in S , che coincide con il punto di partenza di H , e B il punto, fisso in S , che coincide con il punto di arrivo di H . Con i miei regoli misuro la posizione $\mathbf{x}_A = (x_A^1, x_A^2, x_A^3)$ di A , quella \mathbf{x}_B di B , e quindi calcolo il vettore spostamento $\Delta\mathbf{x} = \mathbf{x}_B - \mathbf{x}_A$ compiuto da H . Rimane da misurare il tempo impiegato a percorrere il tragitto $\Delta t = t_B - t_A$. Quindi abbiamo bisogno di due misure di tempo effettuate in luoghi diversi, e che in generale differiscono anche dall'origine del sistema di coordinate O , in cui idealmente si trova l'osservatore con l'orologio campione. Un modo di procedere può essere questo: costruiamo altri 2 orologi identici all'orologio campione, e li sincronizziamo a questo in O . Quindi li trasportiamo (lentamente, sperando di non alterarne il battito) uno in A e l'altro in B . Facciamo registrare al primo orologio il tempo t_A in cui H transita per A , ed al secondo orologio il tempo t_B in cui H arriva in B . Quindi ci facciamo trasmettere i due valori e calcoliamo Δt . Otteniamo così il risultato cercato per la velocità media: $\mathbf{v} = \Delta\mathbf{x}/\Delta t$.

In questa misura di intervallo di tempo abbiamo assunto che gli orologi non vengano alterati nel trasporto, e che funzionino allo stesso modo nei punti A e B (e O). È proprio così? Da un punto di vista logico, non ci sarebbe alcuna contraddizione se i due orologi procedessero a ritmo diverso durante lo spostamento, oppure mentre sono in punti diversi. Tuttavia, l'*ipotesi* che gli orologi continuino a procedere allo stesso ritmo è ben confermata dall'esperienza. Per esempio, riportando gli orologi in A ed in B di nuovo nel punto O , constatiamo che essi segnano ancora lo stesso tempo dell'orologio campione, entro la precisione degli strumenti comuni.

Questo ha portato i fisici dei secoli scorsi a concludere che il tempo scorre in modo uniforme nei vari punti dello spazio, e che è possibile definire in ogni tale punto un tempo correttamente sincronizzato con l'orologio campione dell'osservatore in O . Matematicamente, possiamo rappresentare l'insieme degli istanti temporali come uno spazio affine unidimensionale \mathcal{M}_t . La scelta di un orologio equivale a fissare un punto in \mathcal{M}_t che rappresenta il "tempo zero", ed un vettore che rappresenta come verso lo scorrere del tempo dal passato al futuro, e come lunghezza la scelta dell'unità di misura temporale.

1.1.8 Coordinate spazio-temporali

Quindi, possiamo rappresentare un osservatore S , o equivalentemente, un SDR, come una collezione di regoli che suddividono lo spazio in un reticolato rigido di cubi di lato uguale al regolo campione, in ogni punto del quale si trova un orologio sincronizzato con l'orologio campione.

Cerchiamo di rappresentare in un unico grafico lo spazio \mathcal{M}_s ed il tempo \mathcal{M}_t mediante il prodotto cartesiano $\mathcal{M}_t \times \mathcal{M}_s$, come in fig. 1.3. Ci servirebbero 4 dimensioni (1 temporale e 3 spaziali). Limitiamoci quindi a 1 o 2 dimensioni spaziali, per poter rappresentare su un foglio questo grafico. Ogni punto di questo grafico rappresenta un **evento**, cioè qualcosa che accade in una determinata posizione e ad un determinato istante, per esempio lo scoppio di una bomba. Il concetto di evento è fondamentale in relatività, così come il concetto di punto lo è in geometria. In un dato SDR, Ogni evento è caratterizzato da una quaterna di numeri reali $(t, x^1, x^2, x^3) = (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3$, chiamate "coordinate spazio-temporali". Cambiando SDR, allo stesso evento corrispondono in generale coordinate spazio-temporali diverse.

L'insieme degli eventi la cui posizione coincide con un punto A fissato, descrive una linea parallela all'asse temporale, che chiamiamo **linea universo** del punto A . Fissato invece un tempo t_0 , l'insieme degli eventi che avvengono all'istante t_0 forma uno spazio 3D di eventi simultanei, chiamato **superficie di simultaneità**, per analogia con un grafico con 2 dimensioni spaziali in cui eventi tra loro simultanei formano una superficie.

Ogni oggetto puntiforme che si muove nello spazio individua una linea universo continua che possiamo parametrizzare in funzione del tempo:

$$(t, x^1(t), x^2(t), x^3(t)) \equiv (t, \mathbf{x}(t)) \quad (1.27)$$

in cui $\mathbf{x}(t)$ è la posizione dell'oggetto all'istante t .

Un modo di pensare una superficie di simultaneità è quello di una fotografia istantanea del sistema che stiamo considerando, per esempio delle palle su di un tavolo da biliardo. Il diagramma spazio-temporale consiste allora di una successione di fotografie ad istanti successivi, messe in ordine una sopra l'altra.

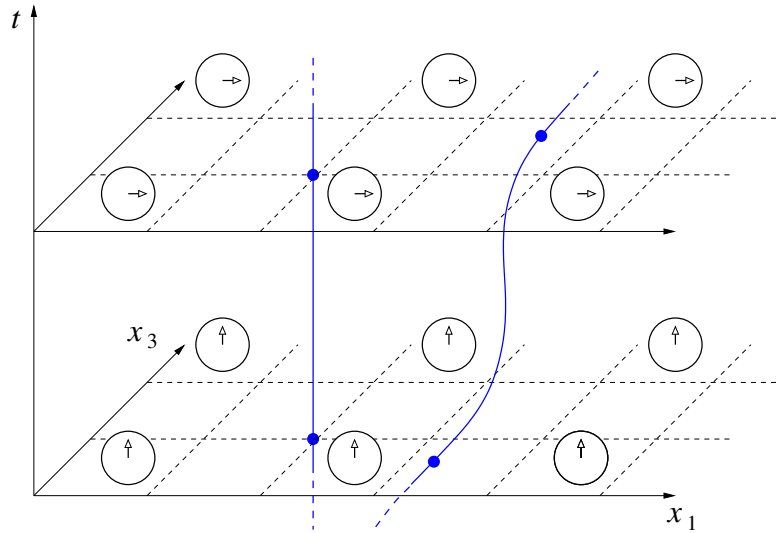


Figure 1.3: *Diagramma spazio-temporale per un osservatore S . Gli (iper)-piani orizzontali rappresentano superfici di simultaneità a tempo costante; il tempo scorre lungo l'asse verticale; le linee blu rappresentano le linee d'universo di due corpi: uno in quiete rispetto ad S (linea verticale), l'altro in moto generico.*

Un altro osservatore individua un differente SDR S' ed attribuisce agli eventi nuove coordinate (t', \mathbf{x}') . Ci proponiamo di determinare la relazione che esiste tra le coordinate spazio-temporali di S' e quelle di S .

1.1.9 Il tempo assoluto

Partiamo alle misure di tempo. Un altro modo di misurare il tempo impiegato da H per andare da A a B potrebbe essere quello di dotare H di un orologio identico a quello campione, che registri i tempi di passaggio in A e B . In altre parole, stiamo considerando un osservatore S' solidale ad H , con un orologio identico a quello di S . In principio non c'è nulla che ci garantisca che l'orologio in moto (rispetto ad S) scandisca il tempo esattamente come gli orologi in quiete (O, A, B). In altre parole, non è detto a priori che il $\Delta t'$ misurato da S' coincida con il Δt misurato da A e B . Quindi non siamo sicuri di poter usare $\Delta t'$ al posto di Δt nella definizione di velocità $\mathbf{v} \equiv \Delta \mathbf{x} / \Delta t$.

È di fondamentale importanza sapere se orologi in moto relativo scandiscano il tempo allo stesso modo o no. Un modo per verificare se orologi in vari stati di moto procedano con la stessa frequenza consiste nel

- 1) sincronizzarli in un punto A ,
- 2) farli muovere lungo percorsi diversi, con velocità diverse,
- 3) riportarli infine assieme in un punto B (che potrebbe anche coincidere con il punto di partenza A).

L'esperienza comune insegna che gli orologi sono ancora sincronizzati, entro gli errori di misura. Se ne conclude che il tempo scorre allo stesso modo per tutti gli osservatori. Questo ha portato a concepire il tempo come un'entità assoluta, indipendente da chi la osserva. In formule, l'intervallo di tempo misurato tra due eventi P e Q è lo stesso per ogni osservatore: $\Delta t'_{PQ} = t'_Q - t'_P = t_Q - t_P = \Delta t_{PQ}$. Valendo questa uguaglianza per ogni P e Q , ne deduciamo che $t' = t + c_0$ ove c_0 è una costante (che vale 0 per orologi sincronizzati, vale un'ora per orologi di due fusi orari adiacenti, ecc.). Vedremo tuttavia che misure di maggiore precisione riserveranno delle sorprese, le cui conseguenze segnalano l'esigenza di una nuova fisica (relativistica).

1.1.10 Lo spazio assoluto

Resta da stabilire se anche lo spazio sia qualcosa di assoluto, indipendente dall'osservatore. In altre parole, ci poniamo la seguente domanda: la lunghezza di un oggetto misurata da due osservatori diversi è la stessa o no? Dall'omogeneità e isotropia dello spazio, osservatori diversi ma in quiete relativa misurano la stessa lunghezza. Consideriamo ora due osservatori S ed S' in moto relativo, per esempio S solidale al suolo ed S' solidale ad un

treno in corsa, che vogliano misurare quant'è lungo il treno. Avendo definito le misure di lunghezza come misure statiche (con regoli fermi), per S' è semplice effettuare la misura: pone in successione tanti regoli identici al regolo campione dalla coda alla testa del treno, e li conta. 125 precisi. $x'_B - x'_A = 125$ m.

Anche S deve porre in successione tanti regoli (fermi rispetto al suolo, ed identici a quelli di S'), in modo che il primo inizi con la coda del treno e l'ultimo termini alla testa del treno. Ma siccome il treno si muove, la posizione della coda x_A dipende dall'istante t_A . E la posizione x_B della testa del treno dipenderà dall'istante t_B a cui la guardo. Se lascio t_A e t_B arbitrari, p.es. t_A alle 12:00:00 e t_B alle 13:00:00, il treno da Firenze è arrivato a Bologna, e mi ci vogliono 120000 regoli. Chiaramente questa procedura non ha senso: è ovvio che devo considerare $t_A = t_B$, cioè devo misurare la *distanza spaziale di due eventi simultanei* (questa precisazione ovvia sarà importante tenerla presente nella teoria della relatività). Se quindi S , con i suoi orologi sincronizzati, riesce a determinare che al medesimo istante t_A la coda del treno si trovava in x_A e la testa in x_B , può comodamente porre in successione i suoi regoli tra x_A ed x_B . Quanti gliene servono? Neanche a farlo apposta, 125! $x_B - x_A = 125$ m.

Si stabilisce così empiricamente che gli intervalli spaziali, intesi come misure di distanza tra eventi simultanei, non dipendono dall'osservatore. Questo ha portato a concepire anche lo spazio come un'entità assoluta, indipendente da chi la osserva. Anche in questo caso, vedremo che misure di maggiore precisione riserveranno delle sorprese, le cui conseguenze segnalano l'esigenza di una nuova fisica (relativistica).

Notiamo che c'è una differenza importante tra l'assolutezza del tempo e dello spazio: mentre tutti gli osservatori inerziali sono d'accordo sul fatto che due eventi avvengano allo stesso istante o meno, non tutti gli osservatori inerziali sono d'accordo sul fatto che due eventi avvengano nello stesso luogo. Solo la distanza tra eventi *simultanei* è assoluta, non la distanza tra eventi qualsiasi. La teoria della relatività eliminerà questa differenza.

1.1.11 Trasformazioni di coordinate spazio-temporali

Possiamo riassumere quanto detto finora nelle seguenti 2 proposizioni:

- 1) Dati due eventi P e Q , tutti gli osservatori sono d'accordo sull'intervallo di tempo che separa P da Q : $\Delta t = \Delta t'$. In particolare, eventi simultanei per un osservatore ($\Delta t = 0$) sono simultanei per qualsiasi altro osservatore ($\Delta t' = 0$). La simultaneità è un concetto assoluto.
- 2) Dati due eventi P e Q simultanei, tutti gli osservatori sono d'accordo sulla distanza spaziale tra P e Q : $|\Delta \mathbf{x}| = |\Delta \mathbf{x}'|$.

Indicando con (t, x, y, z) le coordinate spazio-temporali che S attribuisce ai vari eventi, e con (t', x', y', z') le coordinate in S' , abbiamo: $t_Q - t_P = t'_Q - t'_P$ per ogni P e Q eventi. Se $P = O'$ (origine di S'), $t'_Q = t_Q - t_{O'}$ per ogni Q , pertanto

$$t' = t + t_0, \quad (t_0 = -t_{O'}) . \quad (1.28)$$

Inoltre, fissato un istante t (superficie di simultaneità), l'invarianza delle misure di distanza tra coppie di eventi simultanei ci dice che $|\mathbf{x}_Q - \mathbf{x}_P| = |\mathbf{x}'_Q - \mathbf{x}'_P|$ se $t_P = t_Q$. Pertanto la trasformazione di coordinate che lega \mathbf{x} ad \mathbf{x}' appartiene al gruppo euclideo:

$$\mathbf{x}' = R(t)\mathbf{x} + \mathbf{b}(t), \quad (R(t)^T R(t) = \mathbf{1} \quad \forall t) . \quad (1.29)$$

Il fatto che R ed \mathbf{b} dipendano da t significa che la relazione tra le \mathbf{x} e le \mathbf{x}' può variare nel tempo (si pensi a due SDR in moto relativo); tuttavia, ad ogni istante fissato, questa trasformazione è euclidea.

1.1.12 Sistemi inerziali

Abbiamo visto che, dal punto di vista statico, tutti i SDR sono equivalenti. Se però studiamo il moto dei corpi in relazione alle cause (ossia alle forze) responsabili del moto, cioè dal punto di vista dinamico, sappiamo che ci sono dei SDR particolari. Infatti, come Galileo aveva osservato, un corpo libero, cioè non soggetto a forze

delle quali sia possibile rintracciare la causa, si muove di moto rettilineo uniforme nel laboratorio (incluiamo tra i moti uniformi anche lo stato di quiete: $\mathbf{v} = 0$). Questa proprietà di moto uniforme dei corpi liberi non può certo valere se acceleriamo il SDR, o se lo mettiamo in rotazione. I SDR in cui i corpi liberi si muovono di moto rettilineo uniforme vengono chiamati SDR inerziali.

1.1.13 Il gruppo di Galileo

Ci proponiamo ora di trovare le possibili trasformazioni di coordinate tra SDR inerziali, compatibili con l'ipotesi di spazio e tempo assoluti. Consideriamo un SDR inerziale S con coordinate (t, \mathbf{x}) , nel quale un corpo libero H si muove di moto uniforme $\mathbf{x}_H(t) = \mathbf{u}t + \mathbf{c}$, ossia con velocità $\frac{d}{dt}\mathbf{x}(t) = \mathbf{u}$ costante e perciò con accelerazione nulla $\frac{d^2}{dt^2}\mathbf{x}(t) = 0$. Se il SDR S' è inerziale, anche nelle coordinate (t', \mathbf{x}') il moto di H dovrà essere uniforme. Dalle equazioni (1.28) e (1.29) abbiamo

$$t' = t + t_0, \quad \mathbf{x}'_H(t') = R(t)\mathbf{x}_H(t) + \mathbf{b}(t) = R(t)[\mathbf{u}t + \mathbf{c}] + \mathbf{b}(t). \quad (1.30)$$

Osservando che $dt = dt'$, l'accelerazione in S' vale

$$\frac{d^2}{dt'^2}\mathbf{x}'_H(t') = \frac{d^2 R(t)}{dt^2}[\mathbf{u}t + \mathbf{c}] + 2\frac{dR(t)}{dt}\mathbf{u} + \frac{d^2 \mathbf{b}(t)}{dt^2}, \quad (1.31)$$

che deve essere nulla per tutti i valori di \mathbf{c} ($\implies \frac{d^2 R(t)}{dt^2} = 0$) e di \mathbf{u} ($\implies \frac{dR(t)}{dt} = 0$), per cui $R(t)$ deve essere costante mentre $\mathbf{b}(t)$ può al più dipendere linearmente dal tempo: $\mathbf{b}(t) = \mathbf{v}t + \mathbf{b}$. Pertanto, la più generale trasformazione di coordinate tra due SDR inerziali ha la forma

$$t' = t + t_0, \quad \mathbf{x}' = R\mathbf{x} + \mathbf{v}t + \mathbf{b}, \quad (1.32)$$

in cui ora R è una matrice ortogonale costante, mentre \mathbf{v} e \mathbf{b} sono vettori-colonna costanti. Questa trasformazione si chiama **trasformazione di Galileo** — la possiamo indicare con $\{t_0, R, \mathbf{v}, \mathbf{b}\}$ — e dal punto di vista passivo descrive un SDR S che si muove con *velocità di trascinamento* \mathbf{v} costante rispetto ad S' (ovvero, S' si muove con velocità $-R^{-1}\mathbf{v}$ rispetto ad S). L'orientazione degli assi di S' rispetto a quelli di S è invece fissa, ossia non ci può essere una velocità di rotazione relativa tra due SDR inerziali. È facile verificare che anche la trasformazione inversa della (1.32) ha la stessa forma

$$\{t_0, R, \mathbf{v}, \mathbf{b}\}^{-1} = \{-t_0, R^{-1}, -R^{-1}\mathbf{v}, R^{-1}(\mathbf{v}t_0 - \mathbf{b})\}, \quad (1.33)$$

così come la composizione di due trasformazioni di Galileo è ancora di Galileo

$$\{t_1, R', \mathbf{v}', \mathbf{b}'\} \circ \{t_0, R, \mathbf{v}, \mathbf{b}\} = \{t_1 + t_0, R'R, R'\mathbf{v} + \mathbf{v}', R'\mathbf{b} + \mathbf{v}'t_0 + \mathbf{b}'\}. \quad (1.34)$$

Quindi l'insieme delle trasformazioni di Galileo forma un gruppo, detto appunto **gruppo di Galileo**, che collega le coordinate di tutti i possibili SDR inerziali.

Ogni trasformazione di Galileo si può pensare come combinazioni di 3 tipi di trasformazioni:

- *rotazioni spaziali* attorno all'origine, descritte dalle matrici di rotazione R , che dipendono da 3 parametri reali (p. es., i 3 angoli di rotazione di Eulero);
- *traslazioni spaziali e temporali* descritte dal vettore \mathbf{b} e dal numero t_0 , e quindi parametrizzate da 4 numeri reali;
- *spinte* con velocità \mathbf{v} , parametrizzate da 3 numeri reali.

$$\mathbf{x} \rightarrow R\mathbf{x} \rightarrow R\mathbf{x} + \mathbf{b} \rightarrow R\mathbf{x} + \mathbf{b} + \mathbf{v}t \equiv \mathbf{x}' \quad (1.35)$$

$$t \rightarrow t + t_0 \equiv t'. \quad (1.36)$$

Quindi il gruppo di Galileo dipende da 10 parametri reali.

1.1.14 Addizione delle velocità

Un'immediata conseguenza delle trasformazioni di Galileo, o se vogliamo, dell'aver assunto il tempo e lo spazio come due entità assolute ed indipendenti, è la legge di addizione delle velocità. Se il punto materiale H si muove nel SDR S con legge oraria $\mathbf{x}_H(t)$, nell'altro SDR S' la traiettoria è data da $\mathbf{x}'_H(t') = R\mathbf{x}_H(t) + \mathbf{v}t + \mathbf{b}$, e quindi la velocità misurata da S' all'istante t' è

$$\mathbf{u}'(t') \equiv \frac{d}{dt'}\mathbf{x}_H(t') = R\frac{d}{dt}\mathbf{x}_H(t) + \mathbf{v} = R\mathbf{u}(t) + \mathbf{v}, \quad (1.37)$$

ossia S' rileva una velocità pari alla somma della velocità di H misurata da S' (opportunamente ruotata nel nuovo SDR) più la velocità di trascinamento \mathbf{v} di S rispetto ad S' . Nel caso di SDR con assi paralleli ($R = \mathbf{I}$) ed orologi sincronizzati ($t = t'$), la precedente equazione diventa semplicemente

$$\mathbf{u}'(t) = \mathbf{u}(t) + \mathbf{v}. \quad (1.38)$$

Intuitivamente, la legge di addizione delle velocità è la derivata rispetto al tempo (assoluto!) della addizione di vettori $\underline{O'H(t)} = \underline{O'O(t)} + \underline{O(t)H(t)}$ nello spazio affine euclideo (assoluto!), le cui proprietà metriche sono percepite in modo identico dai due osservatori.

1.1.15 Covarianza delle leggi della dinamica

Stabiliamo ora la *covarianza* delle leggi della dinamica per trasformazioni di Galileo. La prima legge (d'inerzia) è già stata esaminata e sfruttata per ricavare la forma delle trasformazioni di Galileo, ossia le trasformazioni di coordinate tra SDR inerziali.

La seconda legge della dinamica esprime l'effetto delle forze sullo stato di moto dei corpi. Indichiamo con \mathbf{F} le tre componenti della risultante delle forze che agiscono sul corpo H , misurate nel SDR inerziale S , e con

$$\mathbf{a}_H(t) \equiv \frac{d^2}{dt^2}\mathbf{x}_H(t) \quad (1.39)$$

le componenti dell'*accelerazione*. Questa legge afferma che (omettendo per concisione la dipendenza da H e t)

$$\mathbf{F} = m\mathbf{a}, \quad (1.40)$$

ove m è una quantità chiamata *massa inerziale* del corpo H .

Che la forza sia una grandezza vettoriale lo deduciamo immediatamente dalla sua definizione operativa: come misura dell'allungamento della molla del dinamometro, e quindi, dalla misura istantanea di un vettore spostamento. Pertanto, negli altri SDR le componenti della forza si trasformano come le differenze di coordinate a tempo fissato:

$$\mathbf{F}' = R\mathbf{F}. \quad (1.41)$$

Un esempio esplicito è dato dalla forza di gravitazione di Newton del corpo M sul corpo N

$$\mathbf{F}_{MN} = G \frac{m_M m_N}{|\mathbf{x}_M - \mathbf{x}_N|^3} (\mathbf{x}_M - \mathbf{x}_N). \quad (1.42)$$

La legge di trasformazione delle accelerazioni la ricaviamo dalla legge di addizione delle velocità (1.37), derivando una seconda volta rispetto al tempo:

$$\mathbf{a}' = R\mathbf{a}. \quad (1.43)$$

Quindi, forza ed accelerazione trasformano allo stesso modo, e se facciamo l'ipotesi che la massa inerziale non dipenda dallo stato del moto di un corpo, sia cioè un *invariante*

$$m' = m \quad (1.44)$$

allora vediamo che la seconda legge della dinamica vale in qualsiasi SDR inerziale:

$$\mathbf{F}' = m\mathbf{a}'. \quad (1.45)$$

La terza legge della dinamica (principio di azione e reazione), che coinvolge solo forze (la forza $\mathbf{F}_{MN}(t)$ che il corpo M esercita sul corpo N è uguale e contraria alla forza $\mathbf{F}_{NM}(t)$ che N esercita su M), è di verifica immediata, per la linearità della trasformazione indotta da R e per il fatto che la simultaneità non dipende dal SDR :

$$\mathbf{F}_{MN}(t) = -\mathbf{F}_{NM}(t) \implies \mathbf{F}'_{MN}(t') = R\mathbf{F}_{MN}(t) = R(-\mathbf{F}_{NM}(t)) = -R\mathbf{F}_{NM}(t) = -\mathbf{F}'_{NM}(t') . \quad (1.46)$$

Il fatto che entrambi i membri delle equazioni della dinamica (1.40) e (1.46) si trasformino nello stesso modo qualora si cambi SDR, si esprime dicendo che tali equazioni sono *covarianti* per trasformazioni di Galileo. Equivalentemente, si dice che le leggi della dinamica sono *invarianti in forma* sotto trasformazioni di Galileo, in quanto esse si scrivono allo stesso modo in tutti i sistemi inerziali.

1.1.16 I postulati della meccanica Newtoniana

Possiamo riassumere quanto detto finora nei seguenti 3 postulati alla base della meccanica Newtoniana:

M1) *Lo spazio vuoto è omogeneo ed isotropo; il tempo è omogeneo.*

Questo significa che non ci sono punti né direzioni privilegiati nello spazio vuoto, così come non ci sono istanti di tempo privilegiati. In altre parole, un sistema isolato si può realizzare in modo identico in qualsiasi punto dello spazio, con qualsiasi orientazione ed in qualunque momento, e il suo comportamento sarebbe lo stesso in ognuna di queste varianti.

Osservazione: In un sistema non isolato le proprietà M1 non sono generalmente valide. Nel caso in cui il sistema sia un grave (soggetto alla forza di gravità), la direzione verso il centro della Terra è privilegiata, e la quantità di moto in questa direzione non si conserva. Nel caso in cui il sistema sia un oscillatore forzato da una forza esterna non costante, i vari istanti di tempo non sono equivalenti tra loro, e l'energia dell'oscillatore non si conserva.

Osservazione: La conservazione della quantità di moto (\implies omogeneità dello spazio), del momento angolare (\implies isotropia) e dell'energia (\implies omogeneità del tempo) ci dicono che queste ipotesi sono ben fondate sperimentalmente.

M2) *Intervalli temporali e lunghezze spaziali sono gli stessi in tutti i sistemi di riferimento in moto rettilineo uniforme uno rispetto all'altro.*

M3) *Principio di relatività Galileiana:*

Le leggi della meccanica hanno la stessa forma in tutti i sistemi di riferimento inerziali.

Equivalentemente, non è possibile distinguere tra SDR inerziali con esperimenti di meccanica. Notare bene che questo principio si limita alle sole leggi della meccanica, compresa la legge di gravitazione universale. Le leggi dell'elettromagnetismo e dell'ottica sono a parte, e costituiscono il nostro prossimo argomento di studio.

1.2 Le equazioni dell'elettromagnetismo

Scopo di questa sezione è capire se le leggi dell'elettromagnetismo soddisfino al principio di relatività Galileiano. Cominciamo analizzando due aspetti fisici dell'elettromagnetismo: l'induzione prima e la propagazione delle onde poi.

Un tipico esempio di induzione elettromagnetica è il passaggio di corrente in un circuito elettrico immerso in un campo magnetico variabile, prodotto per esempio da un magnete in movimento, come in fig. 1.4a. I fenomeni dell'induzione elettromagnetica sono sempre collegati ad una variazione nel tempo del flusso magnetico. È un fatto sperimentale che tali fenomeni dipendono solo dalla velocità relativa tra inducente e indotto e non dalla velocità dell'una o dell'altra parte singolarmente. In altri termini, il fenomeno dell'induzione magnetica non

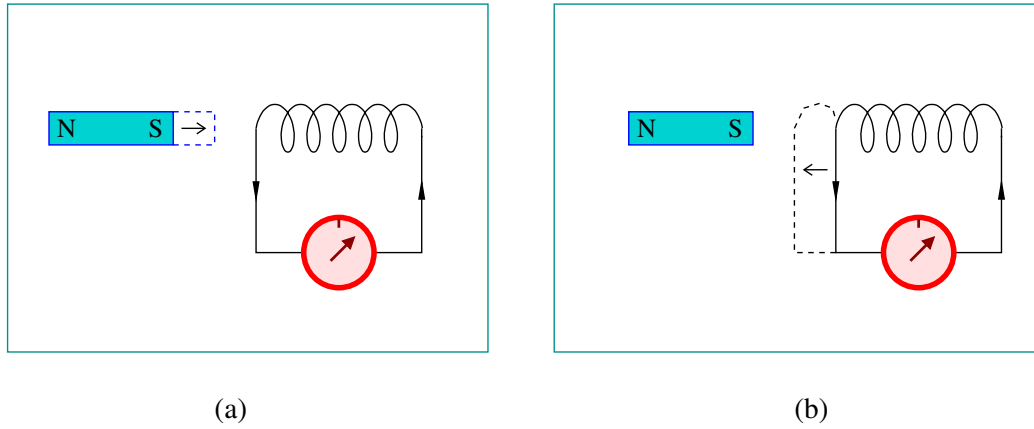


Figure 1.4: *Induzione magnetica: (a) circuito indotto fermo, magnete inducente in moto verso destra; (b) magnete inducente fermo, circuito indotto in moto verso sinistra.*

permette di definire un moto assoluto né un SDR privilegiato. Per l'induzione quindi la situazione è simile a quella espressa dal principio di relatività Galileiano.

Considerando invece la propagazione delle onde elettromagnetiche, giungiamo alla conclusione opposta: se nel SDR S valgono le equazioni di Maxwell, in cui la luce nel vuoto si propaga con velocità c in tutte le direzioni, nel SDR S' che si muove lungo l'asse x con velocità v la luce sarà vista propagarsi con velocità $c - v$ in direzione x (per la legge di addizione di velocità (1.38)), con velocità $c + v$ in direzione opposta e velocità intermedie nelle altre direzioni, ossia in modo non isotropo. Quindi esiste un unico SDR (a meno di rotazioni e traslazioni) in cui la velocità della luce è uguale in tutte le direzioni. Solo in questi SDR in quiete tra loro, che possiamo considerare privilegiati, e che possiamo determinare sperimentalmente, valgono le equazioni di Maxwell.

Questa situazione ben si accorda con l'idea che esista un qualche mezzo, una qualche sostanza che funga da supporto al campo elettromagnetico, le cui “vibrazioni” corrispondono alle onde elettromagnetiche, così come l'aria è il supporto materiale delle onde sonore. A questo mezzo materiale, ipotizzato nella seconda metà dell'800, che copriva tutto lo spazio e penetrava tutti i corpi, fu dato il nome di *etere luminifero*. Solo nei SDR in quiete rispetto all'etere valgono le equazioni di Maxwell, e solo in questi la velocità della luce è la stessa in tutte le direzioni.

Viene così a cadere il principio di relatività Galileiano per le leggi dell'elettromagnetismo.

1.2.1 Trasformazione dei campi tra SDR inerziali

Ora cerchiamo di analizzare la situazione dal punto di vista matematico. Lo scopo è vedere se le equazioni dell'elettromagnetismo possano essere le stesse per due osservatori inerziali in moto relativo con velocità \mathbf{v} ed assi a due a due paralleli, le cui coordinate, velocità e forze siano legate dalle leggi di trasformazione Galileiane

$$t' = t \quad (1.47a)$$

$$\mathbf{x}' = \mathbf{x} - \mathbf{v}t \quad (1.47b)$$

$$\mathbf{u}' = \mathbf{u} - \mathbf{v} \quad (1.47c)$$

$$\mathbf{F}' = \mathbf{F} . \quad (1.47d)$$

Qui il punto di vista è quello passivo, cioè (t, \mathbf{x}) e (t', \mathbf{x}') sono le coordinate dello stesso evento per i due SDR.

Le equazioni dell'elettromagnetismo si riassumono nella forza di Lorentz e nelle equazioni di Maxwell

$$\mathbf{F} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \quad (1.48a)$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \quad (1.48b)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (1.48c)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.48d)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} \quad (1.48e)$$

in cui i campi elettrico $\mathbf{E}(t, \mathbf{x})$ e magnetico $\mathbf{B}(t, \mathbf{x})$ dipendono dalla posizione e dal tempo, mentre $\mathbf{u}(t)$ e q sono la velocità e la carica di una data particella.

Vogliamo pertanto stabilire se esista una legge di trasformazione dei campi

$$\mathbf{E}(t, \mathbf{x}) \rightarrow \mathbf{E}'(t', \mathbf{x}') \quad (1.49a)$$

$$\mathbf{B}(t, \mathbf{x}) \rightarrow \mathbf{B}'(t', \mathbf{x}') \quad (1.49b)$$

(ed analoghe per le sorgenti ρ e \mathbf{J}) tale per cui continuino a valere le equazioni di Maxwell

$$\mathbf{F}' = q'(\mathbf{E}' + \mathbf{u}' \times \mathbf{B}') \quad (1.50a)$$

$$\nabla' \cdot \mathbf{E}' = \frac{\rho'}{\varepsilon_0} \quad (1.50b)$$

$$\nabla' \times \mathbf{E}' = -\frac{\partial \mathbf{B}'}{\partial t'} \quad (1.50c)$$

$$\nabla' \cdot \mathbf{B}' = 0 \quad (1.50d)$$

$$\nabla' \times \mathbf{B}' = \mu_0 \mathbf{J}' + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}'}{\partial t'} \quad (1.50e)$$

ove

$$\nabla'_i \equiv \frac{\partial}{\partial x'^i} . \quad (1.51)$$

Ribadiamo l'uso del punto di vista passivo: i campi \mathbf{E} ed \mathbf{E}' in (1.49a) sono valutati dai 2 osservatori diversi S ed S' nello stesso evento, cioè nello stesso luogo allo stesso istante, di coordinate (t, \mathbf{x}) e (t', \mathbf{x}') rispettivamente. Idem per i campi \mathbf{B} e \mathbf{B}' in (1.49b). Supponiamo che il valore della carica elettrica non dipenda dalla velocità della carica, $q = q'$, in accordo con il principio di conservazione della carica tremendamente ben confermato sperimentalmente. Inoltre, assumiamo universali le costanti ε_0 e μ_0 (altrimenti potremmo distinguere i due SDR). Se dunque esiste una qualche trasformazione (1.49) per cui siano soddisfatte le (1.50), allora il principio di relatività Galileiano sarebbe soddisfatto anche dalle leggi dell'elettromagnetismo.

Prendiamo in esame la forza di Lorentz (1.48a) per vedere come debbano trasformare i campi sotto trasformazioni di Galileo. Siano \mathbf{E} e \mathbf{B} i campi elettrico e magnetico misurati nel SDR S in un generico evento A . Quale valore per il campo elettrico \mathbf{E}' si misurerà nel SDR S' in corrispondenza dello stesso evento A ? Seguiamo la procedura operativa: S' mette una carica “di prova” q in A , ferma (in S'), e misura la forza di Lorentz a cui è sottoposta. Essendo $\mathbf{u}' = 0$, si ha che $\mathbf{F}' = q\mathbf{E}'$, ossia non c'è contributo magnetico, e quindi $\mathbf{E}' = \mathbf{F}'/q$. D'altra parte, S' misura una forza di Lorentz pari a $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, ove \mathbf{v} è la velocità di S' rispetto ad S , e quindi la velocità della carica q in S .

Dall'eguaglianza tra le forze $\mathbf{F} = \mathbf{F}'$ (cfr. eq. (1.47d)), troviamo che

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B} . \quad (1.52)$$

Questa è una relazione tra i campi, e quindi vale tra due qualsiasi SDR in moto relativo con velocità \mathbf{v} , indipendentemente dalla presenza o meno di cariche di prova.

Per valutare il campo magnetico \mathbf{B}' usiamo una carica di prova che si muova in S' con velocità \mathbf{u}' , così da essere sottoposta ad una forza

$$\mathbf{F}' = q(\mathbf{E}' + \mathbf{u}' \times \mathbf{B}') \stackrel{(1.52)}{=} q(\mathbf{E} + \mathbf{v} \times \mathbf{B} + \mathbf{u}' \times \mathbf{B}') \quad (\text{in } S') \quad (1.53a)$$

$$\mathbf{F} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \stackrel{(1.47c)}{=} q(\mathbf{E} + \mathbf{v} \times \mathbf{B} + \mathbf{u}' \times \mathbf{B}) \quad (\text{in } S) \quad (1.53b)$$

Sfruttando ancora la relazione $\mathbf{F} = \mathbf{F}'$ ricaviamo, dalla differenza delle (1.53),

$$\mathbf{u}' \times (\mathbf{B}' - \mathbf{B}) = 0 \quad (1.54)$$

per ogni scelta della velocità \mathbf{u}' della carica di prova in S' , scelta che lascia inalterati S ed S' e quindi anche i campi \mathbf{B} e \mathbf{B}' . Ne consegue che

$$\mathbf{B}' = \mathbf{B}. \quad (1.55)$$

Riassumendo, affinché la formula della forza di Lorentz sia invariante in forma per le trasformazioni di Galileo (1.47), i campi devono trasformare nel seguente modo:

$$\mathbf{E}'(t', \mathbf{x}') = \mathbf{E}(t, \mathbf{x}) + \mathbf{v} \times \mathbf{B}(t, \mathbf{x}) \quad (1.56a)$$

$$\mathbf{B}'(t', \mathbf{x}') = \mathbf{B}(t, \mathbf{x}). \quad (1.56b)$$

Non è difficile verificare che le equazioni di Maxwell omogenee (senza sorgenti) sono invarianti in forma per la trasformazione (1.56), essendo $\partial/\partial x'^i = \partial/\partial x^i$ e $\partial/\partial t' = \partial/\partial t + \mathbf{v} \cdot \nabla$:

$$\nabla' \cdot \mathbf{B}' = \nabla \cdot \mathbf{B} = 0 \quad (1.57)$$

$$\nabla' \times \mathbf{E}' + \frac{\partial \mathbf{B}'}{\partial t'} = \nabla \times (\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \left[\frac{\partial \mathbf{B}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{B} \right] \quad (1.58)$$

$$= \mathbf{v}(\nabla \cdot \mathbf{B}) - (\mathbf{v} \cdot \nabla) \mathbf{B} + (\mathbf{v} \cdot \nabla) \mathbf{B} = 0. \quad (1.59)$$

Nell'ultimo passaggio si è usata l'identità differenziale $\nabla \times (\mathbf{v} \times \mathbf{B}) = \mathbf{v}(\nabla \cdot \mathbf{B}) - (\mathbf{v} \cdot \nabla) \mathbf{B}$. In particolare, è invariante in forma la legge dell'induzione di Faraday. Le equazioni con le densità di carica e corrente invece non mantengono la loro forma, neanche nel vuoto in cui $\rho = 0$, $\mathbf{J} = 0$. Per esempio

$$\nabla' \cdot \mathbf{E}' = \nabla \cdot (\mathbf{E} + \mathbf{v} \times \mathbf{B}) = -\nabla \cdot (\mathbf{B} \times \mathbf{v}) = -\mathbf{v} \cdot (\nabla \times \mathbf{B}) = -\varepsilon_0 \mu_0 \mathbf{v} \cdot \frac{\partial \mathbf{E}}{\partial t} \neq 0.$$

In conclusione, non è possibile stabilire una legge di trasformazione per i campi elettrico e magnetico che lasci invariate in forma le equazioni di Maxwell e la forza di Lorentz, se le trasformazioni delle coordinate sono quelle di Galileo.

È facile verificare che, se eliminiamo la “corrente di spostamento” $\varepsilon_0 \mu_0 \partial_t \mathbf{E}$ dall'eq. (1.48e), allora le equazioni di Maxwell nel vuoto³ sono invarianti in forma sotto le trasformazioni (1.47) e (1.56). In effetti, è proprio la corrente di spostamento il responsabile della propagazione delle onde.

1.2.2 Etere...?

Quindi, se sono valide le ipotesi di spazio e tempo assoluti, le equazioni di Maxwell possono essere valide solamente in un particolare SDR (ed in quelli in quiete rispetto ad esso): il SDR a riposo rispetto all'etere.

Capite bene che i SDR in quiete rispetto all'etere sono privilegiati: solo in essi valgono le equazioni di Maxwell. E noi, la nostra Terra, il nostro sistema solare, come siamo messi rispetto all'etere? E soprattutto: esiste l'etere? Se, e ripeto se, l'etere esistesse, dovrebbe essere possibile con esperienze di elettromagnetismo e di ottica misurare la velocità di un qualunque sistema S rispetto all'etere, misurando in S la velocità della luce c_S in varie direzioni: troveremmo risultati diversi compresi nell'intervallo $[c - v, c + v]$ ove v è la velocità di S rispetto all'etere. Dopo la formulazione delle equazioni di Maxwell, furono fatti svariati esperimenti per cercare di rivelare questo “vento

³Vale l'invarianza in forma delle equazioni di Maxwell senza corrente di spostamento anche in presenza di cariche e correnti, assumendo per queste ultime la legge di trasformazione $\rho'(t', \mathbf{x}') = \rho(t, \mathbf{x}) - (\mathbf{v}/c) \cdot \mathbf{J}(t, \mathbf{x})$, $\mathbf{J}'(t', \mathbf{x}') = \mathbf{J}(t, \mathbf{x})$.

d'etere" che ci investe. È infatti inverosimile che la Terra, che ruota su se stessa ed intorno al Sole, sia sempre in quiete rispetto all'etere, a meno di non voler tornare ad un sistema tolemaico geocentrico. Siccome la Terra ruota attorno al Sole ad una velocità v di circa 30 Km/s, ci aspettiamo che misure della velocità della luce in diverse direzioni sulla Terra diano delle differenze dello stesso ordine di grandezza.⁴

1.2.3 L'esperimento di Michelson-Morley

L'esperimento più significativo in questa serie di tentativi di rilevare il moto della terra rispetto all'etere fu svolto da Michelson e Morley (1887). L'idea consiste nel confrontare il tempo impiegato dalla luce nel percorrere due cammini uguali, uno parallelo e l'altro ortogonale alla velocità della Terra rispetto all'etere. In modo estremamente schematico l'interferometro era realizzato come in fig. 1.5a.

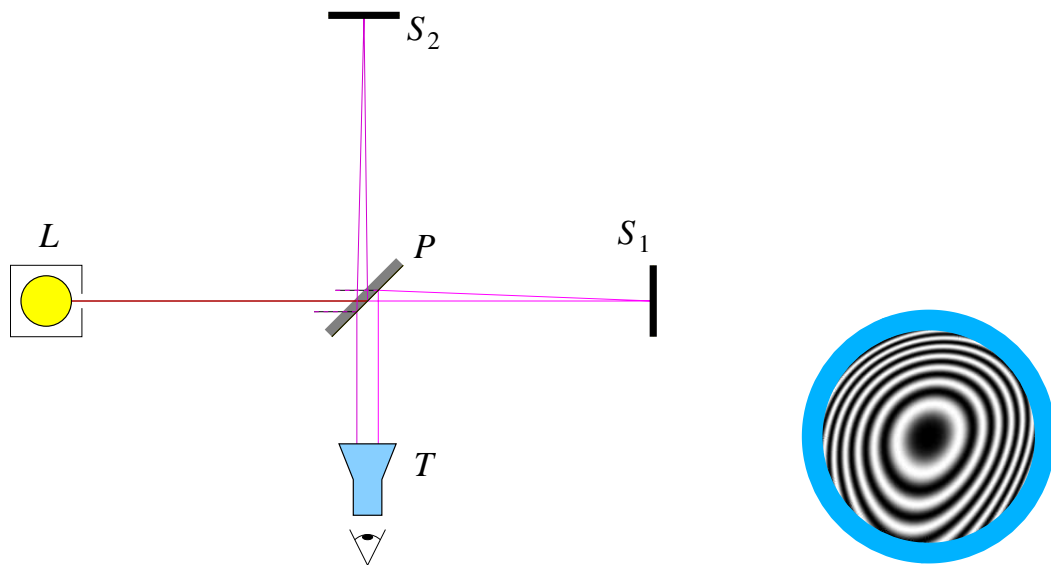


Figure 1.5: (a) Schema dell'interferometro di Michelson; (b) disegno delle frange d'interferenza che si osservano dal telescopio T .

Un raggio di luce proveniente dalla sorgente monocromatica L viene diviso dallo specchio semitrasparente P in due raggi mutuamente perpendicolari.

Il raggio 1 viene riflesso dallo specchio S_1 e ritorna in P ; qui una sua parte viene ulteriormente riflessa in modo che raggiunga il telescopio T .

Il raggio 2 viene riflesso dallo specchio S_2 e ritorna in P ; qui una sua parte riesce a passare e ad entrare nel telescopio T , dove *interferisce* con il raggio 1.

I due raggi compiono lo stesso percorso nel tratto iniziale da L a P e nel tratto finale da P a T . La differenza di cammino ottico tra i due raggi si ha nei viaggi di andata e ritorno da P a ciascuno dei due specchi. I due specchi S_1 ed S_2 sono posti ad una distanza l_1 ed l_2 da P pressapoco uguali, in modo che i due raggi compiano dei cammini ottici all'incirca uguali.

Anche se l'apparato fosse a riposo rispetto all'etere, dovremmo osservare delle frange d'interferenza in T come in fig. 1.5b, a causa delle inevitabili differenza nei due bracci PS_1 e PS_2 .

⁴Queste differenze di velocità di propagazione dovrebbero perciò essere rilevabili potendo disporre di una precisione $v/c \simeq 10^{-4}$ nelle misure. Misure dirette di $c_S = l/t$, ove t è il tempo impiegato da un segnale luminoso a percorrere la distanza l , richiedono precisione di 10^{-4} sia per l che per t . Se questa precisione non era difficile da ottenere per misure di distanza di laboratorio ($10^{-1} \lesssim l \lesssim 10^3$ metri), era praticamente impossibile per i tempi corrispondenti ($3 \cdot 10^{-10} \lesssim t \lesssim 3 \cdot 10^{-6}$). Esisteva già però un modo molto preciso per misurare tempi, o meglio, differenze di tempi che due fasci di luce impiegano per compiere due percorsi diversi: il metodo interferometrico. Tuttavia, negli interferometri i fasci di luce partono ed arrivano nello stesso punto, e siccome gli effetti del primo ordine in v/c si elidono tra fasci con medesimi punto di partenza e di arrivo, le differenze di tempo di percorrenza dovuti al moto rispetto all'etere sono del secondo ordine, cioè $v^2/c^2 \sim 10^{-8}$. Ciò nonostante, la precisione degli apparati interferometrici si può spingere fino a questi valori.

Supponiamo ora che l'apparato sia disposto con il braccio PS_1 parallelo alla direzione della velocità \mathbf{v} della terra rispetto all'etere, mentre l'altro braccio è posto in direzione ortogonale a \mathbf{v} .

Nel tratto di andata da P ad S_1 , il raggio 1 viaggia ad una velocità $c - v$ rispetto al laboratorio, mentre viaggia a velocità $c + v$ nel tratto di ritorno. Nel complesso, il raggio 1 impiega un tempo

$$t_1 = \frac{l_1}{c - v} + \frac{l_1}{c + v} = \frac{2l_1}{c \left(1 - \frac{v^2}{c^2}\right)}. \quad (1.60)$$

Per quanto riguarda il raggio 2, esso si muove sempre perpendicolarmente alla velocità dell'etere $-\mathbf{v}$, e quindi con modulo della velocità v_2 costante, tanto nel tratto di andata che in quello di ritorno. Questa velocità v_2 deve essere tale che nel SDR dell'etere il raggio si propaghi a velocità c . Siccome \mathbf{v}_2 è ortogonale alla velocità $-\mathbf{v}$ dell'etere rispetto al laboratorio, si deve avere

$$v_2^2 + v^2 = c^2 \quad \Rightarrow \quad v_2 = \sqrt{c^2 - v^2}, \quad (1.61)$$

e quindi

$$t_2 = \frac{2l_2}{\sqrt{c^2 - v^2}} = \frac{2l_2}{c\sqrt{1 - \frac{v^2}{c^2}}}. \quad (1.62)$$

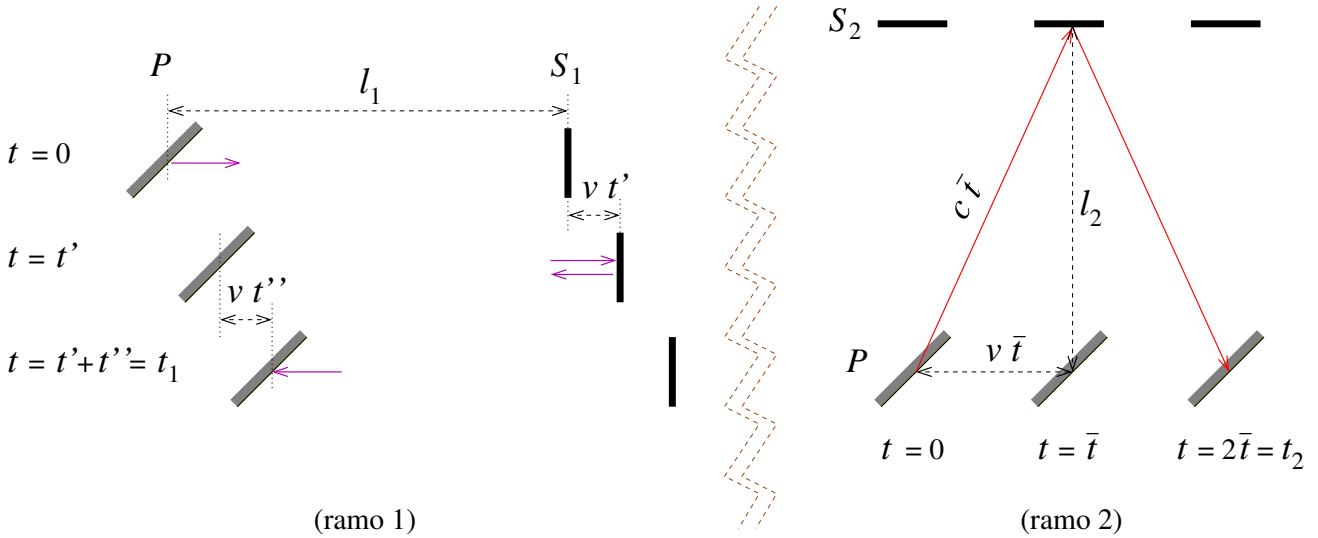


Figure 1.6: Percorsi dei raggi di luce nell'interferometro di Michelson visti dal SDR in quiete rispetto all'etere.

Equivalentemente, possiamo vedere la cosa nel SDR dell'etere in cui la luce viaggia a velocità c in ogni direzione. Come indicato nella fig. 1.6, nel tempo di andata t' da P a S_1 , lo specchio S_1 si allontana dal punto di uscita del raggio 1 da P di una distanza vt' , mentre la luce percorre una distanza ct' ; dovendo essere

$$l_1 + vt' = ct' \quad \text{si ha} \quad t' = \frac{l_1}{c - v}. \quad (1.63)$$

Nel tratto di ritorno, di durata t'' è lo specchio semitrasparente P che si avvicina al punto in cui S_1 ha riflesso il raggio 1 di una distanza vt'' . In questo caso, dovendo essere

$$l_1 - vt'' = ct'' \quad \text{si ha} \quad t'' = \frac{l_1}{c + v}, \quad (1.64)$$

e si ottiene $t_1 = t' + t''$ come nell'eq. (1.60).

Passiamo ora a calcolare, ancora nel SDR dell'etere, il tempo t_2 che impiega il raggio 2 a compiere il percorso P - S_2 - P . Nel tratto di andata da P a S_2 , di durata \bar{t} , lo specchio si sposta di una distanza $v\bar{t}$ perpendicolarmente alla retta P - S_2 , per cui la luce deve percorrere una distanza $c\bar{t}$ tale che

$$(c\bar{t})^2 = l_2^2 + (v\bar{t})^2 \quad \text{per cui} \quad \bar{t} = \frac{l_2}{\sqrt{c^2 - v^2}}. \quad (1.65)$$

Lo stesso tempo viene impiegato dal raggio 2 per tornare da S_2 a P , per cui $t_2 = 2\bar{t}$ e riotteniamo il risultato (1.62).

