

# Relatività (speciale)

L. Fatibene, G. Magnano

Dipartimento di Matematica - Università di Torino (Italy)

## 1. Spaziotempo Newtoniano e gruppo di Galilei

Durante il corso, abbiamo fin qui sviluppato il formalismo Lagrangiano per descrivere il moto dei sistemi meccanici soggetti a vincoli olonomi. Per un po', ora utilizziamo quegli strumenti per analizzare cosa possiamo dire sui moti delle *particelle libere nello spazio*.

Questo sembra un problema banale, già risolto con strumenti molto meno sofisticati. In un certo senso è vero, ma ora possiamo finalmente affrontare il problema da un punto di vista fondamentale, discutendo le strutture che sono implicite nell'analisi tradizionale, sia a livello cinematico sia a livello dinamico.

Questo ci porterà a chiarire i punti deboli del formalismo Newtoniano, basato su spazio e tempo assoluti e osservatori inerziali e, alla fine, alla descrizione relativistica, intendendo con questo la descrizione in relatività speciale (SR).

Cominciamo con chiarire cosa si intende per *osservatore*. Volendo descrivere il moto di un punto materiale nello spazio, abbiamo bisogno di un certo numero di convenzioni.

Siccome dovremo descrivere l'evoluzione della posizione nel tempo abbiamo bisogno di un *orologio* e di un *sistema di riferimento spaziale* che ci consenta di esprimere la posizione del punto in termini di (3) numeri  $x^i$ , o  $(x, y, z)$ .

Non abbiamo mai prestato molta attenzione a cosa sia (o come sia realizzato) un orologio ma ci si dovrebbe chiedere perché si utilizza un cronometro piuttosto che i battiti del cuore o l'altezza di un gesso lasciato cadere da una certa altezza. In tutti i casi si tratta di fenomeni che consentono di ordinare gli eventi con un parametro che potremmo chiamare *tempo*.

Ci sembra ovvio che il tempo di un cronometro sia *uniforme* (almeno *più uniforme* del battito del cuore) e lineare (almeno *più lineare* di quello associato al gesso che cade). Tuttavia non è difficile convincersi che tale convinzione è mantenuta in assenza di una vera definizione di cosa sia *uniforme* e che una tale definizione necessariamente presupporrebbe un altro *tempo* con cui confrontarsi.

Ha certamente senso dire che un cronometro è più o meno uniforme se confrontato con (quindi *rispetto a*) un orologio atomico. Un po' meno in termini assoluti.

In ogni caso abbiamo bisogno di un orologio, anche se uno valesse l'altro almeno l'orologio più fondamentale va scelto.

Anche fosse per mantenerci conformi alla tradizione, scegliamo un cronometro standard.

Per il riferimento spaziale abbiamo bisogno di fare quello che in generale facciamo scegliendo le coordinate Lagrangiane, cioè di un sistema di coordinate sullo spazio delle configurazioni. Per un punto materiale libero, lo spazio delle configurazioni è  $\mathbb{R}^3$  (almeno come varietà differenziabile).

Essendo  $\mathbb{R}^3$  anche uno spazio affine, un riferimento spaziale è formato da un'origine  $O \in \mathbb{R}^3$  e una base  $e_i$  applicata all'origine. Quindi indichiamo un riferimento spaziale con  $(O, e_i)$  ed esso definisce un sistema di coordinate Cartesiane centrate in  $O$  con assi nella direzione dei vettori  $e_i$ .

Indichiamo con  $x^i$  (con  $i = 1, 2, 3$ ) le coordinate Cartesiane così definite. Per essere espliciti, se consideriamo un punto  $P \in \mathbb{R}^3$  ad esso associamo un *vettore posizione*  $\vec{r} = P - O$  che, come

tutti i vettori applicati nell'origine, può essere scritto come combinazione lineare dei vettori della base, cioè  $\vec{r} = x^i e_i$ . In questo caso, banalmente associamo al punto  $P$  le posizioni  $x^i$ .

Quindi un *evento*, cioè qualcosa che succede in una certa posizione ad un certo istante, è etichettato da 4 coordinate  $(t, x^i)$  definite da un *osservatore*, che in buona sostanza coincide con la scelta dell'orologio e del riferimento spaziale.

Già a questo punto abbiamo 2 scelte: possiamo considerare posizione e tempo come 2 quantità diverse, come abbiamo fatto sin qui e descrivere l'evoluzione del sistema come l'evoluzione della posizione nel tempo, oppure considerare quello che succede a tempo e posizione insieme, cioè definendo  $(t, x^i)$  come 4 coordinate sullo spazio degli eventi, che è chiamato *spaziotempo*, come se fossero 4 coordinate Lagrangiane. Le due descrizioni dei fenomeni (come moto nello spazio o insieme di eventi nello spaziotempo) sono equivalenti, almeno in principio.

Anche se alla fine la descrizione sullo spaziotempo si rivelerà più fondamentale andiamo per gradi.

Se consideriamo un osservatore piuttosto standard (che sceglie un ordinario cronometro come orologio e una base ortonormale (?) centrata in un certo punto  $O$  come sistema di riferimento spaziale) che osservi un punto materiale lanciato in aria, in un certo istante  $t = 0$ , con una certa posizione iniziale  $x_o^i$  e velocità iniziale  $v_o^i$  (non necessariamente verticale) sappiamo che descriverà il moto del sistema mediante una curva parametrizzata in  $\mathbb{R}^3$  descritta da

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^3 : t \mapsto x^i(t) = x_o^i + v_o^i t - \frac{1}{2} g^i t^2 \quad (1.1)$$

(se uno degli assi, diciamo l'asse  $z$ , è verticale diretto verso l'alto e  $g^i = (0, 0, g)$  è la solita accelerazione di gravità in direzione dell'asse  $z$ ).

Oppure, equivalentemente, mediante il grafico della curva parametrizzata, che è una traiettoria nello spaziotempo

$$[\gamma] = \{(t, x^i)\} \subset \mathbb{R} \times \mathbb{R}^3 \quad (1.2)$$

Attenzione: non si deve confondere il grafico della curva  $\gamma$ , che è un sottoinsieme dello spaziotempo, con l'immagine di  $\gamma$  che è invece un sottoinsieme dello spazio  $\mathbb{R}^3$ .

Nello spaziotempo, il sottoinsieme contiene informazioni complete su dove passa il punto e come viene percorso il moto. Al contrario nello spazio si perde l'informazione sulla legge di percorrenza. Ad esempio, se si considera un moto rettilineo uniforme, questo è descritto da una retta nello spaziotempo (perché è moto rettilineo uniforme) e anche da una retta nello spazio (perché è moto rettilineo). Se si considera un moto rettilineo accelerato, come nell'esempio sopra, Questo è sempre descritto da una retta nello spazio (che però non è percorsa uniformemente) ma è descritto da una parabola nello spaziotempo, più precisamente dalla parabola

$$[\gamma] = \{(t, x^i) : x^i = x_o^i + v_o^i t - \frac{1}{2} g^i t^2\} \subset \mathbb{R} \times \mathbb{R}^3 \quad (1.3)$$

L'insieme di eventi che rappresentano il moto di un punto materiale nello spazio tempo forma una traiettoria nello spaziotempo, detta *linea di universo* del punto materiale.