Possiamo a questo punto calcolare la differenza di tempo Δt tra i raggi 1 e 2, dovuta al moto dell'apparato sperimentale nell'etere:

$$\Delta t = t_1 - t_2 = \left[\frac{2l_1}{c \left(1 - \frac{v^2}{c^2}\right)} - \frac{2l_2}{c \sqrt{1 - \frac{v^2}{c^2}}} \right]. \quad (1.66)$$

Per osservare uno spostamento delle frange di interferenza, ruotiamo l'interferometro di 90° attorno all'asse verticale, scambiando così il ruolo dei due bracci: $P-S_1$ è ora perpendicolare a \mathbf{v} mentre $P-S_2$ ne è parallelo. In questo caso il tempo t'_1 necessario alla luce per percorrere il percorso $P-S_1-P$ si calcola con il procedimento per il braccio ortogonale alla velocità dell'etere ($t'_1 = t_2|_{l_2 \rightarrow l_1}$), mentre t'_2 con il procedimento per il braccio parallelo ($t'_2 = t_1|_{l_1 \rightarrow l_2}$). In questa configurazione, la differenza dei tempi tra i due raggi di luce si scrive come l'eq. (1.66) ma con i denominatori scambiati:

$$\Delta t' = t'_1 - t'_2 = \left[\frac{2l_1}{c \sqrt{1 - \frac{v^2}{c^2}}} - \frac{2l_2}{c \left(1 - \frac{v^2}{c^2}\right)} \right]. \quad (1.67)$$

Quindi, con la rotazione dell'interferometro, si dovrebbe osservare uno spostamento delle frange di interferenza, in numero pari a

$$\Delta n = \frac{\Delta t - \Delta t'}{\lambda/c} = \frac{2(l_1 + l_2)}{\lambda} \left[\frac{1}{\left(1 - \frac{v^2}{c^2}\right)} - \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \right], \quad (1.68)$$

in cui λ/c è il tempo impiegato dalla luce per percorrere un tratto pari ad una lunghezza d'onda, per cui $\Delta t = \lambda/c$ corrisponderebbe allo spostamento di una frangia. Essendo $v \ll c$, possiamo calcolare Δn sviluppando l'eq. (1.68) in serie di v/c . Il primo termine non nullo si trova al secondo ordine, e dà

$$\Delta n = \frac{2(l_1 + l_2)}{\lambda} \left[\left(1 + \frac{v^2}{c^2}\right) - \left(1 + \frac{1}{2} \frac{v^2}{c^2}\right) + \mathcal{O}\left(\frac{v^4}{c^4}\right) \right] \simeq \frac{l_1 + l_2}{\lambda} \frac{v^2}{c^2}. \quad (1.69)$$

Nell'esperimento originale, Michelson e Morley utilizzarono un sistema di specchi a riflessione multipla, cosicché le lunghezze equivalenti dei due bracci dell'interferometro risultarono $l_1 \simeq l_2 \simeq 11\text{m}$; con luce di lunghezza d'onda $\lambda = 550\text{nm}$ e velocità della Terra rispetto all'etere $v \simeq 30\text{Km/s}$, si sarebbe dovuto osservare uno spostamento di $\Delta n \simeq 0.4$ frange. Il risultato dell'esperimento fu che non venne osservato alcuno spostamento di frange, nonostante che il valore aspettato fosse due ordini di grandezza superiore alla sensibilità dell'apparato (in pratica, si riusciva ad apprezzare uno spostamento di 0.004 frange).

Nota: Evidentemente, anche una piccolissima differenza $l_1 - l_2$ tra la lunghezza dei due bracci dell'ordine di frazioni della lunghezza d'onda λ produrrebbe uno spostamento delle frange rispetto alla "situazione a riposo" $v = 0$, $l_1 = l_2$ in cui $\Delta n = 0$. La rotazione di 90° dell'interferometro elimina questa incertezza sulla lunghezza dei bracci, in quando lo spostamento delle frange è proporzionale alla somma, e non alla differenza tra le due lunghezze, come si vede nell'eq. (1.68). In altre parole, la differenza di tempo Δt tra l'arrivo dei due raggi nella configurazione iniziale è dovuta a due fattori: 1) la velocità della Terra rispetto all'etere (quantificata nella differenza tra i due denominatori in t_1 e in t_2); 2) la differenza di lunghezza dei due bracci. Nella rotazione dell'interferometro, l'effetto su $\Delta t'$ dovuto alla velocità della Terra si inverte, mentre quello dovuto alle differenze di lunghezza resta uguale. Pertanto, nella differenza $\Delta t - \Delta t'$ la differenza delle lunghezze tra i due bracci si elide (quasi esattamente). Questa osservazione, oltre a non porre il grosso problema costruttivo di due bracci esattamente uguali, fa sì che eventuali piccole variazioni indipendenti di l_1 ed l_2 che si potrebbero avere a seguito della rotazione, o comunque al passare del tempo, non causano uno spostamento apprezzabile delle frange.

Per ridurre comunque al minimo qualsiasi deformazione dell'apparato, questo venne montato su un blocco di pietra di oltre un metro di diametro, il quale a sua volta galleggiava in un bagno di Mercurio. In questo modo era possibile ruotare il sistema evitando vibrazioni o deformazioni indesiderate.

1.2.4 Tentativi di salvare l'ipotesi dell'etere

La conclusione che se ne sarebbe potuto trarre era che la Terra fosse ferma rispetto all'etere. Se l'esperimento fosse stato fatto tre secoli prima, molti avrebbero interpretato il risultato come una conferma del fatto che la

Terra si trovi ferma al centro dell'Universo. Ma già nel XIX secolo nessuno ci credeva più. Molte idee furono però avanzate per salvare l'ipotesi dell'etere.

Una delle idee più semplici sosteneva che i corpi materiali, come la Terra, trascinassero con sé l'etere nelle loro vicinanze, così come una palla che si muove nell'aria trascina con sé un sottile strato di aria in prossimità della sua superficie. Questa spiegazione però era in contrasto con l'osservazione dell'aberrazione stellare.

Un'altra ipotesi era quella che riteneva che la velocità della luce vale c non rispetto all'etere, ma rispetto alla sorgente che la emette, come dei proiettili sparati da un fucile. Questo fatto però avrebbe prodotto degli effetti molto strani, non osservati, nel moto apparente delle stelle doppie.

FitzGerald e Lorentz spiegarono il risultato negativo di Michelson-Morley sostenendo che i corpi che si muovono con velocità v rispetto all'etere subiscono una contrazione di un fattore $\sqrt{1 - v^2/c^2}$ nel senso del moto. Se così fosse, il raggio di luce parallelo alla velocità della terra rispetto all'etere avrebbe un po' meno distanza da percorrere, ed arriverebbe perfettamente in fase con il fascio ortogonale. Questa contrazione era giustificata calcolando, tramite le equazioni di Maxwell, il campo elettromagnetico prodotto da una carica in moto, che risulta essere non più a simmetria sferica, bensì deformato nel senso del moto. Se tutta la materia (atomi) è tenuta assieme dalla forza elettromagnetica, allora tutto si deforma nella stessa misura in direzione del moto, e nessuno si accorge della contrazione.

Altre idee furono avanzate per conciliare l'ipotesi di esistenza dell'etere con i risultati sperimentali, nessuno dei quali riusciva a rivelarne la presenza.

1.2.5 Principio di relatività di Einstein

Einstein nel 1905 risolse il problema in modo radicale: elevando a principio i risultati negativi di ricerca dell'etere, postulò che l'etere non esiste. Se non esiste un SDR privilegiato per l'elettromagnetismo, allora non solo le leggi della meccanica, ma tutte le leggi della fisica sono invarianti in forma nel passare da un SDR inerziale ad un altro (*principio di relatività di Einstein*).

Poiché le leggi dell'elettromagnetismo non sono invarianti in forma per trasformazioni di Galileo, ne segue che se si ammette il principio di relatività di Einstein ci si trova di fronte ad un dilemma:

- o si modificano le leggi dell'elettromagnetismo (cioè i fenomeni elettromagnetici sono descritti da equazioni diverse da quelle di Maxwell e Lorentz)
- o si modificano le leggi di trasformazione di Galileo (cioè le leggi di trasformazione di coordinate tra osservatori inerziali non sono quelle del gruppo di Galileo).

1.2.6 Postulati della relatività

I tentativi di modificare le leggi dell'elettromagnetismo si rivelarono infruttuosi. Non restò quindi che percorrere la seconda strada e tentare di modificare la legge di trasformazione di Galileo.

Il carattere rivoluzionario di tale scelta sta nel fatto che le leggi di trasformazione di Galileo sembrano a prima vista ovvie e naturali. Fu merito di Einstein aver messo in evidenza che esse invece si basano su alcune ipotesi a priori: l'omogeneità ed isotropia dello spazio e del tempo (postulato M1) e l'assolutezza degli intervalli spaziali e temporali (postulato M2), come abbiamo visto nella sez. 1.1.16. In particolare, l'ipotesi M2 sembra ovvia perché sembra confermata dall'esperienza quotidiana. Ma l'esperienza quotidiana ha a che fare con velocità piccole rispetto a quelle della luce e quindi tutto ciò che possiamo concludere è che M2 è valida approssimativamente per velocità piccole rispetto a quelle della luce.

Einstein accetta M1, ma sostituisce M2 con l'ipotesi confermata dalle esperienze di elettromagnetismo ed ottica:

R2) *La velocità della luce nel vuoto è la stessa in tutti i sistemi di riferimento inerziali.*

In particolare, tale velocità (che indichiamo con c) non dipende dalla direzione di propagazione della luce, né dallo stato di moto della sorgente che l'ha emessa.

Chiaramente M2 è incompatibile con R2. Consideriamo per esempio la propagazione di un segnale luminoso emesso dalla lampada L e ricevuto dalla fotocellula F , solidali con l'osservatore S . Indichiamo con A l'evento corrispondente all'emissione del raggio di luce da parte di L e con B l'evento corrispondente alla sua ricezione da parte di F . Se l è la distanza tra lampada e fotocellula misurata da S , e τ è il tempo che intercorre tra emissione e ricezione, avremo $c = l/\tau$. Graficamente (fig. 1.7a), le linee universo di L ed F osservate da S sono rette parallele all'asse temporale (essendo questi oggetti in quiete rispetto ad S), mentre la linea universo del raggio luminoso corrisponde ad una retta di pendenza $\Delta t/\Delta x = 1/c$, avendo indicato con x l'asse che congiunge L con F .

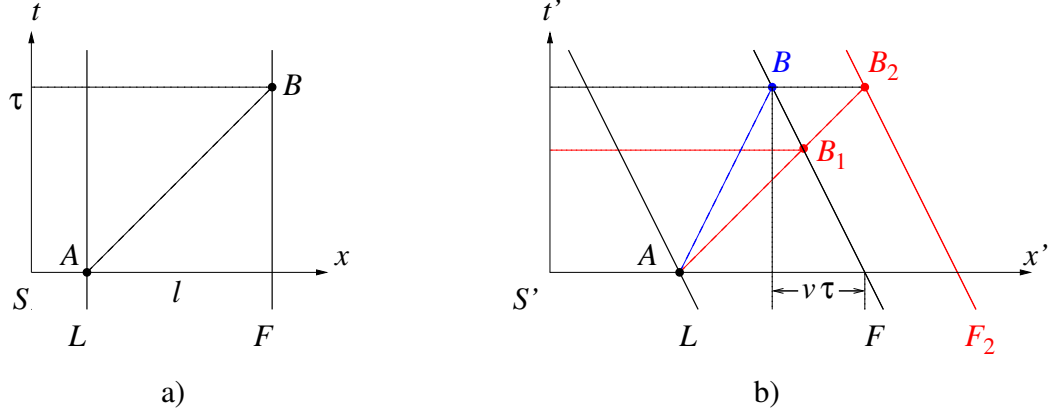


Figure 1.7: Diagrammi spazio-temporali di un raggio di luce emesso da una lampadina L e ricevuto da una fotocellula F in due SDR S ed S' in moto relativo. a) L ed F sono in quiete rispetto ad S , il raggio luminoso è indicato dalla linea nera tratteggiata; b) diagramma nel SDR S' in moto a velocità v rispetto ad S : in blu è riportata la traiettoria del raggio di luce secondo i postulati della meccanica Newtoniana; in rosso la traiettoria secondo la fisica relativistica.

Per l'osservatore S' in moto rispetto ad S a velocità v nella direzione x , la lampada L e la fotocellula F si muovono con velocità $-v$ lungo il suo asse x' , pertanto le linee universo di L ed F in S' sono rappresentate da due rette con pendenza $\Delta t'/\Delta x' = -1/v$, sulle quali troviamo gli eventi A e B di emissione e ricezione.

Se vale M2, l'intervallo impiegato dalla luce per andare da A a B misurato da S' è uguale a $\Delta t' = \tau$. In questo lasso di tempo, la fotocellula che si trovava a distanza l al momento dell'emissione del lampo, si è spostata verso sinistra di un tratto $v\tau$, per cui la distanza percorsa dal raggio di luce in S' è pari a $\Delta x' = l - v\tau$ (fig. 1.7b, linee nere). Quindi il segnale luminoso si è propagato con velocità

$$c' = \frac{\Delta x'}{\Delta t'} = \frac{l - v\tau}{\tau} = \frac{l}{\tau} - v = c - v \quad (1.70)$$

in accordo con la legge di composizione delle velocità, ed in contrasto con R2.

Se invece vale R2, anche in S' il raggio di luce è descritto da una retta con pendenza $1/c$ (fig. 1.7b, linea rossa tratteggiata). Ma allora, se la distanza tra L ed F fosse uguale ad l , l'evento di ricezione si troverebbe in B_1 , ed il tempo di propagazione misurato da S' sarebbe $\tau' < \tau$. Viceversa, se assumiamo che il tempo di propagazione sia lo stesso per entrambi gli osservatori, l'evento di ricezione si troverebbe in B_2 , e quindi la linea universo della fotocellula (F_2 in figura) dovrebbe trovarsi ad una distanza dalla lampadina $l' > l$ maggiore di quella osservata da S' . In ogni caso, le ipotesi di M2 non sono soddisfatte.

Riassumiamo quindi i postulati della fisica relativistica:

R1) Lo spazio vuoto è omogeneo ed isotropo; il tempo è omogeneo.

R2) La velocità della luce nel vuoto è la stessa in tutti i sistemi di riferimento inerziali.

R3) Principio di relatività einsteiniana:

Tutte le leggi della fisica hanno la stessa forma in tutti i sistemi di riferimento inerziali.

Per quanto riguarda le leggi dell'elettromagnetismo, il postulato di relatività einsteiniana R3 implica che le equazioni di Maxwell sono valide in qualunque SDR inerziale. Il principio di inerzia (prima legge della meccanica) è implicitamente assunto valido nella definizione di sistemi inerziali. Le altre leggi della meccanica andranno invece modificate, come mostreremo nei capitoli seguenti.

Chapter 2

Cinematica relativistica

In questo capitolo, partendo dai postulati della fisica relativistica, mostreremo come lo spazio ed il tempo non siano grandezze assolute ma dipendano dall'osservatore che li misura. In particolare, il tempo e lo spazio dovranno essere considerati non più come due grandezze indipendenti, ma due “aspetti” della medesima entità fisica, lo *spazio-tempo*.

Abbiamo già avuto modo di osservare nella sez. 1.2.6 l'incompatibilità dei postulati della meccanica Newtoniana con quello della invarianza della velocità della luce. Mostriamo ora come questa invarianza comporti una radicale revisione delle nozioni di spazio e di tempo, confrontando la propagazione di segnali luminosi in SDR diversi.

2.1 Relatività della simultaneità

Una conseguenza sorprendente dell'invarianza della velocità della luce riguarda il concetto di simultaneità di due eventi.

Consideriamo un osservatore S dotato di un radar, ossia un emettitore di segnali elettromagnetici, dei quali può rivelare il riflesso da parte di oggetti distanti. All'istante $t = 0$ il radar, che coincide con l'origine delle coordinate di S , emette un segnale che si propaga in tutte le direzioni. Indichiamo con A questo evento: Supponiamo che ad una certa distanza l lungo l'asse x in direzioni opposte siano posti due specchi h_1 e h_2 che riflettono il segnale radar. Siano B_1 e B_2 gli eventi corrispondenti alla riflessione. I due segnali riflessi, avendo percorso distanze uguali alla stessa velocità, vengono rivelati dal radar allo stesso istante in corrispondenza dell'evento D .

$$t_A = 0, \quad x_A = 0 \quad (2.1)$$

$$t_{B_1} = t_{B_2} = \frac{l}{c} \equiv \tau \quad x_{B_1} = -l, \quad x_{B_2} = +l \quad (2.2)$$

$$t_D = \frac{2l}{c} = 2\tau, \quad x_D = 0 \quad (2.3)$$

In particolare, deduciamo che gli eventi B_1 e B_2 di riflessione sono simultanei, essendo avvenuti dopo lo stesso tempo τ successivo all'emissione. Graficamente, la situazione è presentata in fig. 2.1a, in cui rappresentiamo solamente gli assi x e t . Se scegliamo l'unità di misura del tempo in modo tale che 1 cm sull'asse t corrisponda al tempo impiegato dalla luce a percorrere la distanza spaziale di 1 cm, ($c = 1$) le linee universo dei raggi di luce sono composte da segmenti inclinati di 45° rispetto agli assi, ossia con pendenza ± 1 a seconda che si propaghino in verso positivo o negativo lungo l'asse x . Con questa scelta, i due segnali radar chiudono un quadrato.

Come viene vista questa sequenza di eventi da un osservatore S' che si muova lungo l'asse x con velocità $v > 0$? Scegliamo l'origine dei tempi e delle coordinate di S' in corrispondenza dell'evento di emissione A

$$t'_A = 0, \quad x'_A = 0$$

e l'asse x' corrispondente alla traiettoria rettilinea degli specchi h_1 e h_2 .¹

¹Siccome l'origine O' di S' coincide con la posizione del radar a $t = 0$, e si muove a velocità v lungo l'asse individuato dagli

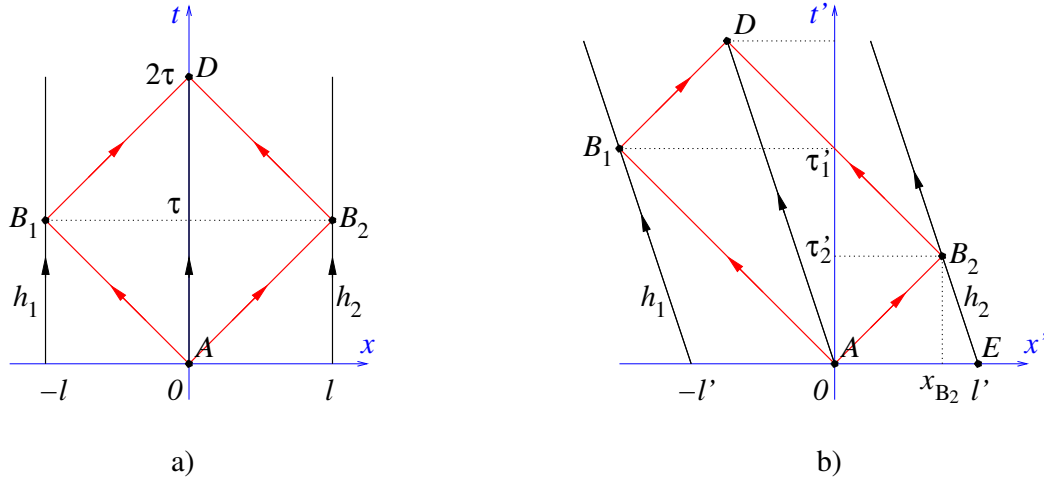


Figure 2.1: *Diagrammi spazio-temporali di emissione (A), riflessione (B_1 e B_2) e ricezione (D) di un segnale radar (in rosso) da parte di due specchi h_1 e h_2 equidistanti dalla sorgente. a) l'osservatore S giudica i due eventi di riflessione simultanei; b) per l'osservatore S' l'evento B_1 avviene successivamente all'evento B_2 .*

Evidentemente, per S' il radar (così come tutti gli oggetti solidali con S) si muove con velocità $-v$ lungo l'asse x' .² Pertanto, nel diagramma spazio-temporale di S' gli eventi di emissione e ricezione giacciono sulla linea universo del radar rappresentata da una retta inclinata con pendenza $-1/v$. Lo stesso dicasi per le linee universo dei due specchi. I segnali radar, che postuliamo muoversi a velocità c anche rispetto ad S' , dovranno ritrovarsi in D dopo essere stati riflessi. Per ricostruire le traiettorie dei due impulsi radar — che dovranno essere di moto rettilineo uniforme nel piano (t', x') nei tratti tra radar e specchi — tracciamo le due rette a 45° dall'evento A (segnale emesso), e due rette a 45° che arrivino in D (segnali riflessi). Le intersezioni di queste rette determinano gli eventi di riflessione B_1 e B_2 nel diagramma spazio-temporale di S' . Da semplici considerazioni di geometria applicate al rettangolo AB_1DB_2 — i triangoli rettangoli AB_1D e DB_2A sono congruenti — ed alle sue proiezioni sugli assi t' ed x' , ricaviamo facilmente le seguenti relazioni:

$$t'_{B_1} \equiv \tau'_1 = t'_D - t'_{B_2} \quad (2.4)$$

$$t'_{B_2} \equiv \tau'_2 = t'_D - t'_{B_1} \quad (2.5)$$

$$(2.4)+(2.5) \implies t'_D = \tau'_1 + \tau'_2 \quad (2.6)$$

$$-x'_{B_1} = x'_{B_2} - x'_D = c\tau'_1 \quad (2.7)$$

$$x'_{B_2} = x'_D - x'_{B_1} = c\tau'_2 \quad (2.8)$$

$$(2.8)-(2.7) \implies x'_D = c(\tau'_2 - \tau'_1) \quad (2.9)$$

(fig. 2.1)

Dall'ultima uguaglianza (2.9) ricaviamo $0 = v(\tau'_1 + \tau'_2) + c(\tau'_2 - \tau'_1)$ e infine

$$\frac{\tau'_2}{\tau'_1} = \frac{c-v}{c+v} < 1 \implies \tau'_1 > \tau'_2. \quad (2.10)$$

Insomma, come è evidente dalla fig. 2.1, per S' la riflessione sullo specchio h_1 avviene successivamente a quella sullo specchio h_2 , a differenza di quanto osservato da S secondo il quale le riflessioni sono simultanee.

L'opinione così generalmente accettata che la simultaneità di due eventi sia un fatto assoluto, indipendente dall'osservatore, viene così a cadere: due eventi possono essere simultanei per un osservatore (S) e non esserlo per un altro (S'). Il concetto di simultaneità è relativo (al sistema di riferimento).

specchi, O' passerà prima per h_1 e poi per h_2 . Quindi, dal punto di vista di S' , entrambi gli specchi si muovono su rette parallele passanti per l'origine O' , e quindi sulla stessa retta. Questa individua l'asse x' .

²Infatti, dalla supposta equivalenza tra due SRD inerziali, i moduli delle velocità relative con cui ciascuno dei due SDR vede l'altro devono essere uguali (altrimenti quello che misurasse una velocità relativa maggiore sarebbe in qualche modo privilegiato rispetto all'altro). I valori delle singole componenti ed il loro segno invece dipendono dalla scelta di orientazione dei rispettivi assi.

È facile trarre una conseguenza ancora più sbalorditiva dal ragionamento precedente: un terzo osservatore S'' che si muova con velocità $-v$ rispetto a S lungo l'asse x giudicherebbe l'evento B_1 precedente a B_2 . Infatti, invertendo il segno della velocità relativa v nelle deduzioni fatte sopra, otterremmo ³

$$\frac{\tau_2''}{\tau_1''} = \frac{c - (-v)}{c + (-v)} > 1 \implies \tau_1'' < \tau_2''.$$

In altre parole, per diversi osservatori passanti per B_1 , lo stesso evento B_2 avviene nel presente (per S), nel passato (per S') oppure nel futuro (per S'').

Capite bene quali conseguenze paradossali comporti l'assunzione del postulato R2 di invarianza della velocità della luce, e quanto difficile sia stato accettarlo (e quale genio sia stato necessario per proporgli). Solamente una lunga serie di tentativi falliti di misurare lo stato di moto rispetto all'etere, che non hanno rivelato la minima differenza di velocità della luce, ha portato i fisici a prendere in seria considerazione questo postulato.

Come vedremo, superate le difficoltà dovute alla rivoluzione concettuale di spazio e di tempo, le conseguenze sperimentali predette dalla fisica relativistica sono state così precisamente verificate dalle esperienze che ormai non vi è più alcun dubbio che i postulati della relatività debbano soppiantare quelli della meccanica Newtoniana ed essere alla base di qualsiasi teoria fisica.⁴

2.2 Invarianza delle distanze ortogonali al moto

Nella procedura della sez. 2.1 abbiamo ricavato delle relazioni tra le distanze ed i tempi di vari eventi osservati in SDR in moto relativo. Per poter quantificare con precisione che relazione ci sia tra le coordinate (t, x) e (t', x') (i piani lungo i quali si svolge il moto relativo di S e S'), analizziamo innanzitutto cosa succede a misure di distanza effettuate in direzione ortogonale alla velocità relativa.

Come prima, supponiamo che il SDR inerziale S' si muova con velocità uniforme \mathbf{v} rispetto al SDR inerziale S , in modo che ad un certo istante le due origini O ed O' coincidano. Prendiamo questo evento A come origine dei tempi sia di S che S' : $(t_A, \mathbf{x}_A) = (0, \mathbf{0}) = (t'_A, \mathbf{x}'_A)$. La direzione di moto di O' rispetto ad S identifichi l'asse x (in direzione e verso), e la direzione di moto di O rispetto ad S' identifichi l'asse x' (in direzione ma verso opposto).

Consideriamo ora in S un filo rettilineo parallelo all'asse x posto ad una distanza d da esso. Possiamo immaginare che su questo filo siano poste delle palline, alcune $\{q_1, q_2, \dots\}$ in quiete (solidali ad S), alcune $\{m_1, m_2, \dots\}$ che scorrono con velocità \mathbf{v} (solidali ad S'). Ci chiediamo come S' veda questo filo, cioè la traiettoria lungo la quale si muovono tutte le palline.

1) Innanzitutto, è ragionevole supporre che S' veda le palline $\{m_1, m_2, \dots\}$ ferme, muovendosi queste con la stessa velocità di S' ,⁵ per cui tutto il filo è visto da S' nella stessa posizione al passare del tempo.

2) Che forma avrà il filo in S' ? Essendo S' un SDR inerziale, vedrà le palline $\{q_1, q_2, \dots\}$ in moto rettilineo uniforme, pertanto il filo giace lungo una retta.

3) Com'è messa questa retta in S' ? Non può intersecare l'asse x' , altrimenti quando O passa per quel punto dell'asse x' la sua distanza dal filo sarebbe nulla, mentre il filo è a distanza fissa d da O . Può essere il filo sghembo rispetto all'asse x' ? Se così fosse, S' vedrebbe le palline $\{q_1, q_2, \dots\}$ prima avvicinarsi e poi allontanarsi da O . In particolare, per S' ci sarebbe un punto x'_A dell'asse x' di minima distanza dal filo con le palline. Allora due osservatori solidali ad S' posti lungo l'asse x' (e visti muoversi lungo l'asse x da S), uno in posizione x'_A ed un'altro in posizione $x'_B \neq x'_A$ vedrebbero il filo a distanze diverse, mentre in S entrambi sono alla stessa distanza d dal filo. Questo è in contrasto con l'ipotesi di omogeneità dello spazio. Deduciamo così che S' vede il filo posto in linea retta parallelamente all'asse x' , analogamente a quanto vede S .

4) Ci chiediamo infine: a quale distanza d' da O' si trova il filo in S' ? Dalla simmetria della condizione del filo e delle palline su di esso tra S ed S' non abbiamo motivo alcuno di ritenere che $d > d'$ o $d < d'$. Se così

³Il diagramma spazio-temporale in S'' si ottiene riflettendo quello di S' rispetto all'asse dei tempi e scambiando i pedici $1 \leftrightarrow 2$ in modo da avere h_1 a sinistra e h_2 a destra.

⁴Naturalmente, la storia della fisica ci insegna che niente è definitivo, e quindi c'è spazio per ulteriori sviluppi dei nuovi postulati, ma queste sono cose che vedrete più avanti.

⁵ S' non può ruotare attorno ad x' , altrimenti non sarebbe un SDR inerziale.

fosse, un osservatore potrebbe distinguere tra i due SDR, assegnando a quello che vede il filo più vicino un ruolo speciale. Questo però è in conflitto con l'indistinguibilità dei SDR inerziali, quindi $d = d'$. Ne concludiamo che le distanze ortogonali alla direzione del moto relativo sono invarianti per i due osservatori.

Un'altra dimostrazione può essere la seguente: Consideriamo ancora due sistemi di riferimento, S costituito da una galleria ferroviaria, ed S' solidale ad un treno che entra in galleria a velocità v (ammettiamo anche che la dimensione trasversale dei vagoni in quiete sia soltanto infinitesimamente inferiore a quella della galleria). Dimosteremo che i vagoni non subiscono alcuna alterazione nella loro dimensione trasversale, chiunque sia ad eseguire le misure. Daremo la dimostrazione per assurdo, cioè dimostreremo che si arriva a un'assurdità ammettendo che la dimensione trasversale dei vagoni in corsa subisca alterazioni, in diminuzione o in aumento. Giudichiamo il fenomeno dal punto di vista del sistema S (galleria): la galleria è ferma ed il treno è in moto a grande velocità. Se ammettiamo per ipotesi che la dimensione trasversale dei vagoni diminuisca con la velocità, il treno scorre nella galleria senza problemi, e tanto più facilmente quanto maggiore è la sua velocità. Ora esaminiamo il fenomeno dal punto di vista del sistema S' (treno): il treno è fermo e la galleria scorre a grande velocità. Se ammettiamo per ipotesi che la dimensione trasversale della galleria diminuisca al crescere della velocità, questa cessa di scorrere (treno e galleria si incastrano). Si ha un'assurdità perché non può essere che uno stesso fenomeno avvenga o non avvenga per due osservatori posti in sistemi di riferimento diversi. Allora, siccome ammettere che la dimensione trasversale degli oggetti in movimento diminuisce con la velocità porta a un assurdo, quell'ammissione dev'essere respinta. Allo stesso modo si dimostra che non si può ammettere che la dimensione trasversale degli oggetti in movimento subisca dilatazione. Quindi rimane invariata.

2.3 Dilatazione dei tempi

Ci proponiamo ora di calcolare, nel SDR S' , il tempo $\Delta t' = t'_D$ impiegato dai due segnali radar per raggiungere il rivelatore in D dal momento dell'emissione in A . Sfruttiamo l'invarianza delle distanze perpendicolari al moto ricavata nella sez. 2.2. Posizioniamo quindi in S uno specchio h_3 , fermo a distanza l dal radar, e parallelo a v (per esempio sull'asse y , come in fig. 2.2a). Evidentemente, il segnale emesso dal radar in A e riflesso in B_3 da h_3 sarà rivelato in D assieme ai due segnali precedenti, quindi dopo un tempo $\Delta t = 2\tau$ dall'emissione.

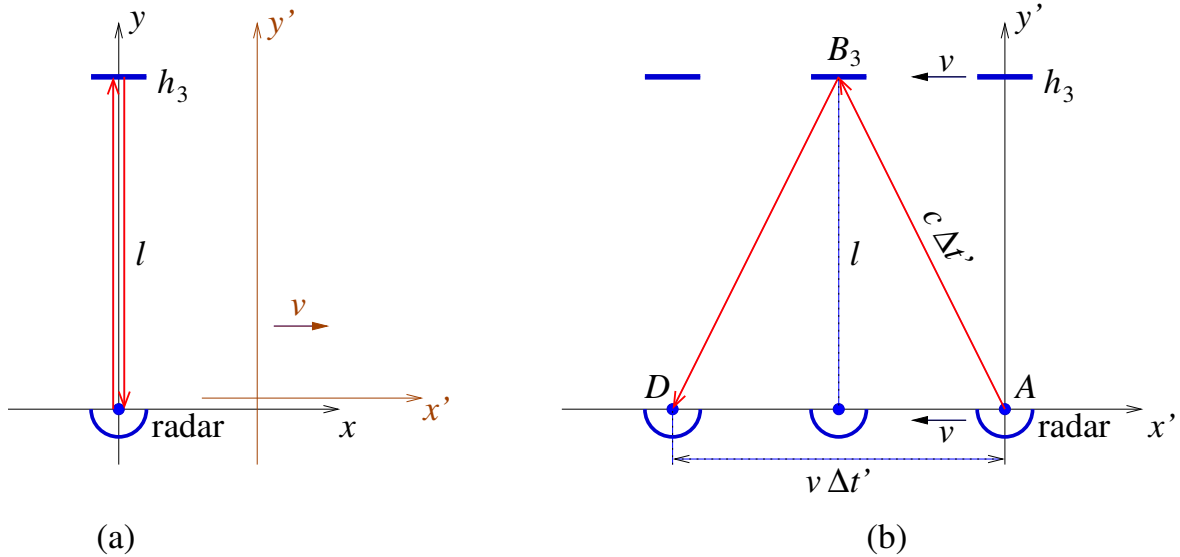


Figure 2.2: *Diagramma spaziale dell'emissione, riflessione e ricezione del segnale radar in direzione trasversale al moto: (a) nel SDR S solidale al radar ed allo specchio h_3 ; (b) nel SDR S' in moto con velocità v rispetto ad S .*

Osservando questo processo dal punto di vista di S' , il segnale è ricevuto dopo un tempo $\Delta t' = t'_D$, e quindi ad una distanza $|x'_D| = v\Delta t'$ dal punto di emissione, dopo aver percorso il tratto AB_3D come in fig. 2.2b. Siccome la distanza di h_3 dall'asse x' vale l , ed il triangolo AB_3D è isoscele per le leggi della riflessione, abbiamo che $x'_{B_3} =$

$-v\Delta t'/2$, $y'_{B_3} = l$. Deduciamo così che questo segnale ha percorso in S' una distanza pari a $2\sqrt{(v\Delta t'/2)^2 + l^2}$. Ma, dato che la velocità della luce vale c anche in S' , si ha anche

$$c\Delta t' = 2\sqrt{(v\Delta t'/2)^2 + l^2} \quad (2.11)$$

che, risolta per $\Delta t'$ dà

$$\Delta t' = \frac{2l}{\sqrt{c^2 - v^2}} = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (2.12)$$

Ciò significa che per S' l'intervallo di tempo tra l'emissione ed il ricevimento dei segnali radar è più lungo dell'intervallo di tempo misurato da S . Questo risultato sorprendente segue necessariamente dalle ipotesi fatte: dall'indistinguibilità dei SDR inerziali abbiamo dedotto l'invarianza delle distanze ortogonali al moto. Quindi in S' è più lungo il percorso compiuto dal terzo segnale radar. Se in più si assume che la velocità della luce sia identica nei due SDR, non resta che una possibilità: $\Delta t' > \Delta t$.

Bisogna rendersi subito conto che, anche per le velocità più elevate a cui si riescono a far viaggiare le persone o comunque degli strumenti di misura macroscopici, diciamo $v \sim 10^3$ m/s, il fattore di dilatazione temporale

$$\gamma \equiv \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \simeq 1 + \frac{v^2}{2c^2} \quad (v \ll c) \quad (2.13)$$

differisce assai poco dall'unità, $\gamma - 1 \simeq 10^{-11}$, molto meno di una parte su un miliardo, ed è impercettibile ai comuni strumenti di misura. Non sfugge invece ai moderni orologi atomici montati sui satelliti dei GPS, che hanno un'accuratezza dell'ordine di 50 ns (corrispondente ad un'accuratezza di 15 metri sulla posizione assoluta).⁶ Questi, viaggiando a velocità di circa 4 Km/s rispetto agli orologi a Terra, ogni ora (3600 s) perdono su per giù 300 ns, ossia dopo 10 minuti hanno già perso il sincronismo necessario all'accuratezza di 15 metri.⁷

2.4 Contrazione delle lunghezze

Il risultato precedente sulla dilatazione dei tempi ci permette di specificare completamente le misure effettuate dall'osservatore S' nell'esperimento con i radar della sez. 2.1. Infatti, dalle eq. (2.12) e (2.9) ricaviamo rispettivamente

$$\tau'_1 + \tau'_2 \stackrel{(2.6)}{=} t'_D \stackrel{(2.12)}{=} \frac{t_D}{\sqrt{1 - \frac{v^2}{c^2}}} \stackrel{(2.3)}{=} \frac{2\tau}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.14)$$

$$\tau'_1 - \tau'_2 \stackrel{(2.9)}{=} \frac{v}{c}(\tau'_1 + \tau'_2) \quad (2.15)$$

e quindi

$$\tau'_1 = \frac{1}{2} \left(1 + \frac{v}{c}\right) (\tau'_1 + \tau'_2) = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}} \tau \quad (2.16)$$

$$\tau'_2 = \frac{1}{2} \left(1 - \frac{v}{c}\right) (\tau'_1 + \tau'_2) = \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}} \tau. \quad (2.17)$$

Si vede anche che $\tau'_2 < \tau < \tau'_1$.

La distanza tra il radar e gli specchi misurata da S' la possiamo ottenere calcolando la distanza tra l'evento A e l'evento E dello specchio h_2 simultaneo ad A in S' (vedi fig. 2.1b). In pratica, basta aggiungere all'ascissa dell'evento B_2 lo spostamento dello specchio h_2 nell'intervallo di tempo τ'_2 :

$$l' = x'_{B_2} + v\tau'_2 = (c + v)\tau'_2 = c \left(1 + \frac{v}{c}\right) \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}} \tau = \sqrt{1 - \frac{v^2}{c^2}} l \quad (2.18)$$

⁶L'accuratezza sulla posizione differenziale è ben maggiore: a livello dei centimetri.

⁷Al ritardo dovuto alla dilatazione dei tempi va aggiunto l'anticipo dovuto allo spostamento gravitazionale: la teoria della relatività generale prevede che orologi vicini a sorgenti di campo gravitazionale (Terra) procedano più lentamente di quelli lontani (satellite). L'effetto gravitazionale è un anticipo di circa 1900 ns all'ora per gli orologi sui satelliti. Togliendo il ritardo di 300 ns dovuto agli effetti del moto, resta un anticipo netto di 1600 ns ogni ora.

Otteniamo così un altro risultato sorprendente: le distanze degli oggetti in movimento sono contratte di un fattore $\sqrt{1 - v^2/c^2}$ nella direzione del moto.

A scanso di equivoci, quanto mostrato finora non significa che i tempi misurati in S' sono dilatati rispetto a quelli misurati in S , né che le distanze misurate in S' sono contratte rispetto a quelle misurate in S . Ciò sarebbe infatti una palese violazione dell'indistinguibilità dei SDR inerziali (S' sarebbe distinguibile da S per il fatto di dare luogo a misure di distanze più corte). Oppure si potrebbe obiettare che se S' misura distanze più corte di S , scambiando il ruolo di S con S' troveremmo che S misura distanze più corte di S' , quindi $l < l' < l$ che è assurdo. Quello che invece abbiamo mostrato va interpretato nel seguente modo:

Tempo: l'intervallo di tempo misurato tra due eventi da un osservatore rispetto al quale questi eventi avvengono nello stesso luogo è sempre minore dell'intervallo di tempo tra gli stessi eventi misurati da un altro osservatore in moto relativo (per il quale i due eventi avvengono in luoghi diversi). Consideriamo per esempio la successione di eventi costituita dagli scatti delle lancette di un orologio. Gli osservatori in moto rispetto all'orologio lo vedranno procedere più lentamente di quanto non veda un osservatore solidale ad esso. E tanto maggiore è la velocità degli osservatori (rispetto all'orologio), tanto più lento sarà visto scandire i secondi. In particolare, un orologio (e qualsiasi altra cosa) che viaggiasse alla velocità della luce sarebbe “congelato” nel tempo, come se per esso il tempo non passasse mai.

Lunghezza: le dimensioni che un osservatore attribuisce ad un oggetto in quiete non sono le stesse di quelle attribuite da un osservatore in moto rispetto all'oggetto: le dimensioni nella direzione del moto subiscono una contrazione tanto maggiore quanto maggiore è la velocità dell'oggetto rispetto all'osservatore. In particolare, un oggetto che si muovesse alla velocità della luce sarebbe contratto ad un foglio perpendicolare alla velocità.

Un'ultima osservazione: queste stranezze di orologi che procedono a rallentare e oggetti compressi a seconda della velocità non sono dovute a “deformazioni elastiche” causate da qualche sollecitazione, né sono “effetti ottici” dovuti alla velocità finita di propagazione della luce. È il tempo vero e proprio che scorre in modo diverso per osservatori in moto relativo, così come è lo spazio reale che, non più assoluto, cambia di estensione da un osservatore all'altro. Questi concetti saranno più chiari quando interpreteremo il tempo e lo spazio come “proiezioni ortogonali” di un'unica entità, lo spazio-tempo, nell'ambito delle trasformazioni di Lorentz.

2.5 Trasformazioni di Lorentz

In questa sezione ci proponiamo di ricavare le trasformazioni di coordinate tra due SDR inerziali in moto relativo. Questo ci permetterà di inquadrare in un contesto unitario la relatività del tempo e dello spazio che abbiamo enunciato nelle sezioni precedenti.

È bene osservare ancora una volta che l'invarianza della velocità della luce non ci permette di trasferire le proprietà metriche (intervalli di tempo, lunghezze, angoli) da un SDR ad un altro, dal momento che il postulato sullo spazio ed il tempo assoluti viene a cadere. Quindi, nel definire le leggi di trasformazione che collegano coordinate cartesiane in SDR inerziali ci dovremo basare esclusivamente sui postulati della relatività R1-3, ed in particolare sul moto dei corpi liberi.

2.5.1 Deduzione delle trasformazioni di Lorentz

Consideriamo dunque due SDR inerziali S ed S' . Questo significa che ogni corpo in moto rettilineo ed uniforme in S è in moto rettilineo ed uniforme anche in S' , e viceversa. La più generale trasformazione tra coordinate cartesiane di SDR inerziali compatibile con l'omogeneità dello spazio ed il tempo è una trasformazione affine

$$\begin{aligned} t' &= a_{00}t + a_{01}x + a_{02}y + a_{03}z + b_0 \\ x' &= a_{10}t + a_{11}x + a_{12}y + a_{13}z + b_1 \\ y' &= a_{20}t + a_{21}x + a_{22}y + a_{23}z + b_2 \\ z' &= a_{30}t + a_{31}x + a_{32}y + a_{33}z + b_3 \end{aligned} \tag{2.19}$$

in cui i coefficienti a_{ij} e b_j sono costanti (indipendenti da t, x, y, z) che dipendono dalle caratteristiche dei SDR S ed S' (scelta delle origini, orientazione degli assi), e dalla loro velocità relativa \mathbf{v} .⁸

Supponiamo quindi che il SDR inerziale S' si muova con velocità uniforme \mathbf{v} rispetto al SDR inerziale S . Per evidenziare le conseguenze fisiche dovute al moto relativo, scegliamo le origini delle coordinate e le orientazioni degli assi in modo da semplificare il più possibile la trasformazione (2.19).

Scelta delle origini delle coordinate Sia A l'evento che coincide con l'origine degli assi O di S a $t = 0$: $(t_A, \mathbf{x}_A) = (0, \mathbf{0})$. Scegliamo l'origine delle coordinate e dei tempi di S' su tale evento, in modo che $(t'_A, \mathbf{x}'_A) = (0, \mathbf{0})$. Con questa scelta, deve essere

$$b_0 = b_1 = b_2 = b_3 = 0. \quad (2.21)$$

e la trasformazione (2.19) diventa lineare.

Scelta degli assi di S La traiettoria di O' vista da S è rettilinea e passante per O : la si identifichi con l'asse x (in direzione e verso); gli assi y e z , perpendicolari tra loro ed alla direzione x del moto, li orientiamo a piacere. Per consuetudine, facciamo in modo che la terna (x, y, z) sia destrorsa. In termini delle coordinate, la traiettoria di O' data da $x' = y' = z' = 0$ viene descritta da S così: $x = vt$, $y = z = 0$, con $v = v_x$. Quindi

$$x' = 0 \quad \text{ma anche, per} \quad x' = a_{10}t + a_{11}vt = (a_{10} + v a_{11})t \quad \implies \quad a_{10} + v a_{11} = 0 \quad (2.22a)$$

$$y' = 0 \quad (2.19) \text{ e } (2.21) \quad y' = a_{20}t + a_{21}vt = (a_{20} + v a_{21})t \quad \implies \quad a_{20} + v a_{21} = 0 \quad (2.22b)$$

$$z' = 0 \quad z' = a_{30}t + a_{31}vt = (a_{30} + v a_{31})t \quad \implies \quad a_{30} + v a_{31} = 0 \quad (2.22c)$$

Scelta degli assi di S' La traiettoria di O vista da S' è rettilinea e passante per O' : la si identifichi con l'asse x' (in direzione ma verso opposto). In valore assoluto, la velocità di S rispetto ad S' deve essere uguale a quella di S' rispetto ad S , per l'indistinguibilità dei SDR inerziali. Quindi alla traiettoria $x = y = z = 0$ corrisponde la traiettoria

$$t' = a_{00}t \quad \text{ma anche, per} \quad x' = a_{10}t = \frac{a_{10}}{a_{00}}t' \quad \implies \quad a_{10} = -v a_{00} \quad (2.23a)$$

$$y' = 0 \quad (2.19) \text{ e } (2.21) \quad y' = a_{20}t \quad \implies \quad a_{20} = 0 \quad (2.23b)$$

$$z' = 0 \quad z' = a_{30}t \quad \implies \quad a_{30} = 0 \quad (2.23c)$$

Mettendo assieme le eq. (2.22,2.23), ricaviamo

$$a_{00} = a_{11} \quad (2.24a)$$

$$a_{21} = 0 \quad (2.24b)$$

$$a_{31} = 0 \quad (2.24c)$$

Per orientare gli assi y', z' trasversi al moto, consideriamo in S un corpo h in moto rettilineo uniforme lungo l'asse y con velocità $u > 0$ e passante per A : $\mathbf{x}_h(t) = (0, ut, 0)$. Questo corpo compie in S' una traiettoria rettilinea che interseca (a $t' = 0$) l'asse x' ma che non coincide con quest'asse: $\mathbf{x}'_h(t) = (a_{10}t + a_{12}ut, a_{22}ut, a_{32}ut)$. Questa

⁸A rigore, se non si assume l'omogeneità dello spazio e del tempo, anche le trasformazioni proiettive del tipo

$$\begin{pmatrix} t' \\ \mathbf{x}' \end{pmatrix} = \frac{1}{rt + \mathbf{k} \cdot \mathbf{x} + q} \left[A \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} + \begin{pmatrix} t_0 \\ \mathbf{x}_0 \end{pmatrix} \right] \quad (2.20)$$

con A matrice non singolare 4×4 , $r, q \in \mathbb{R}$, $\mathbf{k} \in \mathbb{R}^3$ trasformano moti rettilinei uniformi in moti rettilinei uniformi. Tuttavia queste trasformazioni sono singolari nel punto (t, \mathbf{x}) in cui il denominatore dell'eq. (2.20) si annulla — a meno che $r = 0$ e $\mathbf{k} = 0$, nel qual caso la (2.20) si riduce alla (2.19). Questo evento rappresenterebbe un luogo ed un istante privilegiati, in contraddizione con il postulato R1. Equivalentemente, richiedendo che la trasformazione di coordinate non presenti singolarità per poter essere ovunque invertibile, si ottengono le trasformazioni affini (2.19) come uniche candidate a legare SDR inerziali.

retta di S' , assieme all'asse x' , individua un piano. Scegliamo l'asse y' in modo che esso appartenga a questo piano, sia ortogonale ad x' e diretto nel verso del moto di h . In questo modo annulliamo il coefficiente

$$a_{32} = 0 . \quad (2.25)$$

Allo stesso modo, osservando un corpo k che in S si muova lungo l'asse z , $\mathbf{x}_k(t) = (0, 0, ut)$, orientiamo in S' l'asse z' in modo che la traiettoria di k sia contenuta nel piano (x', z') . Questa scelta ci permette di annullare il coefficiente

$$a_{23} = 0 . \quad (2.26)$$

Dall'isotropia dello spazio deduciamo che i versori \hat{y}' e \hat{z}' devono essere ortogonali. Se per assurdo l'angolo tra \hat{y}' e \hat{z}' fosse minore di 90° , allora l'angolo tra \hat{y}' e $-\hat{z}'$ sarebbe maggiore di 90° . Ne deduciamo che:

o1) partendo dalla coppia ortogonale (\hat{y}, \hat{z}) in S la precedente procedura di orientazione determina una coppia (\hat{y}', \hat{z}') ad angolo acuto in S' ;

o2) partendo dalla coppia ortogonale $(\hat{p}, \hat{q}) \equiv (\hat{z}, -\hat{y})$ in S , la stessa procedura di orientazione determina una coppia $(\hat{p}', \hat{q}') \equiv (\hat{z}', -\hat{y}')$ ad angolo ottuso in S' ;

Ma in S le due coppie (\hat{y}, \hat{z}) e (\hat{p}, \hat{q}) si ottengono tramite una rotazione di 90° attorno all'asse x .⁹ Quindi, per l'isotropia dello spazio, partendo da coppie di versori allo stesso angolo α in S , dovremmo ottenere coppie di versori allo stesso angolo $\beta(\alpha)$ in S' , e quindi non è possibile ottenere in un caso un angolo acuto ed in un altro un angolo ottuso: $\alpha = 90^\circ \iff \beta = 90^\circ$. Dovrebbe essere intuitivo che, avendo scelto la terna (x, y, z) destrorsa, anche la terna (x', y', z') è destrorsa.

La situazione fino a questo punto è la seguente:

$$\begin{aligned} t' &= a_{00}t + a_{01}x + a_{02}y + a_{03}z \\ x' &= -va_{00}t + a_{00}x + a_{12}y + a_{13}z \\ y' &= a_{22}y \\ z' &= a_{33}z \end{aligned} \quad (2.27)$$

Avendo fissato origini ed orientazioni relative tra i due SDR, l'unico parametro da cui possono dipendere i coefficienti a_{ij} rimasti è la velocità relativa $v = v_x$, e non dalla scelta dell'evento A di "sincronizzazione" (per l'omogeneità dello spazio-tempo) né dalla scelta della direzione di \hat{y} (per l'isotropia):

$$a_{ij} = a_{ij}(v) . \quad (2.28)$$

In particolare, le coordinate di S si potranno esprimere in funzione di quelle di S' con le stesse formule (2.27) scambiando $t \leftrightarrow t'$, \dots , $z \leftrightarrow z'$, $a_{ij}(v) \rightarrow a_{ij}(-v)$.

Passiamo ora a dimostrare che anche i coefficienti di y e z in t' ed x' si annullano. Infatti, scegliendo \hat{y} e \hat{z} in modo diverso, ruotati di 180° rispetto a prima — e di conseguenza anche \hat{y}' e \hat{z}' vengono ruotati di 180° rispetto a prima, dovendo preservare la procedura di orientazione descritta sopra — gli a_{ij} restano immutati, come anche le coordinate t, x, t', x' , mentre i valori delle coordinate y, z, y', z' cambiano di segno. Si avrebbe pertanto

$$\begin{aligned} t' &= a_{00}t + a_{01}x - a_{02}y - a_{03}z \\ x' &= -va_{00}t + a_{00}x - a_{12}y - a_{13}z \\ -y' &= -a_{22}y \\ -z' &= -a_{33}z . \end{aligned} \quad (2.29)$$

I sistemi (2.27) e (2.29) sono compatibili per ogni y, z solamente se

$$a_{02} = a_{03} = 0 \quad (2.30)$$

$$a_{12} = a_{13} = 0 . \quad (2.31)$$

⁹Per di più, avevamo scelto in modo completamente arbitrario l'orientazione iniziale di \hat{y} e \hat{z} .

Abbiamo visto nella sez. (2.2) che le distanze ortogonali al moto sono invarianti per cambi di SDR, per cui $a_{22} = a_{33} = 1$. Possiamo dimostrarlo in maniera più rigorosa nel seguente modo. Se $y' = a_{22}(v)y$, vale

$$y = \frac{y'}{a_{22}(v)} \quad (2.32)$$

ma anche

$$y = a_{22}(-v)y' \quad (2.33)$$

dato che S si muove con velocità $-v$ rispetto ad S' , quindi

$$a_{22}(-v) = \frac{1}{a_{22}(v)} \implies a_{22}(v)a_{22}(-v) = 1. \quad (2.34)$$

Mostriamo ora che $a_{22}(v)$ non dipende dal segno di v . A questo scopo invertiamo il verso dell'asse x (e anche dell'asse z , in modo da poter pensare la trasformazione come una rotazione di 180° attorno all'asse y che mantiene destrorsa la terna degli assi). Invertiamo anche gli assi x' e z' , in modo da mantenere valida la procedura di orientazione descritta sopra. La velocità di S' rispetto ad S cambia di segno, mentre le coordinate y, y' restano invariate. Quindi

$$y' = a_{22}(-v)y \implies a_{22}(v) = a_{22}(-v), \quad (2.35)$$

e dall'eq. (2.34)

$$[a_{22}(v)]^2 = 1 \implies a_{22}(v) = +1 \quad \forall v, \quad (2.36)$$

ove il segno $+$ deriva dall'aver scelto y ed y' concordi.¹⁰ Lo stesso risultato si ottiene considerando la coordinata trasversa z al posto di y :

$$a_{33} = 1. \quad (2.37)$$

I ragionamenti svolti finora, basati puramente sul postulato di isotropia ed omogeneità dello spazio e del tempo (R1), e sull'equivalenza dei SDR inerziali (R3), ci hanno permesso di scomporre la matrice a_{ij} di cambio di coordinate in due blocchi: quello delle coordinate "longitudinali" (t, x) e quello delle coordinate trasverse (y, z) :

$$A \equiv ||a_{ij}||_{i,j=0,1,2,3} = \left(\begin{array}{cc|cc} \gamma & \alpha & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right), \quad \text{ove} \quad \begin{cases} \gamma \equiv a_{00} \\ \alpha \equiv a_{01} \end{cases}. \quad (2.38)$$

In particolare, le coordinate trasverse sono lasciate inalterate essendo collegate dall'identità. Inoltre, per ogni t fissato, i 3 assi spaziali di S sono paralleli a quelli di S' (vedi fig. 2.3), essendo le rispettive coordinate legate da una matrice diagonale

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} \gamma vt \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} vt \\ 0 \\ 0 \end{pmatrix}. \quad (2.39)$$

Osservazione: Anche le trasformazioni di Galileo hanno questa struttura. Infatti, fino a questo punto abbiamo sfruttato dei postulati che sono validi anche nella fisica classica: (M1)=(R1) ed (M3)⊂(R3).

Per collegare le coordinate (t, x) con le (t', x') , ci serve il postulato dell'invarianza della velocità della luce. Consideriamo un'onda luminosa emessa a $t = t' = 0$ dall'origine comune di S e S' . In virtù della costanza della velocità della luce in ogni sistema inerziale, le equazioni dei fronti d'onda nei due SDR saranno

$$x^2 + y^2 + z^2 = (ct)^2, \quad x'^2 + y'^2 + z'^2 = (ct')^2 \quad (2.40)$$

¹⁰In pratica, l'inversione degli assi x, z, x', z' che porta all'eq. (2.35) corrisponde semplicemente al fatto di eseguire la procedura di orientazione degli assi partendo da S' invece che da S , la cui unica differenza è di avere asse x (x') concorde (discorde) alla velocità di O (O'). Per il principio di indistinguibilità dei SDR inerziali, vale il sistema (2.29) con gli stessi coefficienti, da cui segue l'eq. (2.35).

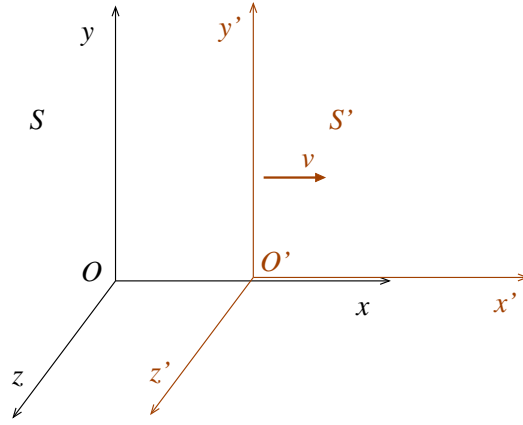


Figure 2.3: Assi spaziali dei riferimenti inerziali S ed S' in moto relativo, come si presentano ad S all'istante t .

Sottraendo membro a membro queste equazioni e ricordando che $y' = y$, $z' = z$ troviamo

$$c^2 t^2 - x^2 = c^2 t'^2 - x'^2 \quad (2.41)$$

$$\stackrel{(2.38)}{=} c^2 (\gamma^2 t^2 + \alpha^2 x^2 + 2\alpha\gamma tx) - \gamma^2 (v^2 t^2 + x^2 - 2vtx) .$$

Potendo x e t assumere valori arbitrari (nella regione $c^2 t^2 \geq x^2$), nell'uguaglianza tra il primo e terzo membro devono essere separatamente uguali i coefficienti di t^2 , x^2 e tx :

$$c^2 = \gamma^2 (c^2 - v^2) \quad \Rightarrow \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.42)$$

$$-1 = c^2 \alpha^2 - \gamma^2 \quad (2.43)$$

$$0 = 2\gamma(c^2 \alpha + v\gamma) \quad \Rightarrow \quad \alpha = -\frac{v}{c^2} \gamma \quad (2.44)$$

(γ dovrà essere positivo, poiché gli assi di x e x' sono concordi, così come gli assi temporali).

Riassumendo, tra i due SDR si ha la **trasformazione di Lorentz**

$$t' = \gamma \left(t - \frac{v}{c^2} x \right) \quad (2.45a)$$

$$x' = \gamma (x - vt) \quad (2.45b)$$

$$y' = y \quad (2.45c)$$

$$z' = z \quad (2.45d)$$

La trasformazione inversa si può ricavare facilmente,

$$t = \gamma \left(t' + \frac{v}{c^2} x' \right) \quad (2.46a)$$

$$x = \gamma (x' + vt') \quad (2.46b)$$

$$y = y \quad (2.46c)$$

$$z = z \quad (2.46d)$$

e come si vede corrisponde al sistema (2.45) scambiando v in $-v$. Infatti, per l'osservatore solidale ad S' , il sistema S si muove in direzione x' con velocità $-v$.

2.5.2 Limite non-relativistico

È importante verificare che, nel limite di piccole velocità, la trasformazione di coordinate relativistica riproduca la legge di trasformazione classica. Infatti, per piccole velocità $v \ll c$, il fattore $\gamma \rightarrow 1$ e la (2.45) si riduce a

$$t' = t \quad (2.47a)$$

$$x' = x - vt \quad (2.47b)$$

$$y' = y \quad (2.47c)$$

$$z' = z, \quad (2.47d)$$

che è la trasformazione di Galileo della meccanica classica.

In generale, le equazioni della meccanica classica si dovranno ottenere dalle corrispondenti equazioni relativistiche ogniqualvolta le velocità di tutti i corpi in gioco siano molto minori della velocità della luce. Equivalentemente, la dinamica Newtoniana si dovrà ottenere facendo il limite $c \rightarrow +\infty$ a fissato valore delle velocità dei corpi (limite non-relativistico).

2.5.3 Trasformazioni di Lorentz proprie non-omogenee

La trasformazione di coordinate (2.45) correla SDR molto particolari, che hanno origini coincidenti $O = O'$, assi paralleli e concordi tra loro, con velocità relativa diretta lungo uno di questi assi. Per questo è detta *trasformazione di Lorentz speciale*. Ciò nonostante, tutte le conseguenze fisiche sullo spazio e sul tempo, dovute ai postulati della relatività, sono contenute in quelle formule. Infatti, per ottenere le eq. (2.45) a partire dalla più generale trasformazione tra SDR inerziali (2.19), abbiamo solamente vincolato la scelta dell'origine di O' e l'orientazione degli assi spaziali di S ed S' . Questa relazione reciproca tra SDR inerziali può essere ottenuta a partire da due qualsiasi SDR inerziali Q ed Q' con velocità relativa \mathbf{v} , operando una traslazione nel tempo e nello spazio (omogenei) di Q ed Q' in modo da sovrapporre le origini con $O = O'$, e ruotando nello spazio (isotropo) i SDR così ottenuti in modo da sovrapporre gli assi di Q a quelli di S e quelli di Q' a quelli di S' (non sono necessarie inversioni degli assi avendo supposto in tutti i casi terne destrorse).

In altre parole, la più generale trasformazione di coordinate tra SDR inerziali orientati si ottiene combinando la trasformazione di Lorentz speciale (2.45) con traslazioni e rotazioni statiche (indipendenti dal tempo). Queste trasformazioni vengono chiamate *trasformazioni di Lorentz proprie*, omogenee se $O = O'$, non-omogenee nel caso di origini non coincidenti.¹¹

Lo stesso succede in meccanica classica: la più generale trasformazione di Galileo (1.32) tra SDR inerziali orientati ed in moto relativo a velocità \mathbf{v} può essere ottenuta dall'eq. (2.47) combinata con traslazioni e rotazioni statiche.

Le trasformazioni di Lorentz proprie possono essere interpretate anche dal punto di vista *attivo*. Parlando alla buona, se le x^μ sono le coordinate in S di un sistema fisico h , le nuove coordinate x'^μ (sempre in S) dello stesso sistema fisico ma in moto lungo l'asse x con velocità $-v$ sono date dalla trasformazione di Lorentz (2.45). Più precisamente, se in S ho un apparato h composto da N elementi che si muovono lungo traiettorie $\mathbf{x}_i(t) : i = 1, \dots, N$, e se a questo apparato do una spinta $h \rightarrow \tilde{h}$ in direzione x con velocità $-v$ in modo che l'osservatore \tilde{S} veda \tilde{h} esattamente come S vedeva h (ossia $\tilde{\mathbf{x}}_i(\tilde{t}) = \mathbf{x}_i(\tilde{t})$), le nuove traiettorie degli elementi di \tilde{h} viste da S sono date da $\mathbf{x}'_i(t')$ tramite le eq. (2.45). Ribadiamo che, nel punto di vista attivo, (t', \mathbf{x}') sono misure di tempo e di spazio fatte da S sul sistema \tilde{h} .

2.5.4 Unicità della velocità invariante

Dal procedimento seguito risulta chiaro che, se si ammette l'esistenza di un tipo di segnale con velocità costante \bar{c} in ogni riferimento, questo può essere usato al posto della luce, dando luogo ad una trasformazione di coordinate con la stessa forma della (2.45), ma con velocità \bar{c} al posto di c . Tuttavia, poiché solo una trasformazione può

¹¹Le trasformazioni di Lorentz non omogenee sono chiamate anche trasformazioni di Poincaré.

essere valida, cioè o la (2.45) o quest'ultima, ne segue che questo segnale dovrà propagarsi alla velocità c e ciò sarà vero per ogni tipo di segnale con queste caratteristiche (per esempio le onde gravitazionali).

2.5.5 La velocità della luce come velocità limite

Osserviamo che per $v \rightarrow c$ la trasformazione è singolare, nel senso che il fattore γ diventa infinito. Ciò significa che un sistema di riferimento non si potrà muovere rispetto ad un altro sistema con velocità uguale o superiore a quella della luce. Ora, un sistema di riferimento si può pensare costituito da corpi materiali, per cui ne segue che una particella materiale non può muoversi con velocità uguale o maggiore di quella della luce, rispetto ad un qualsiasi sistema di riferimento. Confermeremo questa affermazione invocando il principio di causalità (sez. 2.7) e studiando la dinamica relativistica (sez. 5.3).

2.5.6 Contrazione delle lunghezze

Dalla legge di trasformazione di Lorentz (2.45), possiamo ricavare alcune conseguenze che riguardano il confronto di misure effettuate nei due sistemi di riferimento S ed S' in moto relativo. In particolare, vogliamo dedurre nuovamente la contrazione delle lunghezze e la dilatazione dei tempi per oggetti in moto.

Consideriamo un regolo a riposo in S' disposto parallelamente all'asse delle x , individuato al solito dal vettore velocità relativa \mathbf{v} . Gli estremi del regolo in S' abbiano le coordinate x'_1 e x'_2 rispettivamente. La lunghezza del regolo misurata in S' è perciò data da

$$l_0 = x'_2 - x'_1 \quad (2.48)$$

Chiamiamo questa quantità *lunghezza a riposo* del regolo.

Ponendoci in S , è naturale definire come lunghezza del regolo la quantità $l = x_2 - x_1$ dove le misure dei due estremi x_1 ed x_2 sono effettuate allo stesso istante $t_1 = t_2 = t$. Usando allora le eq. (2.45) si ha

$$x'_1 = \gamma(x_1 - vt_1), \quad x'_2 = \gamma(x_2 - vt_2) \quad (2.49)$$

da cui, sottraendo membro a membro, ricaviamo

$$l_0 = x'_2 - x'_1 = \gamma(x_2 - x_1) = \gamma l \quad (2.50)$$

e quindi

$$l = \frac{l_0}{\gamma} = l_0 \sqrt{1 - \frac{v^2}{c^2}}. \quad (2.51)$$

Questo risultato, già ricavato nell'eq. (2.18), esprime il fatto che un oggetto in moto è accorciato nella direzione del moto stesso.

Se invece il regolo viene orientato perpendicolarmente alla velocità \mathbf{v} , la sua lunghezza rimane invariata, in quanto le coordinate trasverse y, z non sono modificate nella trasformazione.

Quindi, se si considera un corpo esteso di volume V_0 misurato a riposo in S' , il volume V misurato in S è diminuito di un fattore γ :

$$V = \frac{V_0}{\gamma} = V_0 \sqrt{1 - \frac{v^2}{c^2}}, \quad (2.52)$$

a causa della contrazione del corpo che avviene in direzione del moto.

2.5.7 Dilatazione dei tempi

Consideriamo ora un orologio a riposo in S' posto in un punto di ascissa x' , che misura l'intervallo di tempo $\Delta t_0 = t'_2 - t'_1$ tra due eventi che avvengono agli istanti t'_1 e t'_2 nello stesso punto x' . Usando le trasformazioni di Lorentz inverse (2.46) si ottiene

$$t_1 = \gamma(t'_1 + vx'/c^2), \quad t_2 = \gamma(t'_2 + vx'/c^2) \quad (2.53)$$

e quindi, sottraendo membro a membro,

$$\Delta t \equiv t_2 - t_1 = \gamma(t'_2 - t'_1) = \gamma \Delta t_0 = \frac{\Delta t_0}{\sqrt{1 - \frac{v^2}{c^2}}} . \quad (2.54)$$

Questo risultato, già ricavato nell'eq.(2.12), esprime il fatto che gli intervalli di tempo scanditi da un oggetto in movimento sono più lunghi degli stessi intervalli misurati nel SDR in cui l'oggetto è in quiete.

2.6 Il tempo proprio

La dipendenza dall'osservatore degli intervalli di tempo, ossia la mancanza di un tempo assoluto, fa sorgere il problema di definire un tempo “naturale” o “intrinseco” per i vari processi che possono aver luogo. Inoltre, finora ci siamo limitati a misure di tempo in osservatori inerziali. In questa sezione ci occuperemo dello scorrere del tempo per osservatori in stato di moto arbitrario (ad esempio di come invecchi un essere vivente in moto accelerato).

Sappiamo che certi processi hanno una durata di vita all'incirca definita: pensiamo agli esseri viventi (qualche ora per i batteri, qualche decennio per le persone), oppure alle particelle instabili (frazioni di secondo). Ci sono poi fenomeni con durata molto precisa, come le oscillazione delle onde elettromagnetiche emesse da transizioni atomiche. Se dunque la durata tipica di un certo processo vale τ , si pone il problema di definire il SDR rispetto al quale tale tempo va misurato. Infatti, siccome il fattore γ può essere arbitrariamente grande per $v \rightarrow c$, la dipendenza della durata di un processo può variare anche di molti ordini di grandezza tra un SDR e l'altro. Se ad esempio un futuro astronauta di 30 anni, con un'aspettativa di vita di ancora 50 anni, compirà una missione spaziale su un'astronave che rispetto alla Terra viaggia a velocità molto prossima a quella della luce ($\gamma \simeq 100$) su di un'orbita circolare, rispetto a chi l'astronauta vivrà i suoi rimanenti 50 anni? Rispetto agli osservatori sulla Terra? Rispetto agli osservatori (non inerziali!) sull'astronave? Rispetto a qualcun'altro? Se sulla Terra il viaggio si conclude dopo 500 anni dalla partenza, c'è speranza che dall'astronave l'astronauta esca vivo?

Dal postulato della relatività non c'è un SDR privilegiato che abbia l'orologio biologico giusto per ogni situazione. Un osservatore *inerziale* vedrà i processi (sia fisici che biologici) del suo laboratorio evolversi con gli stessi tempi che un altro osservatore vedrà per i processi del proprio laboratorio. È quindi naturale che ogni osservatore inerziale adotti il tempo dei propri orologi (in quiete rispetto a lui) per misurare la durata della propria vita o di altri fenomeni del suo laboratorio. Siamo quindi portati ad introdurre il concetto di **tempo-proprio** τ di un corpo come il tempo misurato da un osservatore in quiete rispetto al corpo stesso. In un generico SDR S con coordinata temporale t , se il corpo h in esame si muove con velocità costante \mathbf{u} , l'intervallo di tempo-proprio tra due eventi relativi ad h è dato da

$$\Delta\tau \equiv \Delta t_0 \stackrel{(2.54)}{=} \sqrt{1 - \frac{\mathbf{u}^2}{c^2}} \Delta t . \quad (2.55)$$

Per corpi in moto accelerato, non esiste un SDR inerziale rispetto al quale il corpo sia sempre in quiete. Quello che possiamo supporre (e verificare poi con gli esperimenti) è che, in un intervallo di tempo infinitesimo dt in S (nel quale la velocità di h è all'incirca costante), il tempo proprio di h corrisponda all'intervallo di tempo misurato da un osservatore inerziale istantaneamente in quiete rispetto ad h . In altre parole, si assume che l'accelerazione a cui il corpo è soggetto non alteri il ritmo del suo orologio. Questa ipotesi va sotto il nome di **principio di località**. Assumiamo quindi che l'intervallo di tempo-proprio infinitesimo $d\tau$ sia connesso con l'intervallo di tempo dt misurato da un generico osservatore S sia

$$d\tau = \sqrt{1 - \frac{\mathbf{u}^2}{c^2}} dt . \quad (2.56)$$

in cui \mathbf{u} è la velocità istantanea di h rispetto ad S .

Per un corpo che si muova lungo una linea-universo Γ arbitraria tra gli eventi A a B , descritta in S dalla curva $t \mapsto \mathbf{x}(t)$, l'intervallo di tempo-proprio è dato dall'integrale del tempo-proprio infinitesimo:

$$\Delta\tau = \int_{\Gamma} d\tau = \int_{t_A}^{t_B} \sqrt{1 - \frac{\mathbf{u}(t)^2}{c^2}} dt , \quad \mathbf{u}(t) \equiv \frac{d\mathbf{x}(t)}{dt} , \quad (2.57)$$

avendo indicato con t_A e t_B le coordinate temporali di A e B in S .

La prima osservazione, di carattere matematico, è che il tempo-proprio è un *invariante* relativistico, ossia non dipende dal sistema di riferimento inerziale S scelto per parametrizzare la curva Γ e per calcolare l'integrale (2.57). Infatti, ogni misura di tempo-proprio infinitesimo è effettuata in SDR inerziali istantaneamente in quiete rispetto al corpo h , quindi le misure di $d\tau$ sono univocamente definite. Non ha senso perciò parlare di proprietà di trasformazione per il tempo-proprio. Sarà la sua relazione con il tempo misurato da un orologio fisso in un SDR inerziale a cambiare, al cambiare di tale SDR. E questa relazione dipende, istante per istante, dalla velocità relativa tra orologio e SDR, e non dall'accelerazione (assoluta) dell'orologio.

La seconda osservazione, di carattere fisico, è che il tempo-proprio *definito* dalle eq. (2.57, 2.56) sperimentalmente corrisponde all'intervallo di tempo misurato da un orologio solidale ad h .

Esempio: La dilatazione dei tempi viene verificata ogni giorno nei laboratori di fisica delle alte energie, dove fasci di particelle instabili di vite medie conosciute τ_0 vengono trasportati, prima di decadere, su distanze molte volte superiori al limite galileiano della distanza di decadimento $c\tau_0$. Per esempio, presso il laboratorio dell'Acceleratore Nazionale Enrico Fermi, dei pioni carichi con velocità di $v = 0,999\,999\,75\,c$ vengono prodotti e trasportati su una distanza di 300 metri con una perdita inferiore al 3% per effetto dei decadimenti. Con una vita media $\tau_0 = 2,56 \cdot 10^{-8}$ sec, la distanza media di decadimento galileiana sarebbe $c\tau_0 = 7,7$ metri, e senza l'effetto di dilatazione del tempo solo una frazione $e^{300/7,7} \simeq 10^{-17}$ dei pioni prodotti potrebbe sopravvivere. Ma alla velocità v a cui i pioni si muovono, il fattore di dilatazione temporale vale $\gamma \simeq 1400$, la vita media osservata nel laboratorio è $\gamma\tau_0$ ed il cammino medio di decadimento nel laboratorio è in realtà $c\gamma\tau_0 \simeq 11$ Km!

Esercizio: Nel SDR dei pioni, in cui la vita media vale τ_0 , come si spiega che quasi tutti riescono a raggiungere la fine del tubo lungo 300 metri?

Paradosso dei gemelli Si può dimostrare che, fissati due eventi A e B sulla linea universo di un osservatore S , l'intervallo di tempo proprio più lungo $\Delta\tau_{\max}$ misurato tra tali eventi compete all'unico osservatore inerziale S_I che attraversa tali eventi. Tutti gli altri osservatori che attraversano A e B lungo linee universo differenti da quella di S_I non sono inerziali e misureranno intervalli di tempo proprio $\Delta\tau < \Delta\tau_{\max}$. Ad esempio, due gemelli che da A siano partiti per due viaggi diversi, quando si ritrovano in B possono essere invecchiati uno più dell'altro. In particolare, se uno dei due non ha subito accelerazioni, sarà sicuramente più vecchio del gemello che ha subito accelerazioni.

Al limite, un osservatore che si potesse muovere a cavallo di un raggio di luce opportunamente guidato, misurerebbe un intervallo di tempo proprio nullo.

2.7 Ordine temporale degli eventi: causalità

Nella sez. 2.1 abbiamo visto che gli intervalli di tempo tra due eventi non solo possono essere più o meno lunghi tra diversi SDR, ma che addirittura è possibile che l'ordine temporale degli stessi si inverta: un evento A che per l'osservatore S precede l'evento B , per l'osservatore S' può essere successivo a B . Questo sembrerebbe mettere in crisi il principio di causalità (la causa deve precedere l'effetto), su cui si basa tutta la fisica (e non solo la fisica).

Cerchiamo la via d'uscita da questo paradosso studiando sotto quali condizioni l'evento A , che in S precede l'evento B , possa essere osservato da S' come posteriore a B . Al solito, non è restrittivo, mediante traslazioni, far coincidere l'evento A con le origini O ed O' dei due SDR, ed orientare i loro assi in modo che la velocità relativa giaccia sull'asse x di S ed x' di S' . Infatti, abbiamo visto che solo il moto relativo può alterare le relazioni temporali, non le traslazioni e rotazioni rigide.

Per ipotesi quindi $t_B \equiv t > 0$, e ci chiediamo quando possa essere $t'_B \equiv t' < 0$. Indicando $x_B \equiv x$, dall'eq. (2.45a) abbiamo

$$t' = \gamma(t - \frac{v}{c^2}x) < 0 \quad \Longleftrightarrow \quad t < \frac{v}{c^2}x < \frac{x}{c} \leq \frac{|x|}{c}, \quad (2.58)$$

dovendo essere $v < c$ e $x \leq |\mathbf{x}|$. In pratica, l'ordine temporale degli eventi A e B può essere invertito solamente se l'intervallo temporale t tra i due eventi è minore del tempo $|\mathbf{x}|/c$ che impiega la luce a percorrere la distanza tra tali eventi. Equivalentemente, l'ordine temporale degli eventi A e B può essere invertito solamente se la distanza spaziale $|\mathbf{x}|$ tra i due eventi è maggiore della distanza ct che la luce percorre nel tempo t .

Questa condizione non è solo necessaria ma anche sufficiente: supponiamo che $0 < t < |\mathbf{x}|/c$. Orientiamo S in modo che l'evento B avvenga sull'asse x positivo, così che $x = |\mathbf{x}|$, e consideriamo un SDR S' in moto lungo x con velocità v . Dall'eq. (2.58) vediamo che t' è negativo purché si abbia

$$t < \frac{v}{c^2} |\mathbf{x}| \quad \Longleftrightarrow \quad v > c \frac{ct}{|\mathbf{x}|} \equiv \bar{v}, \quad (2.59)$$

Siccome $ct/|\mathbf{x}| < 1$ per ipotesi, la velocità di soglia \bar{v} (oltre la quale t' diventa negativo) è inferiore a c , quindi accessibile ad S' .

In conclusione, l'ordine temporale di una data coppia di eventi dipende dall'osservatore solamente se $c|\Delta t| < |\Delta \mathbf{x}|$ in almeno un SDR (e, di conseguenza, in tutti i SDR, come dimostreremo nella sez. 3.1).

Possiamo capire graficamente come varia la variabile $t' = t'_B$ al cambiare della velocità v di S' riprendendo in esame l'eq. (2.41), che ci dice in sostanza che

$$c^2 t'^2 - x'^2 = \text{costante} = c^2 t^2 - x^2. \quad (2.60)$$

Questa è l'equazione di una iperbole equilatera nel piano (ct', x') , con asintoti le bisettrici $ct' = \pm x'$. A seconda che la costante sia positiva ($|ct| > |\mathbf{x}|$) o negativa ($|ct| < |\mathbf{x}|$), le iperboli intersecano o meno l'asse x' che corrisponde a $t' = 0$. Se per esempio le coordinate dell'evento B in S coincide con il punto B_1 in fig. 2.4, al variare di v le coordinate ct' ed x' percorrono il ramo di iperbole punteggiato e non si ha inversione della coordinata temporale. Se invece l'evento B coincide in S con il punto B_2 , allora all'aumentare della velocità v le coordinate (ct', x') percorrono il ramo di iperbole tratteggiato, fino ad oltrepassare l'asse delle x' nella regione di t' negativi.

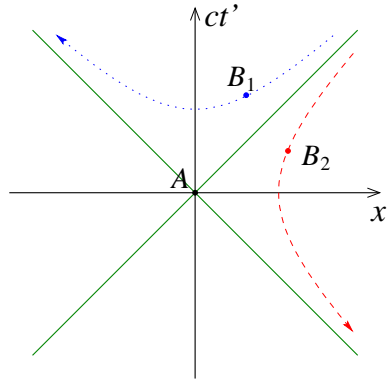


Figure 2.4: *Rappresentazione grafica delle coordinate (ct', x') degli eventi A e B descritti nel testo, al variare della velocità v . Il punto B_1 rappresenta le coordinate di B in S nel caso in cui $ct > |\mathbf{x}|$, la linea blu punteggiata corrisponde al luogo delle coordinate (ct', x') che B_1 può assumere (la freccia indica il verso di velocità crescenti). Il punto B_2 rappresenta le coordinate di B in S nel caso in cui $ct < |\mathbf{x}|$, per il quale si può avere inversione dell'ordine temporale, come indicato dalla linea rossa tratteggiata. In verde sono rappresentate le bisettrici $ct' = \pm x'$ che separano le regioni di ordinamento temporale assoluto da quelle di ordinamento relativo.*

Se assumiamo la validità del principio di causalità, cioè che la causa deve precedere l'effetto, segue che nessun agente fisico (o segnale) che trasporti informazioni può propagarsi a velocità V superiori a quelle della luce $V > c$, poiché, se così fosse, tra i due eventi A (causa) e B (effetto) connessi dal segnale a velocità V si avrebbe

$$t_B - t_A \equiv \Delta t = \frac{|\Delta \mathbf{x}|}{V} < \frac{|\Delta \mathbf{x}|}{c}. \quad (2.61)$$

Ma allora esisterebbero SDR in cui l'ordine temporale degli eventi sarebbe invertito: $t'_B < t'_A$, e per questi la causa precederebbe l'effetto!

2.8 Composizione delle velocità

Dalle trasformazioni di Lorentz sulle coordinate possiamo ricavare la legge di trasformazione delle velocità tra SDR inerziali. Questa, nel caso particolare della trasformazione speciale (2.45), si può ottenere per differenziazione:

$$dt' = \gamma(dt - \frac{v}{c^2}dx) \quad (2.62a)$$

$$dx' = \gamma(dx - vdt) \quad (2.62b)$$

$$dy' = dy \quad (2.62c)$$

$$dz' = dz \quad (2.62d)$$

Se ora un oggetto attraversa l'evento A con velocità $\mathbf{u} \equiv d\mathbf{x}/dt$ nel SDR S , la velocità misurata nel SDR S' dello stesso oggetto mentre attraversa lo stesso evento è data da $\mathbf{u}' \equiv d\mathbf{x}'/dt'$ con

$$u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} \quad (2.63a)$$

$$u'_y = \frac{u_y}{\gamma \left(1 - \frac{u_x v}{c^2}\right)}, \quad \gamma = 1 / \sqrt{1 - \frac{v^2}{c^2}} \quad (2.63b)$$

$$u'_z = \frac{u_z}{\gamma \left(1 - \frac{u_x v}{c^2}\right)}, \quad (2.63c)$$

che si riducono a quelle Galileiane (1.38) per $v/c \rightarrow 0$:

$$u'_x = u_x - v \quad (2.64a)$$

$$u'_y = u_y \quad (2.64b)$$

$$u'_z = u_z \quad (2.64c)$$

Le trasformazioni di velocità inverse si ottengono dal sistema (2.63) cambiando il segno di v :

$$u_x = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}} \quad (2.65a)$$

$$u_y = \frac{u'_y}{\gamma \left(1 + \frac{u'_x v}{c^2}\right)} \quad (2.65b)$$

$$u_z = \frac{u'_z}{\gamma \left(1 + \frac{u'_x v}{c^2}\right)}, \quad (2.65c)$$

È evidente dalle eq. (2.65) che in relatività la legge di composizione delle velocità non è più lineare: se un corpo si muove in S' a velocità \mathbf{u}' ed S' si muove rispetto ad S a velocità \mathbf{v} , non è più vero che il corpo si muove rispetto a S alla velocità $\mathbf{u}' + \mathbf{v}$.

Questo fatto, a prima vista paradossale, segue semplicemente dal fatto che osservatori diversi misurano distanze spaziali e temporali diverse per le stesse coppie di eventi. La non-linearità della legge di composizione delle velocità è necessaria per garantire che la velocità della luce non venga superata da nessun corpo in qualsiasi SDR.

Per esempio, se $\mathbf{u}' = (c, 0, 0)$ è la velocità di un raggio di luce in S' , dall'eq. (2.65a) si vede che il raggio viaggia con velocità $\mathbf{u} = (c, 0, 0)$ anche in S . Sempre in riferimento all'eq. (2.65a), non è difficile convincersi che, fissata $v \in]-c, c[$, le componenti u_k come funzione di $u'_x \in [-c, c]$ sono monotone e comprese nell'intervallo $[-c, c]$.

Notiamo infine che le componenti u_k sono funzioni simmetriche delle variabili u'_x e v , come nel caso non-relativistico.

Esercizio: Consideriamo tre SDR S , S' ed S'' con assi paralleli tra loro. Pensiamo S in quiete, S' in moto con velocità v_1 lungo x ed S'' in moto con velocità v_2 lungo y . Se una particella p_1 si muove rispetto ad S' con

velocità v_2 lungo y' , mentre un'altra particella p_2 si muove rispetto ad S'' con velocità v_1 lungo x'' , stabilire se le due particelle si muovono con la stessa velocità rispetto ad S . In altre parole, c'è da stabilire se la composizione di due velocità $v_1\hat{x}$ e $v_2\hat{y}$ dipende o meno dall'ordine con cui si esegue la composizione.

Esercizio: Determinare la velocità relativa di due corpi che, rispetto ad un SDR S , si muovono con velocità \mathbf{v}_1 e \mathbf{v}_2 (minori di c). Verificare che il risultato è simmetrico nello scambio $\mathbf{v}_1 \leftrightarrow \mathbf{v}_2$ e controllare la formula ottenuta nei casi particolari di: velocità parallele; $\mathbf{v}_2 = 0$; $v_2 = c$.

Suggerimento: orientare i SDR solidali alle particelle in modo opportuno ed usare come parametri i moduli delle velocità v_1 e v_2 e l'angolo θ tra le stesse.

Chapter 3

La struttura dello spazio-tempo

3.1 Invarianza dell'intervallo

Nel capitolo precedente abbiamo ricavato le trasformazioni di coordinate tra due SDR inerziali in moto relativo, avendo scelto con cura l'origine delle coordinate e l'orientazione degli assi in modo da rendere il più semplice possibile le leggi di trasformazione. Questa scelta non pregiudica la generalità delle conseguenze fisiche che si possono dedurre. Infatti, è evidente che per un osservatore inerziale, scegliere un particolare evento come origine delle coordinate piuttosto che un altro equivale ad operare una semplice traslazione statica (nello spazio e nel tempo). Similmente, la scelta di un'orientazione degli assi rispetto ad un'altra scelta corrisponde ad una semplice rotazione statica delle coordinate spaziali. Dal postulato R1, non c'è ragione alcuna per preferire o sfavorire una particolare origine (omogeneità) o una particolare orientazione (isotropia). Inoltre, sempre per R1, le traslazioni e le rotazioni statiche non alterano la geometria dello spazio e del tempo. Pertanto, la più generale trasformazione di coordinate tra SDR inerziali si ottiene combinando la trasformazione (2.45) con usuali traslazioni e rotazioni statiche in S o S' .

Si può fare lo stesso ragionamento in cinematica classica: partendo dalla più semplice trasformazione di Galileo (2.47), applicando traslazioni e rotazioni statiche possiamo ottenere qualsiasi trasformazione del gruppo di Galileo (1.32).¹ In meccanica Newtoniana però avevamo un certo numero di grandezze fisiche che venivano preservate dalle trasformazioni di Galileo: gli intervalli temporali, le distanze spaziali, e gli angoli tra vettori. Questo fatto era espresso dal postulato M2.

Nel caso relativistico, tutte queste quantità sono relative, dipendendo dal SDR in cui si misurano. Ci chiediamo se esistano delle quantità dello spazio-tempo che siano preservate da tutte le trasformazioni relativistiche. Prendiamo spunto dal postulato R2 che per l'appunto sostituisce M2: l'invarianza della velocità della luce. Se un segnale luminoso collega due eventi A e B , per un qualsiasi osservatore deve valere $c|t_B - t_A| = |\mathbf{x}_B - \mathbf{x}_A|$ e quindi

$$c^2(t_B - t_A)^2 = |\mathbf{x}_B - \mathbf{x}_A|^2. \quad (3.1)$$

Naturalmente, il valore dei due membri dipende dall'osservatore. Invece c è la stessa per tutti. Quindi, per coppie di eventi collegati da un raggio di luce, possiamo scrivere l'equazione

$$(c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 = 0 \quad \text{ove} \quad \Delta t \equiv t_B - t_A \quad \text{ecc.} \quad (3.2)$$

valida in ogni SDR inerziale. Per eventi generici (non collegati da un segnale luminoso), la quantità

$$s_{AB}^2 \equiv (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \quad (3.3)$$

non è nulla. Il fatto notevole è che s_{AB}^2 è una grandezza che non dipende dal SDR in cui si calcola. Infatti, sotto traslazioni tutte le differenze di coordinate $\Delta\#$ non variano; sotto rotazioni il Δt non varia e si conserva anche

¹Includendo le riflessioni degli assi, possiamo collegare, tanto nel caso classico che relativistico, anche SDR con terne di assi sia destrorse che sinistrorse.

la somma $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$; infine, sotto la trasformazione speciale di Lorentz (2.45) abbiamo che Δy e Δz non variano, né varia la combinazione $(c\Delta t)^2 - (\Delta x)^2$ dalla quale (nella forma dell'eq. (2.41)), avevamo dedotto proprio la (2.45).² Se vogliamo controllare esplicitamente, dalla (2.45) ricaviamo che

$$\Delta t' = \gamma(\Delta t - \frac{v}{c^2}\Delta x) \quad (3.4)$$

$$\Delta x' = \gamma(\Delta x - v\Delta t) \quad (3.5)$$

$$\implies (c\Delta t')^2 - (\Delta x')^2 = \gamma^2[(\Delta t)^2(c^2 - v^2) + (\Delta x)^2(\frac{v^2}{c^2} - 1)] = (c\Delta t)^2 - (\Delta x)^2. \quad (3.6)$$

In conclusione, la quantità s^2 è un *invariante relativistico*, che quindi merita un nome tutto suo: viene chiamato **intervallo** spazio-temporale tra gli eventi A e B .

3.1.1 Coordinate omogenee

Cominciamo così ad intravedere la struttura metrica dello spazio-tempo relativistico: tempi e distanze sono sì relative ai SDR che li misurano, ma sono strettamente interdipendenti, in modo tale che gli intervalli spazio-temporali tra tutti i punti (eventi) dello spazio-tempo siano delle quantità assolute, indipendenti dall'osservatore.

Il significato geometrico dell'invarianza dell'intervallo è più trasparente se usiamo quattro coordinate con le stesse dimensioni (coordinate omogenee). Vediamo che in s^2 , a fianco dei $(\Delta x^i)^2$, il tempo compare nella combinazione $(c\Delta t)^2$. Introduciamo la coordinata $x^0 \equiv ct$ che ha le dimensioni di una lunghezza. In termini delle quattro coordinate omogenee

$$x^0 \equiv ct, \quad x^1 \equiv x, \quad x^2 \equiv y, \quad x^3 \equiv z \quad (3.7)$$

l'intervallo si scrive

$$s^2 = (\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2. \quad (3.8)$$

Se non fosse per i segni '−', la forma quadratica (3.8) sarebbe simile al modulo quadro della distanza euclidea in \mathbb{R}^4 .

È istruttivo ricavare le trasformazioni di Lorentz dalla richiesta di invarianza dell'intervallo, paragonandole alle rotazioni. Iniziamo da queste ultime, considerando per semplicità due SDR con le origini coincidenti e prendendo l'evento A proprio in tale origine comune, in modo che $\Delta x^\mu = x^\mu \quad \forall \mu = 0, 1, 2, 3$ e analogamente per le x' .

Nel caso di una rotazione spaziale attorno all'asse z , sia la coordinata $x^3 = z$ che la coordinata temporale $x^0 = ct$ restano invariate. Invece le coordinate x^1 ed x^2 si “mescolano” in modo tale però da lasciare invariata la combinazione $(x^1)^2 + (x^2)^2 = \text{cost.}$ Questa equazione descrive delle circonferenze nel piano (x^1, x^2) che possono essere parametrizzate mediante le usuali funzioni trigonometriche, e si ha:

$$\begin{aligned} x'^0 &= x^0 \\ x'^1 &= \cos(\theta)x^1 + \sin(\theta)x^2 \\ x'^2 &= -\sin(\theta)x^1 + \cos(\theta)x^2 \\ x'^3 &= x^3, \end{aligned}$$

Nel caso di una spinta lungo l'asse x^1 abbiamo visto che x^2 ed x^3 non variano, perciò l'invarianza dell'intervallo implica che $(x^0)^2 - (x^1)^2 = \text{cost.}$ Questa equazione nel piano (x^0, x^1) rappresenta delle iperboli equilatera con gli assi coincidenti agli assi coordinati. Tali iperboli si possono parametrizzare mediante le funzioni iperboliche nel seguente modo:

$$x'^0 = \cosh(\eta)x^0 - \sinh(\eta)x^1 \quad (3.9a)$$

$$x'^1 = -\sinh(\eta)x^0 + \cosh(\eta)x^1 \quad (3.9b)$$

$$x'^2 = x^2 \quad (3.9c)$$

$$x'^3 = x^3. \quad (3.9d)$$

²Nell'eq. (2.41) non ci sono i Δ poiché avevamo preso come evento di riferimento (emissione dell'onda luminosa) l'origine comune O dei due SDR.

Abbiamo così riottenuto la trasformazione di Lorentz (2.45), solo che ora, al posto della velocità relativa v , i coefficienti di trasformazione a_{ij} sono espressi in termini di seni e coseni iperbolici dell'angolo iperbolico η , comunemente chiamato **rapidità**.

Innanzitutto è facile verificare che la trasformazione di coordinate (3.9) conserva l'intervallo, per la ben nota proprietà delle funzioni iperboliche $\cosh^2 - \sinh^2 = 1$. Questo modo di scrivere la trasformazione ha il merito di mettere in evidenza la natura dei cambi di SDR in moto come pseudo-rotazioni, o “rotazioni iperboliche” spazio-temporali, che preservano una particolare “distanza” (l'intervallo in questo caso), così come nella geometria euclidea le rotazioni usuali preservano la distanza usuale.

Quindi lo spazio ed il tempo non sono altro che proiezioni di eventi su assi spazio-temporali la cui orientazione dipende anche dallo stato di moto del SDR. In altre parole, spazio e tempo sono aspetti diversi di un'unica entità fisica, lo **spazio-tempo**; in relatività spazio e tempo non possono più essere considerati indipendentemente l'uno dall'altro.

Per ricavare la relazione tra velocità v e rapidità η , introduciamo il parametro di velocità adimensionale $\beta \equiv v/c$, in termini del quale le eq. (2.45) diventano

$$x'^0 = \gamma(x^0 - \beta x^1), \quad \beta \equiv \frac{v}{c}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad (3.10a)$$

$$x'^1 = \gamma(-\beta x^0 + x^1) \quad (3.10b)$$

$$x'^2 = x^2 \quad (3.10c)$$

$$x'^3 = x^3. \quad (3.10d)$$

confrontando queste espressioni con le eq. (3.9) ed identificando i vari coefficienti troviamo

$$\sinh(\eta) = \frac{\beta}{\sqrt{1 - \beta^2}}, \quad \cosh(\eta) = \frac{1}{\sqrt{1 - \beta^2}}, \quad \tanh(\eta) = \beta \quad (3.11)$$

3.2 Lo spazio di Minkowski

3.2.1 Quadri-vettori

L'intima relazione che sussiste tra tempo e spazio, ci porta naturalmente a concepire lo spazio-tempo come un'entità a 4 dimensioni, un insieme continuo di eventi caratterizzati da quattro coordinate. Non essendoci eventi privilegiati (R1), l'ambiente matematico più adatto per descrivere questo insieme è uno spazio affine \mathcal{M} a 4 dimensioni. I vettori dello spazio vettoriale \mathcal{V} soggiacente vengono chiamati **quadri-vettori**, e connettono le coppie di eventi in \mathcal{M} :

$$A, B \in \mathcal{M} \implies B - A \equiv \underline{AB} = \underline{u} \in \mathcal{V} \iff B = A + \underline{u}. \quad (3.12)$$

Un dato osservatore (SDR) inerziale assegna ad ogni evento una quaterna³ di coordinate x^μ : $\mu = 0, 1, 2, 3$. Questo corrisponde ad avere scelto un evento $O \in \mathcal{M}$ come origine, ed una quaterna di quadri-vettori $\underline{e}_\mu \in \mathcal{V}$ in modo tale che

$$A = O + \sum_{\mu=0}^3 \underline{e}_\mu x_A^\mu, \quad \underline{u} = \sum_{\mu=0}^3 \underline{e}_\mu u^\mu. \quad (3.13)$$

3.2.2 Classificazione dei quadri-vettori

Dato un quadri-vettore $\underline{u} = \underline{AB} \in \mathcal{V}$, l'intervallo tra A e B è anche chiamato “quadrato” del quadri-vettore, e viene indicato con⁴

$$\underline{u}^2 \equiv s_{AB}^2 = (u^0)^2 - \mathbf{u}^2 \in \mathbb{R}, \quad (3.14)$$

³Usiamo indici greci $\mu, \nu \dots$ per indicare le 4 componenti spazio-temporali. Quando vorremo indicare solamente le componenti spaziali, useremo indici latini i, j, k, \dots .

⁴Si presenta ora una possibile ambiguità nel simbolo u^2 : come “quadrato” del vettore, oppure come componente y del vettore. Il contesto in cui appare il simbolo permette di distinguere di quale caso si tratti.

dove u^0 indica la componente temporale di \underline{u} mentre $\mathbf{u} = (u^1, u^2, u^3)$ ne indica le tre componenti spaziali, in un qualche SDR, come indicato dall'eq. (3.13). Il fatto che l'intervallo sia invariante fa sì che la (3.14) sia una buona definizione — cioè indipendente dal SDR. Il fatto che l'intervallo non sia definito positivo ci permette di classificare i quadri-vettori — che d'ora in poi chiameremo semplicemente vettori — in tre tipi distinti:

Vettori di tipo-luce

Sono quelli con quadrato nullo $\underline{u}^2 = 0$. In questo caso componente temporale e spaziale hanno la stessa “lunghezza”: $|u^0| = |\mathbf{u}|$ in ogni SDR (anche se u^0 e \mathbf{u} in generale cambiano da un SDR all'altro). L'insieme di tutti i vettori di tipo luce viene chiamato **cono-luce**. “Cono” in quanto in uno spazio-tempo con 2 dimensioni spaziali ed 1 temporale, la sua forma sarebbe proprio quella di un cono a 2 falde, come visualizzato in fig. 3.1; “luce” poiché questi vettori uniscono coppie di eventi collegate da raggi di luce, ossia per i quali $c|\Delta t| = |\Delta \mathbf{x}|$.

Vettori di tipo-tempo

Sono quelli con quadrato positivo $\underline{u}^2 > 0$ come il versore temporale $\underline{e}_0 = (1, 0, 0, 0)$. Si ha evidentemente $|u^0| > |\mathbf{u}|$. Questi si trovano all'interno del cono-luce, in due regioni disgiunte, dovendo essere o $u^0 > 0$ o $u^0 < 0$. I vettori di tipo-tempo connettono coppie di eventi tra i quali possiamo mandare segnali (di velocità inferiore a c), essendo $|c\Delta t| > |\Delta \mathbf{x}|$. Abbiamo visto nella sez. 2.7 che l'ordine temporale di queste coppie di eventi non può essere invertito dalle trasformazioni di SDR, quindi $c\Delta t = u^0$ non può cambiare segno. I vettori con $u^0 > 0$ si dicono rivolti verso il futuro, e formano l'interno del cono-luce futuro; quelli con $u^0 < 0$ si dicono rivolti verso il passato, e formano l'interno del cono-luce passato.

Mostriamo che per un qualsiasi vettore \underline{u} di tipo-tempo è sempre possibile trovare un SDR in cui le sue componenti spaziali sono nulle: $\mathbf{u}' = 0$. Consideriamo infatti una rotazione spaziale che allinei la parte spaziale di \underline{u} lungo l'asse x : $u^\mu = (u^0, |\mathbf{u}|, 0, 0)$. A questo punto, una trasformazione di Lorentz (3.10) con $\beta = -|\mathbf{u}|/u^0 \in]-1, 1[$ annulla anche la componente 1 (come succede al vettore \underline{AB}_1 lungo la linea punteggiata blu in fig. 2.4), per cui in questo nuovo SDR si ha $u'^\mu = (\pm\sqrt{u^2}, 0, 0, 0)$, con segno $+$ se $u^0 > 0$ e viceversa.

Vettori di tipo-spazio

Sono quelli con quadrato negativo $\underline{u}^2 < 0$ come i versori spaziali $\underline{e}_1 = (0, 1, 0, 0)$ ecc., per i quali vale $|u^0| < |\mathbf{u}|$. I vettori di tipo-spazio si trovano all'esterno del cono-luce, e formano un insieme connesso. I vettori di tipo-spazio connettono coppie di eventi tra i quali non possiamo mandare segnali, essendo $|c\Delta t| < |\Delta \mathbf{x}|$. Per questi vettori, possiamo trovare delle trasformazioni di Lorentz che fanno cambiare segno a u^0 (vedi sez. 2.7). Dunque per un vettore di tipo-spazio il segno della componente temporale non è un invariante. In questo caso si possono trovare dei SDR in cui la parte temporale è nulla: $u^0 = 0$. Procedendo come prima con una rotazione spaziale che riduca u^μ alla forma $u^\mu = (u^0, |\mathbf{u}|, 0, 0)$ e passando ad un nuovo SDR con velocità $\beta = -u^0/|\mathbf{u}| \in]-1, 1[$, dall'eq. (3.10) otteniamo $u'^\mu = (0, \sqrt{-u^2}, 0, 0)$, come succede alle componenti del vettore \underline{AB}_2 lungo la linea tratteggiata rossa in fig. 2.4.