## Diagrammi nello spaziotempo

È spesso utile descrivere la realtà fisica mediante grafici nello spaziotempo (che hanno il vantaggio di rappresentare staticamente i moti nello spazio, lo svantaggio di essere in generale in dimensione 4, quindi praticamente spesso si ricorre a grafici semplificati in dimensione più bassa, sottintendendo alcune dimensioni spaziali). Ad esempio per il moto di un grave descritto sopra, possiamo assumere che la velocità iniziale sia verticale, di modo che sia  $y(t) = y_o$ , sia  $x(t) = x_o$  siano costanti e possiamo concentrarci sul piano  $(t, z)$  nello spaziotempo in cui il moto è descritto dall'equazione

$$z = z_o + v_o t - \frac{1}{2} g t^2 \quad (1.4)$$

che corrisponde alla traiettoria nello spaziotempo

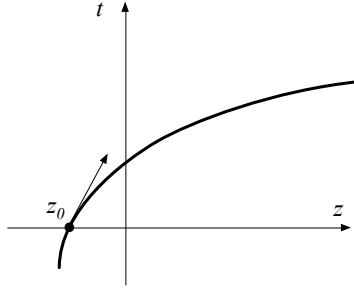


Figura 1: Linea di universo di un moto rettilineo uniformemente accelerato.

Un punto  $(t_1, z_1)$  su tale curva registra che al tempo  $t_1$  la posizione del punto materiale era  $z_1$ . Il moto nello spazio avviene lungo l'asse  $z$  ed è quindi rettilineo, ma sullo spaziotempo è descritto da una parabola perché si tratta di un moto rettilineo accelerato.

A seconda dei libri, il tempo è in genere rappresentato come ordinata o come ascissa, per noi sarà sempre un'ordinata, (quindi il tempo “scorre” dal basso verso l'alto).

Un vettore  $v$  tangente alla curva è un vettore nello spaziotempo, quindi in questo caso ha una componente lungo  $t$  e una lungo  $z$ . Non va pertanto confuso con una velocità fisica, che in questo caso ha un'unica componente visto che il moto è lungo l'asse spaziale  $z$ . La velocità fisica, come vedremo è piuttosto legata alla sua inclinazione.

## Cinematica Newtoniana relativa

Prima di proseguire dobbiamo analizzare cosa succede quando due osservatori diversi descrivono la stessa scena usando convenzioni diverse.

Per confrontare la descrizione del moto di uno o più punti materiali nello spazio fisico (spazio affine euclideo tridimensionale) da parte di due diversi osservatori, dobbiamo conoscere la legge di trasformazione che connette le coordinate nello spazio fisico usate dal primo osservatore con quelle usate dal secondo.

Supponiamo che il primo osservatore, che chiameremo  $X$ , usi coordinate cartesiane ortonormali che denoteremo con  $x^i$  ( $i = 1, 2, 3$ ). Il secondo osservatore,  $Y$ , userà un diverso sistema di coordinate ortonormali  $y^j$ . Un cambiamento di coordinate del tipo  $\hat{\Phi} : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : x^i \mapsto y^j(x^i)$  non

rappresenta, tuttavia, la relazione che intercorre fra osservatori in moto uno rispetto all'altro, in quanto un punto in quiete (= con coordinate costanti nel tempo) per l'osservatore  $X$  lo sarebbe anche per l'osservatore  $Y$ .

La trasformazione fra osservatori in moto relativo dovrà essere quindi dipendente dal tempo; ossia sarà una famiglia (parametrizzata dal tempo  $t$ ) di trasformazioni nello spazio affine tridimensionale. Ci si attende che le distanze e gli angoli nello spazio rappresentino proprietà fisiche indipendenti dall'osservatore: sotto questa ipotesi, per ogni istante  $t$  la trasformazione di coordinate deve essere una trasformazione ortogonale (più esattamente: ortogonale speciale, ossia che non cambi l'orientamento della terna di riferimento), con un'eventuale traslazione dell'origine, eventualmente anch'essa dipendente dal tempo. In altri termini, deve essere rappresentata da

$$x^i = A_j^i(t)y^j + C^i(t) \quad (1.5)$$

dove  $A_j^i(t) \in SO(3)$  per ogni  $t$ ; il vettore  $C^i(t) \in \mathbb{R}^3$  rappresenta in ogni istante la posizione dell'origine delle coordinate del sistema  $X$ , vista dall'osservatore  $Y$ .

Se, in particolare, la matrice  $A$  è costante, allora il moto relativo dei due osservatori è puramente traslatorio (anche quando gli assi dei due riferimenti non sono paralleli fra loro); se, in aggiunta, le  $C^i(t)$  sono polinomi di primo grado del tempo, allora è un moto traslatorio uniforme.

La trasformazione (1.5) si può vedere come una trasformazione di coordinate nello *spaziotempo*.

La retta dei tempi è a sua volta uno spazio affine euclideo: non ha un'origine definita intrinsecamente, ma si suppone che l'intervallo temporale fra due eventi non dipenda dall'osservatore.

Quindi la più generale trasformazione di coordinate fisicamente sensata nello spaziotempo sarà del tipo  $\Phi : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R} \times \mathbb{R}^3 : (t, x^i) \mapsto (t', y^j)$ , con

$$\begin{cases} t' = t + c^0 \\ y^i = A_j^i(t)x^j + C^i(t) \end{cases} \quad (1.6)$$

Ovviamente,  $c^0$  è una costante, visto che l'origine dei tempi di un osservatore rispetto all'altro è fissata una volta per tutte.

In linea di principio, potremmo considerare trasformazioni più generali del tempo, ad esempio  $t' = \alpha t + c^0$  o addirittura  $t' = t'(t)$ , ma questo sembra essere quello che si intende dicendo che i due osservatori usano 2 orologi identici per misurare i tempi, eventualmente sono inizializzati in momenti diversi.

Per l'osservatore  $X$ , il sottospazio affine tridimensionale definito da  $t = t_0$  contiene tutti gli eventi che avvengono simultaneamente. Esso coincide con il sottospazio  $t' = t'_0 = t_0 + c^0$  (solo con un'altra costante  $t = t'_0$ ): la proiezione dello spaziotempo sulla retta temporale, che associa ad ogni evento l'istante in cui l'evento stesso si verifica, è indipendente dall'osservatore.

Per contro, la proiezione degli eventi sullo spazio fisico tridimensionale (che associa ad ogni evento la posizione spaziale in cui questo avviene, e ad ogni linea di universo la traiettoria del moto corrispondente nello spazio fisico) dipende dall'osservatore. Per rendersene conto basta pensare che la linea di universo di un punto materiale in quiete rispetto a un osservatore si proietta interamente, per quell'osservatore, su un unico punto dello spazio tridimensionale; per un secondo osservatore in moto rispetto al precedente, invece, lo stesso punto materiale non sarebbe in quiete: da  $x^i = \text{costante}$ , attraverso le (1.6), non segue  $y^i = \text{costante}$ , e la traiettoria osservata non coincide con un unico punto.