3.2.3 Tensore metrico

Vogliamo adesso introdurre un prodotto scalare (\cdot, \cdot) tra coppie di vettori che “induca” la forma quadratica (3.8), ossia tale che

$$s_{AB}^2 = (\underline{AB}, \underline{AB}) , \quad (3.15)$$

così come nello spazio euclideo si ha $|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}$. Chiaramente questo prodotto scalare non sarà definito positivo, nel senso che $(\underline{u}, \underline{u})$ potrà avere segno positivo o negativo. Richiediamo comunque che questo prodotto scalare sia reale, simmetrico e lineare nei due argomenti:

$$(\underline{u}, \underline{v}) = (\underline{v}, \underline{u}) \in \mathbb{R} , \quad (\underline{u}, \lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2) = \lambda_1 (\underline{u}, \underline{v}_1) + \lambda_2 (\underline{u}, \underline{v}_2) \quad \forall \lambda_1, \lambda_2 \in \mathbb{R} . \quad (3.16)$$

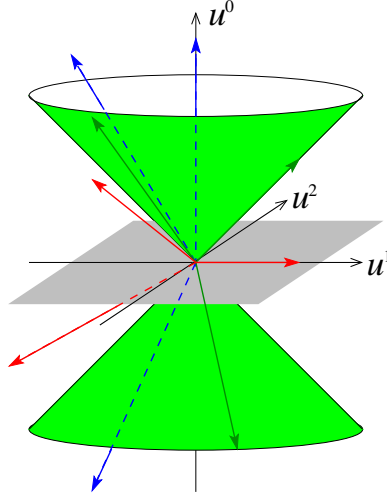


Figure 3.1: *Classificazione dei vettori dello spazio-tempo. In verde, vettori di tipo-luce, l'insieme dei quali forma il cono-luce; in blu, vettori di tipo-tempo; in rosso, vettori di tipo-spazio.*

In termini delle coordinate x^μ in S , l'intervallo del vettore \underline{OA} si esprime

$$s_{OA}^2 = (\underline{OA}, \underline{OA}) = \left(\sum_{\mu} \underline{e}_{\mu} x^{\mu}, \sum_{\nu} \underline{e}_{\nu} x^{\nu} \right) = \sum_{\mu\nu} x^{\mu} x^{\nu} (\underline{e}_{\mu}, \underline{e}_{\nu}) = \sum_{\mu} (x^{\mu})^2 (\underline{e}_{\mu}, \underline{e}_{\mu}) + 2 \sum_{\mu > \nu} x^{\mu} x^{\nu} (\underline{e}_{\mu}, \underline{e}_{\nu}), \quad (3.17)$$

in cui abbiamo sfruttato prima la linearità e poi (in $\sum_{\mu > \nu}$) la simmetria del prodotto scalare. Confrontando con la definizione (3.8)

$$s_{OA}^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2, \quad (3.18)$$

le due relazioni concordano per ogni evento A , ossia per ogni valore di x^μ , se e solo se

$$(\underline{e}_0, \underline{e}_0) = 1 \quad (3.19a)$$

$$(\underline{e}_i, \underline{e}_i) = -1, \quad (i = 1, 2, 3) \quad (3.19b)$$

$$(\underline{e}_{\mu}, \underline{e}_{\nu}) = 0, \quad (\mu \neq \nu) \quad (3.19c)$$

Rispetto a questo prodotto scalare i quattro vettori di base sono mutuamente ortogonali, come è giusto che sia in un SDR con coordinate indipendenti. Tuttavia c'è una differenza cruciale rispetto al prodotto scalare euclideo: i tre “versori” spaziali hanno quadrato negativo, mentre quello temporale ha quadrato positivo. Questa differenza di segno è ciò che distingue lo spazio dal tempo.⁵

Avendo specificato il prodotto scalare per ogni coppia dei vettori di base, possiamo definire, per linearità, il prodotto scalare di una qualsivoglia coppia di vettori

$$(\underline{u}, \underline{v}) = \left(\sum_{\mu} \underline{e}_{\mu} u^{\mu}, \sum_{\nu} \underline{e}_{\nu} v^{\nu} \right) = \sum_{\mu\nu} u^{\mu} v^{\nu} (\underline{e}_{\mu}, \underline{e}_{\nu}) = \sum_{\mu\nu} u^{\mu} v^{\nu} g_{\mu\nu}, \quad (3.20)$$

in cui abbiamo introdotto il **tensore metrico**

$$g_{\mu\nu} \equiv (\underline{e}_{\mu}, \underline{e}_{\nu}) \quad (3.21)$$

che appunto sintetizza l'informazione del prodotto scalare dei versori del SDR S . Il nome “tensore” si usa per designare gli enti che realizzano operazioni lineari, com'è per l'appunto il prodotto scalare che abbiamo introdotto.

Un modo conveniente di rappresentare il prodotto scalare è quello matriciale, in cui le coordinate spazio-temporali vengono disposte in un vettore colonna a 4 componenti, ed il tensore metrico viene rappresentato come

⁵Naturalmente avremmo potuto definire (ed alcuni autori lo fanno) un prodotto scalare con segno opposto, ossia tale che $(\underline{u}, \underline{u}) = -s^2$. Anche in questo caso però l'intervallo di \underline{e}_0 ha segno opposto all'intervallo dei versori spaziali.

una matrice quadrata 4×4 in cui, al solito, il primo indice (μ) è l'indice riga ed il secondo (ν) è l'indice colonna:

$$x \equiv ||x^\mu|| \equiv \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \quad G \equiv ||g_{\mu\nu}|| \equiv \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (3.22)$$

$$(\underline{u}, \underline{v}) = \sum_{\mu\nu} u^\mu g_{\mu\nu} v^\nu = u^T G v \quad (3.23)$$

$$s_{AB}^2 = (\underline{AB}, \underline{AB}) = \sum_{\mu\nu} \Delta x^\mu g_{\mu\nu} \Delta x^\nu = (\Delta x)^T G \Delta x, \quad (\Delta x^\mu \equiv x_B^\mu - x_A^\mu). \quad (3.24)$$

Esplicitando il prodotto scalare in termini delle componenti temporali e spaziali si ha

$$(\underline{u}, \underline{v}) = u^0 v^0 - (u^1 v^1 + u^2 v^2 + u^3 v^3) = u^0 v^0 - \mathbf{u} \cdot \mathbf{v}, \quad (3.25)$$

dove $\mathbf{u} \cdot \mathbf{v}$ indica il prodotto scalare euclideo tra 3-vettori.

Per come è stato definito, il tensore metrico dipende dal SDR usato. Tuttavia, in ogni SDR cartesiano la formula che esprime l'intervallo s_{AB}^2 ha la forma (3.3), come abbiamo mostrato nella sez. 3.1. Pertanto, in qualsiasi SDR la matrice del tensore metrico g' ha la forma (3.22), ossia

$$g'_{\mu\nu} \equiv (e'_\mu, e'_\nu) = g_{\mu\nu} \quad \Longleftrightarrow \quad G' = G. \quad (3.26)$$

In conclusione, per una generica coppia di eventi A e B , l'intervallo in termini delle coordinate in due qualsiasi SDR S ed S' si scrive

$$s_{AB}^2 = \sum_{\mu\nu} \Delta x^\mu g_{\mu\nu} \Delta x^\nu = \sum_{\mu\nu} \Delta x'^\mu g_{\mu\nu} \Delta x'^\nu \quad (3.27)$$

con lo stesso tensore metrico $g_{\mu\nu}$.

Esercizio: Qual è il tensore metrico del prodotto scalare standard nello spazio euclideo?

Convenzione di Einstein Osserviamo che le sommatorie in μ, ν coinvolgono sistematicamente coppie di indici uno dei quali in alto e l'altro in basso. Questo permette di semplificare la notazione: sottointenderemo sempre la sommatoria sulle coppie di indici uguali, uno in alto e l'altro in basso.

Invarianza del prodotto scalare L'invarianza dell'intervallo si traduce nell'invarianza del prodotto scalare di un vettore per se stesso:

$$\underline{u}^2 = (\underline{u}, \underline{u}) = u^\mu g_{\mu\nu} u^\nu = u'^\mu g_{\mu\nu} u'^\nu \quad \forall \underline{u} \in \mathcal{V}, \quad \forall S, S'. \quad (3.28)$$

Ci chiediamo se il prodotto scalare di due vettori distinti definito nell'eq. (3.20) dipenda o meno dal SDR usato per calcolare le componenti u^μ e v^ν . Il fatto è che possiamo definire il prodotto scalare tra vettori in modo intrinseco, ossia svincolato dalla scelta di un SDR. Abbiamo infatti, per le proprietà di linearità e simmetria (3.16),

$$(\underline{u} + \underline{v}, \underline{u} + \underline{v}) = (\underline{u}, \underline{u}) + (\underline{v}, \underline{v}) + 2(\underline{u}, \underline{v}) \quad \Longleftrightarrow \quad (\underline{u}, \underline{v}) = \frac{1}{2} [(\underline{u} + \underline{v})^2 - \underline{u}^2 - \underline{v}^2] \quad (3.29)$$

per ogni coppia di vettori $\underline{u}, \underline{v} \in \mathcal{V}$. Dal momento che la combinazione lineare di quadrati nell'ultimo membro non dipende dal SDR (per l'invarianza dell'intervallo), anche il prodotto scalare $(\underline{u}, \underline{v})$ non dipende dal SDR.

Possiamo quindi affermare che il prodotto scalare (3.20) è un invariante relativistico, ossia

$$(\underline{u}, \underline{v}) = u^\mu g_{\mu\nu} v^\nu = u'^\mu g_{\mu\nu} v'^\nu \quad (3.30)$$

per ogni coppia di vettori $\underline{u}, \underline{v} \in \mathcal{V}$ ed in qualsiasi SDR si calcolino le componenti di tali vettori.

3.2.4 Diagrammi spazio-temporali per trasformazioni di Lorentz speciali

Analizziamo ora le relazioni tra due osservatori inerziali distinti S ed S' nel contesto dello spazio di Minkowski. Ognuno di essi sarà dotato del proprio sistema di riferimento: $S = (O, \underline{e}_\mu)$ e $S' = (O', \underline{e}'_\mu)$. La trasformazione tra i SDR coinvolge una matrice invertibile M di cambio di base, secondo le relazioni

$$\underline{e}'_\mu = \underline{e}_\nu M^\nu{}_\mu, \quad \underline{O}'O = \underline{e}'_\mu a^\mu \quad (3.31)$$

del tutto analoghe a quelle euclidee che si leggono nelle eq. (1.11) e (1.22). È anche immediato verificare che la trasformazione tra le coordinate x^μ in S di un generico evento A e le coordinate x'^ν in S' dello stesso evento, sono legate dalla relazione

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (3.32)$$

ove $\Lambda = M^{-1}$. Nel caso della trasformazione di Lorentz speciale (3.10) vale

$$\underline{e}'_0 = \cosh(\eta) \underline{e}_0 + \sinh(\eta) \underline{e}_1 = \gamma(\underline{e}_0 + \beta \underline{e}_1) \quad (3.33a)$$

$$\underline{e}'_1 = \sinh(\eta) \underline{e}_0 + \cosh(\eta) \underline{e}_1 = \gamma(\beta \underline{e}_0 + \underline{e}_1) \quad (3.33b)$$

$$\underline{e}'_2 = \underline{e}_2 \quad (3.33c)$$

$$\underline{e}'_3 = \underline{e}_3 \quad (3.33d)$$

In fig. 3.2 sono rappresentati i vettori relativi al sottospazio (x^0, x^1) nei due SDR.

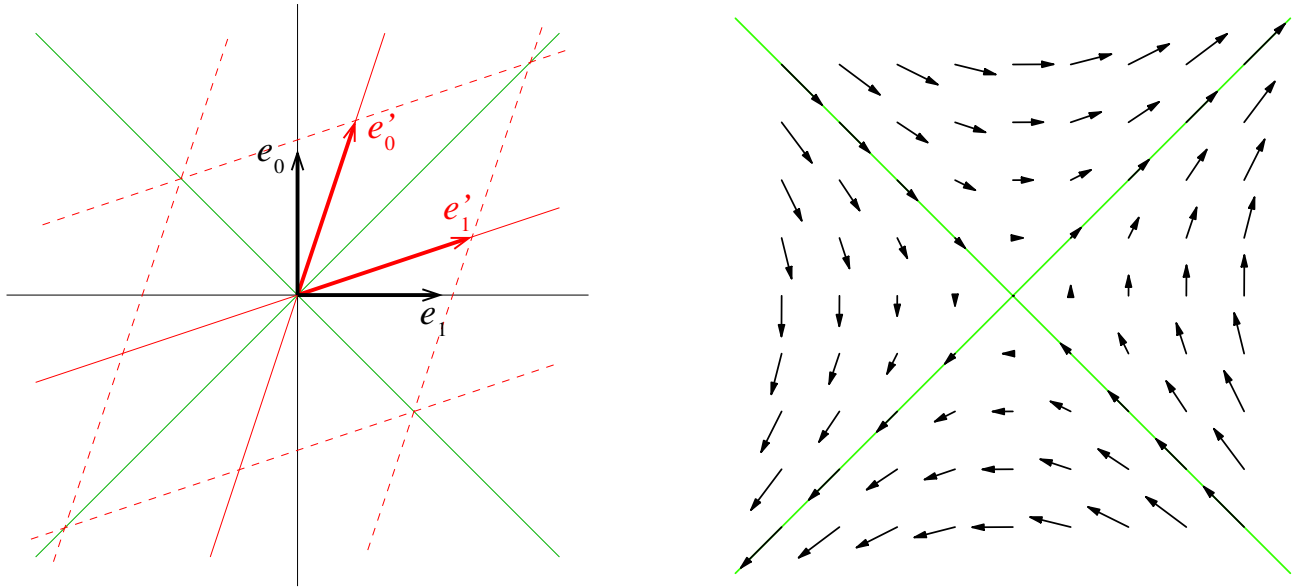


Figure 3.2: *Effetto di una trasformazione speciale di Lorentz nello spazio di Minkowski. (a) In nero sono rappresentati i vettori di base e gli assi coordinati del SDR S . Assi e vettori di base di S' sono in rosso, assieme ad alcune linee a x'^1 costante (punteggiate) e x'^0 costante (tratteggiate). Queste ultime indicano superfici di simultaneità per S' . Le bisettrici in verde indicano il cono-luce, comune a tutti i SDR inerziali. (b) Flusso della trasformazione speciale di Lorentz.*

3.3 Trasformazioni generali tra SDR inerziali

Nel cap. 2 abbiamo ricavato le trasformazioni di Lorentz come trasformazioni tra SDR inerziali soddisfacenti ai postulati R1 (omogeneità ed isotropia) ed R2 (invarianza della velocità della luce), mostrando poi nella sez. 3.1 che tali trasformazioni preservano l'intervallo tra eventi dello spazio-tempo. Viene spontaneo chiedersi se l'invarianza dell'intervallo basti per caratterizzare le trasformazioni di Lorentz.

Vogliamo quindi caratterizzare tutte le possibili trasformazioni tra SDR inerziali che preservano l'intervallo, sempre nelle ipotesi di omogeneità ed isotropia R1. Tali trasformazioni dovranno soddisfare i seguenti requisiti:

1) *Linearità* : (più precisamente, affinità)

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} + a^{\mu} , \quad (3.34)$$

ove Λ^{μ}_{ν} e a^{μ} sono numeri reali indipendenti dalle x^{ν} ;

2) *Invarianza dell'intervallo*:

$$(\Delta x'^0)^2 - (\Delta \mathbf{x}')^2 = (\Delta x^0)^2 - (\Delta \mathbf{x})^2 , \quad (3.35)$$

ove si intende $\Delta x^{\mu} \equiv x_B^{\mu} - x_A^{\mu}$. Questo requisito abbiamo visto essere equivalente all'invarianza del prodotto scalare:

$$u^{\mu} g_{\mu\nu} v^{\nu} = u'^{\mu} g_{\mu\nu} v'^{\nu} \quad (3.36)$$

ove $u^{\mu} \equiv x_B^{\mu} - x_A^{\mu}$ e $v^{\nu} \equiv x_C^{\nu} - x_D^{\nu}$.

Questa seconda condizione farà sì che, nell'espressione della trasformazione lineare (3.34), gli a^{μ} possano essere arbitrari, cancellandosi nelle differenze; gli elementi delle matrici Λ invece dovranno obbedire a precise limitazioni. Per trovare queste limitazioni, procediamo usando la notazione matriciale.

La condizione (3.34) diventa

$$x' = \Lambda x + a \quad \Longleftrightarrow \quad x'^T = x^T \Lambda^T + a^T , \quad (3.37)$$

(Λ matrice reale 4×4 ed a vettore colonna reale a 4 componenti) da cui possiamo subito ricavare

$$\Delta x' = \Lambda \Delta x \quad \Longleftrightarrow \quad \Delta x'^T = \Delta x^T \Lambda^T . \quad (3.38)$$

Essendo le componenti dei vettori delle differenze di coordinate, si ha

$$v' = \Lambda v , \quad u'^T = u^T \Lambda^T . \quad (3.39)$$

La condizione (3.36) sull'invarianza del prodotto scalare si riscrive

$$u'^T G v' = u^T G v . \quad (3.40)$$

Sostituendo in questa equazione le espressioni per u'^T e v' ottenute nell'eq. (3.38), la condizione di invarianza dell'intervallo si scrive

$$u^T \Lambda^T G \Lambda v = u^T G v , \quad (3.41)$$

la quale deve valere per ogni u e v . Questo può essere vero solamente se vale l'equazione tra matrici

$$\Lambda^T G \Lambda = G \quad \Longleftrightarrow \quad g_{\mu\nu} \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} = g_{\alpha\beta} . \quad (3.42)$$

Questa relazione esprime matematicamente il fatto che la trasformazione di coordinate (3.37) preserva il prodotto scalare G (ossia l'intervallo), ed è la relazione che cercavamo per caratterizzare le matrici Λ .

Si osservi l'analogia tra la condizione di "pseudo-ortogonalità" (3.42) e la condizione di ortogonalità $R^T \mathbb{I} R = \mathbb{I}$ delle matrici di rotazione R nello spazio euclideo \mathbb{R}^3 , in cui il tensore metrico $G = \mathbb{I}$.

3.3.1 Il gruppo di Poincaré

Le trasformazioni (3.34) soddisfacenti alla condizione (3.42) sono dette **trasformazioni di Poincaré**, e costituiscono l'insieme di tutte e sole le trasformazioni affini che preservano l'intervallo. Detto P l'insieme delle trasformazioni di Poincaré, ogni elemento di P sarà specificato dalla coppia $\{\Lambda, a\}$, e scriviamo

$$\{\Lambda, a\} \in P . \quad (3.43)$$

Cerchiamo di investigare la struttura di questo insieme P , studiando la composizione di due trasformazioni di Poincaré $\{\Lambda_1, a_1\} \in P$, $\{\Lambda_2, a_2\} \in P$ che realizzano i cambi di coordinate $x \mapsto x'$ e $x' \mapsto x''$ nell'ordine:

$$x' = \Lambda_1 x + a_1 \quad (3.44)$$

$$x'' = \Lambda_2 x' + a_2 = \Lambda_2 \Lambda_1 x + \Lambda_2 a_1 + a_2 . \quad (3.45)$$

Definiamo quindi in P la legge di composizione interna (una sorta di prodotto)

$$\{\Lambda_2, a_2\} \{\Lambda_1, a_1\} \equiv \{\Lambda_2 \Lambda_1, \Lambda_2 a_1 + a_2\} . \quad (3.46)$$

Vogliamo dimostrare la proprietà di chiusura di questo prodotto, cioè che

$$\{\Lambda_2 \Lambda_1, \Lambda_2 a_1 + a_2\} \in P . \quad (3.47)$$

Dall'eq. (3.45) è chiaro che la trasformazione composta $x \mapsto x''$ è affine; è anche intuibile che essa preserva l'intervallo, dal momento che è definita come il prodotto di trasformazioni ciascuna delle quali preserva l'intervallo. Per dimostrare (3.47) è sufficiente far vedere che la matrice $\Lambda_2 \Lambda_1$ soddisfa la condizione (3.42). Ciò è immediato, infatti

$$(\Lambda_2 \Lambda_1)^T G (\Lambda_2 \Lambda_1) = \Lambda_1^T \Lambda_2^T G \Lambda_2 \Lambda_1 \stackrel{(3.42)}{=} \Lambda_1^T G \Lambda_1 \stackrel{(3.42)}{=} G . \quad (3.48)$$

Facciamo ora vedere che l'insieme P , dotato della legge di composizione interna (3.46) appena definita, costituisce un gruppo, detto **gruppo di Poincaré** :

- a) La legge di composizione è associativa, essendo associativo il prodotto di matrici;
- b) Esiste in P l'elemento neutro, dato da

$$\mathbb{I}_P = \{\mathbb{I}, 0\} , \quad (3.49)$$

essendo \mathbb{I} la matrice identità 4×4 e 0 il 4-vettore nullo. Tale elemento neutro corrisponde ovviamente alla trasformazione identica $x \mapsto x$;

- c) Dato un qualsiasi elemento $\{\Lambda, a\} \in P$ esiste in P l'elemento inverso. Innanzitutto, questo inverso avrà la forma $\{\Lambda^{-1}, b\}$, in quanto nella composizione (3.46) le matrici si moltiplicano, e se vogliamo che la composizione con l'inverso dia l'elemento neutro (3.49), non abbiamo altre possibilità. La matrice Λ è invertibile: moltiplicando per G^{-1} a sinistra ambo i membri dell'eq. (3.42) otteniamo

$$G^{-1} \Lambda^T G \Lambda = G^{-1} G = \mathbb{I} \quad \implies \quad G^{-1} \Lambda^T G = \Lambda^{-1} \quad (3.50)$$

Verifichiamo che Λ^{-1} induce una trasformazione di Poincaré, ossia soddisfa la condizione (3.42): usando $G^T = G^{-1} = G$ si ha

$$(\Lambda^{-1})^T G \Lambda^{-1} = (G \Lambda^T G)^T G \Lambda^{-1} = G^T \Lambda G^T G \Lambda^{-1} = G \Lambda \Lambda^{-1} = G . \quad (3.51)$$

Rimane da determinare il 4-vettore b della trasformazione inversa. Si ha

$$\{\Lambda^{-1}, b\} \{\Lambda, a\} \stackrel{(3.46)}{=} \{\Lambda^{-1} \Lambda, \Lambda^{-1} a + b\} \quad \implies \quad b = -\Lambda^{-1} a . \quad (3.52)$$

Abbiamo così ottenuto

$$\{\Lambda, a\}^{-1} = \{\Lambda^{-1}, -\Lambda^{-1} a\} , \quad \Lambda^{-1} = G \Lambda^T G . \quad (3.53)$$

Ogni elemento di P ammette dunque l'inverso, e questo basta per concludere che P è effettivamente un gruppo.

Esempio: Calcoliamo l'elemento $\{\Lambda, a\}$ di P corrispondente alla trasformazione di Lorentz (3.10): si ha

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad a = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.54)$$

3.3.2 Struttura del gruppo di Poincaré

Particolari trasformazioni del gruppo di Poincaré sono quelle del tipo:

$$\{\mathbb{I}, a\} \iff x' = x + a, \quad (3.55)$$

ovverosia sono traslazioni nello spazio e nel tempo. Detto T l'insieme di tali trasformazioni, non è difficile dimostrare che esso è un sottogruppo abeliano invariante di P , detto **gruppo delle traslazioni**.

Altre particolari trasformazioni del gruppo di Poincaré sono quelle omogenee, ossia del tipo

$$\{\Lambda, 0\} \iff x' = \Lambda x. \quad (3.56)$$

Detto \mathcal{L} l'insieme di tale trasformazioni, non è difficile dimostrare che anch'esso è un sottogruppo (non abeliano né invariante) di P , detto **gruppo di Lorentz**.

3.3.3 Il gruppo di Lorentz

Concentriamoci adesso sul gruppo di Lorentz, i cui elementi potremo più semplicemente indicare con la sola matrice Λ . Dalla condizione $\Lambda^T G \Lambda = G$ ricaviamo la seguente relazione tra i determinanti

$$\det(G) = \det(\Lambda^T) \det(G) \det(\Lambda) = \det(G) (\det \Lambda)^2. \quad (3.57)$$

Poiché $\det G = -1 \neq 0$, segue $(\det \Lambda)^2 = 1$ e quindi

$$\det \Lambda = \begin{cases} +1 \\ -1 \end{cases}. \quad (3.58)$$

Sempre dalla condizione (3.42) possiamo ottenere una limitazione per l'elemento tempo-tempo di Λ :

$$1 = g_{00} = \Lambda^\mu_{0} g_{\mu\nu} \Lambda^\nu_{0} = \sum_{\mu} \Lambda^\mu_{0} g_{\mu\mu} \Lambda^\mu_{0} = (\Lambda^0_0)^2 - \sum_{i=1}^3 (\Lambda^i_0)^2 \quad (3.59)$$

da cui⁶

$$(\Lambda^0_0)^2 = 1 + \sum_{i=1}^3 (\Lambda^i_0)^2 \geq 1 \quad (3.60)$$

e quindi si possono avere due casi:

$$\Lambda_{00} \begin{cases} \geq +1 \\ \leq -1 \end{cases} \quad (3.61)$$

Conformemente alle due alternative (3.58) e (3.61), possiamo suddividere le trasformazioni di Lorentz in 4 classi

$$1) \quad \mathcal{L}^{\uparrow}_{+} : \quad \det \Lambda = +1, \quad \Lambda_{00} \geq +1.$$

A questa classe appartiene la trasformazione identica $\Lambda = \mathbb{I}$. Tutte le trasformazioni di questa classe si possono ottenere “con continuità” a partire dalla trasformazione identica, ossia componendo traslazioni, rotazioni e spinte infinitesime. Non è difficile dimostrare che questa classe è un sottogruppo del gruppo di Lorentz, detto **gruppo proprio di Lorentz**.

$$2) \quad \mathcal{L}^{\uparrow}_{-} : \quad \det \Lambda = -1, \quad \Lambda_{00} \geq +1.$$

⁶Questa condizione corrisponde al fatto che la prima riga della matrice Λ deve costituire un versore di tipo tempo. In generale, la condizione $\Lambda^T G \Lambda = G$ dice che le colonne (e anche le righe) di ogni matrice di Lorentz devono formare un insieme di 4 versori pseudo-ortonormali.

Una tipica matrice di questa classe è

$$\Lambda_{\mathcal{P}} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (3.62)$$

che corrisponde alla cosiddetta operazione di **parità** che cambia il segno delle coordinate spaziali:

$$\begin{aligned} x'^0 &= x^0 \\ \mathbf{x}' &= -\mathbf{x}. \end{aligned} \quad (3.63)$$

Questa classe non può ovviamente costituire un sottogruppo del gruppo di Lorentz, in quanto ad essa non appartiene la matrice unità. Questa seconda classe è disgiunta dalla prima, ed è impossibile passare “con continuità” dall’una all’altra. Gli elementi di \mathcal{L}_-^\uparrow si possono però ottenere da quelli di \mathcal{L}_+^\uparrow moltiplicando questi ultimi per $\Lambda_{\mathcal{P}}$.

Può essere utile a questo punto un paragone con il gruppo delle matrici ortogonali $O(N, \mathbb{R})$ definito come l’insieme delle matrici reali $N \times N$ che soddisfano $R^T R = \mathbb{I}$, e che quindi preservano il prodotto scalare euclideo. Da quest’ultima relazione si ricava $\det R = \pm 1$, e quindi possiamo distinguere due classi in $O(N, \mathbb{R})$:

- 1) il sottoinsieme delle matrici con determinante $+1$, che contiene la matrice identità \mathbb{I}_3 , e che forma il sottogruppo $SO(N, \mathbb{R})$ delle rotazioni proprie, le quali si possono ottenere con continuità a partire dalla matrice identità;
- 2) il sottoinsieme delle matrici con determinante -1 , il quale non formano un sottogruppo di $O(N, \mathbb{R})$, non contenendo l’elemento unità.

Per esempio, per $N = 2$ basta un solo parametro (angolo di rotazione) per descrivere le matrici:

$$R_+ = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \in SO(2), \quad R_- = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ \sin(\phi) & -\cos(\phi) \end{pmatrix} \in O(2) \setminus SO(2). \quad (3.64)$$

La differenza tra metrica euclidea e pseudo-euclidea si riflette nel fatto che, per le matrici ortogonali, vale l’analogo della disuguaglianza (3.60) ma in verso opposto:

$$(R^1_1)^2 \leq 1 \quad \Longleftrightarrow \quad -1 \leq R^1_1 \leq +1 \quad (\text{idem per } R^i_j \quad \forall i, j = 1, \dots, N) \quad (3.65)$$

Questa condizione non genera una distinzione tra gli elementi di $O(N)$.

$$\mathbf{3)} \quad \mathcal{L}_-^\downarrow : \quad \det \Lambda = -1, \quad \Lambda_{00} \leq -1.$$

Una tipica matrice di questa classe è

$$\Lambda_{\mathcal{T}} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (3.66)$$

che corrisponde all’operazione di **inversione del tempo** che cambia il segno della coordinata temporale:

$$\begin{aligned} x'^0 &= -x^0 \\ \mathbf{x}' &= \mathbf{x}. \end{aligned} \quad (3.67)$$

Anche per questa classe valgono considerazioni analoghe alla precedente: in particolare, gli elementi di \mathcal{L}_-^\downarrow si ottengono moltiplicando le matrici del sottogruppo proprio di Lorentz per $\Lambda_{\mathcal{T}}$.

$$\mathbf{4)} \quad \mathcal{L}_+^\downarrow : \quad \det \Lambda = +1, \quad \Lambda_{00} \leq -1.$$

Una tipica matrice di questa classe è

$$\Lambda_{\mathcal{T}} \Lambda_{\mathcal{P}} = -\mathbb{I} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (3.68)$$

che corrisponde a cambiare il segno di tutte e quattro le coordinate:

$$\begin{aligned}x'^0 &= -x^0 \\ \mathbf{x}' &= -\mathbf{x} .\end{aligned}\tag{3.69}$$

Anche per \mathcal{L}_+^\downarrow valgono considerazioni analoghe a quelle fatte per \mathcal{L}_-^\uparrow e \mathcal{L}_-^\downarrow : in particolare, ogni elemento di \mathcal{L}_+^\downarrow si può ottenere da un elemento di \mathcal{L}_+^\uparrow moltiplicandolo per $-\mathbb{I}$.

Esercizio: dimostrare che anche per vettori di tipo-luce il segno della componente temporale è invariante per trasformazioni di Lorentz proprie.

Esercizio: dimostrare che $\Lambda^0_0 = \gamma = 1/\sqrt{1-v^2/c^2}$ ove v è la velocità relativa tra i due SDR collegati da Λ . Mostrare che, per $v \neq 0$, gli elementi Λ^i_0 sotto Λ^0_0 non possono essere tutti nulli, così come gli elementi Λ^0_i a destra di Λ^0_0 .

Soluzione: Se nel SDR S due eventi avvengono nello stesso luogo, $\Delta x^\mu = (\Delta x^0, 0, 0, 0)$. Per l'osservatore S' , l'intervallo di tempo tra gli eventi è $\Delta x'^0 = \Lambda^0_\mu \Delta x^\mu = \Lambda^0_0 \Delta x^0$. Ma dalla dilatazione dei tempi sappiamo anche che $\Delta x'^0 = \gamma \Delta x^0$, da cui deduciamo che $\Lambda^0_0 = \gamma$, indipendentemente dall'orientazione tra i due SDR. Inoltre, se $v \neq 0$, $\gamma > 1$, e dall'eq. (3.60) deduciamo che non tutti gli elementi Λ^i_0 possono essere nulli. Lo stesso dicasi per gli elementi Λ^0_i , in quanto anche la matrice Λ^T è una matrice di Lorentz (soddisfa la condizione (3.42)) se Λ è di Lorentz: $\Lambda^T = G\Lambda^{-1}G$ e quindi $(\Lambda^T)^T G \Lambda^T = \Lambda G G \Lambda^{-1} G = G$. Infine, anche Λ^T collega due SDR con velocità relativa v , essendo $\Lambda^0_0 = \gamma$ comune a Λ ed a Λ^T .

3.3.4 Struttura del gruppo proprio di Lorentz

Concentriamo adesso la nostra attenzione sul gruppo proprio di Lorentz \mathcal{L}_+^\uparrow . Per prima cosa, determiniamo il numero di parametri indipendenti delle matrici di Lorentz Λ . La generica matrice reale 4×4 dipende da 16 parametri reali. Se $\Lambda \in \mathcal{L}_+^\uparrow$ essa deve soddisfare la condizione matriciale (3.42) $\Lambda^T G \Lambda = G$. Quest'ultima non fornisce 16 condizioni indipendenti. Infatti, essendo la matrice $G = G^T$ simmetrica, anche la matrice $\Lambda^T G \Lambda$ è simmetrica, qualsiasi sia Λ :

$$(\Lambda^T G \Lambda)^T = \Lambda^T G^T \Lambda = \Lambda^T G \Lambda\tag{3.70}$$

e quindi l'equazione matriciale (3.42) coincide con la stessa equazione trasposta. Pertanto, la condizione fornita in posizione $\mu\nu$ coincide con quella in posizione $\nu\mu$.

Esplicitamente, fissati μ e ν , vale

$$g_{\mu\nu} = \Lambda^\alpha_\mu g_{\alpha\beta} \Lambda^\beta_\nu \quad \Longleftrightarrow \quad g_{\nu\mu} = \Lambda^\alpha_\nu g_{\alpha\beta} \Lambda^\beta_\mu .\tag{3.71}$$

La dimostrazione è semplice: nella prima uguaglianza si scambiano di posizione gli elementi delle matrici Λ (il prodotto di numeri è commutativo), quindi si rinominano gli indici di sommatoria $\alpha \leftrightarrow \beta$ (il nome degli indici sommati è arbitrario) ed infine si sfrutta la simmetria del tensore metrico. Si ottiene così la seconda uguaglianza.

Ci sono quindi 10 condizioni indipendenti: 4 fornite dagli elementi sulla diagonale principale e 6 forniti dagli elementi sopra (o sotto) la diagonale principale. Di conseguenza, la generica matrice $\Lambda \in \mathcal{L}_+^\uparrow$ è individuata da $16 - 10 = 6$ parametri reali indipendenti.

Può essere utile un paragone con le matrici di trasformazione che preservano la metrica euclidea dello spazio a 2 e a 3 dimensioni. Nello spazio 2-dimensionale euclideo, le matrici di rotazione (proprie) sono individuate da 1 solo parametro reale (p.es., l'angolo di rotazione). Nello spazio 3-dimensionale euclideo, le matrici di rotazione (proprie) sono individuate da 3 parametri reali (p.es., i 3 angoli di rotazione di Eulero). In generale, nello spazio \mathbb{R}^N ad N dimensioni, le matrici reali (pseudo)-ortogonali proprie hanno $N(N-1)/2$ parametri indipendenti.

Nel caso del gruppo proprio di Lorentz, la generica Λ è individuata da 6 parametri reali, associati alle “rotazioni” nei seguenti piani coordinati:

$$xy, \quad yz, \quad zx, \quad tx, \quad ty, \quad tz .\tag{3.72}$$

Nei primi 3 casi, abbiamo le solite **rotazioni** spaziali tridimensionali (proprie), le quali costituiscono un sottogruppo del gruppo proprio di Lorentz: $SO(3) < \mathcal{L}_+^\uparrow$.

Nei secondi 3 casi abbiamo delle “pseudo-rotazioni” che coinvolgono il tempo, e corrispondono alle **trasformazioni speciali di Lorentz**, dette anche **spinte** (*boosts* in inglese). Si badi bene che l’insieme delle spinte non costituisce un sottogruppo di \mathcal{L}_+^\uparrow ; solo gli insiemi delle spinte lungo un dato asse formano dei sottogruppi di \mathcal{L}_+^\uparrow .

In definitiva, ogni trasformazione di Poincaré si può realizzare come risultato di

- una spinta (3 parametri)
- una rotazione (3 parametri)
- una traslazione (4 parametri)
- una eventuale riflessione degli assi spaziali e/o temporale.

Abbiamo quindi dimostrato che le trasformazioni di Poincaré, che comprendono anche il caso delle inversioni spaziali e dell’inversione temporale, sono le più generali trasformazioni che lasciano invariata la forma quadratica dell’intervallo.

Esercizio: Determinare la matrice corrispondente ad una spinta in direzione generica (v^1, v^2, v^3) .

Esercizio: Comporre quattro spinte $T \equiv S_y(v_4) \circ S_x(v_3) \circ S_y(v_2) \circ S_x(v_1)$ con $v_1, v_2 > 0$ in modo che la trasformazione complessiva T preservi lo stato di quiete. Determinare v_3 e v_4 in funzione di v_1 e v_2 , e dimostrare che T è una rotazione attorno all’asse z . Determinare l’angolo ed il verso di rotazione.

3.3.5 Invarianza delle leggi della fisica per trasformazioni di Poincaré

Abbiamo derivato le trasformazioni di Poincaré richiedendo che le trasformazioni di coordinate tra SDR diversi siano affini, che l’intervallo spaziotemporale sia invariante, e che tutti i SDR siano equivalenti (o indistinguibili). Queste condizioni implicano i 3 postulati della relatività, ma nel loro insieme sono più forti, perché ammettono di poter considerare come SDR equivalenti anche quelli che differiscono per il verso del tempo (p.es. collegati da $\Lambda_{\mathcal{T}}$) o per l’inversione delle coordinate spaziali (p.es. collegati da $\Lambda_{\mathcal{P}}$, per cui un guanto sinistro per l’uno va bene come guanto destro per l’altro).

Strettamente parlando, R3 afferma l’uguaglianza delle leggi della fisica per SDR inerziali che possiamo identificare con osservatori per i quali il tempo scorre dal passato verso il futuro. Inoltre è implicita la condizione che un qualsiasi SDR S' si possa raggiungere a partire da un altro S mediante operazioni fisiche concrete che coinvolgono traslazioni, rotazioni, spinte e loro combinazioni in modo continuo, ossia traslando con continuità S , ruotandone gli assi con continuità e variandone la sua velocità fino a farlo coincidere con S' . In altre parole, una persona, osservatore in S , può diventare osservatore in S' modificando con continuità il suo stato.

Tradotto nel formalismo dei gruppi di trasformazione, S ed S' devono essere correlati da una trasformazione di Poincaré connessa con l’elemento identità, e quindi la relativa matrice di Lorentz deve appartenere al gruppo proprio: $\Lambda \in \mathcal{L}_+^\uparrow$. Pertanto R3 afferma che le leggi della fisica sono le stesse per osservatori inerziali collegati da trasformazioni di Poincaré *proprie*.

Ci si può chiedere, e ci si è chiesti, se in realtà le leggi della fisica siano le stesse per trasformazioni di coordinate del gruppo di Poincaré completo (non diciamo “osservatori”, per le condizioni implicite che questo termine può avere come argomentato sopra). In effetti, le equazioni dell’elettromagnetismo — e quelle della dinamica relativistica, come vedremo nel cap. 5 — sono invarianti per il gruppo di Poincaré completo. Infatti tale gruppo porta il nome di Poincaré poiché egli per primo lo determinò come il gruppo di trasformazioni che mantengono invariate in forma le equazioni dell’elettromagnetismo.

Fino al 1957 si riteneva che per tutti i sistemi fisici valesse questo principio di relatività esteso. Nel 1957 però si scoprì sperimentalmente che le leggi della fisica non sono necessariamente invarianti per riflessione degli assi spazio-temporali. Ad esempio, le leggi delle interazioni nucleari deboli non sono invarianti per parità o per inversione del tempo. Si deve pertanto concludere che il principio di relatività di Einstein riguarda i sistemi inerziali collegati da trasformazioni del gruppo *proprio* di Poincaré.

Chapter 4

Quadritensori e calcolo quadritensoriale

4.1 Introduzione

In conformità con quanto detto nel capitolo 1, il nostro programma nei prossimi due capitoli sarà il seguente:

- a) modificare le leggi della meccanica in modo che esse soddisfino al terzo postulato della relatività e concordino con quelle della meccanica classica per velocità piccole rispetto a c ;
- b) verificare che le leggi dell'elettromagnetismo soddisfino al terzo postulato della relatività.

Per sviluppare questo programma è opportuno far uso del calcolo quadritensoriale. Per capire questo punto, facciamo una digressione e chiediamoci: perché le equazioni della fisica classica per un sistema isolato si esprimono in forma (tri-)vettoriale? La ragione è che per un sistema isolato le leggi della fisica devono avere la stessa forma in tutti i SDR che differiscono tra loro per una traslazione o una rotazione degli assi. Come si è già detto, ciò è conseguenza dell'omogeneità ed isotropia dello spazio ed implica che, per un sistema S isolato, tutti i SDR ruotati o traslati sono equivalenti, ossia le leggi della fisica non consentono di distinguere tra SDR traslati o ruotati.¹

Come si è già detto, le leggi di trasformazione da un SDR S ad un altro S' rototraslato rispetto a questo sono

$$\mathbf{x}' = R\mathbf{x} + \mathbf{a}, \quad (R^T R = \mathbb{1}), \quad (4.1)$$

dove R è una matrice ortogonale 3×3 , \mathbf{a} è una terna di numeri reali e le terne di numeri reali \mathbf{x} e \mathbf{x}' rappresentano le coordinate del punto P nel sistema S ed S' rispettivamente. Per definizione, un (tri-)scalare $l \in \mathbb{R}$ è un ente che resta immutato per traslazioni e rotazioni:

$$l' = l. \quad (4.2)$$

Un (tri-)vettore

$$\mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \in \mathbb{R}^3 \quad (4.3)$$

è un ente a 3 componenti che resta immutato per traslazioni, e per rotazioni si trasforma secondo la legge

$$\mathbf{v}' = R\mathbf{v} \quad \Longleftrightarrow \quad v'^i = R^i_j v^j. \quad (4.4)$$

Un *tensore doppio*, o di ordine 2,

$$M = (M^{ij})_{j=1,2,3}^{i=1,2,3} = \begin{pmatrix} M^{11} & M^{12} & M^{13} \\ M^{21} & M^{22} & M^{23} \\ M^{31} & M^{32} & M^{33} \end{pmatrix} \quad (4.5)$$

¹Ciò non è più vero se il sistema *non* è isolato: per esempio, se S è un grave, il SDR con l'asse z diretto verso il centro della Terra è privilegiato.

è un ente a 9 componenti che resta immutato per traslazioni, e per rotazioni si trasforma secondo la legge

$$M'^{ij} = R^i_k R^j_n M^{kn} \quad \Longleftrightarrow \quad M' = R M R^T . \quad (4.6)$$

In modo analogo si possono definire tensori tripli, ecc.. Ora, se per un sistema isolato, le leggi della fisica sono espresse in forma vettoriale, cioè come eguaglianza tra scalari e scalari, vettori e vettori, tensori e tensori, ecc., allora automaticamente esse avranno la stessa forma in tutti i sistemi rototraslati. Per esempio, se in S una legge fisica ha la forma

$$\mathbf{p} = \mathbf{q} , \quad (4.7)$$

dove p e q sono vettori, poiché

$$p^i = q^i \quad \Longrightarrow \quad R^j_i p^i = R^j_i q^i \quad \Longrightarrow \quad \mathbf{p}' = \mathbf{q}' , \quad (4.8)$$

la legge avrà la stessa forma anche in S' e, se ciò vale per tutte le leggi fisiche, S ed S' saranno SDR equivalenti.

In fisica relativistica il nostro scopo è quello di ottenere equazioni (per sistemi isolati) che soddisfino al terzo postulato della relatività R3, cioè che abbiano la stessa forma in tutti i SDR inerziali.

Abbiamo già visto che le leggi di trasformazione tra due sistemi inerziali sono quelle (proprie) di Poincaré. Queste ultime sono simili a quelle delle rototraslazioni con le seguenti differenze:

- esse agiscono in uno spazio a 4 dimensioni anziché a tre dimensioni;
- la metrica in tale spazio è pseudo-euclidea anziché euclidea.

Queste differenze tuttavia non ci impediscono di costruire un calcolo quadri-vettoriale simile a quello tri-vettoriale e concludere che se le leggi della fisica sono espresse in forma quadri-vettoriale o, come si dice, sono *covarianti a vista* (cioè sono espresse da eguaglianze tra scalari e scalari, quadri-vettori e quadri-vettori, ecc.) allora esse saranno automaticamente invarianti in forma in tutti i SDR inerziali. A volte anzi, il fatto che le leggi della fisica per un sistema isolato devono essere covarianti a vista, può essere usato in modo euristico per “intuire” le leggi “corrette”.

4.2 Quadritensori

4.2.1 Quadriscalarari

Un (quadri-)scalare è un ente che non cambia per trasformazioni del gruppo di Poincaré

$$x'^\mu = \Lambda^\mu_\nu x^\nu + a^\nu \quad (\Lambda^T G \Lambda = G) . \quad (4.9)$$

Abbiamo incontrato vari esempi di scalari: l'intervallo s^2 tra due eventi; l'intervallo di tempo proprio $\Delta\tau$ lungo un tratto di linea universo Γ_{AB} assegnata; il quadrato di un quadri-vettore ed il prodotto scalare di due quadri-vettori.

4.2.2 Quadrivettori

Abbiamo già introdotto i quadri-vettori dello spazio-tempo come elementi dello spazio vettoriale \mathcal{V} . Qui ne diamo una definizione alternativa — che si può applicare a grandezze di qualsiasi tipo — in termini delle componenti: un (quadri-)vettore è un ente caratterizzato da 4 componenti

$$V^\mu \quad (\mu = 0, 1, 2, 3) \quad (4.10)$$

dette, in questo caso, **controvarianti**, le quali, rispetto alle trasformazioni del gruppo di Poincaré del SDR, si comportano nel modo seguente:

$$\text{per traslazioni:} \quad V'^\mu = V^\mu \quad (4.11a)$$

$$\text{per trasformazioni di Lorentz:} \quad V'^\mu = \Lambda^\mu_\nu V^\nu . \quad (4.11b)$$

Si noti che le coordinate x^μ non formano un quadri-vettore, in quanto esse vengono modificate dalle traslazioni secondo l'eq. (4.9). Un vero quadrivettore è invece il differenziale delle coordinate dx^μ , che trasforma con la legge

$$dx'^\mu = \Lambda^\mu{}_\nu dx^\nu . \quad (4.12)$$

Analogamente le differenze di coordinate $(x_A^\mu - x_B^\mu)$ formano un quadri-vettore. Con un innocuo abuso di notazione indicheremo sempre con $V^\mu = (V^0, V^1, V^2, V^3) = (V^0, \mathbf{V})$ il quadri-vettore completo.

Definiamo ora un secondo tipo di componenti per i quadri-vettori. Partendo dall'espressione del prodotto scalare, che definiamo dall'analogia espressione per i vettori dello spazio-tempo (3.20)

$$(\underline{U}, \underline{V}) = U^\mu g_{\mu\nu} V^\nu , \quad (4.13)$$

definiamo le componenti **covarianti** del vettore \underline{V} tramite la relazione

$$V_\mu \equiv g_{\mu\nu} V^\nu \quad (4.14)$$

in modo che il prodotto scalare si possa scrivere sommando le componenti controvarianti di un vettore con le componenti covarianti dell'altro:

$$(\underline{U}, \underline{V}) = U^\mu V_\mu . \quad (4.15)$$

L'operazione (4.14) che tramite il tensore metrico fa passare dalle componenti controvarianti a quelle covarianti viene anche detta *abbassamento degli indici*. Dalla forma esplicita del tensore metrico (3.22) è facile vedere che nell'abbassamento degli indici la componente temporale resta immutata, mentre le componenti spaziali cambiano segno:

$$V_\mu = (V_0, V_1, V_2, V_3) = (V^0, -V^1, -V^2, -V^3) = (V^0, -\mathbf{V}) . \quad (4.16)$$

È quindi importante ricordare che le componenti covarianti spaziali hanno segno opposto rispetto alle usuali componenti dei tri-vettori.

Nel prodotto scalare, non importa a quale dei due vettori venga abbassato l'indice: il tensore metrico può essere fatto agire indifferentemente sul primo o sul secondo vettore:

$$(\underline{U}, \underline{V}) = U^\mu V_\mu = U_\mu V^\mu . \quad (4.17)$$

Può forse essere utile vedere esplicitate le precedenti espressioni rispetto alle componenti temporali e spaziali:

$$\begin{aligned} (\underline{U}, \underline{V}) &= U^0 V_0 + U^1 V_1 + U^2 V_2 + U^3 V_3 \\ &= U_0 V^0 + U_1 V^1 + U_2 V^2 + U_3 V^3 \\ &= U^0 V^0 - U^1 V^1 - U^2 V^2 - U^3 V^3 = U^0 V^0 - \mathbf{U} \cdot \mathbf{V} . \end{aligned} \quad (4.18)$$

Il quadrato di un quadri-vettore si può calcolare moltiplicando le componenti covarianti con quelle controvarianti, per esempio

$$ds^2 = (dx, dx) = dx^\mu dx_\mu = dx'^\mu dx'_\mu . \quad (4.19)$$

Cerchiamo adesso di determinare la legge di trasformazione delle componenti covarianti di un quadrivettore a seguito di una trasformazione di Poincaré. In notazione matriciale, conviene raccogliere le componenti covarianti in vettori riga

$$\tilde{V} \equiv (V_0, V_1, V_2, V_3) = V^T G , \quad (4.20)$$

in modo che il prodotto scalare si scriva semplicemente come prodotto matriciale riga per colonna

$$(\underline{U}, \underline{V}) = \tilde{U} V = \tilde{V} U . \quad (4.21)$$

Dalla legge di trasformazione dei vettori controvarianti $x' = \Lambda x$ e dalla condizione (3.42) sulle matrici di Lorentz $\Lambda^T G \Lambda = G$ ricaviamo

$$\tilde{V}' = V'^T G = (\Lambda V)^T G = V^T \Lambda^T G (\Lambda \Lambda^{-1}) = V^T G \Lambda^{-1} = \tilde{V} \Lambda^{-1} \quad (4.22)$$

ossia, in componenti,

$$V'_\mu = V_\nu (\Lambda^{-1})^\nu{}_\mu . \quad (4.23)$$

Quindi, le componenti covarianti trasformano con la matrice di Lorentz inversa, esattamente come i vettori di base (3.31), da cui il nome.

Da un punto di vista più matematico, possiamo interpretare le componenti covarianti di un vettore \underline{V} come segue. Il prodotto scalare induce un isomorfismo tra lo spazio vettoriale \mathcal{V} ed il suo duale \mathcal{V}^* costituito dai funzionali lineari $\mathcal{V} \rightarrow \mathbb{R}$. Precisamente, $g : \mathcal{V} \rightarrow \mathcal{V}^*$ associa al vettore $\underline{V} \in \mathcal{V}$ il funzionale lineare “prodotto scalare per \underline{V} ” che indichiamo con $f_{\underline{V}} \in \mathcal{V}^*$ e tale che

$$f_{\underline{V}}(\underline{U}) = (\underline{U}, \underline{V}) \quad (4.24)$$

Quando introduciamo una base $\{\underline{e}_\mu\}$ in \mathcal{V} , è univocamente determinata la *base duale* $\{e^*_\mu : \mu = 0, 1, 2, 3\}$ in \mathcal{V}^* , tale che $e^*_\mu(\underline{e}_\nu) = \delta^\mu{}_\nu$. Allora è facile verificare che le componenti di $f_{\underline{V}}$ nella base duale sono proprio le componenti covarianti di \underline{V} : $f_{\underline{V}} = V_\mu e^*_\mu$ (somma su μ sottointesa).

Quanto detto sopra si può esprimere in termini del seguente diagramma commutativo:

$$\begin{array}{ccc} \mathcal{V} & \xrightarrow{\Lambda} & \mathcal{V} \\ \downarrow g & & \downarrow g \\ \mathcal{V}^* & \xrightarrow{\Lambda} & \mathcal{V}^* \end{array} \quad , \quad \begin{array}{ccc} V & \xrightarrow{\Lambda} & V' \\ \downarrow g & & \downarrow g \\ \tilde{V} & \xrightarrow{\Lambda} & \tilde{V}' \end{array} \quad \Longleftrightarrow \quad \begin{array}{ccc} V^\nu & \xrightarrow{\Lambda^\mu{}_\nu} & V'^\mu \\ \downarrow g_{\nu\alpha} & & \downarrow g_{\mu\beta} \\ V_\alpha & \xrightarrow{(\Lambda^{-1})^\alpha{}_\beta} & V'_\beta \end{array} \quad (4.25)$$

Se il passaggio dalle componenti controvarianti a quelle covarianti si effettua con il tensore metrico, il passaggio dalle componenti covarianti a quelle controvarianti, detto *innalzamento degli indici*, si effettua con l'inverso del tensore metrico. Definiamo il tensore $g^{\mu\nu}$ di componenti

$$||g^{\mu\nu}|| \equiv G^{-1} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \Rightarrow \quad g^{\mu\nu} g_{\nu\alpha} = (G^{-1}G)^\mu{}_\alpha = \delta^\mu{}_\alpha . \quad (4.26)$$

Allora

$$V^\mu = \delta^\mu{}_\alpha V^\alpha = g^{\mu\nu} g_{\nu\alpha} V^\alpha = g^{\mu\nu} V_\nu . \quad (4.27)$$

Come matrici, il tensore metrico G ed il suo inverso G^{-1} coincidono:

$$G^{-1} = G \quad \Longleftrightarrow \quad g^{\mu\nu} = g_{\mu\nu} . \quad (4.28)$$

Questa coincidenza è dovuta al fatto di aver scelto una base (pseudo-)ortonormale: in una base generica il tensore metrico ed il suo inverso hanno componenti differenti. Anche nell'innalzamento degli indici, la componente temporale resta immutata, mentre le componenti spaziali cambiano segno:

La distinzione tra (quadri-)vettori covarianti e controvarianti, che è assente nel caso euclideo tridimensionale, è chiaramente connessa con la metrica pseudo-euclidea dello spazio-tempo.

Con il tensore metrico inverso si può calcolare l'inversa di una matrice di Lorentz: dall'eq. (3.50) risulta infatti

$$(\Lambda^{-1})^\mu{}_\nu = g^{\mu\alpha} \Lambda^\beta{}_\alpha g_{\beta\nu} \quad \Longleftrightarrow \quad \Lambda^{-1} = G \Lambda^T G . \quad (4.29)$$

A parole, gli elementi di matrice di Λ^{-1} si ottengono trasponendo Λ e cambiando di segno gli elementi con un'indice temporale ed uno spaziale, ossia gli elementi $\Lambda^0{}_i$ e $\Lambda^i{}_0$. Ad esempio,

$$\text{spinta:} \quad \begin{pmatrix} \cosh \eta & -\sinh \eta & 0 & 0 \\ -\sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \cosh \eta & \sinh \eta & 0 & 0 \\ \sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4.30)$$

$$\text{rotazione:} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \quad (4.31)$$

4.2.3 Tensori doppi

In modo analogo si possono definire i quadri-tensori di ordine più elevato, con indici controvarianti, covarianti e misti. I tensori doppi, o anche di ordine 2, sono definiti come enti a 16 componenti che, per trasformazioni di Poincaré si trasformano come segue:

$$\text{tensori doppi controvarianti: } T'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta T^{\alpha\beta} \iff T' = \Lambda T \Lambda^T \in \mathcal{V} \otimes \mathcal{V} \quad (4.32)$$

$$\text{tensori doppi covarianti: } T'_{\mu\nu} = T_{\alpha\beta} (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu \iff T' = (\Lambda^{-1})^T T \Lambda^{-1} \in \mathcal{V}^* \otimes \mathcal{V}^* \quad (4.33)$$

$$\text{tensori doppi misti: } T'^\mu{}_\nu = \Lambda^\mu_\alpha T^\alpha{}_\beta (\Lambda^{-1})^\beta_\nu \iff T' = \Lambda T \Lambda^{-1} \in \mathcal{V} \otimes \mathcal{V}^* . \quad (4.34)$$

Semplici esempi di tensori doppi sono ottenuti mediante le componenti di due quadri-vettori u e v : $U^\mu V^\nu$, $U_\mu V_\nu$, $U^\mu V_\nu$ sono tensori doppi.

Il più importante esempio di tensore doppio finora incontrato è il tensore metrico. Vediamo subito che il tensore metrico gode della proprietà di trasformazione (4.33). Infatti, dalla proprietà del tensore metrico di essere uguale in ogni SDR (cfr. sez. 3.2.3), dobbiamo verificare che

$$G = G' = (\Lambda^{-1})^T G \Lambda^{-1} . \quad (4.35)$$

Ma questa equazione non è altro che la condizione sulle matrici di Lorentz (3.42) moltiplicata a destra e sinistra per le matrici di Lorentz inverse (ricordare che trasposizione ed inversione commutano).

Ancora più immediata è la verifica che il tensore metrico inverso è un tensore nel senso dell'eq. (4.32): infatti, è evidente che l'eq. (3.42) coincide con la legge di trasformazione (4.32) con $G^{-1} = G$ al posto di T .

Il tensore metrico $g_{\mu\nu}$ ed il suo inverso $g^{\mu\nu}$ sono esempi di tensori *costanti*, cioè che non cambiano per trasformazioni del SDR. Il fatto di denotarli con la stessa lettera g è consistente con la proprietà di innalzamento ed abbassamento degli indici, per esempio

$$g^{\mu\nu} = g^{\mu\alpha} g_{\alpha\beta} g^{\beta\nu} . \quad (4.36)$$

Da notare inoltre che il tensore metrico misto coincide con l'identità: $g^\mu{}_\nu = \delta^\mu{}_\nu$.

La notazione fin qui sviluppata per i tensori tratta allo stesso modo indici covarianti e controvarianti, i quali possono essere alzati ed abbassati tramite il tensore metrico. La consistenza di tutta la procedura sta nel fatto che abbassare o alzare uno o più indici di un tensore genera un altro tensore. Ad esempio, se $T = ||T^{\alpha\beta}||$ è un tensore doppio controvariante, anche l'ente \mathcal{T} di componenti

$$\mathcal{T}^\alpha{}_\gamma \equiv T^{\alpha\beta} g_{\beta\gamma} \quad (4.37)$$

è un tensore (misto). Per provarlo, bisogna mostrare che

$$\mathcal{T}'^\mu{}_\rho \equiv T'^{\mu\nu} g_{\nu\rho} \quad (4.38)$$

si ottiene da $\mathcal{T}^\alpha{}_\gamma$ mediante l'eq. (4.34). Si ha infatti

$$\begin{aligned} \mathcal{T}'^\mu{}_\rho &\stackrel{(4.38)}{=} T'^{\mu\nu} g_{\nu\rho} \stackrel{(4.32)}{=} \Lambda^\mu_\alpha \Lambda^\nu_\beta T^{\alpha\beta} g_{\nu\rho} \stackrel{(4.37)}{=} \Lambda^\mu_\alpha \Lambda^\nu_\beta \mathcal{T}^\alpha{}_\gamma g^{\gamma\beta} g_{\nu\rho} \\ &= \Lambda^\mu_\alpha \mathcal{T}^\alpha{}_\gamma (g^{\gamma\beta} \Lambda^\nu_\beta g_{\nu\rho}) \stackrel{(4.29)}{=} \Lambda^\mu_\alpha \mathcal{T}^\alpha{}_\gamma (\Lambda^{-1})^\gamma{}_\rho \end{aligned} \quad (4.39)$$

che corrisponde alla trasformazione (4.34) come si voleva. Dal momento che le componenti $\mathcal{T}^\alpha{}_\gamma$ definiscono un tensore che deriva direttamente dal tensore $T^{\alpha\beta}$ ma è di tipo diverso, scriviamo semplicemente i due tensori con lo stesso simbolo:

$$T^\alpha{}_\gamma \equiv T^{\alpha\beta} g_{\beta\gamma} , \quad (4.40)$$

cioè anche per tensori di ordine 2 o superiori, il tensore metrico abbassa l'indice con cui è sommato. Allo stesso modo, il tensore metrico inverso innalza gli indici con cui è sommato. In sintesi, l'abbassamento e l'innalzamento degli indici “commuta” con le trasformazioni di Lorentz, come indicato dal seguente diagramma commutativo:

$$\begin{array}{ccc} \mathcal{V} \otimes \mathcal{V} & \xrightarrow{\Lambda} & \mathcal{V} \otimes \mathcal{V} \\ \downarrow g & & \downarrow g \\ \mathcal{V} \otimes \mathcal{V}^* & \xrightarrow{\Lambda} & \mathcal{V} \otimes \mathcal{V}^* \end{array} , \quad \begin{array}{ccc} T & \xrightarrow{\Lambda} & T' \\ \downarrow g & & \downarrow g \\ \mathcal{T} & \xrightarrow{\Lambda} & \mathcal{T}' \end{array} \iff \begin{array}{ccc} T^{\alpha\beta} & \xrightarrow{\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta} & T'^{\mu\nu} \\ \downarrow g_{\beta\gamma} & & \downarrow g_{\nu\rho} \\ \mathcal{T}^\alpha{}_\gamma & \xrightarrow{\Lambda^\mu{}_\alpha (\Lambda^{-1})^\gamma{}_\rho} & \mathcal{T}'^\mu{}_\rho \end{array} \quad (4.41)$$

4.2.4 Pseudo-tensori

In alcuni casi si presentano degli enti che per trasformazioni di Poincaré *proprie*, cioè con $\Lambda \in \mathcal{L}_+^\uparrow$, trasformano come tensori, mentre per trasformazioni non proprie, come la parità o l'inversione temporale, acquistano un ulteriore fattore moltiplicativo. Questi enti vengono definiti *pseudo-tensori*, e si distinguono in diverse specie a seconda del fattore moltiplicativo che coinvolgono:

- $\det(\Lambda)$ per i pseudo-tensori di I specie;
- $\text{sign}(\Lambda^0_0)$ per i pseudo-tensori di II specie;
- $\det(\Lambda) \text{sign}(\Lambda^0_0)$ per i pseudo-tensori di III specie.

Ad esempio, per gli scalari abbiamo

$$s' = s \cdot \det(\Lambda) \quad \text{pseudo-scalare di I specie ,} \quad (4.42)$$

$$s' = s \cdot \text{sign}(\Lambda^0_0) \quad \text{pseudo-scalare di II specie .} \quad (4.43)$$

Esempi di pseudo-tensori 3-dimensionali² sono i vettori assiali, ottenuti per esempio tramite il prodotto vettoriale di due vettori ordinari, come il momento angolare $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. Anche il campo magnetico \mathbf{B} è uno pseudo-vettore rispetto al gruppo euclideo.

4.2.5 Tensore di Levi-Civita

Un importante tensore di ordine 4 nello spazio quadri-dimensionale è quello completamente antisimmetrico, definito dalle seguenti relazioni:

$$\varepsilon^{\alpha\beta\mu\nu} \equiv \begin{cases} +1 & \text{se } (\alpha, \beta, \mu, \nu) \text{ è una permutazione pari di } (0, 1, 2, 3) , \\ -1 & \text{se } (\alpha, \beta, \mu, \nu) \text{ è una permutazione dispari di } (0, 1, 2, 3) , \\ 0 & \text{se } (\alpha, \beta, \mu, \nu) \text{ non è una permutazione di } (0, 1, 2, 3) . \end{cases} \quad (4.44)$$

In particolare, $\varepsilon^{0123} = +1$, lo scambio di una qualsiasi coppia di indici cambia il segno dell'elemento, e se due (o più) indici sono uguali allora l'elemento vale 0.

Per dimostrare che $\varepsilon^{\alpha\beta\mu\nu}$ è un tensore, ricordiamo la definizione di determinante di una matrice come somma dei prodotti delle quaterne di elementi, presi uno per ogni riga, in cui ciascun prodotto è moltiplicato per il segno della permutazione:

$$\det \Lambda \equiv \sum_{\alpha, \beta, \mu, \nu} \Lambda^0_\alpha \Lambda^1_\beta \Lambda^2_\mu \Lambda^3_\nu \varepsilon^{\alpha\beta\mu\nu} . \quad (4.45)$$

Sappiamo che scambiare due righe inverte il segno del determinante, mentre se due righe sono uguali il determinante è nullo. Si ha quindi, operando una permutazione sugli indici riga $(0, 1, 2, 3) \rightarrow (\alpha', \beta', \mu', \nu')$

$$\Lambda^{\alpha'}_\alpha \Lambda^{\beta'}_\beta \Lambda^{\mu'}_\mu \Lambda^{\nu'}_\nu \varepsilon^{\alpha\beta\mu\nu} = \det \Lambda \varepsilon^{\alpha'\beta'\mu'\nu'} . \quad (4.46)$$

Siccome per trasformazioni proprie di Lorentz $\det \Lambda = 1$, la precedente equazione ci garantisce che $\varepsilon^{\alpha\beta\mu\nu}$ trasforma come un tensore controvariante di ordine 4, ed è chiamato **tensore di Levi-Civita** o anche tensore totalmente antisimmetrico. È evidente che esso è un tensore costante. Siccome per trasformazioni di parità ed inversione temporale $\det \Lambda = -1$, l'equazione (4.46) implica che il tensore di Levi-Civita è in realtà uno pseudo-tensore di I specie.

Occorre infine fare attenzione alla versione con indici covarianti, che ha segno opposto rispetto al tensore covariante

$$\varepsilon_{\alpha\beta\mu\nu} = -\varepsilon^{\alpha\beta\mu\nu} , \quad (4.47)$$

come si può vedere nel caso di indici 0,1,2,3.

²Nel caso 3-dimensionale, non c'è l'analogo di Λ^0_0 , mentre ha senso considerare il determinante delle matrici ortogonali, che può assumere i due valori ± 1 . In questo contesto gli pseudo-tensori possono essere solo di prima specie.

4.3 Algebra tensoriale

4.3.1 Prodotto scalare

Siano U^μ le componenti controvarianti di un quadri-vettore, V_μ le componenti covarianti di un altro quadri-vettore. Sappiamo già che, nel caso di vettori dello spazio-tempo, il prodotto scalare $U^\mu V_\mu$ è un quadri-scalare, ossia è un invariante per trasformazioni di Poincaré. Questo vale in realtà per ogni coppia di vettori definiti tramite la loro legge di trasformazione (4.11) e (4.23). Infatti

$$V'_\mu U'^\mu = V_\nu (\Lambda^{-1})^\nu{}_\mu \Lambda^\mu{}_\alpha U^\alpha = V_\nu \delta^\nu{}_\alpha U^\alpha = V_\nu U^\nu . \quad (4.48)$$

In particolare, il *quadrato* di U definito da $U^\mu U_\mu$ è un quadri-scalare. In analogia con quanto si è detto per i vettori dello spazio-tempo, diremo che U^μ (ovvero U_μ) è un vettore di tipo tempo, spazio o luce a seconda che $U^\mu U_\mu$ sia maggiore, minore o uguale a zero.

4.3.2 Contrazione tensoriale

Se $T^{\alpha\beta}{}_{\mu\nu}$ sono le componenti di un tensore misto, allora anche le quantità ottenute sommando un indice controvariante con uno covariante (per esempio β e μ)

$$C^\alpha{}_\nu \equiv \sum_\beta T^{\alpha\beta}{}_{\beta\nu} = T^{\alpha 0}{}_{0\nu} + T^{\alpha 1}{}_{1\nu} + T^{\alpha 2}{}_{2\nu} + T^{\alpha 3}{}_{3\nu} \quad (4.49)$$

sono le componenti di un tensore, di ordine minore di 2 unità rispetto a T .

Infatti, nell'esempio precedente,

$$C'^\gamma{}_\sigma \equiv T'^{\gamma\rho}{}_{\rho\sigma} = \Lambda^\gamma{}_\alpha \Lambda^\rho{}_\beta T^{\alpha\beta}{}_{\mu\nu} (\Lambda^{-1})^\mu{}_\rho (\Lambda^{-1})^\nu{}_\sigma . \quad (4.50)$$

La somma in T' sull'indice ρ si ripercuote su una somma tra le matrici di Lorentz

$$(\Lambda^{-1})^\mu{}_\rho \Lambda^\rho{}_\beta = \delta^\mu{}_\beta \quad (4.51)$$

che quindi produce una somma in T tra gli indici β e μ . Otteniamo così

$$C'^\gamma{}_\sigma = \Lambda^\gamma{}_\alpha \delta^\mu{}_\beta T^{\alpha\beta}{}_{\mu\nu} (\Lambda^{-1})^\nu{}_\sigma = \Lambda^\gamma{}_\alpha T^{\alpha\beta}{}_{\beta\nu} (\Lambda^{-1})^\nu{}_\sigma = \Lambda^\gamma{}_\alpha C^\alpha{}_\nu (\Lambda^{-1})^\nu{}_\sigma \quad (4.52)$$

che stabilisce appunto il carattere tensoriale di $C^\alpha{}_\nu$.

Questa operazione di somma di due indici di tipo diverso nello stesso tensore si chiama *contrazione*.

Esempio: La contrazione del tensore metrico misto dà sempre la dimensione dello spazio vettoriale dei vettori su cui opera, in qualsiasi sistema di coordinate su qualsiasi varietà, affine o curva, euclidea o pseudo-euclidea che sia:

$$g^\mu{}_\nu = \delta^\mu{}_\nu \quad \implies \quad g^\mu{}_\mu = \delta^\mu{}_\mu = \dim \mathcal{V} . \quad (4.53)$$

Osservazione: Le componenti del tensore metrico *misto* $g^\mu{}_\nu$ sono sempre (cioè in qualunque base) uguali al simbolo di Kronecker $\delta^\mu{}_\nu$, essendo per definizione $g^\mu{}_\nu \equiv g^{\mu\alpha} g_{\alpha\nu} = (G^{-1}G)^\mu{}_\nu = \mathbb{1}^\mu{}_\nu = \delta^\mu{}_\nu$. Bisogna però fare attenzione ad alzare o abbassare un indice di $\delta^\mu{}_\nu$: è sbagliato interpretare il risultato ancora come una delta di Kronecker: $g_{\alpha\mu} \delta^\mu{}_\nu = g_{\alpha\nu} \neq \delta_{\alpha\mu}$.

Esempio: la contrazione di due tensori di Levi-Civita su 2 coppie di indici dà un tensore di ordine 4:

$$\varepsilon^{\alpha\beta\mu\nu} \varepsilon_{\alpha\beta\rho\sigma} = -2(\delta^\mu{}_\rho \delta^\nu{}_\sigma - \delta^\mu{}_\sigma \delta^\nu{}_\rho) = -2(g^\mu{}_\rho g^\nu{}_\sigma - g^\mu{}_\sigma g^\nu{}_\rho) . \quad (4.54)$$

la contrazione di due tensori di Levi-Civita su 3 coppie di indici dà un tensore di ordine 2:

$$\varepsilon^{\alpha\beta\mu\nu} \varepsilon_{\alpha\beta\mu\sigma} = -6\delta^\nu{}_\sigma = -6g^\nu{}_\sigma . \quad (4.55)$$

la contrazione di due tensori di Levi-Civita su tutte e 4 le coppie di indici dà uno scalare

$$\varepsilon^{\alpha\beta\mu\nu}\varepsilon_{\alpha\beta\mu\nu} = -24. \quad (4.56)$$

Tenendo in conto dell'eq. (4.47), il primo membro della contrazione (4.56) si può interpretare come -1 moltiplicato la somma dei quadrati di tutte le componenti del tensore. Siccome nel tensore di Levi-Civita le componenti possono assumere solo i valori $-1, 0, 1$, ne deduciamo che il numero di componenti non nulle di $\varepsilon^{\alpha\beta\mu\nu}$ è 24 (su 256).

4.3.3 Prodotto diretto

Dati due vettori U^μ e V^ν , il prodotto delle componenti produce un nuovo tensore $T^{\mu\nu} \equiv U^\mu V^\nu$. Infatti

$$T'^{\mu\nu} \equiv U'^\mu V'^\nu \stackrel{(4.11)}{=} \Lambda^\mu{}_\alpha U^\alpha \Lambda^\nu{}_\beta V^\beta = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta U^\alpha V^\beta = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta T^{\alpha\beta}, \quad (4.57)$$

in accordo con la legge di trasformazione (4.32).

Più in generale, dati due tensori T e W , il prodotto delle componenti dei due tensori dà un nuovo tensore i cui indici sono dati da tutti gli indici dei tensori originari. Per esempio, se $T^\mu{}_\nu$ e $W_{\alpha\beta}$ sono tensori, anche

$$Z^\mu{}_{\nu\alpha\beta} \equiv T^\mu{}_\nu W_{\alpha\beta} \quad (4.58)$$

è un tensore, chiamato *prodotto diretto* dei tensori T e W . Che il nuovo oggetto sia un tensore si verifica facilmente osservando che ogni indice si porta dietro la sua matrice di trasformazione: Λ per ogni indice controvariante, Λ^{-1} per ogni indice covariante.

4.3.4 Leggi di trasformazione dei campi

Così come in meccanica classica, i vettori possono essere associati ad un ben preciso oggetto (la velocità di un corpo), oppure possono descrivere quantità vettoriali che si estendono nello spazio, come il campo elettrico. Anche in relatività dovremo considerare vettori — e più in generale tensori — che si possono associare ad un preciso oggetto (una particella) oppure a campi definiti nello spazio-tempo.

Definizione: un **campo tensoriale** T è una funzione che ad ogni punto (evento) A dello spaziotempo \mathcal{M} associa un tensore $T(A)$.

Avendo definito i tensori come oggetti con particolari proprietà di trasformazione per cambio di coordinate, conviene definire i campi tensoriali mediante le componenti rispetto ad un SDR S .

Campo scalare Il campo più semplice è il campo scalare $\Phi : \mathcal{M} \rightarrow \mathbb{R}$ che al punto A associa il numero reale $\Phi(A)$. Per definizione di scalare, il valore di $\Phi(A)$ è indipendente dall'osservatore, e quindi dalla scelta del SDR.

Se consideriamo due osservatori con SDR S ed S' , l'osservatore S potrà pensare il campo scalare come una funzione $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}$ che associa alla quaterna di coordinate x^μ il valore $\phi(x) = \Phi(A)$, mentre l'osservatore S' potrà pensare lo stesso campo come una funzione $\phi' : \mathbb{R}^4 \rightarrow \mathbb{R}$ che associa alla quaterna di coordinate x'^ν il valore $\phi'(x') = \Phi(A)$. Chiaramente x^μ e x'^ν sono le coordinate dello *stesso evento* A nei due SDR S ed S' rispettivamente, legate dalla trasformazione di Poincaré

$$x'^\nu = \Lambda^\nu{}_\mu x^\mu + a^\nu. \quad (4.59)$$

Trattandosi di uno scalare, dovremo imporre che sia

$$\phi'(x') = \phi(x), \quad (4.60)$$

ed essendo $x = \Lambda^{-1}(x' - a)$, avremo

$$\phi'(x') = \phi(\Lambda^{-1}(x' - a)). \quad (4.61)$$

Campo vettoriale Un campo vettoriale è una funzione $V : \mathcal{M} \rightarrow \mathcal{V}$ che ad ogni punto A associa il vettore $V(A) \in \mathcal{V}$ (qui \mathcal{V} è un generico spazio vettoriale quadri-dimensionale). Dal momento che le componenti di un vettore come $V(A)$ dipendono dal SDR, i due osservatori S ed S' assegneranno a $V(A)$ delle componenti diverse, rispettivamente $V'^\mu(x')$ e $V^\nu(x)$, legate dalla relazione

$$V'^\nu(x') = \Lambda^\nu_\mu V^\mu(x) , \quad (4.62)$$

come richiede la condizione (4.11) sui vettori controvarianti. In definitiva, il campo V è un campo vettoriale se, in seguito alla trasformazioni di coordinate (4.59), si ha

$$V'^\nu(x') = \Lambda^\nu_\mu V^\mu(\Lambda^{-1}(x' - a)) . \quad (4.63)$$

Un campo vettoriale di componenti controvarianti deve invece soddisfare

$$V'_\nu(x') = V_\mu(x)(\Lambda^{-1})^\mu_\nu . \quad (4.64)$$

Campo tensoriale In modo analogo si definiscono i campi tensoriali di ordine più elevato. Per esempio, il campo tensoriale misto trasforma con la seguente legge

$$T'^\mu_\nu(x') = \Lambda^\mu_\alpha (\Lambda^{-1})^\beta_\nu T^\alpha_\beta(\Lambda^{-1}(x' - a)) . \quad (4.65)$$

Quadri-gradiente Dato un campo scalare $\phi(x)$, possiamo considerarne il *quadri-gradiente*

$$\partial_\mu \phi(x) \equiv \frac{\partial \phi(x)}{\partial x^\mu} . \quad (4.66)$$

Vogliamo dimostrare che il quadri-gradiente si trasforma come un campo vettoriale controvariante. In corrispondenza alla trasformazione di coordinate (4.59), abbiamo stabilito che deve valere la legge di trasformazione (4.61). Derivando quest'ultima rispetto alla variabile x'^ν otteniamo

$$\frac{\partial \phi'(x')}{\partial x'^\nu} \stackrel{(4.60)}{=} \frac{\partial \phi(x(x'))}{\partial x'^\nu} = \frac{\partial \phi(x)}{\partial x^\mu} \frac{\partial x^\mu}{\partial x'^\nu} \stackrel{(4.68)}{=} \partial_\mu \phi(x) (\Lambda^{-1})^\mu_\nu , \quad (4.67)$$

dove abbiamo usato la regola di differenziazione delle funzioni composte (nel nostro caso della funzione $\phi(x(x'))$) e la relazione

$$\frac{\partial x^\mu}{\partial x'^\nu} \stackrel{(4.59)}{=} \frac{\partial [(\Lambda^{-1})^\mu_\alpha (x'^\alpha - a^\alpha)]}{\partial x'^\nu} = (\Lambda^{-1})^\mu_\alpha \delta^\alpha_\nu = (\Lambda^{-1})^\mu_\nu . \quad (4.68)$$

Pertanto, il quadrigradiente

$$\partial'_\nu \phi'(x') = \partial_\mu \phi(x) (\Lambda^{-1})^\mu_\nu \quad (4.69)$$

trasforma come un campo vettoriale covariante (4.64), e questo giustifica la scrittura ∂_μ con l'indice in basso. In generale, possiamo scrivere

$$\partial'_\nu \equiv \frac{\partial}{\partial x'^\nu} = \frac{\partial x^\mu}{\partial x'^\nu} \frac{\partial}{\partial x^\mu} = (\Lambda^{-1})^\mu_\nu \partial_\mu \quad (4.70)$$

$$\partial'^\nu \equiv \frac{\partial}{\partial x'_\nu} = \frac{\partial x_\mu}{\partial x'_\nu} \frac{\partial}{\partial x_\mu} = \Lambda^\mu_\nu \partial^\mu \quad (4.71)$$

per operatori differenziali che operano su tensori di qualsiasi ordine, ed anche per derivate di ordine superiore.

In particolare, l'operatore differenziale del secondo ordine, chiamato *d'Alembertiano*

$$\square \equiv \partial_\mu \partial^\mu = (\partial_0)^2 - (\partial_1)^2 - (\partial_2)^2 - (\partial_3)^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \quad (4.72)$$

è un invariante. Questo significa che, se $\phi(x)$ è un campo scalare, $\square \phi(x)$ è un campo scalare; se $V^\mu(x)$ è un campo vettoriale, allora $\square V^\mu(x)$ è un campo vettoriale, e così via.

Chapter 5

Meccanica relativistica

5.1 Cinematica del punto materiale

Ci proponiamo adesso di affrontare il problema di modificare le leggi della meccanica Newtoniana per ottenere quelle della meccanica relativistica di un punto materiale, in forma covariante a vista.

Ricordiamo che il *punto materiale* nasce come un'idealizzazione che descrive in modo semplice un corpo di estensione finita, le cui dimensioni sono molto più piccole delle distanze che percorre. Più precisamente, un corpo può essere approssimato con un punto quando le sue dimensioni sono molto più piccole della sensibilità degli strumenti di misura di posizione, o comunque più piccole delle differenze di posizione che siamo interessati a studiare, per cui la sua struttura e conformazione interna possono essere trascurate. L'aggettivo “materiale” si riferisce poi al fatto che questo corpo non è unicamente caratterizzato dalla sua posizione nello spazio, ma possiede altre quantità che lo distinguono e che determinano il suo comportamento dinamico. Di queste quantità la più importante in meccanica è la *massa inerziale*, che quantifica la “resistenza” con cui il corpo si oppone a variazioni della sua velocità; la *carica elettrica* è un'altra quantità fondamentale nello studio dei fenomeni elettromagnetici.

Se in meccanica classica il punto materiale è una comoda astrazione per trattare piccoli corpi estesi, in relatività il punto materiale diventa un concetto fondamentale, sul quale basare tutta la trattazione della materia. In relatività, l'impossibilità di propagare segnali istantaneamente, ma solo con velocità finita minore o uguale a c , rende privo di senso il concetto di corpo rigido. Infatti, nel corpo rigido le distanze tra le sue parti possono essere costanti solo se una perturbazione di un suo punto si ripercuote istantaneamente sulle altre parti. Pertanto, in relatività i corpi estesi, anche di piccole dimensioni, devono essere considerati come un insieme di punti materiali, chiamati anche *particelle*.

Tuttavia, trattando problemi con corpi estesi dei quali non interessano deformazioni o movimenti interni, ma solamente la posizione o la velocità approssimata del corpo nel suo complesso, potremo anche in questo caso usare l'approssimazione di punto materiale per descrivere tale corpo.

Nel generalizzare i concetti classici di velocità \mathbf{v} e accelerazione \mathbf{a} , cercheremo di costruire degli enti u e w che soddisfino a queste due condizioni:

- per velocità $v \ll c$ molto piccole rispetto a quella della luce si devono ottenere le note espressioni non-relativistiche per le componenti spaziali

$$\begin{aligned} u^i &\rightarrow v^i \\ w^i &\rightarrow a^i \end{aligned} \quad (i = 1, 2, 3) ; \quad (5.1)$$

- unendo alle tre componenti dei vettori relativistici una quarta componente u^0 , w^0 , si possano formare dei quadrivettori u^μ e w^μ .

5.1.1 Quadri-velocità

Consideriamo una particella il cui moto in un SDR S sia descritto dalla linea universo

$$x^\mu = x^\mu(\lambda) \quad \text{cioè} \quad \begin{cases} ct = ct(\lambda) \\ \mathbf{x} = \mathbf{x}(\lambda) \end{cases}, \quad (5.2)$$

in cui λ è un parametro reale che parametrizza la linea universo della particella. Naturalmente si potrebbe prendere t come parametro, ma come vedremo tra poco, la scelta più conveniente è un'altra.

In fisica non-relativistica, la velocità si definisce come $\mathbf{v} \equiv d\mathbf{x}/dt$. Questa quantità è un 3-vettore poiché le differenze di coordinate $d\mathbf{x}$ sono 3-vettori, mentre le differenze di tempo sono quantità assolute, cioè 3-scalari.

In fisica relativistica, le differenze di coordinate

$$dx^\mu = x^\mu(\lambda + d\lambda) - x^\mu(\lambda) \quad (5.3)$$

sono 4-vettori, mentre le differenze di tempo non sono 4-scalari (sono proporzionali alla componente dx^0 di un 4-vettore), e quindi la grandezza dx^μ/dt non definisce un 4-vettore. Ricordiamo invece che l'intervallo $ds^2 = dx^\mu dx_\mu$ è un invariante. Se consideriamo particelle che si muovono a velocità inferiore a quella della luce (in un certo SDR e quindi in tutti gli altri), allora il vettore dx^μ è sempre di tipo tempo ($ds^2 > 0$) e quindi, è ben definito il tempo proprio (2.56)

$$d\tau = \frac{ds}{c} \equiv \frac{1}{c} \sqrt{ds^2} = \frac{1}{c} \sqrt{c^2(dt)^2 - (d\mathbf{x})^2} = dt \sqrt{1 - \frac{v^2}{c^2}}, \quad (5.4)$$

anch'esso invariante di Lorentz. In particolare, nel SDR in cui la particella è a riposo, $d\tau = dt$.

Allora l'ente di componenti $dx^\mu/d\tau$ dovrà essere un 4-vettore, in quanto rapporto tra un 4-vettore ed un invariante. Definiamo quindi il 4-vettore velocità, detto anche **quadri-velocità**

$$u^\mu(\tau) \equiv \frac{dx^\mu(\tau)}{d\tau}, \quad (5.5)$$

che ovviamente dipende dal punto della traiettoria in cui viene calcolato, ossia dal parametro della traiettoria (τ nel nostro caso). Verifichiamo che, per piccole velocità $v \ll c$, le componenti spaziali della quadri-velocità diventano le componenti della velocità ordinaria: osservando che $dt = \gamma d\tau$ si ha

$$\mathbf{u} = \frac{d\mathbf{x}}{d\tau} = \frac{d\mathbf{x}}{dt \sqrt{1 - \frac{v^2}{c^2}}} \xrightarrow{\frac{v}{c} \rightarrow 0} \frac{d\mathbf{x}}{dt} = \mathbf{v} \quad (5.6a)$$

$$u^0 = \frac{dx^0}{d\tau} = \frac{c dt}{dt \sqrt{1 - \frac{v^2}{c^2}}} \xrightarrow{\frac{v}{c} \rightarrow 0} c. \quad (5.6b)$$

Possiamo scrivere la quadri-velocità in termini della velocità ordinaria:

$$u^\mu = (u^0, \mathbf{u}) = \gamma(c, \mathbf{v}), \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (5.7)$$

come si vede anche direttamente dalle eq. (5.6).

In pratica, la quadri-velocità è un 4-vettore che, istante per istante, è tangente alla linea universo della particella parametrizzata con il tempo proprio. Le quattro componenti della quadri-velocità non sono indipendenti, dipendendo unicamente dalle tre componenti di \mathbf{v} . Vale infatti

$$u^2 = u^\mu u_\mu = \frac{dx^\mu}{d\tau} \frac{dx_\mu}{d\tau} = \frac{dx^\mu dx_\mu}{(d\tau)^2} = \frac{ds^2}{ds^2/c^2} = c^2 \quad (5.8)$$

ossia il modulo quadro della quadri-velocità è costante. Essendo inoltre strettamente positivo, abbiamo anche dimostrato che la quadri-velocità è sempre un vettore di tipo tempo.

Se due SDR inerziali S ed S' sono legati dalla trasformazione di coordinate

$$x'^\nu = \Lambda^\nu_\mu x^\mu + a^\nu, \quad (5.9)$$

la quadri-velocità misurata da S' è data semplicemente da

$$u'^\nu = \frac{d(\Lambda^\nu_\mu x^\mu(\tau) + a^\nu)}{d\tau} = \Lambda^\nu_\mu \frac{dx^\mu(\tau)}{d\tau} = \Lambda^\nu_\mu u^\mu, \quad (5.10)$$

essendo i coefficienti Λ^ν_μ ed a^ν della trasformazione indipendenti dalle coordinate e quindi indipendenti da τ .

5.1.2 Quadri-accelerazione

L'analogo relativistico del 3-vettore accelerazione $\mathbf{a} \equiv d\mathbf{v}/dt$ lo possiamo definire derivando la quadri-velocità rispetto al tempo proprio. Introduciamo quindi la **quadri-accelerazione**

$$w^\mu(\tau) \equiv \frac{du^\mu(\tau)}{d\tau} = \frac{d^2x^\mu(\tau)}{d\tau^2} \quad (5.11)$$

che è chiaramente un quadri-vettore, essendo il rapporto tra 4-vettore ed un invariante. Inoltre, nel limite di piccole velocità $v \ll c$

$$\mathbf{w} = \gamma \frac{d}{dt}(\gamma \mathbf{v}) = \frac{1}{1 - \frac{v^2}{c^2}} \left[\frac{\frac{1}{c^2} \mathbf{v} \cdot \frac{d\mathbf{v}}{dt}}{1 - \frac{v^2}{c^2}} \mathbf{v} + \frac{d\mathbf{v}}{dt} \right] \xrightarrow{\frac{v}{c} \rightarrow 0} \frac{d\mathbf{v}}{dt} = \mathbf{a} \quad (5.12)$$

le componenti spaziali riproducono l'accelerazione non-relativistica. Anche in questo caso le quattro componenti della quadri-accelerazione non sono indipendenti, ma devono soddisfare ad una precisa relazione. Infatti, siccome nella metrica pseudo-euclidea il “modulo” della quadri-velocità è costante, la variazione infinitesima du di quadri-velocità deve essere ortogonale ad u (esattamente come succede nel caso euclideo):

$$0 = \frac{d(u^\mu u_\mu)}{d\tau} = w^\mu u_\mu + u^\mu w_\mu = 2w^\mu u_\mu \quad (5.13)$$

e quindi si trova

$$(\underline{w}, \underline{u}) = 0. \quad (5.14)$$

In altre parole, nella metrica pseudo-euclidea la quadri-accelerazione risulta sempre ortogonale alla quadri-velocità.

Per un SDR S' legato ad S tramite la trasformazione di coordinate (5.9), le componenti della quadri-accelerazione si ricavano dalla solita legge di trasformazione dei 4-vettori

$$w'^\nu = \Lambda^\nu_\mu w^\mu. \quad (5.15)$$

5.1.3 Quadri-impulso

Vogliamo ora generalizzare alla meccanica relativistica quello che in fisica classica era la *quantità di moto*, spesso chiamata anche *impulso*

$$\mathbf{p}_{nr} = m\mathbf{v}. \quad (5.16)$$

La proprietà più importante dell'impulso non-relativistico sta nel fatto che, per sistemi isolati, l'impulso totale del sistema si conserva, ossia è una costante del moto (terzo principio della dinamica). Questo principio si può enunciare in qualsiasi SDR inerziale: se l'impulso si conserva in un SDR inerziale, allora si conserva in qualsiasi SDR inerziale. Ciò è una diretta conseguenza di due leggi (e mezzo):

- legge di conservazione della massa (la massa è un invariante di Galileo);
- legge di addizione delle velocità.

Consideriamo infatti un sistema isolato che ad un certo momento (iniziale) sia formato da un dato numero di corpi di massa m_i i quali interagiscono tra loro, potendosi anche rompere o fondere. Ad un istante successivo

(finale) il sistema potrà essere composto da un diverso numero di corpi di massa m_f . Ipotizziamo che nel SDR S l'impulso totale e la massa totale si conservino:

$$\mathbf{P}_{(\text{in})} \equiv \sum_i m_i \mathbf{v}_i = \mathbf{P}_{(\text{fi})} \equiv \sum_f m_f \mathbf{v}_f \quad (5.17)$$

$$M_{(\text{in})} \equiv \sum_i m_i = M_{(\text{fi})} \equiv \sum_f m_f . \quad (5.18)$$

Nel SDR S' che si muove a velocità \mathbf{v} rispetto ad S e con assi a lui paralleli, i corpi si muovono a velocità $\mathbf{v}'_i = \mathbf{v}_i - \mathbf{v}$ ecc. (addizione di velocità Galileiana). Si suppone inoltre che la massa non dipenda dal SDR, ossia dallo stato di moto del corpo:

$$\mathbf{v}'_i = \mathbf{v}_i - \mathbf{v} \quad m'_i = m_i \quad (5.19)$$

$$\mathbf{v}'_f = \mathbf{v}_f - \mathbf{v} \quad m'_f = m_f \quad (5.20)$$

S' misurerà quindi i seguenti impulsi:

$$\mathbf{P}'_{(\text{in})} \equiv \sum_i m'_i \mathbf{v}'_i = \sum_i m_i (\mathbf{v}_i - \mathbf{v}) = \mathbf{P}_{(\text{in})} - M_{(\text{in})} \mathbf{v} \quad (5.21)$$

$$\mathbf{P}'_{(\text{fi})} \equiv \sum_f m'_f \mathbf{v}'_f = \sum_f m_f (\mathbf{v}_f - \mathbf{v}) = \mathbf{P}_{(\text{fi})} - M_{(\text{fi})} \mathbf{v} = \mathbf{P}'_{(\text{in})} \quad (5.22)$$

ed anche per lui l'impulso è una quantità conservata — benché diversa da quella misurata da S .

Capiamo bene che, in meccanica relativistica, l'impulso definito dall'eq. (5.16) non può essere una quantità conservata in ogni SDR, dal momento che la legge di composizione delle velocità (2.63) non è additiva: se $\mathbf{P}_{(\text{in})} = \mathbf{P}_{(\text{fi})}$ allora $\mathbf{P}'_{(\text{in})} \neq \mathbf{P}'_{(\text{fi})}$. Chiaramente, per piccole velocità $v, v_i, v_f \ll c$ la differenza tra impulso totale iniziale e finale non è apprezzabile dagli strumenti di misura: entro gli errori sperimentali la conservazione dell'impulso non-relativistico è ben verificata nelle esperienze quotidiane.

Si pone quindi la seguente domanda: esiste una qualche quantità che si conserva esattamente per sistemi con particelle di qualsiasi velocità e in tutti i SDR inerziali che la misurano? Volendo generalizzare una legge di conservazione vettoriale, l'idea è di promuoverla ad una legge di conservazione per 4-vettori.

Cerchiamo quindi di costruire, per ogni particella, un 4-vettore p^μ che nel limite non relativistico $v \ll c$ si riduca all'impulso classico (5.16), $\mathbf{p} \rightarrow \mathbf{p}_{\text{nr}}$, analogamente a quanto abbiamo fatto per la velocità e l'accelerazione. Sfruttando proprio le proprietà della quadri-velocità di una particella di essere un 4-vettore che nel limite non-relativistico si riduce alla velocità classica, definiamo il **quadri-impulso**

$$p^\mu \equiv m u^\mu \stackrel{(5.7)}{=} m \gamma(c, \mathbf{v}) \quad (5.23)$$

in cui m è un parametro non nullo, detto **massa a riposo** della particella, che per definizione consideriamo come una quantità invariante.

Il quadri-impulso così definito è quindi un 4-vettore (essendo il prodotto di un invariante per un quadrivettore), ed in termini della velocità \mathbf{v} si scrive

$$\mathbf{p} = m \mathbf{u} = \frac{m \mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \xrightarrow{\frac{v}{c} \rightarrow 0} m \mathbf{v} \quad (5.24)$$

$$p^0 = m u^0 = \frac{m c}{\sqrt{1 - \frac{v^2}{c^2}}} \xrightarrow{\frac{v}{c} \rightarrow 0} m c . \quad (5.25)$$

Dal limite non-relativistico (5.24) identifichiamo la massa a riposo m con la massa inerziale Newtoniana che usiamo abitualmente quando trattiamo di meccanica classica. Vedremo più avanti però che la massa inerziale, definita come rapporto tra tri-forza e tri-accelerazione, dipende dal SDR in cui si misura. Il parametro m per l'appunto è la massa inerziale misurata nel SDR in cui la particella in esame è a riposo.

Perché il quadri-impulso è un buon candidato per generalizzare la legge di conservazione dell'impulso? Perché è un 4-vettore, e se due 4-vettori sono uguali in un SDR, lo sono anche in tutti gli altri SDR. Spieghiamoci

meglio, riprendendo il sistema di particelle considerato poco fa. L'osservatore S associa ad ogni particella dello stato iniziale il quadri-impulso $p_i^\mu = m_i u_i^\mu$, e ad ogni particella dello stato finale il quadri-impulso $p_f^\mu = m_f u_f^\mu$. (Il numero ed il tipo di particelle finali potrà essere in generale diverso da quello iniziale, se vi sono processi di annichilazione o creazione di particelle). Definiamo il quadri-impulso totale come la somma dei quadri-impulsi dei singoli costituenti:

$$P_{(\text{in})}^\mu \equiv \sum_i p_i^\mu, \quad P_{(\text{fi})}^\mu \equiv \sum_f p_f^\mu. \quad (5.26)$$

Per l'osservatore S' con coordinate

$$x'^\nu = \Lambda^\nu_\mu x^\mu + a^\nu \quad (5.27)$$

si ha $u'^\nu = \Lambda^\nu_\mu u^\mu$ (cfr. eq. (5.10)) e quindi

$$p_i'^\nu = \Lambda^\nu_\mu p_i^\mu \quad (5.28)$$

$$P_{(\text{in})}'^\nu = \sum_i \Lambda^\nu_\mu p_i^\mu = \Lambda^\nu_\mu \sum_i p_i^\mu = \Lambda^\nu_\mu P_{(\text{in})}^\mu. \quad (5.29)$$

A parole, tanto l'impulso di una singola particella come l'impulso totale trasformano con la stessa matrice Λ , come del resto ogni 4-vettore. Lo stesso dicasi per l'impulso dello stato finale:

$$P_{(\text{fi})}'^\nu = \Lambda^\nu_\mu P_{(\text{fi})}^\mu. \quad (5.30)$$

Quindi, se l'osservatore S trova sperimentalmente che $P_{(\text{in})}^\mu = P_{(\text{fi})}^\mu$, allora l'impulso totale si conserva anche in S' , e viceversa:

$$P_{(\text{in})}'^\nu = P_{(\text{fi})}'^\nu \iff \Lambda^\nu_\mu P_{(\text{in})}^\mu = \Lambda^\nu_\mu P_{(\text{fi})}^\mu \iff P_{(\text{in})}^\mu = P_{(\text{fi})}^\mu \quad (5.31)$$

Riassumendo, abbiamo definito per ogni particella un 4-vettore impulso che dipende dalla sua velocità. Abbiamo definito il quadri-impulso totale di un sistema come la somma dei quadri-impulsi dei suoi costituenti. Se il quadri-impulso totale si conserva in un SDR, allora si deve conservare in tutti i SDR, poiché tutti i 4-vettori trasformano allo stesso modo (con Λ) e la legge di trasformazione è lineare (la somma dei trasformati è uguale al trasformato della somma).

Rimane ora da vedere se il quadri-impulso totale si conserva. L'ultima parola sulla conservazione di una data quantità spetta alla verifica sperimentale. Ebbene, le eq. (5.23) e (5.26) definiscono una quantità che sperimentalmente è conservata. Enunciamo quindi la

Legge di conservazione del quadri-impulso: in ogni sistema isolato, il quadri-impulso totale si conserva.

Questa legge è stata verificata innumerevoli volte fino a velocità elevatissime in molti acceleratori di particelle, ed è una legge del moto ben stabilita poiché si accorda con tutti i dati sperimentali.

Torniamo a considerare il quadri-impulso di una singola particella, che d'ora in poi chiameremo semplicemente "impulso". Le sue quattro componenti non sono indipendenti, infatti dalla (5.8) segue subito che il quadrato dell'impulso vale

$$p^2 = p^\mu p_\mu = m^2 c^2 \quad (5.32)$$

indipendentemente dal SDR. Esplicitando questa relazione otteniamo

$$(p^0)^2 - (\mathbf{p})^2 = m^2 c^2 \implies p^0 = +\sqrt{\mathbf{p}^2 + m^2 c^2} \quad (5.33)$$

che ci dà la componente temporale dell'impulso in termini delle componenti spaziali. Abbiamo preso il segno positivo della radice in quanto l'impulso, come la quadri-velocità, è rivolto verso il futuro.

5.1.4 Energia relativistica

Cerchiamo adesso di capire il significato fisico della componente temporale dell'impulso, studiandone il limite di piccole velocità $v \ll c$. Si ha

$$cp^0 \stackrel{(5.25)}{=} \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} = mc^2 \left[1 + \frac{1}{2} \frac{v^2}{c^2} + \mathcal{O}\left(\frac{v^4}{c^4}\right) \right] = mc^2 + \frac{1}{2}mv^2 + \dots \quad (5.34)$$

Dunque, la quantità cp^0 è data dall'energia cinetica classica della particella più un termine costante mc^2 (indipendente dallo stato della particella), più termini di ordine relativo v^2/c^2 rispetto all'energia cinetica classica. Questa relazione con l'energia cinetica ci induce a ipotizzare che la quantità cp^0 sia legata all'energia della particella. Conveniamo allora di chiamare la quantità

$$E \equiv cp^0 = mc^2\gamma \quad (5.35)$$

energia relativistica, o semplicemente energia, della particella. L'impulso si potrà quindi scrivere

$$p^\mu = \left(\frac{E}{c}, \mathbf{p} \right). \quad (5.36)$$

Essendo mc^2 l'energia della particella quando $v = 0$, essa viene detta **energia a riposo** della particella. La rimanente parte di energia legata al moto

$$E_{\text{cin}} \equiv E - mc^2 = mc^2(\gamma - 1) \quad (5.37)$$

la interpretiamo come **energia cinetica relativistica**.

Assegnare un nome così importante ad una quantità dal significato fisico non evidente, è ampiamente giustificato sperimentalmente, in quanto la componente temporale dell'impulso totale cP^0 è una quantità conservata nei sistemi isolati, mentre le sue variazioni corrispondono all'energia che gli apparati sperimentali forniscono o ricevono dalle particelle, come vedremo meglio in seguito — per esempio nella sez. 5.3.3. Per ora, facciamo queste considerazioni:

- In fisica classica, l'energia è sempre definita a meno di una arbitraria costante additiva. Quindi aggiungere la costante mc^2 all'energia cinetica non pone problemi concettuali a livello classico, nonostante questa costante mc^2 sia molto maggiore dell'energia cinetica classica per velocità ordinarie.

- La presenza del termine mc^2 in $E = cp^0$ è comunque necessaria affinché p^0 sia la componente di un 4-vettore. Abbiamo visto che una legge di conservazione 4-vettoriale in un SDR garantisce la stessa legge di conservazione in tutti i SDR. Quindi la conservazione della componente temporale P^0 in un sistema isolato garantisce la conservazione dell'energia relativistica.

- Per quanto appena detto, in fisica relativistica l'energia di una particella è completamente determinata, non sono ammesse costanti arbitrarie aggiuntive.

- La conservazione dell'energia relativistica implica, per piccole velocità, la conservazione dell'energia cinetica classica negli urti elastici (in cui le particelle finali sono identiche alle particelle iniziali):

$$\begin{cases} E_{\text{in}} = E_{\text{fi}} \\ m_i = m_f \quad \forall i, f = 1, 2, \dots \end{cases} \implies \sum_i m_i c^2 + \frac{1}{2} m_i v_i^2 + \dots = \sum_f m_f c^2 + \frac{1}{2} m_f v_f^2 + \dots \quad (5.38)$$

$$\xrightarrow{v \rightarrow 0} \sum_i \frac{1}{2} m_i v_i^2 = \sum_f \frac{1}{2} m_f v_f^2. \quad (5.39)$$

- Negli urti anelastici, in cui alcune particelle finali sono diverse da quelle iniziali, le masse finali possono essere diverse da quelle iniziali. Siccome l'energia relativistica si conserva, questo vuol dire che le differenze di massa si possono convertire in differenze di energia cinetica, in modo tale che P_{tot}^0 si conservi:

$$\Delta P^0 = 0 \implies \Delta E_{\text{cin}} = -(\Delta m)c^2, \quad (5.40)$$

ove

$$\Delta E_{\text{cin}} \equiv \sum_f E_{\text{cin},f} - \sum_i E_{\text{cin},i}, \quad \Delta m \equiv \sum_f m_f - \sum_i m_i. \quad (5.41)$$

In altre parole, nelle reazioni tra particelle non si conservano separatamente la massa e l'energia cinetica, ma la loro combinazione data dall'energia relativistica totale. In particolare, per velocità ordinarie $v \ll c$, a piccolissime variazioni di massa corrispondono enormi variazioni di energia cinetica. Questo è il principio su cui si basa l'energia nucleare, dalle bombe atomiche alle centrali nucleari: convertire una piccola frazione della massa dei nuclei atomici in energia cinetica degli atomi, che si manifesta in enorme produzione di calore.

Quindi, la costante mc^2 è a tutti gli effetti un'energia che la particella possiede, in forma per così dire “congelata” finché la particella mantiene la sua integrità, ma che può essere convertita in altre forme di energia qualora la particella si trasformi.