Da notare che lo spaziotempo, per come l'abbiamo definito finora, ha una struttura naturale di varietà differenziabile (come *varietà prodotto*), ma non è dotato di una struttura affine complessiva. Le trasformazioni (1.6) non sono affini nello spaziotempo, dato che la dipendenza di  $A_j^i(t)$  e  $C^i(t)$  da  $t$  è, fino ad ora, arbitraria.

Tutto questo si può riassumere nelle ipotesi seguenti:

- N1. Lo spaziotempo è il prodotto cartesiano di uno spazio affine unidimensionale (la retta dei tempi) con uno spazio affine tridimensionale (lo spazio fisico);
- N2. la proiezione  $\pi$  dallo spaziotempo sulla retta temporale è indipendente dall'osservatore (questo corrisponde all'idea newtoniana di tempo assoluto);
- N3. la retta dei tempi ha una struttura euclidea intrinseca (invarianza degli intervalli temporali rispetto all'osservatore);
- N4. lo spazio fisico tridimensionale, definito in ciascun istante di tempo come spazio di simultaneità, ossia come controimmagine  $\pi^{-1}(t)$ , ha una struttura euclidea intrinseca (invarianza delle distanze nello spazio rispetto all'osservatore).

Le trasformazioni di coordinate nello spaziotempo che rappresentano un cambiamento di osservatore devono conservare tutte queste strutture, e quindi hanno necessariamente la forma (1.6).

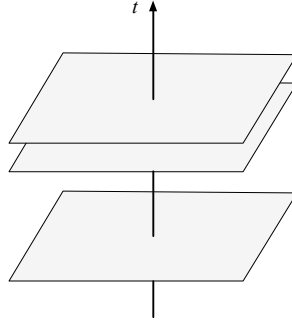


Figura 2: Superfici di isocronia per un osservatore.

Le trasformazioni (1.6) determinano la relazione fra le coordinate di un evento  $P$  osservate dai due osservatori  $X$  e  $Y$ .

Intanto  $P$  giace su un certo piano  $t = t_0 = \pi(P)$  (e quindi  $t' = t'_0 = t_0 + c^0$ ). Poi l'evento  $P$  per l'osservatore  $X$  è associato al vettore spaziale  $\vec{r} = P - O(t_0) = x^i e_i$ , mentre per  $Y$  è associato al vettore  $\vec{r}' = P - O'(t'_0) = y^i e'_i$ .

Se le basi scelte dai 2 osservatori su ogni spazio sono legate dalla relazione  $e_i = A_j^i(t) e'_j$  e  $O - O' = C^i(t) e'_i$ , allora  $P - O' = P - O + O - O'$  e quindi

$$\vec{r}' = y^i e'_i = x^i e_i + C^i e'_i = (x^j A_j^i + C^i) e'_i \quad \Rightarrow \quad \begin{cases} t' = t + c^0 \\ y^i = A_j^i(t) x^j + C^i(t) \end{cases} \quad (1.7)$$

Per un generico moto  $\gamma : t \mapsto P(t)$  nello spazio, le componenti della velocità fisica  $\vec{v} = \frac{d}{dt}(P - O)$  viste dall'osservatore  $X$ , relative alla sua terna di riferimento, saranno le  $\dot{x}^i = dx^i/dt$ , di modo che  $\vec{v} = \dot{x}^i e_i$ . Le  $\dot{y}^i = dy^i/dt'$  saranno invece le componenti della velocità  $\vec{v}' = \frac{d}{dt'}(P - O')$  osservate da  $Y$  e relative alla sua terna di riferimento, di modo che  $\vec{v}' = \dot{y}^i e'_i$ .

Da notare che  $\frac{d}{dt}$  e  $\frac{d}{dt'}$  coincidono, dato che  $t$  e  $t'$  differiscono solo per una costante.

Si trova immediatamente che  $\dot{y}^i = \dot{A}_j^i x^j + A_j^i \dot{x}^j + \dot{C}^i$ , cioè che

$$\vec{v}' = (\dot{A}_j^i x^j + A_j^i \dot{x}^j + \dot{C}^i) \bar{A}_i^k e_k = \Omega_j^k x^j e_k + \vec{v} + \vec{v}_O \quad (1.8)$$

dove abbiamo posto  $\Omega_j^k := \bar{A}_i^k \dot{A}_j^i$  e  $\vec{v}_O := \dot{C}^i e'_i = \frac{d}{dt'}(O - O')$  per la velocità di  $O$  misurata da  $Y$ .

Data la matrice ortogonale  $A_i^j(t)$ , possiamo definire la mappa lineare  $\Phi_t : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : x^i e_i \mapsto x^i \bar{A}_k^j \dot{A}_i^k e_j = x^i \Omega_i^j e_j$ , che ha come matrice dei coefficienti  $\Omega_i^j := \bar{A}_k^j \dot{A}_i^k$ .

La matrice  $\Omega_i^j$  è antisimmetrica perché  $\dot{A}_i^j$  è ortogonale. Infatti, abbiamo  ${}^t A A = A {}^t A = \mathbb{I}$  e derivando si ottiene

$${}^t \dot{A} A + {}^t A \dot{A} = {}^t ({}^t A \dot{A}) + {}^t A \dot{A} = {}^t \Omega + \Omega = 0 \quad (1.9)$$

Come è noto, per due matrici qualsiasi  ${}^t(AB) = {}^t B {}^t A$ .

Siccome  $\Omega$  è una matrice tridimensionale antisimmetrica, è nella forma

$$\Omega = \begin{pmatrix} 0 & -\omega^3 & \omega^2 \\ \omega^3 & 0 & -\omega^1 \\ -\omega^2 & \omega^1 & 0 \end{pmatrix} \quad (1.10)$$

per cui possiamo definire un vettore  $\vec{\omega} = \omega^i e_i$  e per ogni  $\vec{r} = x^i e_i$  possiamo confrontare

$$\Omega X = \begin{pmatrix} 0 & -\omega^3 & \omega^2 \\ \omega^3 & 0 & -\omega^1 \\ -\omega^2 & \omega^1 & 0 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \omega^2 x^3 - \omega^3 x^2 \\ \omega^3 x^1 - \omega^1 x^3 \\ \omega^1 x^2 - \omega^2 x^1 \end{pmatrix} = \vec{\omega} \times \vec{r} \quad (1.11)$$

Il vettore  $\vec{\omega}$  è detto *velocità angolare* dell'osservatore  $X$  vista dall'osservatore  $Y$  e abbiamo  $\Omega_j^k x^j e_k := \vec{\omega} \times \vec{r}$ .