Esempio: la bomba atomica sganciata su Nagasaki conteneva 7 Kg di Plutonio, di cui il 15%, cioè circa 1 Kg, ha subito la fissione nucleare. Circa una parte su mille della massa del Plutonio che ha reagito si è trasformata in energia, principalmente termica e di radiazioni: 1 grammo di massa si è convertito in 10^{14} J.

Nota: Sappiamo dalla meccanica classica che la conservazione dell'impulso è una conseguenza dell'invarianza per traslazioni spaziali, ossia dell'omogeneità dello spazio, mentre la conservazione dell'energia è conseguenza dell'invarianza per traslazioni temporali, ossia dell'omogeneità dello spazio. L'omogeneità dello spazio-tempo implica che anche in meccanica relativistica ci dovranno essere 4 quantità conservate. Il fatto di avere scoperto che spazio e tempo sono aspetti di un'unica entità, ci fa capire che impulso ed energia relativistici devono essere unite in modo analogo, ossia essere aspetti diversi di un unico ente: l'impulso. Come lo spazio ed il tempo si “mescolano” cambiando SDR, così impulso ed energia si mescolano nel cambiare SDR, secondo le leggi di trasformazione dei 4-vettori.

Nota: È importante osservare che, a livello relativistico, la conservazione dell'energia E è una conseguenza della conservazione dell'impulso spaziale \mathbf{p} . Supponiamo infatti che l'impulso spaziale sia conservato in due differenti SDR in moto relativo, collegati dalla trasformazione di Lorentz Λ :

$$\Delta \mathbf{P} \equiv \sum_f \mathbf{p}_f - \sum_i \mathbf{p}_i = 0 \quad (5.42a)$$

$$\Delta \mathbf{P}' \equiv \sum_f \mathbf{p}'_f - \sum_i \mathbf{p}'_i = 0 \quad (5.42b)$$

$$p'^{\nu} = \Lambda^{\nu}_{\mu} p^{\mu} . \quad (5.42c)$$

L'operatore differenza Δ e l'operatore somma Σ sono lineari, e quindi “trasparenti” alle matrici di Lorentz che altro non fanno che costruire combinazioni lineari delle componenti dei vettori su cui agiscono. In altre parole, somme e differenze di 4-vettori sono 4-vettori, per cui abbiamo

$$\Delta P'^k = \sum_f \Lambda^k_{\mu} p'_f{}^{\mu} - \sum_i \Lambda^k_{\mu} p'_i{}^{\mu} = \Lambda^k_{\mu} \left(\sum_f p_f^{\mu} - \sum_i p_i^{\mu} \right) = \Lambda^k_{\mu} \Delta P^{\mu} \quad \forall k = 1, 2, 3 . \quad (5.43)$$

Ora sfruttiamo l'ipotesi (5.42) che l'impulso spaziale si conservi in entrambi i SDR. Allora, nell'eq. (5.43) il primo membro si annulla, e nella sommatoria ad ultimo membro si annullano i termini con $\mu = 1, 2, 3$:

$$0 = \Lambda^k_0 \Delta P^0 \quad \forall k = 1, 2, 3 . \quad (5.44)$$

Ma dato che almeno uno dei tre coefficienti Λ^k_0 è non nullo,¹ deve essere

$$0 = \Delta P^0 = \frac{1}{c} \Delta E = \frac{1}{c} \left(\sum_f E_f - \sum_i E_i \right) , \quad (5.45)$$

ossia l'energia è conservata nel SDR S (e quindi anche in S' , ripetendo il ragionamento con $p = \Lambda^{-1}p'$). Generalizzando, se l'impulso spaziale è conservato in ogni SDR, anche l'energia è conservata in ogni SDR.

5.1.5 Trasformazione di Lorentz speciale per energia ed impulso

Osserviamo che, dall'eq. (5.23), la velocità della particella si può ricavare dal rapporto tra impulso spaziale ed energia:

$$\frac{\mathbf{v}}{c} = \frac{\mathbf{p}}{p^0} \quad \Longleftrightarrow \quad \mathbf{v} = c^2 \frac{\mathbf{p}}{E} . \quad (5.46)$$

¹In caso contrario la trasformazione di Lorentz sarebbe al più una rotazione, senza velocità relativa tra S ed S' , come mostrato nell'esercizio alla fine della sez. 3.3.3.

Per particelle di massa $m > 0$ si ha che $|\mathbf{p}|$ è sempre minore di p^0 (vedi eq. (5.33)), pertanto la velocità delle particelle con massa è sempre minore della velocità della luce.

Se un secondo SDR S' si muove rispetto ad S alla stessa velocità \mathbf{v} della particella, è naturale pensare che la particella sia in quiete rispetto ad S' , e che quindi il suo impulso sia diretto lungo l'asse temporale t' : $p'^\mu = (mc, \mathbf{0})$. Questo si può dimostrare in modo semplice nel caso in cui S ed S' abbiano gli assi paralleli tra loro e la velocità relativa $\mathbf{v} = v\hat{x}$ sia diretta lungo l'asse x . Per SDR siffatti la trasformazione di Lorentz speciale (5.28) di un generico impulso p^μ è

$$E' = \gamma(E - vp^1) \quad (5.47a)$$

$$p'^1 = \gamma\left(p^1 - \frac{v}{c^2}E\right) \quad (5.47b)$$

$$p'^2 = p^2 \quad (5.47c)$$

$$p'^3 = p^3. \quad (5.47d)$$

Applicando questa trasformazione all'impulso di una particella in moto alla stessa velocità di S' , dall'eq. (5.46) si ha $\mathbf{p} = (Ev/c^2, 0, 0)$ e quindi $\mathbf{p}' = (0, 0, 0)$. Ribadiamo che quanto appena detto vale solo per particelle con massa $m > 0$ per le quali $v < c$.

5.1.6 Particelle di massa nulla

Il caso di un raggio di luce si può inquadrare nello schema sviluppato, introducendo un nuovo ed importante tipo di particelle.

Volendo applicare le definizioni di quadri-velocità e di impulso ad un raggio luminoso, ci rendiamo subito conto che c'è un problema: il tempo proprio di una linea universo di un raggio di luce è nullo:

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt = 0 \quad \text{per} \quad v = c \quad (5.48)$$

e quindi la definizione (5.5) non ha senso, dando luogo ad una quadri-velocità infinita.

Tuttavia, l'espressione dell'impulso (5.23) rimane finita per $v \rightarrow c$ se si fa tendere la massa a zero in modo inversamente proporzionale al fattore $\gamma \rightarrow \infty$. La nozione di particella di luce (fotone) assumerà un preciso significato fisico nella teoria della meccanica quantistica. Per ora possiamo affermare che ha senso estrapolare i concetti di energia ed impulso a particelle di massa nulla. Infatti, le eq. (5.32) e (5.33) hanno senso anche per $m = 0$, nel qual caso l'impulso è un quadri-vettore di tipo-luce, e vale

$$m = 0 \quad \Longleftrightarrow \quad p^0 = |\mathbf{p}| \quad \Longleftrightarrow \quad |\mathbf{v}| = c. \quad (5.49)$$

Dunque, una particella di massa nulla si muove sempre alla velocità della luce, indipendentemente dalla sua energia, dal suo impulso e dal SDR in cui è osservata. In particolare, una particella a massa nulla può avere energie arbitrariamente piccole (ma non può avere energia uguale a zero). Anche per particelle a massa nulla l'energia e l'impulso dipendono dal SDR; valgono in particolare le trasformazioni di Lorentz (5.47).

5.1.7 Difetto di massa nei nuclei: equivalenza tra massa ed energia di legame

L'equivalenza tra massa ed energia risulta particolarmente evidente nei nuclei atomici, nei quali l'energia di legame è così grande da provocare differenze di massa significative. Un nucleo atomico è composto da protoni e neutroni. Si è soliti indicare con A il numero totale di protoni e neutroni (numero di massa), e con Z il numero di protoni (numero atomico). Indichiamo con²

$$m_e = 0,000549 \text{ a.m.u.} \simeq 0,511 \text{ MeV} \quad (5.50)$$

$$m_p = 1,007276 \text{ a.m.u.} \simeq 938,272 \text{ MeV} \quad (5.51)$$

$$m_n = 1,008665 \text{ a.m.u.} \simeq 939,565 \text{ MeV} \quad (5.52)$$

²La definizione di unità di massa atomica (a.m.u.) è riportata nell'app. B, assieme alle relazioni tra le più comuni unità di misura di massa ed energia.

le masse degli elettroni, dei protoni e dei neutroni liberi, e con $M_{A,Z}$ la massa del nucleo atomico in esame nello stato fondamentale, cioè di minima energia, solitamente stabile. Il *difetto di massa* è definito dalla differenza tra la somma delle masse dei costituenti del nucleo e la massa nucleare effettiva

$$\Delta m \equiv Zm_p + (A - Z)m_n - M_{A,Z} . \quad (5.53)$$

Il difetto di massa Δm corrisponde ad un'energia Δmc^2 che può essere interpretata come *energia di legame* (negativa). Tanto più grande è il difetto di massa, tanto più il nucleo è stabile. Infatti, per decomporre un nucleo nei suoi costituenti occorre fornirgli un'energia almeno uguale all'energia di legame.

Com'è noto, sia i processi di *fusione* nucleare, nei quali nuclei leggeri si fondono in un nucleo più pesante, sia i processi di *fissione* nucleare, nei quali un nucleo si rompe in frammenti, sono basati sul difetto di massa. Questo è possibile perchè si può vedere sperimentalmente che in entrambi i casi i processi vanno nel senso di aumentare il difetto di massa: gli stati iniziali hanno un difetto di massa minore degli stati finali. Consideriamo ad esempio la reazione nucleare



dove i reagenti sono un protone (nucleo di Idrogeno) ed un isotopo del Litio, mentre i prodotti sono due isotopi dell'Elio, l'Elio-3 e l'Elio-4.

Le masse di questi elementi espresse in unità di massa atomica sono³

$$\begin{array}{ll} \text{massa del} & {}^1_1\text{H} = 1,007825 \\ " & " \quad {}^6_3\text{Li} = 6,015122 \\ " & " \quad {}^3_2\text{He} = 3,016029 \\ " & " \quad {}^4_2\text{He} = 4,002603 . \end{array} \quad (5.55)$$

Se calcoliamo i difetti di massa troviamo

$$\begin{aligned} \Delta m({}^1_1\text{H}) &= 0 \\ \Delta m({}^6_3\text{Li}) &= 0,034348 \text{ a.m.u.} \simeq 32,00 \text{ MeV} \\ \Delta m({}^3_2\text{He}) &= 0,008286 \text{ a.m.u.} \simeq 7,72 \text{ MeV} \\ \Delta m({}^4_2\text{He}) &= 0,030377 \text{ a.m.u.} \simeq 28,30 \text{ MeV} \end{aligned} \quad (5.56)$$

da cui si ha il difetto di massa complessivo

$$\Delta m_{\text{tot}} = \Delta m({}^1_1\text{H}) + \Delta m({}^6_3\text{Li}) - \Delta m({}^3_2\text{He}) - \Delta m({}^4_2\text{He}) = -0,004315 \text{ a.m.u.} \simeq -4,02 \text{ MeV} \quad (5.57)$$

che è negativo e quindi la reazione avviene con produzione di energia.

Lo stesso risultato si ottiene facendo il bilancio delle masse, poiché le masse dei costituenti si elidono nella differenza tra il contributo del primo membro della reazione e il secondo. Il risultato sarà con il segno opposto:

$$m({}^1_1\text{H}) + m({}^6_3\text{Li}) - m({}^3_2\text{He}) - m({}^4_2\text{He}) = +0,004315 \text{ a.m.u.} \simeq 4,02 \text{ MeV} . \quad (5.58)$$

Nel grafico (5.1) è riportata l'energia di legame media per nucleone (cioè il difetto di massa diviso il numero dei nucleoni) in funzione del numero atomico.⁴

5.2 Sistemi a molte particelle

La conservazione dell'impulso ci permette di trattare diversi processi di interazione relativistica, senza conoscere le forze che li hanno provocati, purché siano note le caratteristiche principali dei corpi nello stato finale.

³Queste sono masse "atomiche", che comprendono le masse degli elettroni. Però nel calcolo del difetto di massa ciò non influisce, poiché tanti sono gli elettroni prima e dopo la reazione, per cui il loro contributo si elide tra primo e secondo termine della reazione. I dati sono ripresi da <http://periodni.com/it>.

⁴L'immagine è ripresa da http://en.wikipedia.org/wiki/Binding_energy.

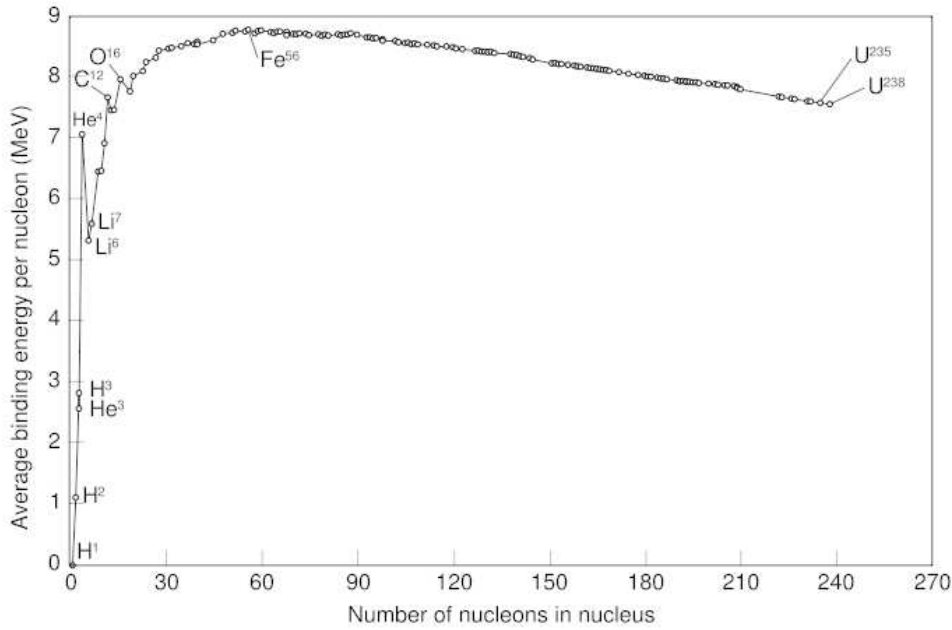


Figure 5.1: *Energia di legame media per nucleone, in funzione del numero atomico A . L'elemento più stabile è il ferro ${}^{56}_{26}\text{Fe}$. Salendo la curva da sinistra (Idrogeno) si hanno i processi di fusione nucleare, mentre salendo da destra (Uranio) si hanno i processi di fissione.*

5.2.1 Decadimenti: equivalenza tra massa ed energia cinetica

I decadimenti sono processi in cui lo stato iniziale è rappresentato da una sola particella di massa M , che spontaneamente si disintegra producendo due o più particelle diverse da quella iniziale.

Notiamo innanzitutto l'impossibilità che una particella si trasformi in una sola particella con massa m diversa dalla massa iniziale M . Infatti, nel SDR in cui M è a riposo, l'impulso iniziale è $P_{\text{in}}^\mu = M(c, \mathbf{0})$, mentre l'impulso finale sarà $P_{\text{fi}}^\mu = m\gamma(v)(c, \mathbf{v})$, ove \mathbf{v} è la velocità della particella finale. Dall'uguaglianza $P_{\text{in}}^\mu = P_{\text{fi}}^\mu$ segue che $\mathbf{v} = \mathbf{0}$, quindi $\gamma(v) = 1$ ed infine $m = M$.

Il processo di decadimento più semplice è quello che produce due particelle nello stato finale. Un tipico esempio è quello del mesone K^0 che decade in due mesoni π^+ e π^- — qui gli apici indicano la carica elettrica — secondo la reazione

$$K^0 \rightarrow \pi^+ \pi^- . \quad (5.59)$$

La massa del mesone K^0 vale⁵ $M_K \simeq 500 \text{ MeV}/c^2 \simeq 8,9 \cdot 10^{-25} \text{ g}$, e la sua energia a riposo $E_K = M_K c^2 \simeq 500 \text{ MeV} \simeq 8,0 \cdot 10^{-11} \text{ J}$ costituisce l'energia totale del sistema nel SDR in cui il K^0 è in quiete. I due mesoni π prodotti hanno massa uguale tra loro, pari a $m_\pi \simeq 140 \text{ MeV}/c^2 \simeq 2,5 \cdot 10^{-25} \text{ g}$. La differenza di massa tra stato iniziale e stato finale, pari a $\Delta M \simeq 220 \text{ MeV}/c^2 \simeq 3,9 \cdot 10^{-25} \text{ g}$ viene convertita in energia cinetica dei prodotti di decadimento: $110 \text{ MeV} \simeq 1,75 \cdot 10^{-11} \text{ J}$ per ognuno. Ciascuno dei due mesoni si ritrova con un'energia $E_\pi \simeq 250 \text{ MeV} \simeq 4 \cdot 10^{-11} \text{ J}$, e quindi ciascuno si allontana dal punto di produzione ad una velocità $v = c\sqrt{1 - (m_\pi c^2/E_\pi)^2} \simeq 0,83 c$.

In generale, le particelle prodotte possono avere massa diversa, che indichiamo con m_1 ed m_2 . Nel SDR del centro di massa in cui la particella iniziale di massa M ed impulso P^μ è a riposo, si ha

$$P^\mu = (Mc, \mathbf{0}) , \quad p_1^\mu = (E_1/c, \mathbf{p}_1) , \quad p_2^\mu = (E_2/c, \mathbf{p}_2) . \quad (5.60)$$

La conservazione dell'impulso

$$P^\mu = p_1^\mu + p_2^\mu \quad (5.61)$$

⁵Nell'app. B sono riportate le relazioni tra le più comuni unità di misura di massa ed energia.

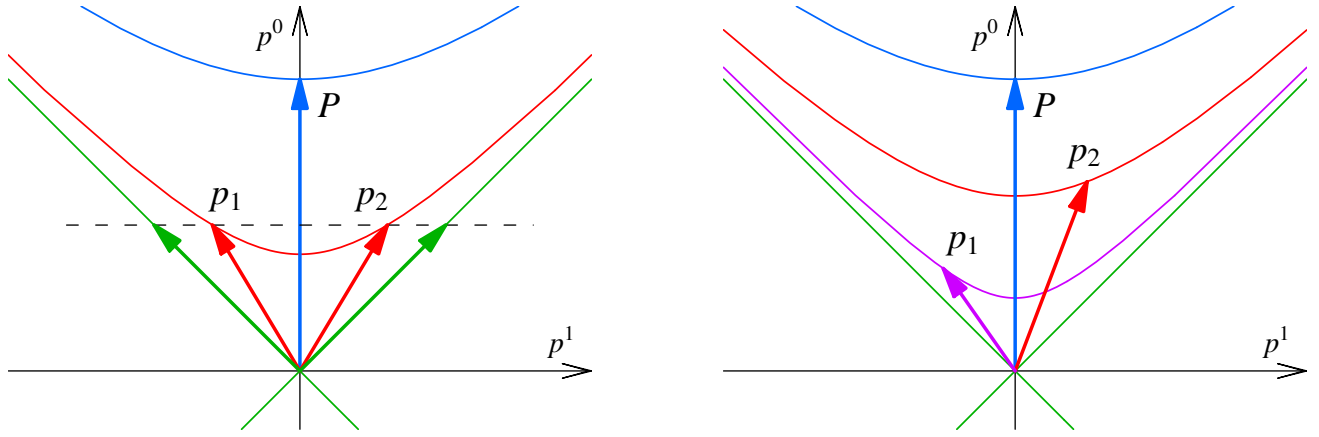


Figure 5.2: *Diagramma energia-impulso per il decadimento in due particelle: (a) particelle finali di uguale massa (diversa da zero in rosso, uguale a zero in verde); (b) particelle finali con massa diversa.*

implica che $\mathbf{p} \equiv \mathbf{p}_1 = -\mathbf{p}_2$, e quindi le due particelle si allontanano in direzione opposta. Naturalmente, per l'isotropia dello spazio, la direzione di decadimento è casuale e la distribuzione di probabilità è uniforme nell'angolo solido. Le energie delle particelle prodotte sono univocamente determinate e si possono ricavare così: dall'eq. (5.61) segue che

$$p_1^\mu = P^\mu - p_2^\mu, \quad p_2^\mu = P^\mu - p_1^\mu. \quad (5.62)$$

Prendendo il quadrato di queste espressioni si ha

$$m_1^2 c^2 = (M^2 + m_2^2) c^2 - 2M E_2 \quad (5.63)$$

$$m_2^2 c^2 = (M^2 + m_1^2) c^2 - 2M E_1 \quad (5.64)$$

$$\begin{aligned} E_1 &= \frac{M^2 + m_1^2 - m_2^2}{2M} c^2 \\ E_2 &= \frac{M^2 + m_2^2 - m_1^2}{2M} c^2. \end{aligned} \quad (5.65)$$

Quindi la particella con massa maggiore si porta via più energia dell'altra. Di conseguenza, essendo l'impulso spaziale delle due particelle uguale in modulo, dall'eq. (5.46) deduciamo che la particella con velocità maggiore è quella con massa minore. Essendo note l'energia e le masse delle particelle, l'impulso spaziale si può calcolare con la formula (5.33) che dà

$$\frac{\mathbf{p}^2}{c^2} = \frac{(M^2 - m_1^2 - m_2^2)^2 - 4m_1^2 m_2^2}{4M^2}, \quad (5.66)$$

giustamente simmetrica in m_1 ed m_2 .

La somma delle masse delle particelle prodotte non può eccedere la massa della particella iniziale. Infatti, ricordando che l'energia di una qualsiasi particella soddisfa $E = mc^2 \gamma \geq mc^2$, dalla conservazione dell'energia si ha

$$Mc^2 \stackrel{(5.60)}{=} E_1 + E_2 \geq m_1 c^2 + m_2 c^2, \quad (5.67)$$

da cui ricaviamo la condizione cinematica per il decadimento

$$M \geq m_1 + m_2. \quad (5.68)$$

Solitamente le particelle prodotte dei laboratori non sono ferme. Sia \mathbf{V} la velocità della particella che decade in volo. In questo caso, le energie delle particelle finali dipendono dall'angolo di emissione θ rispetto a \mathbf{V} . Consideriamo una delle due particelle prodotte nel decadimento. Indichiamo con il pedice "cm" le quantità nel SDR del centro di massa precedente, e senza pedice le analoghe quantità nel sistema del laboratorio. Orientiamo

l'asse x lungo \mathbf{V} e l'asse y nel piano di decadimento. Si ha $\mathbf{p} = (p \cos \theta, p \sin \theta, 0)$. La trasformazione di Lorentz tra i due SDR dà

$$E_{\text{cm}} = \frac{E - Vp \cos \theta}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad (5.69)$$

da cui

$$\cos \theta = \frac{E - E_{\text{cm}} \sqrt{1 - \frac{V^2}{c^2}}}{(V/c) \sqrt{E^2 - (mc^2)^2}}, \quad (5.70)$$

in cui ricordiamo che E_{cm} è una quantità nota, data da una delle eq. (5.65).

Possiamo visualizzare graficamente il decadimento analizzando le componenti dell'impulso nel SDR del laboratorio. Dalle formule di trasformazione di Lorentz segue

$$p^x = \frac{p_{\text{cm}} \cos \theta_{\text{cm}} + \frac{V}{c^2} E_{\text{cm}}}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad p^y = p_{\text{cm}} \sin \theta_{\text{cm}}. \quad (5.71)$$

Isolando $p_{\text{cm}} \cos \theta_{\text{cm}}$ e sommando quadraticamente con $p_{\text{cm}} \sin \theta_{\text{cm}}$ otteniamo

$$\left(p^x \sqrt{1 - \frac{V^2}{c^2}} - E_{\text{cm}} \frac{V}{c^2} \right)^2 + (p^y)^2 = p_{\text{cm}}^2 \quad (5.72)$$

che è l'equazione di un'ellisse rispetto alle variabili (p^x, p^y) , di semiasse $(p_{\text{cm}}/\sqrt{1 - \frac{V^2}{c^2}}, p_{\text{cm}})$ e di centro $Q = (E_{\text{cm}} \frac{V}{c^2} / \sqrt{1 - \frac{V^2}{c^2}}, 0)$.

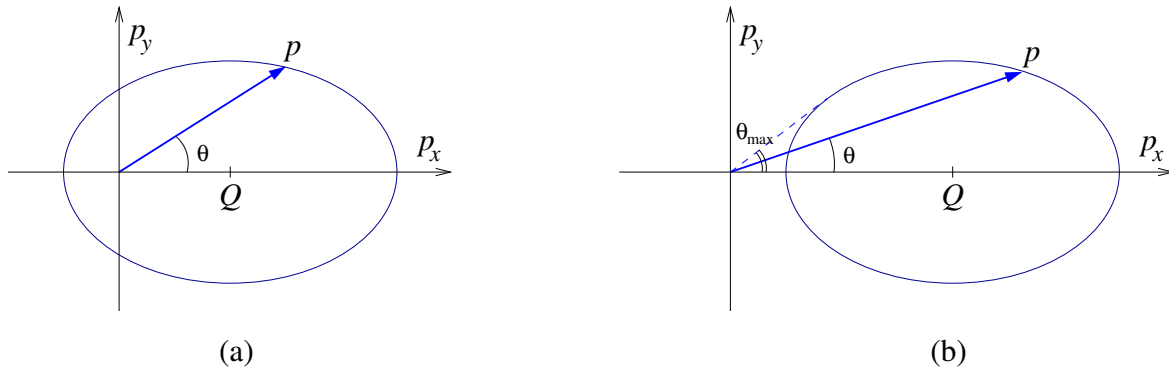


Figure 5.3: *Ellissi di decadimento, luoghi dei punti generati dai possibili impulsi della particella prodotta nel decadimento binario: (a) velocità V nel LAB della particella iniziale minore della velocità nel CM della particella finale v_{cm} ; (b) situazione opposta $V > v_{\text{cm}}$.*

Se la velocità nel CM della particella finale $v_{\text{cm}} = c^2 p_{\text{cm}} / E_{\text{cm}}$ è maggiore della velocità relativa V tra CM e LAB, $v_{\text{cm}} > V$, allora $Q_x <$ semiasse- x , ossia l'origine degli impulsi sta all'interno dell'ellisse (fig. 5.3a), e la particella prodotta può essere emessa in qualsiasi direzione. Se invece $v_{\text{cm}} < V$ l'origine degli impulsi sta al di fuori dell'ellisse (fig. 5.3b), e quindi la particella è emessa all'interno di un cono con apertura angolare θ_{max} dato da

$$\sin \theta_{\text{max}} = \frac{p_{\text{cm}} \sqrt{1 - \frac{V^2}{c^2}}}{mV}. \quad (5.73)$$

Esercizio: Determinare l'angolo θ_{max} .

Soluzione: Risolvendo l'eq. (5.70) rispetto all'energia E , otteniamo l'equazione di secondo grado

$$E^2 \left(1 - \frac{V^2}{c^2} \cos^2 \theta \right) - 2EE_{\text{cm}} \sqrt{1 - \frac{V^2}{c^2}} + E_{\text{cm}}^2 \left(1 - \frac{V^2}{c^2} \right) + m^2 V^2 c^2 \cos^2 \theta = 0. \quad (5.74)$$

L'angolo massimo corrisponde a due soluzioni coincidenti, che si hanno quando il discriminante dell'equazione si annulla. La condizione di annullamento del discriminante determina l'angolo massimo come nell'eq. (5.73).

5.2.2 Sistema del centro di massa

Nel descrivere un sistema fisico a molti corpi, anche in meccanica relativistica, così come in meccanica classica, può essere conveniente separare il moto del sistema nel suo complesso dal moto relativo dei suoi componenti, introducendo il *sistema di riferimento del centro di massa* (CM), in cui l'impulso spaziale totale è nullo.

Dimostriamo che, dato un qualsiasi sistema di particelle libere, è sempre possibile trovare un SDR nel quale l'impulso spaziale totale sia nullo. Abbiamo osservato nella sez. (5.1.5) che un quadri-vettore (non banale) può avere componenti spaziali nulle se e solo se è di tipo tempo. Infatti, se l'impulso totale $P^\mu = (P^0, \mathbf{P})$ è di tipo tempo, $|\mathbf{P}| < P^0$, allora una spinta (passiva, cioè sul SDR) di velocità

$$\mathbf{v}/c = \mathbf{P}/P^0 \quad (5.75)$$

trasforma P^μ in $P'^\mu = (P'^0, \mathbf{0})$ con componenti spaziali nulle.

Quindi, per dimostrare l'esistenza del sistema di CM è sufficiente far vedere che l'impulso totale di N particelle è sempre di tipo tempo. Indichiamo con $p_i^\mu : i = 1, \dots, N$ gli impulsi delle particelle in un SDR S arbitrario, le quali sappiamo essere di tipo-tempo o di tipo-luce: $p_i^2 = m_i^2 c^2 \geq 0$. Allora

$$\begin{aligned} P^2 &= g_{\mu\nu} P^\mu P^\nu = \sum_{i,j=1}^N g_{\mu\nu} p_i^\mu p_j^\nu = \sum_i g_{\mu\nu} p_i^\mu p_i^\nu + \sum_{i \neq j} g_{\mu\nu} p_i^\mu p_j^\nu \\ &= \sum_i m_i^2 c^2 + \sum_{i \neq j} (p_i^0 p_j^0 - |\mathbf{p}_i| |\mathbf{p}_j| \cos \theta_{ij}) . \end{aligned} \quad (5.76)$$

Sempre per il fatto che i vettori sono di tipo-tempo o di tipo-luce rivolti verso il futuro, abbiamo $p_i^0 \geq |\mathbf{p}_i|$, quindi

$$p_i^0 p_j^0 - |\mathbf{p}_i| |\mathbf{p}_j| \cos \theta_{ij} \geq 0 , \quad (5.77)$$

da cui deduciamo $P^2 \geq 0$. Vediamo che l'impulso totale può essere di tipo luce ($P^2 = 0$) solamente se tutte le particelle sono di tipo luce ($m_i = 0 \quad \forall i$) e se tutte le particelle si muovono nella stessa direzione ($\theta_{ij} = 0 \quad \forall i, j$). In tutti gli altri casi, l'impulso totale è di tipo tempo.

Indichiamo ora con S_{cm} il SDR del CM (da identificare con S' nelle formule precedenti), che si muove con velocità \mathbf{v} rispetto ad S . In questi due SDR l'impulso totale si scrive

$$P_{\text{cm}}^\mu = \left(\frac{E_{\text{cm}}}{c}, \mathbf{0} \right) , \quad P^\mu = \left(\frac{E}{c}, \mathbf{P} \right) , \quad (5.78)$$

in cui, dall'eq. (5.75),

$$\mathbf{P} = \frac{E \mathbf{v}}{c^2} . \quad (5.79)$$

Dal quadrato dell'impulso totale ricaviamo

$$E_{\text{cm}}^2 = c^2 P^\mu P_\mu = E^2 - c^2 |\mathbf{P}|^2 = E^2 - \left(\frac{E \mathbf{v}}{c} \right)^2 = \left(1 - \frac{v^2}{c^2} \right) E^2 \quad (5.80)$$

da cui otteniamo

$$E = \gamma E_{\text{cm}} , \quad \mathbf{P} = \gamma \frac{E_{\text{cm}} \mathbf{v}}{c^2} , \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} . \quad (5.81)$$

Possiamo riassumere le due precedenti equazioni così:

$$P^\mu = \left(\frac{E}{c}, \mathbf{P} \right) = \gamma E_{\text{cm}} \left(1, \frac{\mathbf{v}}{c} \right) = \frac{E_{\text{cm}}}{c^2} \gamma(c, \mathbf{v}) . \quad (5.82)$$

Confrontando la precedente espressione con la definizione di impulso (5.23), deduciamo che il sistema composto dalle N particelle ha una massa inerziale a riposo data da

$$M \equiv \frac{E_{\text{cm}}}{c^2} , \quad (5.83)$$

la quale è maggiore della somma delle masse a riposo delle singole particelle:

$$E_{\text{cm}} = \sum_i E_{\text{cm},i} \stackrel{(5.33)}{=} \sum_i c \sqrt{m_i^2 c^2 + \mathbf{p}_{\text{cm},i}^2} \geq c^2 \sum_i m_i \quad (5.84)$$

e quindi

$$M \geq \sum_i m_i . \quad (5.85)$$

In questa equazione il segno di uguale si ha solo se tutte le particelle hanno impulso nullo nel CM, ovvero se si muovono con la stessa velocità \mathbf{v} rispetto ad S .

L'importante conclusione che ne possiamo trarre è che l'energia cinetica delle particelle nel CM

$$E_{\text{cm,cin}} \equiv \sum_i (E_{\text{cm},i} - m_i c^2) = E_{\text{cm}} - \sum_i m_i c^2 = (M - \sum_i m_i) c^2 \quad (5.86)$$

contribuisce alla massa inerziale del sistema fisico, essendo

$$M = \sum_i m_i + \frac{E_{\text{cm,cin}}}{c^2} . \quad (5.87)$$

Nel caso di particelle interagenti, il risultato si può generalizzare affermando che l'energia interna del sistema contribuisce alla sua massa inerziale, come abbiamo visto nella sez. 5.1.7.

Per esempio, un corpo caldo ha massa maggiore rispetto a quando è freddo; un oggetto in rotazione ha massa maggiore di quando è fermo; una molla in tensione ha una massa maggiore di quando è rilassata; una batteria carica ha massa maggiore di quando è scarica. Empiricamente non ce ne rendiamo conto in quanto le differenze di massa in questi casi sono impercettibili ai comuni strumenti di misura.

5.2.3 Urti tra particelle

Se nei decadimenti avviene una trasformazione di massa in energia, è anche possibile il processo opposto, in cui si ha una conversione di energia in massa. Infatti, riuscendo a concentrare una sufficiente quantità di energia in una piccola regione di spazio, per esempio facendo collidere due particelle, si riescono a produrre ulteriori particelle. Un risultato sensazionale degli esperimenti di collisione ad alte energie è stato la produzione di particelle di tipo nuovo e sconosciuto, anche molto più pesanti dei protoni.

Maggiore è l'energia che si riesce a fornire alle particelle incidenti, maggiore è la massa delle particelle che si riescono a produrre. Supponiamo di avere a disposizione delle particelle di massa m che riusciamo ad accelerare fino ad energia E . Ci chiediamo quale sia la massa massima delle particelle che riusciamo a produrre. Sia M la massa di una delle nuove particelle. Nel CM del sistema, l'energia totale E_{cm} deve essere almeno pari all'energia della particella M , la quale è almeno pari a Mc^2 . Quindi per poter produrre tale particella deve essere

$$E_{\text{cm}} \geq Mc^2 . \quad (5.88)$$

Se indichiamo con $p_1^\mu = (E_1/c, \mathbf{p}_1)$ e $p_2^\mu = (E_2/c, \mathbf{p}_2)$ gli impulsi delle due particelle incidenti che facciamo scontrare, nel CM si deve avere $\mathbf{p}_2 = -\mathbf{p}_1$ e quindi

$$E_{\text{cm}} = E_1 + E_2 = c(p_1^0 + p_2^0) = c\sqrt{(p_1 + p_2)^2} . \quad (5.89)$$

Pertanto, in un urto di due particelle, la condizione per poter produrre una particella (o un sistema di particelle) di massa M si scrive in modo invariante come segue:

$$(p_1 + p_2)^2 \geq (Mc)^2 \quad (5.90)$$

I primi apparati per provocare questi urti consistevano in un fascio di particelle con il quale bombardare un bersaglio fisso. In questo caso si ha, per particelle iniziali identiche,

$$p_1^\mu = (E/c, p, 0, 0) , \quad p_2^\mu = (mc, 0, 0, 0) \quad \implies \quad (p_1 + p_2)^2 = 2m^2 c^2 + 2Em = 2m(E + mc^2) \quad (5.91)$$

che determina la massima massa finale

$$M \leq \frac{\sqrt{(p_1 + p_2)^2}}{c} = \frac{\sqrt{2m(E + mc^2)}}{c} . \quad (5.92)$$

Al crescere dell'energia E delle particelle nel fascio, la massa prodotta (e quindi l'energia nel CM) cresce al più come la radice quadrata di E . Questo perché con E cresce anche l'energia cinetica delle particelle prodotte, e solo una minima parte dell'energia disponibile nel laboratorio si può convertire in massa.

Se invece si fanno collidere due particelle identiche una contro l'altra alla stessa energia, il LAB diventa il sistema del CM:

$$p_1^\mu = (E/c, p, 0, 0) , \quad p_2^\mu = (E/c, -p, 0, 0) \quad \implies \quad (p_1 + p_2)^2 = \left(\frac{2E}{c}\right)^2 \quad (5.93)$$

per cui

$$M \leq \frac{2E}{c^2} . \quad (5.94)$$

Quindi, in un collisore, la massa producibile cresce linearmente con l'energia delle particelle iniziali, poiché tutta l'energia disponibile può essere convertita in massa. Per questo motivo, gli acceleratori di alta energia vengono realizzati generando due fasci di particelle che vengono fatti scontrare frontalmente.

Esempio: La scoperta del quark *top* ($m_t \simeq 175 \text{ GeV} = 175 \cdot 10^9 \text{ eV}$) è stata possibile facendo scontrare un fascio di protoni contro un fascio di antiprotoni ($m_p = m_{\bar{p}} \simeq 0,94 \text{ GeV}$) in cui ogni (anti-)protone aveva un'energia di 1000 GeV, oltre mille volte superiore all'energia a riposo. Da notare che il quark top può essere generato solo in coppie $t\bar{t}$, e che il protone e l'antiprotone sono composti da “partoni”, ciascuno dei quali porta con sé solo una frazione dell'energia del protone. Negli urti che producono i top, solo una coppia di partoni è attiva. Questo richiede che l'energia nel CM dei partoni attivi sia maggiore di 350 GeV. A questo scopo è necessaria un'energia nel CM dei protoni decisamente maggiore. Il valore di 2000 GeV è sufficiente per produrre le coppie di top.

5.3 Dinamica del punto materiale

Dopo aver introdotto le grandezze cinematiche relativistiche u e w che descrivono il moto dei corpi, passiamo a considerare le cause del moto: le forze.

In fisica non relativistica, la dinamica dei punti materiali è espressa dal secondo e terzo principio. Il secondo principio della dinamica afferma che la forza è la causa di variazione della velocità, e quindi dell'impulso

$$\mathbf{F} = m \frac{d\mathbf{v}}{dt} = \frac{d\mathbf{p}_{\text{nr}}}{dt} . \quad (5.95)$$

Il terzo principio afferma che se il corpo A esercita una forza sul corpo B ad un dato istante t , simultaneamente il corpo B esercita una forza uguale e contraria sul corpo A , qualsiasi sia la distanza tra i due corpi:

$$\mathbf{F}_{AB}(t) = -\mathbf{F}_{BA}(t) , \quad \iff \quad \frac{d\mathbf{P}_{\text{nr}}^{(\text{tot})}}{dt} = 0 . \quad (5.96)$$

Immediata conseguenza del principio di azione e reazione è la conservazione dell'impulso totale per sistemi isolati.

La generalizzazione della dinamica alla relatività in termini di forze è una questione molto delicata, soprattutto per quanto riguarda il secondo principio. Infatti, in relatività non possiamo più parlare di simultaneità assoluta, né di azione a distanza (le interazioni non possono propagarsi a velocità infinita, ma solo con velocità minore o uguale a c).

Il modo più semplice e rigoroso di sviluppare la dinamica relativistica parte dal principio di conservazione dell'impulso relativistico, già enunciato nella sezione precedente. A questo principio di conservazione affianchiamo il principio di località, che afferma che le interazioni si possono propagare solo tra punti dello spazio “infinitamente” vicini, ed a velocità non superiore a c . Quest'ultimo principio porta a formulare le teorie relativistiche delle forze in termini di *campi*, mediatori delle forze di qualsiasi natura: gravitazionale, elettromagnetica, nucleare, ecc.. Vedremo nel prossimo capitolo come formulare l'interazione delle particelle con il campo elettromagnetico. Per ora, ci limitiamo a considerare le interazioni che possono instaurarsi tra due corpi che vengano in contatto tra loro (forze di contatto), oppure le interazioni prodotte da un dato campo di forze su una certa particella.

5.3.1 Quadri-forza

Procedendo con l'analogia formale tra formule 3-vettoriali e 4-vettoriali, definiamo la **quadri-forza** che agisce su una particella nel seguente modo:

$$f^\mu \equiv \frac{dp^\mu}{d\tau} . \quad (5.97)$$

Chiaramente, la quadri-forza è un 4-vettore, essendo definita come un 4-vettore diviso un 4-scalare. Questa generalizzazione della seconda legge della meccanica è quindi valida in ogni SDR inerziale, dal momento che anche la quadri-forza trasforma come tutti i 4-vettori (cfr. eq. (5.15)).

La quadri-forza, così come la quadri-accelerazione, è ortogonale al quadri-impulso. Infatti, essendo il quadrato del quadri-impulso costante nel tempo, si ha:

$$p^\mu(\tau)p_\mu(\tau) = m^2c^2 \implies 0 = \frac{dp^\mu p_\mu}{d\tau} = 2\frac{dp^\mu}{d\tau}p_\mu = 2f^\mu p_\mu \implies f^\mu p_\mu = 0 \quad (5.98)$$

Nel caso di particelle con massa $m > 0$, vale $f^\mu = mw^\mu$, ove w^μ è la quadri-accelerazione (5.11). In questo caso, l'ortogonalità tra quadri-forza e quadri-impulso è anche una immediata conseguenza dell'ortogonalità tra quadri-accelerazione e quadri-velocità, come ci dice l'eq. (5.14). Di conseguenza, non tutte e 4 le componenti della quadri-forza sono indipendenti. Infatti, possiamo ricavare la componente temporale della quadri-forza in funzione delle componenti spaziali e della velocità:

$$0 = f^\mu u_\mu = f^0 u^0 - \mathbf{f} \cdot \mathbf{u} = \gamma(f^0 c - \mathbf{f} \cdot \mathbf{v}) \quad (5.99)$$

$$\implies f^0 = \frac{\mathbf{f} \cdot \mathbf{v}}{c} . \quad (5.100)$$

Se in un certo SDR è assegnato un campo di quadri-forza come funzione delle coordinate spazio-temporali (ed eventualmente funzione anche della quadri-velocità), è possibile determinare il moto di una particella di massa $m > 0$, posizione e velocità iniziali date, risolvendo il sistema differenziale del primo ordine $\tau \mapsto (x^\mu, u^\mu)$

$$\frac{dx^\mu(\tau)}{d\tau} = u^\mu(\tau) \quad (5.101)$$

$$\frac{du^\mu(\tau)}{d\tau} = \frac{1}{m} f^\mu(x^\nu(\tau), u^\nu(\tau)) \quad (5.102)$$

date le condizioni iniziali $x^\mu(0) = x_{\text{in}}^\mu$, $u^\mu(0) = u_{\text{in}}^\mu$.

Chiaramente, per poter risolvere il precedente sistema, bisogna determinare la quadri-forza per i vari sistemi che vogliamo studiare: particella in campo elettromagnetico, oppure in un campo gravitazionale, oppure soggetta a sollecitazioni meccaniche.

5.3.2 Forza di Newton

Cerchiamo ora di esprimere il concetto classico di *forza Newtoniana*, che pensiamo definito operativamente dallo sforzo misurato da un dinamometro, in termini delle quantità relativistiche fin qui introdotte. Pensiamo al seguente apparato, schematizzato in fig. 5.4: una guida meccanica che deflette la traiettoria di una particella. La particella arriva con velocità tangente al primo tratto della guida; successivamente nel tratto curvo viene deflessa, ed infine esce con velocità ed impulso diversi da quelli iniziali. Poiché supponiamo di poter misurare la velocità della particella in ogni punto della guida, in ogni punto conosciamo il suo impulso e quindi la *quadri-forza* che agisce sulla particella. Ci chiediamo ora quale *forza Newtoniana* \mathbf{F}_{gp} la guida abbia esercitato sulla particella. È più facile rispondere a questa domanda: quale forza \mathbf{F}_{pg} ha esercitato la particella sulla guida? Basterebbe aver misurato la forza \mathbf{F}_{sg} che i supporti hanno esercitato sulla guida in modo da mantenerla ferma, per contrastare la forza esercitata dalla particella. Per calcolare la forza, analizziamo le due interazioni particella-guida e guida-supporto come se avvenissero in fasi separate:

1) Un elemento di guida di massa ΔM , a seguito dell'interazione nel tempo Δt con la particella, acquista un impulso Δp_g uguale e contrario alla variazione di impulso della particella:

$$\Delta p_g^\mu + \Delta p_p^\mu = 0 \quad (5.103)$$

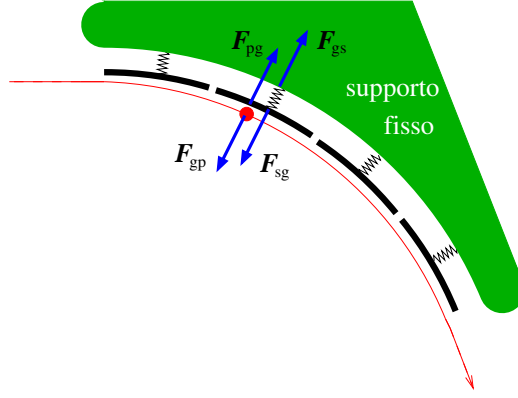


Figure 5.4: *Deflessione di una particella (pallina rossa) da parte di una guida (archi neri, ciascuno di massa ΔM). Le forze in gioco sono indicate dalle frecce blu. Si suppone che i tratti della guida acquistino delle quantità di moto (uguale ed opposto all'impulso impresso alla particella) che vengono successivamente assorbite dal supporto fisso (verde).*

Supponiamo ogni elemento di guida di massa sufficientemente grande, così che la velocità da esso acquistata sia molto piccola: piccola rispetto a c ; piccola in modo che durante il tempo Δt la guida non si sposti apprezzabilmente dalla sua posizione iniziale. Si ha allora, per la (5.24)

$$\Delta \mathbf{p}_g \simeq \Delta M \Delta \mathbf{v}_g \quad (5.104)$$

e si ha $|\Delta v_g| \ll c$ purché $|\Delta \mathbf{p}_g| \ll \Delta M c$.

2) Il supporto, per frenare la guida, deve esercitare su di essa una forza \mathbf{F}_{sg} . Essendo la guida quasi ferma, nell'interazione supporto-guida possiamo applicare i principi della dinamica Newtoniana:

$$-\mathbf{F}_{gs} = \mathbf{F}_{sg} = -\Delta M \frac{\Delta \mathbf{v}_g}{\Delta t} \simeq -\frac{\Delta \mathbf{p}_g}{\Delta t}, \quad (5.105)$$

in cui le forze vanno intese come forze medie nel tempo Δt .

Mettiamo ora assieme i due processi particella-guida e guida-supporto, e consideriamo la guida molto ben legata al supporto, in modo che la risposta \mathbf{F}_{sg} del supporto sia pressoché simultanea alla sollecitazione \mathbf{F}_{pg} della particella. Vediamo che la guida gioca un ruolo da mediatore praticamente inerte (ferma inizialmente, ferma alla fine), per cui è ragionevole pensare che il supporto abbia esercitato sulla particella (tramite la guida) la forza necessaria alla deflessione

$$\mathbf{F}_{gp} = \mathbf{F}_{sg} \stackrel{(5.105)}{=} -\frac{\Delta \mathbf{p}_g}{\Delta t} \stackrel{(5.103)}{=} \frac{\Delta \mathbf{p}_p}{\Delta t}. \quad (5.106)$$

Viceversa, è ragionevole pensare che la particella abbia esercitato sul supporto (tramite la guida) una forza uguale ed opposta $\mathbf{F}_{pg} = \mathbf{F}_{gs}$.

Notate che non abbiamo supposto la guida liscia: la particella poteva essere stata frenata o accelerata nella direzione del moto. Non abbiamo nemmeno fatto ipotesi sulla grandezza della velocità della particella: quanto abbiamo detto è valido sia per piccole velocità $v_p \ll c$ sia per velocità relativistiche $v_p \lesssim c$.

Questo ragionamento avrebbe l'ambizione di convincerci che la tri-forza Newtoniana, come appare per esempio nelle formule della gravitazione o dell'elettromagnetismo, va identificata con la variazione dell'impulso relativistico spaziale nell'intervallo di tempo unitario

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}, \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (5.107)$$

ove il tempo t è quello dell'osservatore che sta misurando la forza.

L'apparente complicazione nell'avere introdotto la guida nell'azione esercitata dal supporto (laboratorio) sulla particella ha lo scopo di trasformare la variazione di impulso della particella $\Delta \mathbf{p}_p$ (in moto relativistico) in una uguale variazione di impulso $\Delta \mathbf{p}_g$ di un oggetto "lento"; questo secondo impulso può essere quindi misurato o

calcolato con i metodi classici, ed in particolare può essere messo in relazione alle forze a noi note della meccanica classica.

Un ulteriore modo di vedere la cosa è questo: se un corpo A esercita nel tempo Δt una forza \mathbf{F}_A , il suo effetto su di un corpo B non relativistico è $\Delta \mathbf{p}_B \simeq m_B \Delta \mathbf{v}_B = \mathbf{F}_A \Delta t$. L'effetto della stessa forza su un oggetto relativistico C lo si può calcolare invocando la conservazione del impulso, trasferendo l'impulso $\mathbf{F}_A \Delta t$ prima ad un oggetto non relativistico B , e quindi trasferendo l'impulso acquistato da B completamente a C : $\Delta \mathbf{p}_C = \Delta \mathbf{p}_B = \mathbf{F}_A \Delta t$. Questa relazione deve valere anche nel caso di azione diretta di A su C , ossia

$$\mathbf{F}_A \Delta t = \Delta \mathbf{p}_C . \quad (5.108)$$

In fondo, come abbiamo già accennato, il mediatore delle interazioni alla fine è un campo, quindi A non sa chi ha beneficiato del suo sforzo, se B o C . Nel limite di intervalli di tempo infinitesimi, dall'eq. (5.108) deduciamo la relazione (5.107) tra forza Newtoniana e la variazione di impulso spaziale.⁶

Dalle equazioni (5.97) e (5.107) è immediato trovare la relazione tra quadri-forza f^μ e forza Newtoniana \mathbf{F} :

$$f^i = \frac{dp^i}{d\tau} = \frac{dp^i}{dt} \frac{dt}{d\tau} = \gamma F^i \quad (i = 1, 2, 3) \quad (5.109)$$

$$f^0 \stackrel{(5.100)}{=} \frac{\mathbf{f} \cdot \mathbf{v}}{c} = \gamma \frac{\mathbf{F} \cdot \mathbf{v}}{c} \quad (5.110)$$

e quindi

$$f^\mu = \gamma \left(\frac{\mathbf{F} \cdot \mathbf{v}}{c}, \mathbf{F} \right), \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (5.111)$$

ove v è la velocità della particella sulla quale stiamo valutando l'azione della forza.

5.3.3 Teorema delle forze vive

Sappiamo che in meccanica classica il lavoro della risultante delle forze su un corpo è uguale alla variazione di energia cinetica del corpo (teorema delle forze vive). In dinamica relativistica il lavoro della forza di Newton

$$\mathbf{F} \cdot d\mathbf{x} = \mathbf{F} \cdot \mathbf{v} dt \stackrel{(5.110)}{=} \frac{c}{\gamma} f^0 dt \stackrel{(5.97)}{=} c \frac{dp^0}{d\tau} \frac{dt}{\gamma} = c \frac{dp^0}{d\tau} d\tau \stackrel{(5.35)}{=} dE . \quad (5.112)$$

è uguale alla variazione di energia totale, e quindi alla variazione di energia cinetica relativistica. Questo risultato suffraga l'interpretazione di cp^0 come energia relativistica della particella.

Supponendo che la forza \mathbf{F} agisca su una particella inizialmente ferma, il lavoro compiuto

$$L \equiv \int \mathbf{F} \cdot d\mathbf{x} = E - mc^2 \quad (5.113)$$

è uguale all'energia cinetica (5.37) della particella.

Complessivamente, la definizione (5.97) delle tre componenti spaziali della quadri-forza fornisce la generalizzazione della legge di Newton. La definizione della quarta componente dà poi il teorema delle forze vive che, così come in meccanica classica, si può riguardare come una conseguenza delle prime tre.

Ora che ci siamo convinti che cp^0 è l'energia della particella, osserviamo che essa cresce illimitatamente per particelle con massa $m > 0$ se la velocità della particella tende alla velocità della luce. Infatti $E = mc^2 \gamma$ e, poiché il fattore di Lorentz γ tende ad infinito per $v \rightarrow c$, otteniamo

$$\lim_{v \rightarrow c} E = +\infty . \quad (5.114)$$

Quindi, per quanto lavoro si faccia, per quanta energia si dia ad un corpo con massa, non gli si potrà mai fare raggiungere la velocità della luce. La velocità della luce è una velocità limite, irraggiungibile per corpi con massa, raggiungibile (anzi, obbligatoria) per particelle a massa nulla.

⁶Se disturba il fatto di avere usato delle uguaglianze approssimate dovute al limite non-relativistico, si riacquista la fiducia osservando che l'imprecisione può essere resa piccola a piacere aumentando la massa (quindi diminuendo la velocità) del corpo B (la guida nella discussione precedente).

5.3.4 Forza e accelerazione

In meccanica classica, se è nota la forza \mathbf{F} che agisce su una particella di massa m , il moto della particella si può ricavare calcolando l'accelerazione tramite la formula di Newton $\mathbf{a} = \mathbf{F}/m$. In meccanica relativistica, la formula di Newton non è più valida. Tuttavia, è possibile anche in questo caso ricavare l'accelerazione conoscendo la forza e la velocità della particella. Dall'espressione della forza (5.107) ricaviamo

$$\mathbf{F} = \frac{d(m\gamma\mathbf{v})}{dt} = m \left[\mathbf{v} \frac{d\gamma}{dt} + \gamma \frac{d\mathbf{v}}{dt} \right]. \quad (5.115)$$

Esprimendo γ in termini della velocità \mathbf{v} della particella si ha

$$\frac{d\gamma}{dt} = \frac{d \left(1 - \frac{\mathbf{v} \cdot \mathbf{v}}{c^2} \right)^{-1/2}}{dt} = \left(1 - \frac{\mathbf{v} \cdot \mathbf{v}}{c^2} \right)^{-3/2} \frac{1}{c^2} \mathbf{v} \cdot \frac{d\mathbf{v}}{dt} \quad (5.116)$$

da cui ricaviamo la forza in funzione dell'accelerazione

$$\mathbf{F} = m\gamma \left[\mathbf{a} + \gamma^2 \frac{\mathbf{v} \cdot \mathbf{a}}{c^2} \mathbf{v} \right], \quad \mathbf{a} \equiv \frac{d\mathbf{v}}{dt}. \quad (5.117)$$

Esprimendo invece γ in termini dell'energia $E = mc^2\gamma$ si ha

$$\frac{d\gamma}{dt} \stackrel{(5.35)}{=} \frac{1}{mc^2} \frac{dE}{dt} \stackrel{(5.112)}{=} \frac{\mathbf{F} \cdot \mathbf{v}}{mc^2} \quad (5.118)$$

da cui ricaviamo l'accelerazione in funzione della forza

$$\mathbf{a} = \frac{1}{m\gamma} \left[\mathbf{F} - \frac{\mathbf{F} \cdot \mathbf{v}}{c^2} \mathbf{v} \right]. \quad (5.119)$$

Ci si rende subito conto che la forza e l'accelerazione non hanno la stessa direzione, se non nei due casi particolari di moto rettilineo ($\mathbf{v} \parallel \mathbf{a}$) o di accelerazione ortogonale alla velocità ($\mathbf{v} \perp \mathbf{a}$), come nel moto circolare uniforme. In particolare, isolando le componenti di \mathbf{F} ed \mathbf{a} parallele e perpendicolari alla velocità troviamo

$$\mathbf{F}_{\parallel} = m\gamma \left[\mathbf{a}_{\parallel} + \frac{\gamma^2 v^2}{c^2} \mathbf{a}_{\parallel} \right] = m\gamma^3 \mathbf{a}_{\parallel} \quad (5.120)$$

$$\mathbf{F}_{\perp} = m\gamma \mathbf{a}_{\perp}. \quad (5.121)$$

Possiamo interpretare quest'ultimo risultato dicendo che se una forza agisce in direzione perpendicolare alla velocità del corpo, la massa inerziale efficace vale $m\gamma$, mentre se agisce parallelamente alla velocità, la massa inerziale efficace vale $m\gamma^3$.

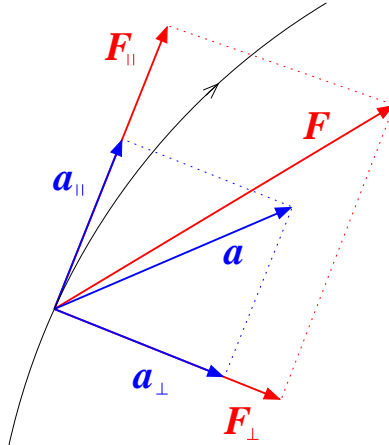


Figure 5.5: Per velocità elevate, l'accelerazione non risulta parallela alla forza.

5.3.5 Moto in campo di forza omogeneo e costante (campo elettrico)

Ci proponiamo di risolvere il problema del moto di una particella di massa m in un campo di forze omogeneo e costante. Fisicamente questo sistema può essere realizzato ponendo una carica elettrica q in un campo elettrico \mathbf{E} uniforme e costante, come vedremo nel prossimo capitolo. Supponiamo quindi che, in qualsiasi istante ed in qualsiasi punto del SDR S si trovi la particella, su di essa agisca la forza di Newton \mathbf{F} , che senza perdita di generalità possiamo pensare diretta nel verso positivo dell'asse x . Orientiamo poi l'asse y in modo che l'impulso iniziale $\mathbf{p}(0)$ giaccia nel piano (x, y) . Dall'eq. (5.107) otteniamo

$$\frac{dp_x}{dt} = F \quad (5.122)$$

$$\frac{dp_y}{dt} = 0 \quad (5.123)$$

$$\frac{dp_z}{dt} = 0 \quad (5.124)$$

che integrata nel tempo dà

$$p_x(t) = p_x(0) + Ft \quad (5.125)$$

$$p_y(t) = p_y(0) \equiv p_y \quad (5.126)$$

$$p_z(t) = 0, \quad (5.127)$$

da cui deduciamo immediatamente che il moto si svolge nel piano (x, y) . Traslando l'origine del tempo all'istante $t_0 = -p_x(0)/F$ in cui p_x si annulla, possiamo eliminare la costante $p_x(0)$ dalle formule. L'energia relativistica della particella si scrive

$$E(t) = c\sqrt{\mathbf{p}^2 + (mc)^2} = c\sqrt{(Ft)^2 + p_y^2 + m^2c^2} \quad (5.128)$$

da cui possiamo ricavare le componenti della velocità

$$\mathbf{v}(t) = c^2 \frac{\mathbf{p}(t)}{E(t)}. \quad (5.129)$$

Osserviamo che, al crescere del tempo $t \rightarrow \infty$, sia l'impulso che l'energia della particella crescono sempre di più, $|\mathbf{p}(t)|, E(t) \rightarrow +\infty$. Tuttavia la velocità della particella rimane limitata, tende alla velocità della luce $|\mathbf{v}(t)| \rightarrow c$ senza però mai raggiungerla:

$$\frac{|\mathbf{v}(t)|}{c} \stackrel{(5.129)}{=} \frac{c|\mathbf{p}(t)|}{E(t)} \stackrel{(5.128)}{=} \frac{\sqrt{(Ft)^2 + p_y^2}}{\sqrt{(Ft)^2 + p_y^2 + m^2c^2}} \xrightarrow{t \rightarrow \infty} 1^- \quad (5.130)$$

Integrando nel tempo la velocità, determiniamo la posizione all'istante t generico:

$$x(t) = \int \frac{cFt}{\sqrt{(Ft)^2 + p_y^2 + m^2c^2}} dt = \frac{c}{F} \sqrt{(Ft)^2 + p_y^2 + m^2c^2} + C_x \quad (5.131)$$

$$y(t) = \int \frac{cp_y}{\sqrt{(Ft)^2 + p_y^2 + m^2c^2}} dt = \frac{cp_y}{F} \operatorname{arcsinh} \left(\frac{Ft}{\sqrt{p_y^2 + m^2c^2}} \right) + C_y \quad (5.132)$$

ove C_x e C_y sono costanti di integrazione determinate dalla posizione iniziale della particella. Esprimendo $x(t)$ rispetto ad $y(t)$ otteniamo la traiettoria della particella

$$x(t) - C_x = \frac{c\sqrt{p_y^2 + m^2c^2}}{F} \cosh \left(\frac{y(t) - C_y}{cp_y/F} \right), \quad (5.133)$$

che è una catenaria. Nel limite non relativistico $c \rightarrow \infty$, la traiettoria si riduce ad una parabola

$$x(t) - C_x = \frac{F}{2mv_y^2} (y(t) - C_y)^2 \quad (5.134)$$

Nel caso in cui la particella non abbia velocità iniziale in direzione y , $p_y = 0$ ed il moto si svolge parallelamente all'asse x :

$$x(t) = \frac{c}{F} \sqrt{(Ft)^2 + m^2 c^2} + C_x , \quad (5.135)$$

e corrisponde al moto con “accelerazione propria” $a_0 = F/m$ costante.

Esercizio: Dimostrare l'ultima affermazione, ossia che nel SDR inerziale istantaneamente commovente alla particella, la quadri-accelerazione è costante.

Chapter 6

Elettrodinamica

6.1 Equazioni dell'elettromagnetismo classico

La teoria dell'elettromagnetismo descrive una ampissima varietà di processi che abbiamo tutti i giorni sotto gli occhi. Gli enti fisici fondamentali di questa descrizione sono:

- due campi vettoriali $\mathbf{E}(t, \mathbf{x})$ e $\mathbf{B}(t, \mathbf{x})$ chiamati rispettivamente *campo elettrico* e *campo magnetico*;
- particelle, o più in generale corpi, dotati di una particolare proprietà chiamata *carica elettrica*.

Le relazioni tra questi enti, ossia l'effetto dei campi sulle cariche e delle cariche sui campi, sono descritte da un insieme di equazioni.¹

La forza di Lorentz

$$\mathcal{F} = q(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}) . \quad (6.1)$$

determina la forza che il campo elettromagnetico esercita su una particella di carica q e velocità \mathbf{v} .

Le 8 equazioni di Maxwell determinano invece il campo elettromagnetico

$$\text{primo gruppo (omogenee)} \quad \left\{ \begin{array}{l} \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \end{array} \right. \quad (6.2a)$$

$$(6.2b)$$

$$\text{secondo gruppo (con sorgenti)} \quad \left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = \rho \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{J}}{c} \end{array} \right. \quad (6.2c)$$

$$(6.2d)$$

Il primo gruppo è costituito da 4 equazioni che eguagliano a zero delle funzioni dei soli campi elettrico e magnetico, determinando così le proprietà “intrinseche” di questi.

Il secondo gruppo è costituito da 4 equazioni che specificano come le densità di carica ρ e di corrente \mathbf{J} influenzano il campo elettromagnetico.

Notiamo che dalle equazioni di Maxwell segue l'equazione di continuità della carica

$$\frac{\partial \rho}{\partial t} \stackrel{(6.2c)}{=} \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} \stackrel{(6.2d)}{=} \nabla \cdot (c \nabla \times \mathbf{B} - \mathbf{J}) = -\nabla \cdot \mathbf{J} , \quad (6.3)$$

essendo la divergenza di un rotore sempre nulla.

La costanza della velocità della luce in tutti i SDR inerziali, il fallimento di rivelare il moto assoluto tramite esperimenti di ottica, ed il fallimento di modificare le equazioni di Maxwell in modo da renderle covarianti rispetto alle trasformazioni di Galileo, ci fa sospettare che le equazioni di Maxwell siano esatte anche dal punto di vista relativistico, siano cioè covarianti per trasformazioni di Lorentz. Se questo è il caso, vorremmo scrivere le equazioni di Maxwell in forma covariante a vista, ossia come una uguaglianza tra tensori.

¹In questo capitolo usiamo le unità di misura razionalizzate di Heaviside-Lorentz. La definizione di queste unità rispetto al sistema MKSA è riassunta in app. D.

La struttura delle equazioni (6.2) ci suggerisce che ciascun gruppo possa essere riscritto sotto forma di due equazioni quadri-vettoriali. Per sviluppare questa idea, studiamo il comportamento delle densità di carica e corrente per trasformazioni di Lorentz.

6.2 Densità di carica e corrente

Stabiliamo innanzitutto un importantissimo risultato sperimentale: *l'invarianza della carica elettrica*. Questo significa che la carica elettrica di un corpo in moto è uguale alla carica dello stesso corpo qualora sia in quiete. Un modo per convincersi di (e stabilire sperimentalmente) questa invarianza, si basa sui seguenti fatti: 1) la carica elettrica è quantizzata; 2) la carica elettrica degli elettroni e dei protoni è uguale (ed opposta); 3) il moto degli elettroni e dei protoni all'interno degli atomi è diverso, e varia tra le specie di atomi. Se la carica elettrica non fosse un invariante relativistico, non potrebbe essere vero che tutte le specie di atomi siano neutri. La neutralità degli atomi è quindi un forte argomento a supporto dell'invarianza della carica elettrica.

Consideriamo adesso nel SDR \underline{S} un elemento di volume cubico $d^3\underline{x}$ in quiete, all'interno del quale sia distribuita uniformemente una carica elettrica $d\underline{Q}$ anch'essa in quiete. A questo corpo infinitesimo siamo soliti associare la densità di carica e di corrente

$$\underline{\rho} = \frac{d\underline{Q}}{d^3\underline{x}}, \quad \underline{\mathbf{J}} = \mathbf{0} \quad (6.4)$$

in cui la densità di corrente è nulla poiché non ci sono cariche in moto. Nel SDR S in moto rispetto ad \underline{S} con velocità $-\mathbf{v}$, l'elemento di volume è in moto a velocità \mathbf{v} , e quindi esso è contratto di un fattore $\sqrt{1 - v^2/c^2}$ nella direzione del moto. Dall'invarianza della carica elettrica $dQ = d\underline{Q}$ deduciamo che la densità di carica in S vale

$$\rho \equiv \frac{dQ}{d^3x} = \frac{d\underline{Q}}{\sqrt{1 - \frac{v^2}{c^2}} d^3\underline{x}} = \gamma \underline{\rho} \quad (6.5)$$

e la densità di corrente vale

$$\mathbf{J} \equiv \rho \mathbf{v} = \gamma \underline{\rho} \mathbf{v}. \quad (6.6)$$

Osserviamo che la densità di carica trasforma come la componente temporale di un quadri-vettore, in quanto acquista un fattore γ rispetto al SDR in cui è a riposo. Proviamo allora a mettere insieme densità di carica e di corrente, nel tentativo di formare un quadri-vettore

$$J^\mu \equiv (c\rho, \mathbf{J}) = \underline{\rho} \gamma(c, \mathbf{v}) \stackrel{(5.7)}{=} \underline{\rho} u^\mu. \quad (6.7)$$

Abbiamo davvero ottenuto un quadri-vettore, chiamato (densità di) **quadri-corrente**, proporzionale alla quadri-velocità u^μ del volumetto di carica, con a fattore la densità di carica $\underline{\rho}$ nel sistema a riposo, che per definizione è un invariante.

6.2.1 La quadri-corrente per un sistema di cariche puntiformi

La definizione della quadri-corrente nella teoria relativistica si formalizza in modo più rigoroso in termini di particelle, in quanto, come osservato all'inizio del cap. 5, in relatività non ha senso parlare di corpi rigidi, e conviene trattare i corpi estesi come formati da particelle puntiformi. Consideriamo dunque un insieme di N particelle, ciascuna con carica elettrica $q_r : r = 1, \dots, N$ il cui moto in S è descritto dalla linea universo

$$\Gamma_r : t \mapsto X_r^\mu(t). \quad (6.8)$$

Per descrivere la densità di carica di particelle puntiformi è necessario ricorrere alla teoria delle distribuzioni. Una carica unitaria in posizione spaziale \mathbf{X} corrisponde alla densità di distribuzione $\rho(\mathbf{x}) = \delta^3(\mathbf{x} - \mathbf{X})$, dove la delta di Dirac tridimensionale è data dal prodotto diretto di 3 delta di Dirac unidimensionali

$$\delta^3(\mathbf{x}) \equiv \delta(x^1)\delta(x^2)\delta(x^3). \quad (6.9)$$

In ogni caso, la proprietà fondamentale che definisce la distribuzione δ^3 è la seguente: data una funzione di prova $\mathbf{x} \mapsto f(\mathbf{x})$ (in particolare continua), si ha

$$\int d^3x f(\mathbf{x}) \delta^3(\mathbf{x} - \mathbf{X}) \equiv f(\mathbf{X}) . \quad (6.10)$$

Quindi, per il sistema di cariche che stiamo considerando, la densità di carica è data da

$$\rho(t, \mathbf{x}) \equiv \sum_{r=1}^N q_r \delta^3(\mathbf{x} - \mathbf{X}_r(t)) , \quad (6.11)$$

e la densità di corrente si otterrà moltiplicando ciascun termine della densità di carica per la velocità \mathbf{v}_r della particella in questione:

$$\mathbf{J}(t, \mathbf{x}) \equiv \sum_{r=1}^N q_r \delta^3(\mathbf{x} - \mathbf{X}_r(t)) \frac{d\mathbf{X}_r(t)}{dt} . \quad (6.12)$$

Sfruttando il fatto che $X_r^0(t) = ct$ e quindi $dX_r^0(t)/dt = c$, possiamo definire con un'unica formula la quadri-corrente nella maniera dell'eq. (6.7):

$$J^\mu(ct, \mathbf{x}) \equiv (c\rho(t, \mathbf{x}), \mathbf{J}(t, \mathbf{x})) = \sum_{r=1}^N q_r \delta^3(\mathbf{x} - \mathbf{X}_r(t)) \frac{dX_r^\mu(t)}{dt} , \quad (6.13)$$

in cui abbiamo indicato la dipendenza temporale di J^μ tramite la coordinata omogenea $ct \equiv x^0$.

Per dimostrare che la precedente equazione definisce un quadri-vettore, consideriamo le variabili X_r^0 indipendenti da t , integrate con una delta di Dirac che le valuti al valore fisico ct

$$\begin{aligned} J^\mu(ct, \mathbf{x}) &= \sum_{r=1}^N q_r \int dX_r^0 \delta(ct - X_r^0) \delta^3(\mathbf{x} - \mathbf{X}_r(X_r^0)) \frac{dX_r^\mu(X_r^0)}{d(X_r^0/c)} \\ &= c \sum_{r=1}^N q_r \int dX_r^0 \delta^4(x - X_r(X_r^0)) \frac{dX_r^\mu(X_r^0)}{dX_r^0} , \end{aligned} \quad (6.14)$$

in cui si sono unite le delta di Dirac nella versione quadri-dimensionale: $\delta^4(x) \equiv \delta(x^0)\delta(x^1)\delta(x^2)\delta(x^3)$. Per ogni particella, l'integrale in X_r^0 può essere interpretato come un integrale lungo la linea universo Γ_r , con elemento di linea infinitesimo dX_r^μ , in quanto l'ultimo integrale nell'eq. (6.14) è invariante per riparametrizzazioni. Formalmente, possiamo semplificare i due differenziali dX_r^0 ottenendo

$$J^\mu(x) = c \sum_{r=1}^N q_r \int_{\Gamma_r} dX_r^\mu \delta^4(x - X_r) \quad (6.15)$$

Il carattere quadri-vettoriale della quadri-corrente è a questo punto una semplice conseguenza della natura scalare delle cariche q_r e della δ^4 e della natura quadri-vettoriale di dX_r^μ : nel SDR S' in cui $x' = \Lambda x + a$ abbiamo

$$\begin{aligned} J'^\mu(x') &= c \sum_r q_r \int_{\Gamma'_r} dX_r'^\mu \delta^4(x' - X'_r) = c \sum_r q_r \int_{\Gamma_r} \Lambda^\mu{}_\nu dX_r^\nu \delta^4(\Lambda(x - X_r)) \\ &= \Lambda^\mu{}_\nu c \sum_r q_r \int_{\Gamma_r} dX_r^\nu \frac{\delta^4(x - X_r)}{|\det(\Lambda)|} = \Lambda^\mu{}_\nu J^\nu(x) , \end{aligned} \quad (6.16)$$

essendo $\det \Lambda = \pm 1$ per trasformazioni di Poincaré.

Verifichiamo che la quadri-corrente soddisfa l'equazione di continuità (6.3): per una sola carica unitaria $q_r = 1$ avremmo che

$$\begin{aligned} \frac{\partial \rho_{\text{un}}}{\partial t} &= \frac{d}{dt} \delta^3(\mathbf{x} - \mathbf{X}_r(t)) = \frac{\partial \delta^3}{\partial x^i}(\mathbf{x} - \mathbf{X}_r(t)) \frac{d}{dt} (-X_r^i(t)) \\ &= -\frac{\partial}{\partial x^i} \left[\delta^3(\mathbf{x} - \mathbf{X}_r(t)) \frac{dX_r^i(t)}{dt} \right] = -\nabla \cdot \mathbf{J}_{\text{un}} . \end{aligned} \quad (6.17)$$

Moltiplicando per la carica elettrica q_r e sommando su tutte le particelle, riproduciamo l'equazione di continuità.

6.2.2 Conservazione della carica elettrica

In termini della quadri-corrente, l'equazione di continuità (6.3) della carica elettrica diventa

$$0 = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = \partial_0 J^0 + \partial_i J^i \quad \implies \quad \partial_\mu J^\mu = 0 \quad (6.18)$$

che è covariante a vista.

Ricordiamo che, per distribuzioni di carica generali, non necessariamente composte da un numero finito di particelle, in particolare per distribuzioni continue, purché contenute ad ogni istante in un volume finito V_t e soddisfacenti l'equazione di continuità (6.18), la carica totale — che si ottiene integrando la componente temporale di J^μ

$$Q(t) \equiv \int d^3x \rho(t, \mathbf{x}) = \frac{1}{c} \int d^3x J^0(t, \mathbf{x}) , \quad (6.19)$$

è costante nel tempo. Infatti si ha

$$\frac{d}{dt} Q(t) = \int_{V_t} d^3x \frac{\partial \rho(t, \mathbf{x})}{\partial t} = - \int_{V_t} d^3x \nabla \cdot \mathbf{J}(t, \mathbf{x}) = - \int_{S_t} dS \hat{\mathbf{n}} \cdot \mathbf{J}(t, \mathbf{x}) = 0 . \quad (6.20)$$

L'ultimo integrale, che esprime il flusso della densità di corrente \mathbf{J} attraverso la superficie S_t che racchiude V_t , è stato ottenuto tramite il teorema della divergenza. Per ipotesi, la corrente è nulla in S_t , essendo diversa da zero solo all'interno di V_t , da cui il risultato.

Si potrebbe dimostrare che Q è scalare di Lorentz come conseguenza del fatto che J^μ è un quadri-vettore a quadridivergenza nulla, così da rendere consistente l'ipotesi di invarianza della carica affermata all'inizio della sez. 6.2, ma per questo rimandiamo a S. Weinberg, citato in bibliografia, pag. 40-41.

6.3 Forma covariante a vista delle equazioni di Maxwell

Stabilito che densità di carica e corrente formano un quadri-vettore, cerchiamo di scrivere i primi membri del secondo gruppo come componenti di un quadri-vettore. La struttura dei primi membri è una somma su derivate prime dei campi. La struttura più semplice di ente covariante a vista che coinvolge somme di derivate prime producendo un quadri-vettore è data da

$$\partial_\mu F^{\mu\nu} = \frac{J^\nu}{c} , \quad (6.21)$$

in cui $F^{\mu\nu}$ è un tensore doppio (da non confondere con il simbolo della forza di Lorentz \mathcal{F}) che contiene le componenti di \mathbf{E} e \mathbf{B} , ed il fattore $1/c$ è stato messo per analogia con l'eq. (6.2d). Per riprodurre il secondo gruppo di equazioni di Maxwell, basta riempire le 4 colonne di $F^{\mu\nu}$ con le opportune componenti dei campi. Cominciamo con la componente $\nu = 0$:

$$\partial_\mu F^{\mu 0} = \frac{J^0}{c} = \rho \stackrel{(6.2c)}{=} \nabla \cdot \mathbf{E} = \partial_i E^i \quad (6.22)$$

è soddisfatta purché

$$F^{00} = 0 , \quad F^{i0} = E^i . \quad (6.23)$$

Le componenti spaziali $\nu = i$:²

$$\partial_\mu F^{\mu i} = \partial_0 F^{0i} + \partial_j F^{ji} = \frac{J^i}{c} \stackrel{(6.2d)}{=} -\frac{1}{c} \frac{\partial E^i}{\partial t} + (\nabla \times \mathbf{B})^i = -\partial_0 E^i + \varepsilon^{ijk} \partial_j B^k \quad (6.24)$$

$$[\text{esplicitamente: } \partial_\mu F^{\mu 1} = \partial_0 F^{01} + \partial_j F^{j1} = \frac{J^1}{c} \stackrel{(6.2d)}{=} -\frac{1}{c} \frac{\partial E^1}{\partial t} + (\nabla \times \mathbf{B})^1 = -\partial_0 E^1 + \partial_2 B^3 - \partial_3 B^2]$$

sono soddisfatte purché

$$F^{0i} = -E^i , \quad F^{ji} = \varepsilon^{ijk} B^k = -\varepsilon^{jik} B^k . \quad (6.25)$$

²Nelle formule che seguiranno, coppie di indici spaziali ripetuti saranno sempre da considerare sommati anche quando non siano uno in alto ed uno in basso. Resta comunque valida la regola che abbassare o alzare un'indice spaziale comporta un cambio di segno. Il tensore di Levi-Civita tri-dimensionale sarà sempre scritto con indici in alto, con la convenzione che $\varepsilon^{123} = +1$. In particolare, avremo la seguente relazione tra i tensori di Levi-Civita tri-dimensionale e quadri-dimensionale: $\varepsilon^{0ijk} = \varepsilon^{ijk}$.

Raccogliendo tutte le componenti del tensore $F^{\mu\nu}$ otteniamo la matrice

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ +E^1 & 0 & -B^3 & +B^2 \\ +E^2 & +B^3 & 0 & -B^1 \\ +E^3 & -B^2 & +B^1 & 0 \end{pmatrix} \quad (6.26)$$

che è evidentemente antisimmetrica:

$$F^{\mu\nu} = -F^{\nu\mu} . \quad (6.27)$$

Siamo così arrivati ad un risultato molto importante: la scrittura del secondo gruppo delle equazioni di Maxwell in forma covariante a vista. In sintesi: definendo il tensore doppio antisimmetrico $F^{\mu\nu}$ come nell'eq. (6.26), il secondo gruppo di equazioni di Maxwell è dato dalla formula (6.21), evidentemente covariante a vista.

La covarianza dell'eq. (6.21) è subordinata al fatto che $F^{\mu\nu}$ trasformi come un tensore doppio controvariante, secondo la definizione (4.32)

$$F'^{\mu\nu}(x') = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta}(x) \quad \Longleftrightarrow \quad F' = \Lambda F \Lambda^T . \quad (6.28)$$

Per quanto riguarda le rotazioni, dobbiamo *controllare* che le componenti E^i e B^i di $F^{\mu\nu}$ trasformino come tri-vettori. Per quanto riguarda le spinte, la legge di trasformazione (6.28) ci permette di *dedurre* le leggi di trasformazione del campo elettrico e magnetico tra SDR in moto relativo. Rimandiamo questa questione alla sez. (6.5).

Passiamo ora a scrivere il primo gruppo delle equazioni di Maxwell in forma covariante a vista. Anche per queste 4 equazioni cerchiamo un'equazione quadri-vettoriale, che coinvolga derivate prime (il tensore ∂_β) ed i campi (il tensore $F^{\mu\nu}$). Cerchiamo quindi di formare un quadri-vettore a partire dal tensore $\partial_\beta F^{\mu\nu}$. Un quadri-vettore lo abbiamo già formato per contrazione degli indici β e μ , per dare luogo all'eq. (6.21). La contrazione degli indici β e ν riproduce lo stesso quadri-vettore, con segno opposto, a causa dell'antisimmetria di $F^{\mu\nu}$. La contrazione degli indici μ e ν (uno preventivamente abbassato) invece dà luogo ad il tensore identicamente nullo, in quanto $F^\mu{}_\mu = 0 \ \forall \mu$ (μ non sommato). Possiamo però sfruttare il tensore completamente antisimmetrico di Levi-Civita (4.44) e contrarne 3 indici con $\partial_\beta F^{\mu\nu}$:

$$K^\alpha \equiv \varepsilon^{\alpha\beta\mu\nu} \partial_\beta F_{\mu\nu} \quad (6.29)$$

definisce un quadri-vettore K^α . Calcoliamone le componenti: quella temporale

$$K^0 = \varepsilon^{0\beta\mu\nu} \partial_\beta F_{\mu\nu} = \varepsilon^{ijk} \partial_i F_{jk} = \varepsilon^{ijk} \partial_i F^{jk} \stackrel{(6.25)}{=} -\varepsilon^{ijk} \varepsilon^{jkm} \partial_i B^m \stackrel{(A.32)}{=} -2\delta_{im} \partial_i B^m = -2\nabla \cdot \mathbf{B} \quad (6.30)$$

$$\begin{aligned} \text{[esplicitamente: } K^0 &= \varepsilon^{0\beta\mu\nu} \partial_\beta F_{\mu\nu} \\ &= \varepsilon^{0123} \partial_1 F_{23} + \varepsilon^{0132} \partial_1 F_{32} + \varepsilon^{0213} \partial_2 F_{13} + \varepsilon^{0231} \partial_2 F_{31} + \varepsilon^{0312} \partial_3 F_{12} + \varepsilon^{0321} \partial_3 F_{21} \\ &= -2(\partial_1 B^1 + \partial_2 B^2 + \partial_3 B^3) = -2\nabla \cdot \mathbf{B} \quad] \end{aligned}$$

è proporzionale alla divergenza di \mathbf{B} , mentre la i -esima componente spaziale

$$\begin{aligned} K^i &= \varepsilon^{i\beta\mu\nu} \partial_\beta F_{\mu\nu} = \varepsilon^{i0jk} \partial_0 F_{jk} + \varepsilon^{ij0k} \partial_j F_{0k} + \varepsilon^{ijk0} \partial_j F_{k0} \stackrel{(6.25)}{=} -\varepsilon^{ijk} \partial_0 (-\varepsilon^{jkm}) B^m + 2\varepsilon^{ijk} \partial_j F_{0k} \\ &\stackrel{(6.23)}{=} 2\delta_{im} \partial_0 B^m + 2\varepsilon^{ijk} \partial_j E^k = 2[\partial_0 B^i + (\nabla \times \mathbf{E})^i] \end{aligned} \quad (6.31)$$

è proporzionale alla i -esima componente della legge di Faraday (6.2b). In sintesi, il primo gruppo di equazioni di Maxwell si scrive

$$\varepsilon^{\alpha\beta\mu\nu} \partial_\beta F_{\mu\nu} = 0 \quad (6.32)$$

che è covariante a vista.

Osservazione: Definendo il *tensore duale elettromagnetico*

$${}^*F^{\alpha\beta} \equiv \frac{1}{2} \varepsilon^{\alpha\beta\mu\nu} F_{\mu\nu} \quad (6.33)$$

il primo gruppo di equazioni si scrive anche

$$\partial_\mu {}^*F^{\mu\nu} = 0. \quad (6.34)$$

Esplicitamente, il tensore duale vale

$${}^*F^{\mu\nu} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ +B^1 & 0 & +E^3 & -E^2 \\ +B^2 & -E^3 & 0 & +E^1 \\ +B^3 & +E^2 & -E^1 & 0 \end{pmatrix} \quad (6.35)$$

e si ottiene dal tensore elettromagnetico rimpiazzando

$$\mathbf{E} \rightarrow \mathbf{B}, \quad \mathbf{B} \rightarrow -\mathbf{E}. \quad (6.36)$$

In altre parole, nel tensore duale i campi elettrico e magnetico hanno ruoli scambiati. Il fatto che nell'eq. (6.21) ci sia una quadri-corrente mentre nell'eq. (6.32) ci sia zero corrisponde al fatto che esistono cariche (e correnti) elettriche ma non esistono cariche (né correnti) magnetiche, almeno per quanto ne sappiamo attualmente. Questo genera una profonda asimmetria tra il campo elettrico e quello magnetico.

Tuttavia, nel vuoto ($J^\mu = 0$) c'è completa simmetria tra il campo elettrico e magnetico: la *trasformazione di dualità* (6.36) non modifica le equazioni del moto dei campi, le quali sono “indistinguibili”

$$\nabla \cdot \mathbf{E} = 0 \quad (6.37a)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (6.37b)$$

$$\nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (6.37c)$$

$$\nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = 0 \quad (6.37d)$$

tenendo conto che il segno dei campi è puramente convenzionale.

6.4 Forma covariante a vista della forza di Lorentz

La forza di Lorentz (6.1) si può riscrivere in termini del tensore di campo $F^{\mu\nu}$ nel seguente modo:

$$\frac{\mathcal{F}^i}{e} = E^i + \left(\frac{\mathbf{v}}{c} \times \mathbf{B} \right)^i = E^i + \frac{1}{c} \varepsilon^{ijk} v^j B^k \stackrel{(6.25)}{=} F^{i0} - \frac{1}{c} F^{ij} v^j = \frac{1}{c} (F^{i0} c + F^{ij} v_j) \quad (6.38)$$

dove, nell'ultima uguaglianza, l'abbassamento dell'indice spaziale della velocità ha prodotto un cambio di segno. Nell'ultimo termine cominciamo ad intravedere una contrazione del tensore di campo con le componenti covarianti della quadri-velocità della particella $u_\nu = \gamma(c, -\mathbf{v})$, non fosse per un fattore γ mancante. Il fattore γ viene introdotto considerando la quadri-forza invece della forza:

$$f^i \stackrel{(5.109)}{=} \gamma \mathcal{F}^i = \frac{e}{c} (F^{i0} \gamma c + F^{ij} \gamma v_j) = \frac{e}{c} F^{i\nu} u_\nu. \quad (6.39)$$

La componente temporale della quadri-forza vale

$$f^0 \stackrel{(5.110)}{=} \frac{\gamma}{c} \mathcal{F} \cdot \mathbf{v} = \frac{\gamma}{c} e \mathbf{E} \cdot \mathbf{v} \stackrel{(6.25)}{=} -\frac{e}{c} F^{0i} u^i = \frac{e}{c} F^{0i} u_i = \frac{e}{c} F^{0\nu} u_\nu. \quad (6.40)$$

Raccogliendo le componenti (6.39,6.40) otteniamo

$$f^\mu = \frac{e}{c} F^{\mu\nu} u_\nu \quad (6.41)$$

che è la forma covariante a vista della forza di Lorentz, e che include la legge della potenza $\mathcal{F} \cdot \mathbf{v} = e \mathbf{E} \cdot \mathbf{v}$ nella componente temporale.

È interessante osservare che la struttura tensoriale (6.41) è il modo più semplice per ottenere una quadri-forza dipendente dalla velocità. Infatti, dato un qualsiasi tensore antisimmetrico $A^{\mu\nu} = -A^{\nu\mu}$, la contrazione

con la quadri-velocità $f^\mu \equiv A^{\mu\nu}u_\nu$ definisce una possibile quadriforza, dal momento che f^μ è ortogonale alla quadri-velocità e quindi all'impulso della particella:

$$f^\mu u_\mu = A^{\mu\nu}u_\nu u_\mu = 0, \quad (6.42)$$

in quanto la contrazione di un tensore antisimmetrico $A^{\mu\nu}$ e di un tensore simmetrico $u_\mu u_\nu$ è identicamente nulla.

Esercizio: ((LL) pag. 79) determinare il moto di una carica in un campo magnetico uniforme e costante (E indicherà l'energia della particella, non il campo elettrico che qui non c'è).

Soluzione:

$$\begin{aligned} \mathbf{B} &= B\hat{z} \\ f^0 = 0 &\implies E = \text{cost} \implies \mathcal{F} \cdot \mathbf{v} = 0 \implies \mathcal{F} = m\gamma \mathbf{a} = \frac{E}{c} m \mathbf{a} \\ \mathcal{F} &= \frac{e}{c} \mathbf{v} \times \mathbf{B} \implies \frac{E}{c^2} \frac{d\mathbf{v}}{dt} = \frac{e}{c} \mathbf{v} \times \mathbf{B} \\ \implies \frac{dv_x}{dt} &= \omega v_y, \quad \frac{dv_y}{dt} = -\omega v_x, \quad \frac{dv_z}{dt} = 0, \quad \text{ove} \quad \omega \equiv \frac{eB}{E} \\ v_x(t) &= v_\perp \cos(\omega t + \alpha), \quad v_y(t) = v_\perp \sin(\omega t + \alpha), \quad v_z(t) = v_\parallel \\ (v_\parallel &= v_z \text{ velocità parallela a } z \text{ costante}) \\ (v_\perp &= \text{modulo della componente della velocità perpendicolare a } z, \text{ costante}) \\ x(t) &= x_0 + r \sin(\omega t + \alpha), \quad \text{ove} \quad r = \frac{v_\perp}{\omega} = \frac{v_\perp E}{eB} = \frac{cp_\perp}{eB} \\ y(t) &= y_0 + r \cos(\omega t + \alpha) \\ z(t) &= z_0 + v_z t. \end{aligned} \quad (6.43)$$

La carica descrive un'elica, il cui asse è diretto lungo il campo magnetico, ed il cui raggio r dipende dall'impulso trasverso al campo, oltre che dalla carica e dalla massa della particella e dall'intensità del campo. La velocità angolare ω nel piano (x, y) dipende dall'energia totale iniziale della particella; nel caso non relativistico ($E \simeq mc^2$) ω è indipendente dalla velocità della particella: $\omega \simeq eB/mc$.

6.5 Le proprietà di trasformazione dei campi

Determiniamo ora come variano le componenti del tensore elettromagnetico $F^{\mu\nu}$ per trasformazioni di Poincaré

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta g_{\mu\nu} = g_{\alpha\beta} \quad (6.44)$$

ovvero, in forma matriciale,

$$x' = \Lambda x + a, \quad \Lambda^T G \Lambda = G, \quad (6.45)$$

che possiamo interpretare indifferentemente dal punto di vista attivo (trasformazione sul sistema fisico), o passivo (cambio di SDR). La richiesta che $F^{\mu\nu}$ sia un campo tensoriale si esprime nella legge di trasformazione (6.28) che riscriviamo:

$$F'^{\mu\nu}(x') = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta}(x) \iff F' = \Lambda F \Lambda^T. \quad (6.46)$$

Ci limiteremo ad analizzare il comportamento delle trasformazioni “fondamentali”, ossia traslazioni, rotazioni proprie, spinte, parità ed inversione temporale. La generale trasformazione di Poincaré si può poi ottenere combinando un insieme opportuno delle trasformazioni suddette.

6.5.1 Trasformazione per traslazioni

Per traslazioni, $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu$, e quindi il tensore $F^{\mu\nu}$ resta invariato:

$$F'^{\mu\nu}(x+a) = F^{\mu\nu}(x). \quad (6.47)$$

Quindi, le sue componenti non variano nella traslazione, e quindi restano invariati i vettori campo elettrico e magnetico:

$$\mathbf{E}'(x+a) = \mathbf{E}(x), \quad \mathbf{B}'(x+a) = \mathbf{B}(x). \quad (6.48)$$

6.5.2 Trasformazione per rotazioni

Per rotazioni proprie, la componente temporale non viene alterata, mentre vengono trasformate le componenti spaziali:

$$\begin{aligned} x'^0 &= x^0 \\ x'^i &= R^i_j x^j, \quad R^T R = \mathbb{I}, \quad \det(R) = +1. \end{aligned} \quad (6.49)$$

La matrice di Lorentz corrispondente ha pertanto la struttura a blocchi

$$\Lambda = \left(\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & R \end{array} \right) \quad (6.50)$$

in cui $\mathbf{0}$ è una colonna di 3 zeri, ed $R \in SO(3, \mathbb{R})$ è la matrice di rotazione 3×3 dell'eq. (6.49). Scomponiamo anche la matrice F in blocchi delle stesse dimensioni:

$$F = \left(\begin{array}{c|c} 0 & -\mathbf{E}^T \\ \hline \mathbf{E} & A \end{array} \right) \quad (6.51)$$

in cui \mathbf{E} è il vettore colonna con le 3 componenti del campo elettrico, mentre A è una matrice 3×3 antisimmetrica che contiene le componenti del campo magnetico

$$A^{ij} = -\varepsilon^{ijk} B^k. \quad (6.52)$$

Mediante il prodotto matriciale (6.46), che si può svolgere a blocchi, otteniamo

$$F' \equiv \left(\begin{array}{c|c} 0 & -\mathbf{E}'^T \\ \hline \mathbf{E}' & A' \end{array} \right) = \left(\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & R^T \end{array} \right) \left(\begin{array}{c|c} 0 & -\mathbf{E}^T \\ \hline \mathbf{E} & A \end{array} \right) \left(\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & R \end{array} \right) = \left(\begin{array}{c|c} 0 & -\mathbf{E}^T R^T \\ \hline R\mathbf{E} & RAR^T \end{array} \right). \quad (6.53)$$

Ne deduciamo che il campo elettrico

$$\mathbf{E}' = R\mathbf{E} \quad \Longleftrightarrow \quad E'^i = R^i_j E^j \quad (6.54)$$

trasforma come un vettore per rotazioni. Inoltre, deduciamo anche che A^{ij} trasforma come un tensore tri-dimensionale doppio per rotazioni:

$$A' = RAR^T \quad \Longleftrightarrow \quad A'^{ij} = R^i_m R^j_n A^{mn}. \quad (6.55)$$

Del resto, se il tri-tensore A^{ij} è ottenuto per contrazione di un tri-tensore (ε^{ijk}) ed un ente a 3 componenti B^k (cfr. eq. (6.52), questo implica necessariamente che B^k è un tri-tensore, quindi un vettore per rotazioni:

$$\mathbf{B}' = R\mathbf{B} \quad \Longleftrightarrow \quad B'^i = R^i_j B^j. \quad (6.56)$$

come è dimostrato nell'app. A.

6.5.3 Trasformazione per spinte

Consideriamo ora una spinta di Lorentz con velocità v lungo l'asse x . Tanto per fissare le idee, prendiamo il punto di vista attivo, in cui le x' ed F' sono le coordinate ed i campi prodotti da un sistema che si sta muovendo

a velocità $\mathbf{v} = v\hat{x}$ rispetto ad un identico sistema specificato dalle coordinate x e dai campi F . Si ha pertanto $x' = \Lambda x$ in cui la matrice Λ ha la struttura a blocchi

$$\Lambda = \left(\begin{array}{c|c} H & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{I} \end{array} \right), \quad H = H^T = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} = \begin{pmatrix} \cosh \eta & \sinh \eta \\ \sinh \eta & \cosh \eta \end{pmatrix} \quad (6.57)$$

in cui tutti i blocchi sono matrici 2×2 : H è la matrice simmetrica che mescola le coordinate t ed x (cfr. eq. (4.30) e anche la sez. 3.1.1), \mathbb{O} è la matrice nulla, e \mathbb{I} è la matrice identità. Scomponiamo anche il tensore di campo in blocchi 2×2 :

$$F = \left(\begin{array}{c|c} \mathcal{E}_1 & -\mathcal{C}^T \\ \hline \mathcal{C} & \mathcal{B}_1 \end{array} \right), \quad (6.58)$$

ove

$$\mathcal{E}_1 = \begin{pmatrix} 0 & -E^1 \\ E^1 & 0 \end{pmatrix}, \quad \mathcal{C} = \begin{pmatrix} E^2 & B^3 \\ E^3 & -B^2 \end{pmatrix}, \quad \mathcal{B}_1 = \begin{pmatrix} 0 & -B^1 \\ B^1 & 0 \end{pmatrix}. \quad (6.59)$$

Mediante il prodotto matriciale (6.46), che si può svolgere a blocchi, otteniamo

$$F' \equiv \left(\begin{array}{c|c} \mathcal{E}'_1 & -\mathcal{C}'^T \\ \hline \mathcal{C}' & \mathcal{B}'_1 \end{array} \right) = \left(\begin{array}{c|c} H & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{I} \end{array} \right) \left(\begin{array}{c|c} \mathcal{E}_1 & -\mathcal{C}^T \\ \hline \mathcal{C} & \mathcal{B}_1 \end{array} \right) \left(\begin{array}{c|c} H^T & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{I} \end{array} \right) = \left(\begin{array}{c|c} H\mathcal{E}_1H & -H\mathcal{C}^T \\ \hline \mathcal{C}H & \mathcal{B}_1 \end{array} \right). \quad (6.60)$$

Svolgendo i prodotti dei vari blocchi troviamo:

$$\mathcal{E}'_1 = H\mathcal{E}_1H = \begin{pmatrix} 0 & -E^1 \\ E^1 & 0 \end{pmatrix} = \mathcal{E}_1 \quad (6.61)$$

$$\mathcal{C}' = \mathcal{C}H = \begin{pmatrix} E^2 \cosh \eta + B^3 \sinh \eta & E^2 \sinh \eta + B^3 \cosh \eta \\ E^3 \cosh \eta - B^2 \sinh \eta & E^3 \sinh \eta - B^2 \cosh \eta \end{pmatrix} \quad (6.62)$$

$$\mathcal{B}'_1 = \mathcal{B}_1 \quad (6.63)$$

avendo sfruttando $\cosh^2 \eta - \sinh^2 \eta = 1$ nell'eq. (6.61). Dalle equazioni (6.61,6.63) deduciamo che le componenti del campo elettrico e magnetico lungo la direzione del moto E^1 e B^1 non variano. L'equazione (6.62) invece ci insegna come trasformano le componenti dei campi perpendicolari alla direzione di moto. Complessivamente

$$\begin{aligned} E'^1 &= E^1 & B'^1 &= B^1 \\ E'^2 &= E^2 \cosh \eta + B^3 \sinh \eta & B'^2 &= B^2 \cosh \eta - E^3 \sinh \eta \\ E'^3 &= E^3 \cosh \eta - B^2 \sinh \eta & B'^3 &= B^3 \cosh \eta + E^2 \sinh \eta. \end{aligned} \quad (6.64)$$

Quindi, per spinte lungo l'asse x , le coppie (E^2, B^3) e $(E^3, -B^2)$ trasformano con una rotazione iperbolica, alla stregua della coppia (x^0, x^1) .

Esercizio: determinare i campi elettrico e magnetico prodotti da una carica elettrica in moto rettilineo uniforme a velocità v lungo l'asse x , operando una trasformazione di Lorentz dei campi prodotti dalla stessa particella quando è ferma.

6.5.4 Trasformazione per inversione degli assi

Consideriamo adesso le trasformazioni di Poincaré discrete, che alterano il segno del determinante di Λ oppure il segno di Λ^0_0 .

Parità L'inversione delle coordinate spaziali è determinata dalla matrice di parità (3.62)

$$\Lambda_{\mathcal{P}} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (6.65)$$

che si può scrivere a blocchi, come nell'eq. (6.50) con $R = -\mathbb{I}_3$, l'opposta della matrice identità tri-dimensionale. Ripetendo la procedura della sez. 6.5.2, otteniamo

$$\mathbf{E}' = -\mathbb{I}\mathbf{E} = -\mathbf{E} \quad (6.66)$$

$$\mathbf{A}' = (-\mathbb{I})\mathbf{A}(-\mathbb{I}) = \mathbf{A} \quad \implies \quad \mathbf{B}' = \mathbf{B} . \quad (6.67)$$

Ne deduciamo che il campo elettrico è un campo vettoriale *polare*, mentre il campo magnetico è uno pseudo-vettore, ovvero un campo vettoriale *assiale*, che non cambia segno per inversione degli assi spaziali.

L'interpretazione fisica per una trasformazione di parità *attiva* è la seguente: se invertiamo fisicamente un sistema, ponendo tutti i suoi corpi in posizione opposta rispetto all'origine degli assi (che lasciamo al loro posto), i valori delle cariche elettriche restano inalterate ma cambiano di posizione, dando luogo ad un campo elettrico invertito rispetto all'originale. Questa inversione però cambia il segno della velocità delle cariche, e quindi della corrente elettrica. Di conseguenza, il campo magnetico, che è dato dal prodotto vettoriale $d\mathbf{B} = (1/c) d\mathbf{I} \times \mathbf{x}/|\mathbf{x}|^2$ non cambia segno.

Inversione temporale L'inversione della coordinata temporale è determinata dalla matrice di inversione temporale (3.66)

$$\Lambda_{\mathcal{T}} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad (6.68)$$

anch'essa a blocchi, e che si può scrivere $\Lambda_{\mathcal{T}} = -\Lambda_{\mathcal{P}}$. Siccome nella trasformazione di $F^{\mu\nu}$ la matrice Λ interviene in due fattori, il segno relativo tra parità e inversione temporale si elide, e quindi otteniamo lo stesso risultato della trasformazione di parità:

$$\mathbf{E}' = -\mathbb{I}\mathbf{E} = -\mathbf{E} \quad (6.69a)$$

$$\mathbf{A}' = (-\mathbb{I})\mathbf{A}(-\mathbb{I}) = \mathbf{A} \quad \implies \quad \mathbf{B}' = \mathbf{B} . \quad (6.69b)$$

Qui le conclusioni sono opposte a quelle che ci saremmo aspettati. Infatti, invertendo segno del tempo, è naturale, conveniente e lecito credere che la carica elettrica sia immutata, e così il campo elettrico. Siccome poi la velocità inverte il verso, ci aspetteremmo che la densità di corrente e quindi il campo magnetico cambino segno. Ma se assumiamo che la carica elettrica sia un invariante per inversioni del tempo, allora la quadri-corrente è uno pseudo-vettore di II specie, quindi anche il tensore $F^{\mu\nu}$ deve essere uno pseudo-tensore di II specie, per cui le equazioni (6.69) sono sbagliate.