La velocità angolare di  $Y$  vista da  $X$  è invece il vettore associato alla matrice  $\dot{A}_k^i \bar{A}_j^k$ .

Il vettore  $\vec{v}_{tr} = \vec{\omega} \times \vec{r} + \vec{v}_O$  è detto *velocità di trascinamento* del punto individuato dal vettore posizione  $\vec{r}$ . Nel caso in cui  $A$  sia costante  $\vec{\omega}$  si annulla, quindi in quel caso la velocità di trascinamento coincide con  $\vec{v}_O$  ed è uguale per tutti i punti.

Derivando ancora rispetto al tempo la (1.8) si ottiene

$$\ddot{y}^i e'_i = \bar{A}_i^k \ddot{A}_j^i x^j e_k + 2\bar{A}_i^k \dot{A}_j^i \dot{x}^j e_k + \ddot{x}^j e_j + \ddot{C}^i e'_i \quad (1.12)$$

da cui si può ottenere con qualche passaggio (cfr. [Benenti] per una dimostrazione alternativa) la relazione fra l'accelerazione  $\vec{a}$  osservata da  $X$  e l'accelerazione  $\vec{a}'$  osservata da  $Y$ :

$$\vec{a}' = \vec{a} + \dot{\omega} \times \vec{r} + \vec{\omega} \times (\vec{\omega} \times \vec{r}) + 2\vec{\omega} \times \vec{v} + \vec{a}_O \quad (1.13)$$

dove abbiamo posto  $\vec{a} := \ddot{x}^j e_j$  per l'accelerazione del punto  $P$  misurata da  $X$ ,  $\vec{a}_O := \ddot{C}^i e'_i$  per l'accelerazione di  $O$  misurata da  $Y$  e  $\dot{\omega} := \dot{\omega}^i e_i$ .

Dobbiamo solo sistemare il termine  $\bar{A}_i^k \ddot{A}_j^i x^j e_k$  e scriverlo in modo intrinseco. Integrando (due volte) per parti e inserendo un fattore  $(A\bar{A}) = \mathbb{I}$ , abbiamo:

$$\begin{aligned} \bar{A}_i^k \ddot{A}_j^i x^j e_k &= \frac{d}{dt} \left( \Omega_j^k \right) x^j e_k - \dot{A}_i^k (A_l^i \bar{A}_h^l) \dot{A}_j^h x^j e_k = \frac{d}{dt} \left( \Omega_j^k \right) x^j e_k + \bar{A}_i^k \dot{A}_l^i \bar{A}_h^l \dot{A}_j^h x^j e_k = \\ &= \dot{\omega} \times \vec{r} + \vec{\omega} \times (\vec{\omega} \times \vec{r}) \end{aligned} \quad (1.14)$$

Il termine  $2\omega \times v$  è detto *accelerazione di Coriolis*, mentre nel termine  $\omega \times (\omega \times r)$  riconosciamo l'*accelerazione centripeta*.

Dalla (1.13) si ricava che  $\vec{a} = \vec{a}'$  per tutti i moti se e solo se  $\vec{\omega} = 0$  e  $\vec{a}_O = 0$ , ossia se il moto di un osservatore rispetto all'altro è traslatorio uniforme:  $A_j^i(t) = \text{costante}$  e  $\vec{v}_{tr} = \text{costante}$ . È a questo punto che possiamo introdurre l'ulteriore proprietà costitutiva della meccanica galileiana e newtoniana, il *principio d'inerzia*:

*N5. Un punto materiale non soggetto a forze si muove di moto rettilineo uniforme.*

Questa proposizione non può essere vera per tutti gli osservatori, dato che in generale  $\vec{a} = 0$  non implica  $\vec{a}' = 0$ . La proprietà N5, come è noto, seleziona la classe dei *sistemi di riferimento inerziali*.

In realtà la N5 definisce operativamente cosa significa che un corpo non è soggetto a forze, per un osservatore. Essa dice che, se un osservatore scopre che un punto materiale si muove di moto rettilineo uniforme, allora tale punto non è soggetto a forze (per quell'osservatore, perché in generale sappiamo che un altro osservatore può vedere forze agire in virtù del proprio moto accelerato).

Ad esempio, se un osservatore in orbita sulla ISS osserva una goccia d'acqua libera, la vede muoversi di moto rettilineo uniforme nel suo sistema di riferimento, proprio in virtù del fatto che entrambi sono co-accelerati e la forza peso che agisce sia sull'osservatore che sulla goccia. La forza peso è esattamente cancellata dalla forza centrifuga, entrambe essendo tutt'altro che trascurabili se osservate da un riferimento inerziale.

*Calcolare la forza peso agente su una persona sulla ISS.*

(Per mostrare che la gravità agisce sui corpi nella ISS, comunque, basta pensare che se il campo gravitazionale della Terra fosse trascurabile a 300Km dalla superficie, la Luna chi la farebbe girare?)

La N5 dice comunque *qualcosa* sugli osservatori. Dice che *se X è inerziale*, allora gli osservatori in moto rettilineo uniforme rispetto a X sono inerziali anch'essi. Se noi conoscessimo un osservatore inerziale, allora N5 selezionerebbe una classe di osservatori inerziali. Peccato che per conoscere *un osservatore inerziale* avremmo bisogno di un criterio per dire che un corpo non è soggetto a forze, indipendente da N5, che non abbiamo. Questo è il punto debole (e fatale) della fisica di Newton che porta inevitabilmente alla relatività generale (GR).

In altre parole, la N5 non determina una classe di osservatori inerziali, determina però delle trasformazioni che mandano osservatori inerziali in osservatori inerziali, cioè le trasformazioni di Galileo che seguono.

La trasformazione che connette due osservatori inerziali avrà dunque la forma

$$\begin{cases} t' = t + c^0 \\ y^i = A_j^i x^j + v_{tr}^i t + c^i \end{cases} \quad (1.15)$$

questa volta con  $A_j^i$ ,  $v_{tr}^i$  e  $c^i$  costanti. Osserviamo che le (1.15) sono *trasformazioni affini nello spaziotempo*. Possiamo complessivamente descrivere le trasformazioni fra osservatori inerziali (*trasformazioni di Galileo*) attraverso una *matrice di Galileo* e una traslazione dell'origine nello spaziotempo:

$$\begin{pmatrix} t' \\ y^1 \\ y^2 \\ y^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ v_{tr}^1 & A_1^1 & A_2^1 & A_3^1 \\ v_{tr}^2 & A_1^2 & A_2^2 & A_3^2 \\ v_{tr}^3 & A_1^3 & A_2^3 & A_3^3 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} + \begin{pmatrix} c^0 \\ c^1 \\ c^2 \\ c^3 \end{pmatrix} \quad (1.16)$$

Come abbiamo visto, per una trasformazione di questo tipo si ha sempre

$$\vec{a} = \vec{a}' \quad \vec{v} = \vec{v}' + \vec{v}_{tr}. \quad (1.17)$$

Una matrice di Galileo è contraddistinta dal fatto che la prima riga è sempre  $(1, 0, 0, 0)$  – la trasformazione deve conservare la proiezione universale  $\pi$  sulla retta dei tempi e le distanze su quest’ultima, come richiesto da  $N2$  e  $N3$  – e dal fatto che la sottomatrice  $A$  deve essere ortogonale,  $A \in SO(3)$ . Poiché una matrice ortogonale ha tre parametri indipendenti, complessivamente una trasformazione di Galileo è definita da dieci parametri.

Dimostrare che le matrici (1.16) formano gruppo, il *gruppo di Galilei*, cioè il prodotto di 2 matrici di Galileo è una matrice di Galileo (e l’identità è una matrice di Galileo) e che l’inversa di una matrice di Galileo è una matrice di Galileo.

Il principio di inerzia, selezionando fra tutte le curve nello spaziotempo quelle che devono essere fisicamente identificate come rette (e cioè le linee di universo di punti non soggetti a forze), introduce una *struttura affine sullo spaziotempo*. Quest’ultima non deriva in alcun modo dalle precedenti assunzioni  $N1$ – $N4$ , che presupponevano una struttura affine sulla retta dei tempi e su ciascuno spazio di simultaneità, ma non nello spaziotempo. Le trasformazioni (1.16) fra osservatori inerziali, oltre a rispettare le ipotesi  $N1$ – $N4$ , conservano anche la struttura affine quadridimensionale introdotta da  $N5$ .

Per contro, lo spaziotempo newtoniano non è dotato di una struttura euclidea (che esiste solo, separatamente, per la retta dei tempi e per gli spazi di simultaneità). Infatti, se si suppone che le coordinate associate a un osservatore inerziale siano ortonormali (il che definirebbe univocamente un tensore metrico, a meno della segnatura), risulta immediatamente che le matrici di Galileo non sono ortogonali: la metrica dello spaziotempo, quindi, risulterebbe diversa per ciascun osservatore inerziale e non potrebbe avere un significato intrinseco.

## 2. Trasformazioni di Lorentz

Nel 1905 Albert Einstein propose una nuova interpretazione *dell’elettrodinamica dei corpi in movimento* che è alla base delle teorie fisiche del XX secolo.

Contrariamente a quello che spesso si dice, le motivazioni di Einstein non erano affatto specificamente connesse con l’esperimento di Michelson e Morley. Einstein ha diverse volte dichiarato di non ricordare neanche se conoscesse tale esperimento nel momento in cui proponeva la SR. Nell’articolo in cui propone la SR, l’autore si riferisce principalmente al fastidio che provoca l’elettrodinamica e la possibilità che essa pare avere di definire un sistema di quiete assoluta (o comunque privilegiato). Ad esempio, “*se infatti il magnete è in moto e il conduttore è a riposo, nei dintorni del magnete esiste un campo elettrico con un certo valore dell’energia, che genera una corrente nei posti dove si trovano parti del conduttore. Ma se il magnete è in quiete e si muove il conduttore, nei dintorni del magnete non esiste alcun campo elettrico, e si ha invece nel conduttore una forza elettromotrice, alla quale non corrisponde nessuna energia, ma che —a parità di moto relativo nei due casi considerati— dà luogo a correnti elettriche della stessa intensità e dello stesso andamento di quelle alle quali dà luogo nel primo caso la forza elettrica.*”

L’autore cita anche genericamente “*come pure i tentativi andati a vuoto di constatare un moto della terra relativamente al “mezzo luminoso” portano alla supposizione che il concetto di quiete assoluta non solo in meccanica, ma anche in elettrodinamica non corrisponda ad alcuna proprietà dell’esperienza.*”

Egli mostrò che assumendo due principi fisici:

- E1. principio di relatività: le leggi secondo le quali evolvono gli stati dei sistemi fisici sono le stesse in due sistemi di coordinate che siano in moto traslatorio uniforme l’uno rispetto*



all'altro;

*E2. costanza della velocità della luce: ogni raggio di luce si muove in ogni sistema di coordinate con la velocità fissa  $c$ , indipendentemente dal fatto che questo raggio di luce sia emesso da un corpo a riposo o in moto,*

si possono spiegare tutti gli effetti sopra descritti, e si può costruire una teoria generale dell'interazione elettromagnetica di corpi in movimento. Nei prossimi capitoli ricostruiremo questo risultato, attraverso passaggi matematici parzialmente diversi da quelli usati da Einstein, ma fisicamente equivalenti.

L'equazione delle onde di d'Alembert (qui utilizzata per descrivere le onde elettromagnetiche o comunque onde che viaggiano alla velocità della luce  $c$ )

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0, \quad (2.1)$$

si può riscrivere nel modo seguente: chiamiamo  $x^i$  ( $i = 1, 2, 3$ ) le tre coordinate nello spazio fisico tridimensionale associate a un osservatore inerziale; chiamiamo  $x^0$  la lunghezza  $ct$ , dove  $c$  è la velocità della luce e  $t$  è la coordinata tempo per il medesimo osservatore. Le  $x^\mu$  ( $\mu = 0, 1, 2, 3$ ) formano quindi un sistema di coordinate nello spaziotempo, associate all'osservatore  $X$ . L'equazione delle onde elettromagnetiche diventa allora

$$\eta^{\mu\nu} \frac{\partial^2 u}{\partial x^\mu \partial x^\nu} = 0, \quad \text{dove} \quad \eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.2)$$

Supponiamo di cambiare coordinate, passando quindi a un secondo osservatore  $Y$ . Per il principio di inerzia (ovvero in conseguenza dell'assunzione *E1*), questa trasformazione deve essere affine:

$$y^\mu = \Lambda^\mu_\alpha x^\alpha + c^\mu \quad \Longleftrightarrow \quad x^\alpha = \bar{\Lambda}^\alpha_\mu y^\mu - \bar{\Lambda}^\alpha_\mu c^\mu \quad (2.3)$$

con  $\Lambda^\mu_\alpha$  e  $c^\mu$  costanti. Nelle nuove coordinate, l'equazione (2.2) diventa

$$\tilde{\eta}^{\alpha\beta} \frac{\partial^2 u}{\partial y^\alpha \partial y^\beta} = 0, \quad \text{con} \quad \eta^{\mu\nu} = \Lambda^\mu_\alpha \tilde{\eta}^{\alpha\beta} \Lambda^\nu_\beta \quad (2.4)$$

Se

$${}^t\Lambda\eta\Lambda = \eta, \quad \text{allora} \quad \tilde{\eta} = \eta \quad (2.5)$$

e l'equazione (2.2) resta invariata, e in questo caso anche l'osservatore  $Y$  vedrà qualunque segnale luminoso propagarsi a velocità  $c$ , come richiesto dall'assunzione *E2*. L'insieme delle matrici  $\Lambda$  che lasciano invariata la forma quadratica  $\eta$  – che ha segnatura  $(3, 1)$  – si denota con  $O(3, 1)$ . Dal punto di vista fisico, siamo interessati solo alle trasformazioni in  $O(3, 1)$  che conservano l'orientamento delle tre coordinate spaziali e della coordinata temporale: il gruppo di matrici che ha questa proprietà si chiama *gruppo di Lorentz (ristretto)* e si denota con  $SO^+(3, 1)$

Supponiamo dunque che la (2.3), con  $\Lambda \in SO^+(3, 1)$ , rappresenti la forma generale della trasformazione fra due osservatori inerziali. Una trasformazione affine con queste proprietà è detta *trasformazione di Poincaré* (o *trasformazione di Lorentz affine*). Dato che le condizioni (2.5) corrispondono a 10 equazioni indipendenti (il sistema è simmetrico), le matrici del gruppo di Lorentz hanno sei parametri indipendenti. Una trasformazione di Poincaré è quindi determinata da dieci parametri.