Viceversa, se assumiamo che $F^{\mu\nu}$ e quindi anche J^μ siano dei tensori anche per inversione temporale, dobbiamo ammettere che la carica elettrica cambia segno, ossia che la carica è uno pseudo-scalare di II specie. In questo caso, le eq. (6.69) sono giuste.

6.6 Invarianti del campo elettromagnetico

Con il tensore di campo $F^{\mu\nu}$ possiamo formare due invarianti:

$$\begin{aligned} F^{\mu\nu} F_{\mu\nu} &= -F^{i0} F^{i0} - F^{0i} F^{0i} + F^{ij} F^{ij} = -2E^i E^i + \varepsilon^{ijk} B^k \varepsilon^{ijl} B^l \\ &= -2E^i E^i + 2\delta_{kl} B^k B^l = -2E^i E^i + 2B^k B^k = -2(\mathbf{E}^2 - \mathbf{B}^2) . \end{aligned} \quad (6.70)$$

$$\begin{aligned} \varepsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu} &= \varepsilon^{0ijk} F_{0i} F_{jk} + \varepsilon^{i0jk} F_{i0} F_{jk} + \varepsilon^{ij0k} F_{ij} F_{0k} + \varepsilon^{ijk0} F_{ij} F_{k0} \\ &= 4\varepsilon^{0ijk} F_{0i} F_{jk} = 4\varepsilon^{ijk} E^i (-\varepsilon^{jkl}) B^l = -4 \cdot 2\delta_{il} E^i B^l = -8\mathbf{E} \cdot \mathbf{B} . \end{aligned} \quad (6.71)$$

Concludiamo che le due combinazioni

$$\mathbf{E}^2 - \mathbf{B}^2 , \quad \mathbf{E} \cdot \mathbf{B} \quad (6.72)$$

sono invarianti per trasformazioni di Lorentz, rispettivamente uno scalare ed uno pseudo-scalare di prima specie. In altre parole, due osservatori in moto relativo potranno misurare, in un dato evento, campi elettrici e magnetici diversi tra loro. Ciò nonostante, le due combinazioni (6.72) saranno le stesse per tali osservatori.

6.7 Il quadri-potenziale

Ricordiamo i seguenti risultati dell'analisi vettoriale tridimensionale:

$$\nabla \cdot \mathbf{V} = 0 \quad \Longleftrightarrow \quad \exists \mathbf{A} : \quad \mathbf{V} = \nabla \times \mathbf{A} \quad (6.73)$$

$$\nabla \times \mathbf{U} = \mathbf{0} \quad \Longleftrightarrow \quad \exists \phi : \quad \mathbf{U} = -\nabla \phi . \quad (6.74)$$

Consideriamo ora il primo gruppo delle equazioni di Maxwell. Dal fatto che la divergenza di \mathbf{B} è sempre nulla, esiste un campo vettoriale \mathbf{A} , detto **potenziale vettore**, tale che

$$\mathbf{B} = \nabla \times \mathbf{A} . \quad (6.75)$$

Sostituendo questa equazione nell'eq. (6.2b) otteniamo

$$\mathbf{0} = \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial}{\partial t} \nabla \times \mathbf{A} = \nabla \times \left(\mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \right) , \quad (6.76)$$

da cui

$$\mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} = -\nabla \phi \quad (6.77)$$

per un qualche campo scalare ϕ detto **potenziale scalare**, e quindi

$$\mathbf{E} = -\nabla \phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} . \quad (6.78)$$

In definitiva, il primo gruppo delle equazioni di Maxwell può essere sostituito dalle equazioni (6.75) e (6.78), che esprimono i campi elettrico e magnetico in termini dei potenziali vettori e scalare.

Cerchiamo ora di trasportare il risultato ottenuto nel formalismo quadri-vettoriale. Introduciamo il seguente campo a 4 componenti, detto **quadri-potenziale**

$$A^\mu \equiv (\phi, \mathbf{A}) . \quad (6.79)$$

Le eq. (6.75) e (6.78) si potranno scrivere allora nella forma

$$B^i = \varepsilon^{ijk} \partial_j A^k = -\varepsilon^{ijk} \partial^j A^k \quad (6.80)$$

$$E^i = -\partial_i A^0 - \partial_0 A^i = \partial^i A^0 - \partial^0 A^i . \quad (6.81)$$

A questo punto possiamo esprimere le componenti del tensore elettromagnetico $F^{\mu\nu}$ in termini del quadri-potenziale:

$$F^{i0} \stackrel{(6.23)}{=} E^i \stackrel{(6.81)}{=} \partial^i A^0 - \partial^0 A^i \quad (6.82)$$

$$F^{ij} \stackrel{(6.25)}{=} -\varepsilon^{ijk} B^k \stackrel{(6.80)}{=} -\varepsilon^{ijk} (-\varepsilon^{kmn}) \partial^m A^n \stackrel{(A.26)}{=} (\delta_{im} \delta_{jn} - \delta_{in} \delta_{jm}) \partial^m A^n = \partial^i A^j - \partial^j A^i . \quad (6.83)$$

$$[\text{esplicitamente:} \quad F^{21} = B^3 = \partial_1 A^2 - \partial_2 A^1 = -\partial^1 A^2 + \partial^2 A^1 \quad]$$

Essendo $F^{\mu\nu}$ un tensore antisimmetrico, potremo scrivere sinteticamente

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu . \quad (6.84)$$

In conclusione, la possibilità di esprimere il tensore elettromagnetico tramite un quadri-potenziale nella forma (6.84) è equivalente al primo gruppo delle equazioni di Maxwell.

Naturalmente, dobbiamo assumere che in seguito a trasformazioni di Poincaré il quadri-potenziale si trasformi come un campo quadri-vettoriale:

$$A'^\mu(x') = \Lambda^\mu{}_\nu A^\nu(x) . \quad (6.85)$$

Ciò è perfettamente equivalente al richiedere che $F^{\mu\nu}$ si trasformi come un tensore doppio, secondo l'eq. (6.28).

Pertanto, possiamo dimenticarci del primo gruppo di equazioni di Maxwell non appena pensiamo il tensore elettromagnetico espresso in termini del quadripotenziale secondo l'eq. (6.84). L'equivalenza tra l'eq. (6.84) e la scrittura covariante a vista del primo gruppo (6.32) si può provare anche per via diretta. Infatti, assumendo (6.84) abbiamo che

$$\varepsilon^{\alpha\beta\mu\nu}\partial_\beta F_{\mu\nu} = \varepsilon^{\alpha\beta\mu\nu}\partial_\beta(\partial_\mu A_\nu - \partial_\nu A_\mu) = 2\varepsilon^{\alpha\beta\mu\nu}\partial_\beta\partial_\mu A_\nu = 0 \quad (6.86)$$

per l'antisimmetria del tensore $\varepsilon^{\alpha\beta\mu\nu}$ rispetto agli indici $\beta\mu$. Il viceversa è una diretta conseguenza del teorema di Stokes dell'analisi quadrivettoriale.

6.8 Invarianza di gauge

Nella sez. 6.7 abbiamo visto che è possibile esprimere i campi elettrico e magnetico — e quindi il tensore di campo $F^{\mu\nu}$ — in termini di un campo quadri-vettoriale A^μ . Osserviamo che, dato il quadri-potenziale A^μ , i campi \mathbf{E} , \mathbf{B} e $F^{\mu\nu}$ risultano univocamente determinati. Non vale però il viceversa: esistono infiniti quadri-potenziali diversi tra loro che determinano lo stesso campo elettromagnetico.

Infatti, sia $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, e definiamo un nuovo quadri-potenziale

$$\hat{A}^\mu(x) \equiv A^\mu(x) + \partial^\mu f(x) \quad (6.87)$$

dove $f(x)$ è un arbitrario campo scalare. Il tensore di campo corrispondente al nuovo quadri-potenziale vale

$$\begin{aligned} \hat{F}^{\mu\nu} &= \partial^\mu \hat{A}^\nu - \partial^\nu \hat{A}^\mu = \partial^\mu A^\nu + \partial^\mu \partial^\nu f - \partial^\nu A^\mu - \partial^\nu \partial^\mu f \\ &= \partial^\mu A^\nu - \partial^\nu A^\mu = F^{\mu\nu}, \end{aligned} \quad (6.88)$$

in cui abbiamo sfruttato il fatto che le derivate commutano: $\partial^\mu \partial^\nu f = \partial^\nu \partial^\mu f$. Possiamo dunque affermare che il campo elettromagnetico e le equazioni dell'elettromagnetismo sono invarianti se sul quadri-potenziale eseguiamo la *trasformazione di gauge* (6.87). Questa importante proprietà è detta *invarianza di gauge*. L'indeterminazione ora evidenziata nella scelta del quadri-potenziale consente sempre di sceglierlo in maniera tale da soddisfare una condizione arbitraria supplementare, determinata da ragioni di comodità.

6.9 Equazioni di campo del quadri-potenziale

Esprimendo il campo elettromagnetico in termini del quadri-potenziale, il primo gruppo delle equazioni di Maxwell è automaticamente soddisfatto. Vediamo ora come si scrive il secondo gruppo delle equazioni di Maxwell in termini del quadri-potenziale. Dalla formula (6.21) otteniamo

$$\frac{J^\nu}{c} = \partial_\mu F^{\mu\nu} \stackrel{(6.84)}{=} \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = (\partial_\mu \partial^\mu) A^\nu - \partial^\nu(\partial_\mu A^\mu). \quad (6.89)$$

L'operatore differenziale

$$\square \equiv \partial_\mu \partial^\mu = \frac{1}{c^2} \partial_t^2 - \nabla^2 \quad (6.90)$$

è il cosiddetto *d'Alembertiano*, l'operatore differenziale dell'equazione delle onde, ed è evidentemente un operatore invariante.

Grazie all'invarianza di gauge, possiamo modificare il quadri-potenziale in modo da annullare la quadri-divergenza $\partial_\mu A^\mu$, e semplificare così l'eq. (6.89). Infatti, in seguito ad una trasformazione di gauge (6.87), la quadri-divergenza del potenziale diventa

$$\partial_\mu \hat{A}^\mu = \partial_\mu A^\mu + \partial_\mu \partial^\mu f = \partial_\mu A^\mu + \square f. \quad (6.91)$$

Quindi basta scegliere la f in modo che sia (gauge di Lorentz)

$$\partial_\mu \hat{A}^\mu = 0 \quad \Longleftrightarrow \quad \square f = -\partial_\mu A^\mu \quad (6.92)$$

per ottenere l'equazione di campo

$$\square \hat{A}^\nu = \frac{J^\nu}{c} . \quad (6.93)$$

Un'immediata conseguenza della precedente equazione è che, in una gauge opportuna, il quadri-potenziale nel vuoto ($J^\mu = 0$) soddisfa l'equazione delle onde. Questa proprietà si ripercuote immediatamente sui campi, per esempio

$$\square E^i \stackrel{(6.81)}{=} \square (\partial^i \hat{A}^0 - \partial^0 \hat{A}^i) = \partial^i \square \hat{A}^0 - \partial^0 \square \hat{A}^i \stackrel{(6.93)}{=} 0 \quad (J^\mu = 0) . \quad (6.94)$$

Pertanto il campo elettrico nel vuoto soddisfa l'equazione delle onde, e questo indipendentemente dalla scelta della gauge. Lo stesso vale per il campo magnetico:

$$\square B^i = 0 \quad (J^\mu = 0) . \quad (6.95)$$

Appendix A

Calcolo tensoriale tri-dimensionale

A.1 Vettori e tensore metrico

Nello spazio affine tri-dimensionale euclideo $(\mathcal{M}, \mathcal{V}, \cdot)$, è conveniente adottare sistemi di riferimento ortonormali, in cui i vettori di base sono a due a due ortogonali e di lunghezza unitaria. In SDR ortonormali $(O, \underline{e}_1, \underline{e}_2, \underline{e}_3)$ il tensore metrico ha la sua forma più semplice:

$$g_{ij}^{(\text{eucl})} \equiv \underline{e}_i \cdot \underline{e}_j = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} . \quad (\text{A.1})$$

Scomponendo il generico vettore $\underline{v} \in \mathcal{V}$ rispetto alla base scelta

$$\underline{v} = \underline{e}_i v^i , \quad (\text{A.2})$$

il prodotto scalare tra vettori si scrive

$$\underline{u} \cdot \underline{v} = u^i v^j \delta_{ij} = u^i v^i , \quad (\text{A.3})$$

Notazioni: le componenti covarianti rispetto alla metrica euclidea, definite da $v_i = \delta_{ij} v^j$, sono identiche alle componenti controvarianti v^i , e quindi non è molto utile introdurle. Quindi, useremo sempre indici in alto per vettori e tensori euclidei: v^i , A^{ij} , ε^{ijk} , riservando la notazione con indici in basso alle componenti covarianti nella metrica pseudo-euclidea quadri-dimensionale: $v_i = -v^i$. Sempre in conformità con la notazione pseudo-euclidea, manteniamo la posizione in basso per gli indici dei vettori di base \underline{e}_i , per il tensore metrico δ_{ij} e per gli indici colonna delle matrici di trasformazione di SDR: Λ^i_j , R^i_j .

Convenzione: per quanto stabilito nella nota precedente, trattando le componenti spaziali, sottintenderemo sempre un indice di somma per ogni coppia di indici ripetuti, anche qualora essi non si presentino uno in alto ed uno in basso.

Usando un SDR ortonormale diverso, si ha

$$\underline{e}'_j = \underline{e}_i (R^{-1})^i_j \quad (\text{A.4})$$

$$v'^i = R^i_j v^j \quad (\text{A.5})$$

ove R è una matrice ortogonale

$$R^i_m R^j_n \delta_{ij} = \delta_{mn} \quad \Longleftrightarrow \quad R^i_m R^i_n = \delta_{mn} \quad \Longleftrightarrow \quad R^T R = \mathbb{I} . \quad (\text{A.6})$$

Equivalentemente, un cambio di coordinate

$$x'^i = R^i_j x^j + a^i \quad (\text{A.7})$$

preserva la forma (A.3) del prodotto scalare euclideo se e solo se la matrice R è ortogonale. In effetti, la prima uguaglianza nell'eq. (A.6) esprime il fatto che il tensore metrico è preservato dal cambio di coordinate, ossia che esso è un tensore costante.

Contraendo i due indici del tensore metrico, formiamo uno scalare, che vale

$$\delta_{ii} = 3. \quad (\text{A.8})$$

A.2 Tensore di Levi-Civita

Chiameremo *tensori* tri-dimensionali di ordine n quegli enti a 3^n componenti $T^{i_1 \dots i_n} : i_k = 1, 2, 3$ che, per trasformazioni di coordinate (A.7), variano le loro componenti secondo la formula

$$T'^{i_1 \dots i_n} = R^{i_1}_{j_1} \dots R^{i_n}_{j_n} T^{j_1 \dots j_n} \quad (\text{A.9})$$

Un importante tensore dello spazio tridimensionale è il tensore di Levi-Civita

$$\varepsilon^{ijk} \equiv \begin{cases} +1 & \text{se } (i, j, k) \text{ è una permutazione pari di } (1, 2, 3) \\ -1 & \text{se } (i, j, k) \text{ è una permutazione dispari di } (1, 2, 3) \\ 0 & \text{se } (i, j, k) \text{ non è una permutazione di } (1, 2, 3) \end{cases} \quad (\text{A.10})$$

completamente antisimmetrico nello scambio di una qualsiasi coppia di indici.

Per prima cosa, vediamo se siamo autorizzati a chiamarlo tensore. La definizione (A.10) non fa riferimento ad alcun SDR, quindi in ogni SDR le componenti di questo ente sono date dal criterio (A.10). Per dimostrare che ε^{ijk} è un tensore, ricordiamo la definizione di determinante di una matrice come somma dei prodotti delle terne di elementi, presi uno per ogni riga, in cui ciascun prodotto è moltiplicato per il segno della permutazione:

$$\det(R) \equiv \sum_{i,j,k} R^1_i R^2_j R^3_k \varepsilon^{ijk}. \quad (\text{A.11})$$

Sappiamo che scambiare due righe inverte il segno del determinante, mentre se due righe sono uguali il determinante è nullo. Si ha quindi, operando una permutazione sugli indici riga $(1, 2, 3) \rightarrow (m, n, l)$

$$\det(R) \varepsilon^{mnl} = R^m_i R^n_j R^l_k \varepsilon^{ijk}. \quad (\text{A.12})$$

Quest'ultima equazione, per $\det(R) = 1$, equivale alla legge di trasformazione (A.9) di un tensore costante. Quindi, per rotazioni proprie, abbiamo dimostrato che il simbolo di Levi-Civita definisce un tensore. Per rotazioni improprie ($\det(R) = -1$), come la trasformazione di parità $R = -\mathbb{I}$, la formula (A.12) differisce per un fattore $\det(R)$ dalla trasformazione (A.9). Si dice in questo caso che ε^{ijk} è uno pseudo-tensore tri-dimensionale.

A.2.1 Prodotto vettoriale

Con il (pseudo-)tensore di Levi-Civita possiamo costruire delle importanti applicazioni lineari.

Dati due vettori u^j e v^k , la terna di componenti

$$w^i \equiv \varepsilon^{ijk} u^j v^k \quad (\text{A.13})$$

definisce un vettore. Abbiamo infatti, per un cambio di coordinate (A.7)

$$w'^m \equiv \varepsilon^{mnl} u'^n v'^l \stackrel{(\text{A.12})}{=} [\det(R) R^m_i R^n_j R^l_k \varepsilon^{ijk}] [R^n_p u^p] [R^l_q v^q]. \quad (\text{A.14})$$

Sfruttando l'ortogonalità delle matrici R , abbiamo che

$$R^n_j R^n_p = \delta_{jp}, \quad R^l_k R^l_q = \delta_{kq} \quad (\text{A.15})$$

e quindi

$$w'^m = \det(R) R^m_i \varepsilon^{ijk} \delta_{jp} \delta_{kq} u^p v^q = \det(R) R^m_i \varepsilon^{ijk} u^j v^k \stackrel{(\text{A.13})}{=} \det(R) R^m_i w^i. \quad (\text{A.16})$$

Quindi, per rotazioni proprie, w^i trasforma come un vettore, mentre per rotazioni improprie acquista un segno meno. Si dice che w^i è uno pseudo-vettore, ossia, un vettore *assiale*.

Calcolando esplicitamente le componenti di w^i , per esempio

$$w^1 = \varepsilon^{123} u^2 v^3 + \varepsilon^{132} u^3 v^2 = u^2 v^3 - u^3 v^2, \quad (\text{A.17})$$

riconosciamo nella definizione (A.13) la formula del prodotto vettoriale:

$$\underline{w} = \underline{u} \times \underline{v}. \quad (\text{A.18})$$

L'interpretazione geometrica del prodotto vettoriale è nota: esso è quel vettore che ha per modulo l'area del parallelogramma individuato dai vettori \underline{u} e \underline{v}

$$|\underline{w}| = |\underline{u}| |\underline{v}| \sin \theta_{\underline{u}\underline{v}}, \quad (\text{A.19})$$

per direzione la normale al piano $\langle \underline{u}, \underline{v} \rangle$, e verso tale che $(\underline{u}, \underline{v}, \underline{w})$ formino una terna destrorsa.

Un'altra interpretazione è questa: costruito il parallelogramma individuato dai vettori \underline{u} e \underline{v} , lo si proietti sul piano coordinato $\langle j, k \rangle$, e se ne calcoli l'area orientata. Quest'area orientata vale w^i — con (i, j, k) permutazione pari di $(1, 2, 3)$.

A.2.2 Prodotto misto

Dati 3 vettori u^i , v^j e w^k , possiamo costruire un invariante tramite il tensore di Levi-Civita:

$$\omega = \varepsilon^{ijk} u^i v^j w^k. \quad (\text{A.20})$$

Procedendo come nel caso precedente, si dimostra che, per un cambio di coordinate (A.7),

$$\omega' = \det(R) \omega, \quad (\text{A.21})$$

ossia ω è uno pseudo-scalare.

Avendo identificato $\varepsilon^{ijk} v^j w^k = (\underline{v} \times \underline{w})^i$, abbiamo che

$$\omega = u^i (\underline{v} \times \underline{w})^i = \underline{u} \cdot (\underline{v} \times \underline{w}) \quad (\text{A.22})$$

corrisponde al prodotto misto dei tre vettori dati. Dall'antisimmetria ed invarianza per permutazioni cicliche degli indici di ε^{ijk} , deduciamo che il prodotto misto è invariante per permutazioni cicliche dei vettori, ed antisimmetrico nello scambio di una coppia dei vettori.

L'interpretazione geometrica del prodotto misto è nota: esso dà il volume orientato del parallelepipedo individuato dai vettori $(\underline{u}, \underline{v}, \underline{w})$.

A.2.3 Contrazioni del tensore di Levi-Civita

Spesso compaiono delle formule in cui compaiono due tensori di Levi-Civita con uno, due o tre indici sommati. Ad esempio, volendo calcolare il doppio prodotto vettoriale

$$\underline{u} \times (\underline{v} \times \underline{w}) \quad (\text{A.23})$$

avremmo

$$[\underline{u} \times (\underline{v} \times \underline{w})]^i = \varepsilon^{ijk} u^j (\underline{v} \times \underline{w})^k = \varepsilon^{ijk} u^j [\varepsilon^{klm} v^l w^m]. \quad (\text{A.24})$$

La somma su k coinvolge solamente

$$\varepsilon^{ijk} \varepsilon^{klm} = \varepsilon^{ijk} \varepsilon^{lmk} \quad (\text{A.25})$$

Fissati i e j , se $i = j$ il risultato è 0. Se $i \neq j$, allora l'unico termine della somma che contribuisce è quello con k diverso sia da i che da j . Allo stesso modo, per ottenere un risultato diverso da 0, è necessario che l ed m siano

diversi tra loro e diversi da k . Quindi $\{i, j\} = \{l, m\}$ come insiemi, e si presentano due casi: 1) $i = l$ e $j = m$, e l'unico termine della somma è il quadrato di ε^{ijk} che dà 1; 2) $i = l$ e $j = l$, e si ottiene il risultato opposto, cioè -1 . Tutto questo si può riassumere nella formula

$$\varepsilon^{ijk}\varepsilon^{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} = \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix}. \quad (\text{A.26})$$

Tornando al doppio prodotto vettoriale, ricaviamo

$$[\underline{u} \times (\underline{v} \times \underline{w})]^i = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})u^j v^l w^m = u^j v^i w^j - u^j v^j w^i = (\underline{u} \cdot \underline{w})v^i - (\underline{u} \cdot \underline{v})w^i \quad (\text{A.27})$$

e quindi

$$\underline{u} \times (\underline{v} \times \underline{w}) = \underline{v}(\underline{u} \cdot \underline{w}) - \underline{w}(\underline{u} \cdot \underline{v}) \quad (\text{A.28})$$

Un metodo più sistematico, che si può applicare anche a tensori di Levi-Civita n -dimensionali, è quello di scrivere tali tensori mediante determinanti di delta di Kronecker. Infatti, la definizione (A.10) è equivalente a

$$\varepsilon^{ijk} = \begin{vmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ \delta_{j1} & \delta_{j2} & \delta_{j3} \\ \delta_{k1} & \delta_{k2} & \delta_{k3} \end{vmatrix} \quad (\text{A.29})$$

da cui, permutando le colonne $(1, 2, 3) \rightarrow (l, m, n)$

$$\varepsilon^{ijk}\varepsilon^{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}. \quad (\text{A.30})$$

Sviluppando il determinante, si può scrivere il prodotto di due tensori di Levi-Civita come somma di prodotti di delta di Kronecker.

La contrazione sugli k ed n si può calcolare agevolmente sviluppando il determinante rispetto all'ultima riga:

$$\begin{aligned} \varepsilon^{ijk}\varepsilon^{lmk} &= \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{ik} \\ \delta_{jl} & \delta_{jm} & \delta_{jk} \\ \delta_{kl} & \delta_{km} & \delta_{kk} \end{vmatrix} = \delta_{kl} \begin{vmatrix} \delta_{im} & \delta_{ik} \\ \delta_{jm} & \delta_{jk} \end{vmatrix} - \delta_{km} \begin{vmatrix} \delta_{il} & \delta_{ik} \\ \delta_{jl} & \delta_{jk} \end{vmatrix} + \delta_{kk} \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix} \\ &= \begin{vmatrix} \delta_{im} & \delta_{il} \\ \delta_{jm} & \delta_{jl} \end{vmatrix} - \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix} + 3 \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix} = \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix}, \end{aligned} \quad (\text{A.31})$$

in accordo con l'eq. (A.26).

La contrazione su due indici è presto fatta:

$$\varepsilon^{ijk}\varepsilon^{ljk} \stackrel{(\text{A.26})}{=} \delta_{il}\delta_{jj} - \delta_{ij}\delta_{jl} = 3\delta_{il} - \delta_{il} = 2\delta_{il}. \quad (\text{A.32})$$

Infatti, se i deve essere diverso da j e da k , e anche l deve essere diverso da j e da k , con $j \neq k$, allora deve essere $i = l$. Fissato $i = l$, per esempio $i = l = 1$, ci sono due termini che contribuiscono: $(j, k) = (2, 3)$ e $(j, k) = (3, 2)$, ciascuno con “peso” 1, da cui il fattore due nella (A.32).

Infine, contraendo tutti e tre gli indici otteniamo lo scalare

$$\varepsilon^{ijk}\varepsilon^{ijk} \stackrel{(\text{A.32})}{=} 2\delta_{ii} = 6. \quad (\text{A.33})$$

Il primo membro si può interpretare come la somma dei quadrati di tutte le componenti del tensore. Siccome nel tensore di Levi-Civita le componenti possono assumere solo i valori $-1, 0, 1$, ne deduciamo che il numero di componenti non nulle di ε^{ijk} è 6 (su 64).

A.3 Tensori doppi antisimmetrici

Un tensore doppio antisimmetrico $A^{ij} = -A^{ji}$ è rappresentabile da una matrice 3×3 antisimmetrica $A = -A^T$. Gli elementi della diagonale principale sono necessariamente nulli, e i 3 elementi sotto la diagonale principale sono l'opposto dei 3 elementi sopra tale diagonale. Quindi un tensore doppio antisimmetrico ha solo 3 parametri indipendenti. Infatti, il più generale tensore doppio antisimmetrico si può rappresentare nella forma

$$A^{ij} = \varepsilon^{ijk} v^k \quad (\text{A.34})$$

con tre opportuni coefficienti v^k . La cosa interessante è che, se in un nuovo SDR rappresentiamo le nuove componenti del tensore A allo stesso modo di prima, cioè

$$A'^{lm} = \varepsilon'^{lmn} v'^n \quad (\text{A.35})$$

la nuova terna di coefficienti v'^n trasforma come uno pseudo-vettore:

$$v'^n = R^n_k v^k \det(R) . \quad (\text{A.36})$$

Infatti

$$\begin{aligned} A'^{lm} &\stackrel{(\text{A.9})}{=} R^l_i R^m_j A^{ij} \stackrel{(\text{A.34})}{=} R^l_i R^m_j \varepsilon^{ijk} v^k = R^l_i R^m_j \varepsilon^{ijk} \delta_{kp} v^p \\ &\stackrel{(\text{A.6})}{=} R^l_i R^m_j \varepsilon^{ijk} R^n_k R^n_p v^p \stackrel{(\text{A.12})}{=} \det(R) \varepsilon'^{lmn} R^n_p v^p \end{aligned} \quad (\text{A.37})$$

che, confrontata con l'eq. (A.35), dimostra la (A.36).

Appendix B

Unità di energia

Nel sistema di unità MKS, l'unità di massa è il chilogrammo (Kg) e l'unità di energia è il joule (J):

$$1 \text{ J} = 1 \text{ Kg} \cdot \text{m}^2/\text{sec}^2 . \quad (\text{B.1})$$

Nel sistema di unità CGS, l'unità di massa è il grammo (g) e l'unità di misura è l'erg (erg):

$$1 \text{ erg} = 1 \text{ g} \cdot \text{cm}^2/\text{sec}^2 = 10^{-7} \text{ J} . \quad (\text{B.2})$$

Per processi molecolari, atomici e subatomici un'unità di energia molto usata è l'elettronvolt (eV), definita come l'energia acquistata da un elettrone nell'attraversare una differenza di potenziale di 1 Volt (V). Tenuto conto che la carica di un elettrone è data da

$$e = 1,602 \cdot 10^{-19} \text{ Coulomb} , \quad (\text{B.3})$$

si ha

$$1 \text{ eV} = 1,602 \cdot 10^{-19} \text{ Coulomb} \cdot 1 \text{ Volt} = 1,602 \cdot 10^{-19} \text{ J} , \quad (\text{B.4})$$

dal momento che $1 \text{ Coulomb} \cdot 1 \text{ Volt} = 1 \text{ Joule}$.

Dall'equivalenza tra massa ed energia, possiamo esprimere le masse in unità di energia: diciamo che una particella ha una massa di $x \text{ eV}/c^2$ se $m \cdot c^2 = x \text{ eV}$. Dal valore della velocità della luce $c = 2,998 \text{ m/sec}$ otteniamo il fattore di conversione

$$1 \text{ eV}/c^2 = 1,7827 \cdot 10^{-36} \text{ Kg} \quad (\text{B.5})$$

Spesso si omette il divisore c^2 dopo l'unità di energia.

Sempre a livello atomico, si usa un'unità di misura di massa tale che la massa dell'atomo sia circa uguale al numero di nucleoni nel nucleo — questo è possibile perché la massa degli elettroni è molto minore di quella del nucleo, ed in quest'ultimo i protoni ed i neutroni hanno massa circa uguale. Dato che l'elemento più importante (almeno in chimica organica) è il carbonio, il cui isotopo stabile $^{12}_6\text{C}$ più abbondante ha 6 neutroni (e 6 protoni), per convenzione si assegna a questo atomo una massa di 12 unità di massa atomiche (a.m.u.). Quindi l'unità di massa atomica è definita come la dodicesima parte della massa dell'atomo di $^{12}_6\text{C}$. Risulta che $(1 \text{ MeV} = 10^6 \text{ eV})$

$$1 \text{ a.m.u.} = 931,494 \text{ MeV} = 1,66054 \cdot 10^{-27} \text{ Kg} \quad (\text{B.6})$$

A titolo di esempio, diamo le masse dei costituenti atomici

$$m_e = 0,000549 \text{ a.m.u.} = 0,511 \text{ MeV} \quad (\text{B.7})$$

$$m_p = 1,007276 \text{ a.m.u.} = 938,272 \text{ MeV} \quad (\text{B.8})$$

$$m_n = 1,008666 \text{ a.m.u.} = 939,565 \text{ MeV} . \quad (\text{B.9})$$

Il fatto che $m_n > m_p + m_e$ fa sì che il neutrone possa decadere in un protone più un elettrone (ed un neutrino di massa a confronto piccolissima). La più pesante particella elementare scoperta finora è il quark top, con massa

$$m_t \simeq 175.000 \text{ MeV} . \quad (\text{B.10})$$

Appendix C

Annichilazione $e^+e^- \rightarrow \gamma\gamma$

Si consideri l'annichilazione $e^+e^- \rightarrow \gamma\gamma$ nel SDR del laboratorio (LAB) in cui l'elettrone è fermo. Nota l'energia E_1 del positrone,

- 1) determinare l'energia di ciascun fotone in funzione dell'angolo di emissione rispetto alla direzione del positrone incidente.

Supponiamo che, nel SDR del centro di massa (CM), i fotoni vengano emessi in modo isotropo. Determinare nel LAB:

- 2) la funzione di distribuzione dell'energia di ciascun fotone;
- 3) la funzione di distribuzione dell'angolo di emissione di ciascun fotone;
- 4) i valori massimo e minimo che può assumere l'angolo di emissione tra i due fotoni.

Abbiamo a disposizione due modi di procedere: studiare la reazione nel SDR del CM, e quindi effettuare una spinta nel SDR del LAB, oppure usare gli invarianti. Procediamo nel secondo modo. Allineiamo l'asse x lungo l'impulso del positrone incidente, e l'asse y nel piano di decadimento. Indichiamo con p_1^μ e p_2^μ gli impulsi delle particelle incidenti, e con p_3^μ e p_4^μ gli impulsi dei fotoni emessi. Sia poi m la massa dell'elettrone e del positrone, mentre i fotoni hanno massa nulla:¹

$$p_1^2 = m^2 = p_2^2, \quad p_3^2 = 0 = p_4^2. \quad (\text{C.1})$$

Nel LAB abbiamo

$$\begin{aligned} p_1^\mu &= (E_1, p_1, 0, 0) & p_3^\mu &= (E_3, E_3 \cos \theta_3, E_3 \sin \theta_3, 0) \\ p_2^\mu &= (m, 0, 0, 0) & p_4^\mu &= (E_4, E_4 \cos \theta_4, E_4 \sin \theta_4, 0) \end{aligned} \quad \longrightarrow \quad (\text{C.2})$$

Consideriamo E_1 e $p_1 = \sqrt{E_1^2 - m^2}$ come parametri noti, mentre $E_3, E_4, \theta_3, \theta_4$ sono variabili di cui determinare le relazioni. Scriviamo i 6 prodotti scalari che si possono formare con questi vettori:

$$p_1 \cdot p_2 = E_1 m \quad (\text{C.3})$$

$$p_1 \cdot p_3 = E_1 E_3 - p_1 E_3 \cos \theta_3 = E_3 (E_1 - p_1 \cos \theta_3) \quad (\text{C.4})$$

$$p_1 \cdot p_4 = E_1 E_4 - p_1 E_4 \cos \theta_4 = E_4 (E_1 - p_1 \cos \theta_4) \quad (\text{C.5})$$

$$p_2 \cdot p_3 = E_3 m \quad (\text{C.6})$$

$$p_2 \cdot p_4 = E_4 m \quad (\text{C.7})$$

$$p_3 \cdot p_4 = E_3 E_4 (1 - \cos \theta_{34}), \quad \theta_{34} \equiv \theta_3 - \theta_4 \quad (\text{C.8})$$

¹Usiamo $c = 1$. In altri termini, la velocità della luce è scelta come unità di misura della velocità. La dimensione della lunghezza è la stessa di quella del tempo. In questo modo anche energia, impulso e massa hanno la stessa dimensione.

1) Scriviamo la conservazione dell'impulso come $p_4^\mu = p_1^\mu + p_2^\mu - p_3^\mu$. Prendendone il quadrato ricaviamo

$$\begin{aligned} 0 = p_4^2 &= (p_1 + p_2 - p_3)^2 = 2m^2 + 2p_1 \cdot p_2 - 2p_1 \cdot p_3 - 2p_2 \cdot p_3 \\ &= 2[m^2 + E_1 m - E_3(E_1 - p_1 \cos \theta_3) - E_3 m] = 2[m(E_1 + m) - E_3(E_1 + m - p_1 \cos \theta_3)] , \end{aligned} \quad (\text{C.9})$$

da cui ricaviamo

$$E_3 = \frac{m(E_1 + m)}{E_1 + m - p_1 \cos \theta_3} \quad (\text{C.10})$$

La stessa equazione vale per l'altro fotone, scambiando il pedice $3 \rightarrow 4$. Essendo i fotoni a massa nulla, l'angolo di emissione θ_3 può assumere qualsiasi valore (lo verificheremo esplicitamente al punto 3), e quindi l'energia è compresa tra

$$E_{\min} = \frac{m(E_1 + m)}{E_1 + m + p_1} = \frac{E_1 + m - p_1}{2} \quad (\text{C.11})$$

ed

$$E_{\max} = \frac{m(E_1 + m)}{E_1 + m - p_1} = \frac{E_1 + m + p_1}{2} \quad (\text{C.12})$$

2) Per rispondere alle altre domande, analizziamo il processo di annichilazione nel CM. Innanzitutto, avendo le due particelle incidenti impulsi uguali ed opposti, ed essendo identiche, esse avranno energie uguali, diciamo E_0 . Allo stesso modo i due fotoni finali avranno energie uguali tra loro, e per la conservazione dell'energia questa sarà ancora E_0 . Quindi, nel CM abbiamo

$$\begin{aligned} p_1^\mu &= (E_0, p_0, 0, 0) & \longrightarrow & \quad p_3^\mu = (E_0, E_0 \cos \phi, E_0 \sin \phi, 0) \\ p_2^\mu &= (E_0, -p_0, 0, 0) & & \quad p_4^\mu = (E_0, -E_0 \cos \phi, -E_0 \sin \phi, 0) \end{aligned} \quad (\text{C.13})$$

ove ϕ è l'angolo tra uno dei due fotoni e il positrone. I prodotti scalari invarianti valgono

$$p_1 \cdot p_2 = E_0^2 + p_0^2 = 2E_0^2 - m^2 \quad (\text{C.14})$$

$$p_1 \cdot p_3 = p_2 \cdot p_4 = E_0(E_0 - p_0 \cos \phi) \quad (\text{C.15})$$

$$p_1 \cdot p_4 = p_2 \cdot p_3 = E_0(E_0 + p_0 \cos \phi) \quad (\text{C.16})$$

$$p_3 \cdot p_4 = 2E_0^2, \quad (\text{C.17})$$

con $p_0^2 = E_0^2 - m^2$. Confrontando i valori dell'invariante $p_1 \cdot p_2$ nei due SDR, ricaviamo

$$E_0^2 = \frac{m(E_1 + m)}{2}, \quad p_0^2 = \frac{m(E_1 - m)}{2}, \quad (\text{C.18})$$

che possiamo considerare come parametri conosciuti. Si noti che l'elettrone in questo SDR del CM si muove a velocità $-\beta$ con

$$\beta = \frac{p_0}{E_0} = \sqrt{\frac{E_1 - m}{E_1 + m}}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}} = \frac{E_0}{m}. \quad (\text{C.19})$$

Questi sono i parametri della spinta che collega i due SDR, che grazie all'eq. (C.19) sono noti, e tramite i quali è conveniente esprimere varie relazioni che deriveremo.

Per calcolare le funzioni di distribuzione dei fotoni, ragioniamo così: rispetto all'angolo di emissione ϕ nel CM i fotoni sono emessi con uguale densità di probabilità \mathcal{P} :

$$\frac{d\mathcal{P}}{d\phi} = \text{cost}, \quad 1 = \mathcal{P}_{\text{tot}} = \int_0^\pi \frac{d\mathcal{P}}{d\phi} d\phi \quad \implies \quad \frac{d\mathcal{P}}{d\phi} = \frac{1}{\pi}. \quad (\text{C.20})$$

Se riusciamo ad esprimere la generica variabile del fotone nel CM, per esempio la sua energia E_3 , in funzione dell'angolo ϕ e di altri parametri fissi, allora avremo

$$\frac{d\mathcal{P}}{dE_3} = \frac{d\mathcal{P}}{d\phi} \frac{d\phi}{dE_3} = \frac{1}{\pi} \left[\frac{dE_3}{d\phi} \right]^{-1} \quad (\text{C.21})$$

Cominciamo proprio dall'energia E_3 . La relazione tra E_3 e l'angolo ϕ la ricaviamo uguagliando le due espressioni per l'invariante $p_2 \cdot p_3$:

$$E_3 = \frac{E_0}{m} (E_0 + p_0 \cos \phi) = E_0 \gamma (1 + \beta \cos \phi) \quad (\text{C.22})$$

da cui ricaviamo

$$\frac{dE_3}{d\phi} = \beta\gamma E_0 \sin \phi . \quad (\text{C.23})$$

Il nostro scopo è quello di esprimere la densità di probabilità dell'energia in funzione dell'energia stessa. Bisogna quindi esprimere ϕ in funzione di E_3 . Sempre dalla (C.22) ricaviamo

$$\cos \phi = \frac{E_3 - \gamma E_0}{\beta\gamma E_0} \quad (\text{C.24})$$

e quindi

$$\frac{d\mathcal{P}}{dE_3} = \frac{1}{\pi\beta\gamma E_0} \left[1 - \left(\frac{E_3 - \gamma E_0}{\beta\gamma E_0} \right)^2 \right]^{-1/2} . \quad (\text{C.25})$$

Vediamo quindi che i valori più probabili dell'energia sono quelli estremi. Non è difficile verificare la condizione di normalizzazione

$$\int_{E_{\min}}^{E_{\max}} \frac{d\mathcal{P}}{dE_3} dE_3 = 1 . \quad (\text{C.26})$$

3) Veniamo all'angolo di emissione del fotone, per esempio θ_3 . Viene spontaneo usare le espressioni per l'invariante $p_1 \cdot p_3$, che danno

$$E_1 - p_1 \cos \theta_3 = m \frac{1 - \beta \cos \phi}{1 + \beta \cos \phi} , \quad (\text{C.27})$$

in cui si sono usate le espressioni (C.22) e (C.19). Esprimendo anche

$$E_1 = \frac{2E_0^2 - m^2}{m} , \quad p_1 = \frac{2E_0 p_0}{m} \quad (\text{C.28})$$

in funzione di E_0 e p_0 , si riesce a semplificare ulteriormente la relazione tra i due angoli:

$$\cos \theta_3 = \frac{\cos \phi + \beta}{1 + \beta \cos \phi} \quad \Longleftrightarrow \quad \cos \phi = \frac{\cos \theta_3 - \beta}{1 - \beta \cos \theta_3} . \quad (\text{C.29})$$

Prendendo l'arcocoseno dell'ultima uguaglianza e derivando rispetto a θ_3 troviamo

$$\frac{d\phi}{d\theta_3} = \frac{1}{\gamma(1 - \beta \cos \theta_3)} \quad (\text{C.30})$$

e quindi

$$\frac{d\mathcal{P}}{d\theta_3} = \frac{1}{\pi\gamma(1 - \beta \cos \theta_3)} \quad (\text{C.31})$$

4) Dall'invariante $p_3 \cdot p_4$ ricaviamo la relazione che lega l'angolo θ_{34} tra i fotoni nel LAB e l'angolo di emissione ϕ nel CM:

$$E_3 E_4 (1 - \cos \theta_{34}) = 2E_0^2 \quad \Longrightarrow \quad \cos \theta_{34} = 1 - \frac{2(1 - \beta^2)}{1 - \beta^2 \cos^2 \phi} , \quad (\text{C.32})$$

in cui si è usata l'eq. (C.22) per E_3 e l'analoga (con $+\beta \rightarrow -\beta$) per E_4 . Essendo il coseno una funzione monotona decrescente nell'intervallo $[0, \pi]$, il valore minimo di θ_{34} si ha per $\cos \phi = 0$, mentre il valore massimo si ha per $\cos \phi = \pm 1$:

$$\cos(\theta_{34,\min}) = 2\beta^2 - 1 = \frac{E_1 - 3m}{E_1 + m} , \quad \theta_{34,\max} = \pi . \quad (\text{C.33})$$

Vediamo che per piccole energie $E_1 \simeq m$, $\beta \simeq 0$, i fotoni sono sempre emessi in direzione pressoché opposte, mentre per grandi energie $E_1 \gg m$, $\beta \simeq 1$ ed i fotoni possono essere emessi entro un angolo molto stretto $\theta_{34,\min} \simeq \sqrt{8m/E_1}$.

Appendix D

Sistema di unità razionalizzate

Ci proponiamo di riscrivere le equazioni dell'elettromagnetismo in un sistema di unità di misura che semplifichi le formule e soprattutto che evidenzi la simmetria tra campo elettrico e magnetico. Le equazioni dell'elettromagnetismo si riassumono nella forza di Lorentz e nelle equazioni di Maxwell

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (\text{D.1})$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \quad \rightarrow \quad \mathbf{F}_{12} = k_1 \frac{q_1 q_2}{r_{12}^2} \hat{r}_{12}, \quad k_1 = \frac{1}{4\pi\varepsilon_0}, \quad (\text{D.2})$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (\text{D.3})$$

$$\nabla \cdot \mathbf{B} = 0 \quad (\text{D.4})$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} \quad (\text{D.5})$$

dalla quale si deduce l'equazione di continuità della corrente elettrica

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (\text{D.6})$$

Come punto di partenza per fissare le unità di misura, prendiamo come grandezze fondamentali indipendenti la lunghezza (L), la massa (M) ed il tempo (T), per le quali abbiamo stabilito delle unità di misura.¹ Consideriamo innanzitutto la forza di Coulomb \mathbf{F}_{12} che la carica elettrica q_1 esercita sulla carica elettrica q_2 . Nel sistema di unità di misura MKSA, detto anche “Sistema Internazionale”, che abbiamo usato nelle eq. (D.1-D.5), vale

$$\mathbf{F}_{12} = k_1 \frac{q_1 q_2}{r_{12}^2} \hat{r}_{12}, \quad k_1 = \frac{1}{4\pi\varepsilon_0}, \quad (\text{D.7})$$

in cui r_{12} è la distanza tra le cariche ed \hat{r}_{12} il versore che da q_1 punta verso q_2 . La presenza ed il valore di ε_0 dipendono dall'aver scelto una certa unità di misura per le cariche. Nulla ci vieta però di definire l'unità di carica a partire da questa stessa formula: diciamo che due cariche uguali² hanno carica unitaria quando, poste alla distanza unitaria (1 cm) esercitano tra loro una forza unitaria (1 dina = 1 g cm s⁻²). In questo modo $k_1 = 1$, e la carica avrebbe le dimensioni $M^{1/2} L^{3/2} T^{-1}$. Questa definizione è adottata nel sistema di unità elettrostatiche (u.e.s.), in cui

$$q_{\text{ues}} = \frac{q_{\text{SI}}}{\sqrt{4\pi\varepsilon_0}} \quad \Rightarrow \quad 1_{\text{ues}} = 3.3356 \times 10^{-10} \text{C}. \quad (\text{D.8})$$

La medesima relazione vale per la densità di carica $\rho \equiv dq/d^3x$.

Di conseguenza, se definiamo il campo elettrico prodotto da q_1 come il rapporto tra la forza che agisce sulla carica di prova q e la carica q stessa, secondo l'equazione (D.1), avremo

$$\mathbf{E}_{\text{ues}} \equiv \frac{\mathbf{F}}{q_{\text{ues}}} = \sqrt{4\pi\varepsilon_0} \frac{\mathbf{F}}{q_{\text{SI}}} = \sqrt{4\pi\varepsilon_0} \mathbf{E}_{\text{SI}}. \quad (\text{D.9})$$

¹Nel sistema MKS, tali unità sono il metro, il chilogrammo ed il secondo; nel sistema CGS vengono scelti il grammo, il centimetro ed il secondo rispettivamente.

²Per stabilire il *rapporto* (e quindi l'eventuale uguaglianza) tra cariche è sufficiente confrontare le forze che esse esercitano su di una particella di prova posta ad una determinata distanza.

La legge di Gauss diventa quindi

$$\nabla \cdot \mathbf{E}_{\text{SI}} = \frac{\rho_{\text{SI}}}{\varepsilon_0} \quad \Longleftrightarrow \quad \nabla \cdot \mathbf{E}_{\text{ues}} = 4\pi\rho_{\text{ues}} : \quad (\text{D.10})$$

è scomparso ε_0 ma è apparso il 4π di fronte alla densità di carica.

Se si lavora principalmente con le equazioni di campo piuttosto che con la legge di Coulomb, conviene lasciare il 4π in quest'ultima, e definire

$$q \equiv \frac{q_{\text{SI}}}{\sqrt{\varepsilon_0}}, \quad k_1 = \frac{1}{4\pi}, \quad \mathbf{E} \equiv \sqrt{\varepsilon_0} \mathbf{E}_{\text{SI}}, \quad \nabla \cdot \mathbf{E} = \rho. \quad (\text{D.11})$$

La densità di corrente, definita come il flusso di carica elettrica attraverso la superficie unitaria per unità di tempo, riscalda come la carica:

$$\mathbf{J} \equiv \frac{\mathbf{J}_{\text{SI}}}{\sqrt{\varepsilon_0}}, \quad (\text{D.12})$$

in modo che l'equazione di continuità della corrente non dipenda dal sistema di unità e mantenga quindi l'espressione (D.6).

Possiamo anche eliminare la permeabilità magnetica del vuoto μ_0 dalle equazioni di Maxwell, ridefinendo il campo magnetico

$$\mathbf{B} \equiv \frac{\mathbf{B}_{\text{SI}}}{\sqrt{\mu_0}}. \quad (\text{D.13})$$

In questo sistema di unità, detto di “Heaviside-Lorentz razionalizzato”, le equazioni di Maxwell assumono la seguente forma:

$$\text{primo gruppo (omogenee)} \quad \left\{ \begin{array}{l} \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \end{array} \right. \quad (\text{D.14})$$

$$\text{secondo gruppo (con sorgenti)} \quad \left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = \rho \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{J}}{c} \end{array} \right. \quad (\text{D.15})$$

$$\quad \quad \quad \left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = \rho \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{J}}{c} \end{array} \right. \quad (\text{D.16})$$

$$\quad \quad \quad \left\{ \begin{array}{l} \nabla \cdot \mathbf{E} = \rho \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{J}}{c} \end{array} \right. \quad (\text{D.17})$$

e la forza di Lorentz si scrive

$$\mathbf{F} = q(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}). \quad (\text{D.18})$$

L'unica costante che compare in queste formule è la velocità della luce nel vuoto $c = 1/\sqrt{\varepsilon_0\mu_0}$. Inoltre, il campo elettrico ed il campo magnetico hanno le stesse dimensioni: $[E] = [B] = M^{1/2}L^{-1/2}T^{-1}$, come si vede subito dall'eq. (D.18).

È interessante notare che, se si elimina la corrente di spostamento, è possibile eliminare la costante c dalle equazioni di Maxwell, ridefinendo $\mathbf{B}_{\text{nuovo}} \equiv \mathbf{B}/c$ e $\mathbf{J}_{\text{nuovo}} \equiv \mathbf{J}/c^2$. Con questa ridefinizione, il termine di spostamento $\partial_t \mathbf{E}$ responsabile della violazione dell'invarianza di Galileo, appare moltiplicato per il fattore $1/c^2$. Ciò significa che le violazioni al principio di relatività galileiano che ci dobbiamo aspettare in esperienze di elettromagnetismo sono di ordine $\mathcal{O}(1/c^2)$.

Bibliografia

Testi di riferimento

- ⟨Mo⟩ C. Möller, *The Theory of Relativity*, Oxford at the Clarendon Press, 1952 (i primi 3 capitoli).
- ⟨Ro⟩ W. G. V. Rosser, *An Introduction to the Theory of Relativity*, London Butterworths, terza ed. 1971 (i primi 5 capitoli).
- ⟨Ri⟩ W. Rindler, *Essential Relativity, Special, General, and Cosmological*, Springer-Verlag, seconda ed. 1977 (i primi 5 capitoli)
- ⟨Fr⟩ A. P. French, *Special relativity*, Van Nostrand, 1988.
- ⟨FB⟩ *La fisica di Berkeley*, decimo volume, Zanichelli, 1973.

Lecture consigliate

- ⟨Ja⟩ J. D. Jackson, “Elettrodinamica Classica”, Zanichelli, 1995 (il capitolo 11).
- ⟨LL⟩ L. D. Landau e E. M. Lifshitz, *Teoria dei campi*, Editori Riuniti Edizioni Mir 1976 (i primi 3 capitoli).
- ⟨We⟩ S. Weinberg, *Gravitation and Cosmology*, John Wiley and Sons, 1972 (i primi quattro paragrafi del secondo capitolo).

Dispense di corsi analoghi

- ⟨Lo⟩ G. Longhi, *Appunti di relatività speciale*.
- ⟨Ca⟩ R. Casalbuoni, *Appunti di relatività speciale*.