Le trasformazioni di Poincaré sono anche trasformazioni di Galileo? È facile verificare che una matrice della forma

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & A_1^1 & A_2^1 & A_3^1 \\ 0 & A_1^2 & A_2^2 & A_3^2 \\ 0 & A_1^3 & A_2^3 & A_3^3 \end{pmatrix}, \quad (2.6)$$

con  $A \in SO(3)$ , che rappresenta una trasformazione di Galileo con  $\vec{v}_{tr} \equiv 0$ , appartiene anche al gruppo di Lorentz. Una trasformazione di questo tipo è detta *rotazione spaziale*, ma è di scarso interesse ai nostri fini dato che rappresenta due osservatori in quiete l'uno rispetto all'altro.

Cerchiamo invece una trasformazione che coinvolga, ad esempio, le coordinate  $x^0$  e  $x^1$ , lasciando invariate le altre, cioè una matrice della forma

$$\Lambda = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.7)$$

La condizione (2.5) diventa in questo caso

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{cases} a^2 - c^2 = 1 \\ ab - cd = 0 \\ d^2 - b^2 = 1 \end{cases} \quad (2.8)$$

In aggiunta, la condizione che la trasformazione conservi l'orientamento implica  $ad - bc = 1$ . La soluzione generale ha la forma  $a = d = \cosh(\theta)$ ,  $b = c = \sinh(\theta)$ , con  $\theta$  numero reale arbitrario. La matrice così ottenuta non è di Galileo (tranne nel caso in cui  $b = 0$  e  $d = 1$ , ossia  $\theta = 0$ , in cui  $\Lambda$  si riduce alla matrice identità).

Quindi questa trasformazione di Lorentz, detta *boost* nella direzione dell'asse  $x$ , è

$$\begin{cases} ct' = c \cosh(\theta) + \sinh(\theta)x & \Rightarrow t' = \cosh(\theta) + \sinh(\theta)\frac{x}{c} \\ x' = c \sinh(\theta)t + \cosh(\theta)x \\ y' = y \\ z' = z \end{cases} \quad (2.9)$$

Questa è la forma in cui le *trasformazioni di Lorentz* sono correntemente presentate nei testi di fisica: un *boost* in una direzione qualsiasi si può sempre ricondurre a questa forma a seguito di una rotazione spaziale.

Le *trasformazioni di Lorentz*, o di *Lorentz-FitzGerald*, non sono altro che trasformazioni di Poincaré senza traslazione affine. Quest'ultima non è significativa ai fini del confronto fra due osservatori inerziali, ma è utile includerla nella legge di trasformazione più generale perché le traslazioni nello spaziotempo (da non confondersi con i *boost*!!) risulteranno rilevanti come possibili simmetrie della dinamica. Storicamente, Henri Poincaré generalizzò in questo modo nel 1905 le trasformazioni lineari introdotte da Lorentz.

Siccome  $\cosh^2(\theta) - \sinh^2(\theta) = 1$ , se poniamo  $\gamma := \cosh(\theta)$  e  $\gamma\beta := \sinh(\theta)$  allora abbiamo  $\gamma^2 = 1 + \gamma^2\beta^2$ , cioè

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad (2.10)$$

Sia  $\beta$  che  $\gamma$  sono fattori adimensionali. Il parametro  $\gamma$  è sempre maggiore di 1, mentre  $\beta \in (-1, 1)$ .  
 Il coefficiente  $\beta$  si scrive anche  $\beta = v/c$  e  $v \in (-c, c)$  è detta *velocità* (scalare) del boost.  
 Un boost di Lorentz si può scrivere, a seconda della convenienza, in termini di  $\theta$  o in termini di  $\beta$  (o  $v$ ). Talvolta si pone  $c = 1$  e si identifica  $\beta = v$ .

Per interpretare il significato di questo boost dobbiamo identificare il moto relativo fra gli osservatori. Un boost di Lorentz con parametro  $\beta$  manda gli assi di un osservatore  $X$  in rette dello spaziotempo, che non coincidono, ovviamente, con gli assi di  $Y$ . L'asse  $x$  è parametrizzato  $\sigma : s \mapsto (t = 0, x = s)$  e dopo il boost di Lorentz diventa  $\sigma(s) = (ct' = \gamma\beta s, x' = \gamma s)$ . Analogamente, l'asse  $t$  è parametrizzato  $\sigma : s \mapsto (t = s, x = 0)$  e dopo il boost di Lorentz diventa  $\sigma(s) = (t' = \gamma s, x' = \gamma\beta cs)$ . Quindi gli assi del riferimento spaziotemporale di  $X$ , rappresentati rispetto al riferimento di  $Y$  sono:

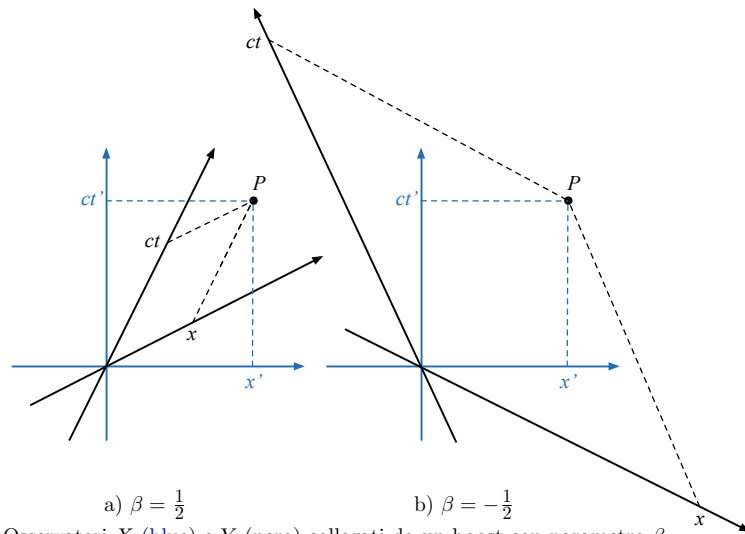


Figura 3: Osservatori  $X$  (blue) e  $Y$  (nero) collegati da un boost con parametro  $\beta$ .

Un evento  $P$  ha coordinate  $(t, x)$  rispetto a  $X$ , e coordinate  $(t', x')$  rispetto a  $Y$ , ma è rappresentato da un unico punto per tutti gli osservatori inerziali. Il boost di Lorentz trasforma precisamente le coordinate di un evento letto da  $X$  nelle coordinate dello stesso evento letto da  $Y$ .

Si noti come l'osservatore  $Y$  veda l'origine del riferimento spaziale di  $X$  muoversi lungo l'asse  $ct$ , cioè con velocità positiva (di medesimo segno scelto per  $\beta$ ). Si noti anche come l'inclinazione di tale asse sia dato da  $\beta$  che è quindi la velocità di  $O$  rispetto a  $Y$ . Il boost quindi descrive il passaggio tra 2 osservatori inerziali in moto rettilineo uniforme con velocità  $\beta$  lungo l'asse  $x$ .

Si noti anche come le trasformazioni di Lorentz non conservino gli angoli sullo spaziotempo che quindi sono quantità non fisiche, in quanto dipendenti dall'osservatore.

Dato che  $\beta \neq 0$ , la trasformazione non conserva la proiezione sull'asse dei tempi: da  $x^0 = ct = \text{costante}$  non segue  $y^0 = ct' = \text{costante}$ , e viceversa. Dunque due eventi simultanei per un osservatore non saranno, in generale, simultanei per l'altro. La *relatività della simultaneità*, che al momento della pubblicazione dell'articolo di Einstein del 1905 suscitò forti controversie, fu chiaramente individuata dallo stesso Einstein come uno degli elementi fondamentali per comprendere la vera natura della “contrazione dello spazio” ipotizzata da FitzGerald e Lorentz.

Poiché la coordinata temporale non è più la stessa per tutti gli osservatori, nella descrizione spaziotemporale della dinamica di una particella non possiamo supporre che esista una parametrizzazione intrinsecamente privilegiata della linea di universo. Per ora, quindi, non facciamo alcuna assunzione sulla parametrizzazione di una linea di universo: supponiamo semplicemente di descrivere quest'ultima come immagine di una curva nello spaziotempo,  $\tau \mapsto x^\mu(\tau)$ . Chiamiamo *quadrivelocità* il vettore tangente a questa curva,  $u^\mu \frac{\partial}{\partial x^\mu}$ , con  $u^\mu = \frac{dx^\mu}{d\tau}$ .

La proiezione della linea di universo nello spazio tridimensionale si ottiene, per un osservatore  $X$ , eliminando il parametro  $\tau$  in modo da ottenere un'equazione  $x^i = x^i(t)$ . Le componenti della velocità nello spazio tridimensionale viste dall'osservatore  $X$  saranno allora

$$v^i = \frac{dx^i}{dt} = c \frac{dx^i}{d\tau} \frac{d\tau}{dx^0} = c \frac{u^i}{u^0} \quad (2.11)$$

Si vede immediatamente che le  $v^i$  non dipendono dalla parametrizzazione usata per la linea di universo: cambiando parametrizzazione,  $\tau \mapsto \sigma(\tau)$ , tutte le componenti della quadrivelocità sono moltiplicate per un medesimo fattore  $\frac{d\tau}{d\sigma}$ , quindi il rapporto fra due componenti resta invariato.

Sotto una trasformazione di Poincaré, le componenti della quadrivelocità si trasformano come di consueto con la mappa tangente:

$$\frac{dy^\mu}{d\tau} = \Lambda^\mu_\nu \frac{dx^\nu}{d\tau} \quad (2.12)$$

Pertanto, per un *boost* nella direzione  $x^1$  avremo

$$\begin{aligned} v'^1 &= c \frac{u'^1}{u'^0} = c \frac{\sinh(\theta) \frac{dx^0}{d\tau} + \cosh(\theta) \frac{dx^1}{d\tau}}{\cosh(\theta) \frac{dx^0}{d\tau} + \sinh(\theta) \frac{dx^1}{d\tau}} = \frac{c \tanh(\theta) + v^1}{1 + \tanh(\theta) v^1/c} \\ v'^2 &= c \frac{u'^2}{u'^0} = c \frac{\frac{dx^2}{d\tau}}{\cosh(\theta) \frac{dx^0}{d\tau} + \sinh(\theta) \frac{dx^1}{d\tau}} = \frac{v^2}{\cosh(\theta)(1 + \tanh(\theta) v^1/c)} \\ v'^3 &= c \frac{u'^3}{u'^0} = \frac{v^3}{\cosh(\theta)(1 + \tanh(\theta) v^1/c)} \end{aligned} \quad (2.13)$$

Da queste trasformazioni delle velocità possiamo vedere cosa succede componendo piccole velocità e velocità prossime a quella della luce.

Se consideriamo un punto materiale  $P$  che rispetto a  $X$  si muove con velocità piccola ( $|v| \ll c$ ) e consideriamo un boost con parametro  $\beta$  piccolo ( $\beta \ll 1$ ), per l'osservatore  $Y$  avremo che  $P$  si muove con velocità

$$\begin{cases} v'^1 = \frac{c\beta + v^1}{1 + \beta v^1/c} \simeq (c\beta + v^1) \left(1 - \beta \frac{v^1}{c}\right) \simeq v_{tr}^1 + v^1 \\ v'^2 = \frac{v^2}{\gamma(1 + \beta v^1/c)} \simeq v^2 \sqrt{1 - \beta^2} \simeq v^2 \left(1 - \frac{1}{2}\beta^2\right) \simeq v^2 \\ v'^3 = \frac{v^3}{\gamma(1 + \beta v^1/c)} \simeq v^3 \sqrt{1 - \beta^2} \simeq v^3 \end{cases} \quad (2.14)$$

che riproduce la composizione delle velocità nel gruppo di Galileo.

Un altro esempio interessante è quando  $P$  si muove alla velocità della luce per  $X$ . In questo caso, se prendiamo la velocità  $\vec{v} = c(u^1, u^2, u^3)$  per qualche versore  $\vec{u}$ , qualunque sia  $\beta$ , abbiamo

$$\begin{cases} v^1 = c \frac{\beta + u^1}{1 + \beta u^1} \\ v^2 = c \frac{u^2}{\gamma(1 + \beta u^1)} \\ v^3 = c \frac{u^3}{\gamma(1 + \beta u^1)} \end{cases} \Rightarrow |\vec{v}'|^2 = c^2 \frac{1 - \beta^2}{(1 + \beta u^1)^2} \left( \frac{(\beta + u^1)^2}{1 - \beta^2} + (u^2)^2 + (u^3)^2 \right) =$$

$$= c^2 \frac{1 - \beta^2}{(1 + \beta u^1)^2} \frac{1 + \beta^2 + 2\beta u^1 - \beta^2(1 - (u^1)^2)}{1 - \beta^2} =$$

$$= c^2 \frac{1 + 2\beta u^1 + \beta^2(u^1)^2}{(1 + \beta u^1)^2} = c^2 \quad (2.15)$$

in accordo con  $E2$ .

La velocità di trascinamento fra i due sistemi è, come si è detto, la velocità osservata da  $Y$  per un punto che sia in quiete rispetto all'osservatore  $X$ . Per un tale punto  $v = 0$ , e quindi si ottiene

$$v_{tr}^1 = c \tanh(\theta) = c\beta \quad v_{tr}^2 = v_{tr}^3 = 0 \quad (2.16)$$

Dunque per questa trasformazione delle velocità si ha  $v_{tr} = (c \tanh(\theta), 0, 0)$ . Il parametro  $\theta$  della trasformazione determina quindi la velocità relativa fra i due osservatori: per questa ragione è detto *rapidità*. Dalle formule già viste, con pochi passaggi, si può osservare che:

- la rapidità  $\theta$  può assumere qualsiasi valore reale, ma poiché la tangente iperbolica assume solo valori compresi nell'intervallo  $(-1, 1)$  ( $\tanh(\theta) \Rightarrow \pm 1$  per  $\theta \Rightarrow \pm\infty$ ), la velocità di trascinamento ha necessariamente modulo minore di  $c$ . In altri termini, *la velocità di trascinamento fra due osservatori inerziali sarà sempre inferiore alla velocità della luce*;
- se si compongono due *boost* nella stessa direzione, le corrispondenti rapidità si sommano:  $\theta'' = \theta + \theta'$ . Le due velocità di trascinamento, invece, si compongono in modo non lineare:  $v_{tr}'' = c \tanh(\tanh^{-1}(v_{tr}/c) + \tanh^{-1}(v_{tr}'/c))$ ;

Nel caso di una trasformazione di Poincaré generica, per trovare la velocità di trascinamento corrispondente si deve applicare la legge generale di trasformazione delle componenti (2.12) al vettore quadrivelocità di un punto che nel sistema  $X$  è in quiete. Tale vettore avrà componenti  $\vec{v} = (\rho, 0, 0, 0)$ , dove  $\rho = dx^0/d\tau$  dipende della parametrizzazione scelta. Dalla (2.12) si ricava che nel riferimento  $Y$  le componenti della quadrivelocità saranno  $(\rho\Lambda_0^0, \rho\Lambda_0^1, \rho\Lambda_0^2, \rho\Lambda_0^3)$ . Usando la (2.11) si trova subito che la velocità di trascinamento ha componenti  $v_{tr}^i = c \frac{\Lambda_0^i}{\Lambda_0^0}$

Si noti che questa relazione vale anche per le trasformazioni di Galileo.

### 3. Spaziotempo einsteiniano ed effetti relativistici

In termini di struttura geometrica dello spaziotempo, il principio di inerzia e i due postulati di Einstein ( $E1$ ,  $E2$ ) conducono alle seguenti assunzioni:

- M1. lo spaziotempo è uno spazio affine quadridimensionale; in esso, le linee di universo di punti materiali non soggetti a forze sono rette;*
- M2. lo spaziotempo è dotato di una forma bilineare simmetrica  $\eta$  di segnatura  $(3, 1)$ , che determina le equazioni di propagazione delle onde elettromagnetiche. Tale forma bilineare definisce una struttura metrica pseudoeuclidea, detta metrica di Minkowski. In*

coordinate cartesiane, la metrica  $\eta$  ha componenti costanti e pertanto le traiettorie geodetiche per tale metrica sono tutte e sole le rette affini.

M3. Le trasformazioni di coordinate nello spaziotempo che connettono due osservatori inerziali sono le trasformazioni affini che conservano la metrica  $\eta$ , ossia le trasformazioni di Poincaré.

Uno spazio affine quadridimensionale con una struttura metrica di segnatura  $(3, 1)$  è detto a sua volta *spazio di Minkowski* e si denota con  $(M = \mathbb{R}^4, \eta)$ .

Noi abbiamo assunto  $\eta = \text{diag}(-1, 1, 1, 1)$ , mentre una parte della letteratura in fisica adotta una convenzione opposta, con una matrice  $\eta$  di segnatura  $(1, 3)$ , ossia  $\eta = \text{diag}(1, -1, -1, -1)$ . Questo introduce delle differenze di segno in alcune formule, ma le due formulazioni sono perfettamente equivalenti. La segnatura  $(1, 3)$  è più usata fra i fisici delle particelle, la  $(3, 1)$  nei lavori di relatività e cosmologia.

Negli USA, a causa della distribuzione geografica dei maggiori gruppi di ricerca nell'uno e nell'altro campo, la  $\eta$  con segnatura  $(+ - - -)$  è detta *West-coast metric*, quella con segnatura  $(- + + +)$  *East-coast metric*.

Inizialmente si era immaginato di introdurre una coordinata temporale immaginaria, e anziché con  $x^0 = ct$  la si indicava con  $x^4 = ict$ : in questo modo la metrica  $\eta$  appare definita positiva, ma la scelta non comporta vantaggi teorici né pratici.

La possibilità di ottenere tanto la metrica di Minkowski quanto la metrica euclidea come restrizioni su sottospazi diversi di un'unica struttura in uno spazio quadridimensionale complesso è stata però sfruttata nella fisica quantistica ed è nota come *rotazione di Wick* (dal nome di un importante fisico teorico torinese che utilizzò in questo senso un metodo fondamentale dell'analisi complessa, il *prolungamento analitico*).

Vediamo ora le conseguenze fenomenologiche del fatto di aver sostituito alla struttura newtoniana descritta da (N1-N5) la struttura descritta da (M1-M3) come modello matematico dello spaziotempo.

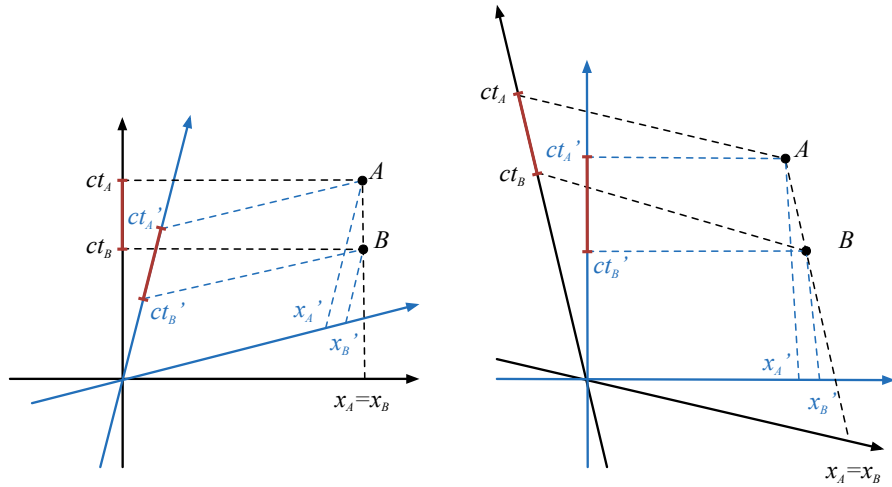


Figura 4: Dilatazione dei tempi. La stessa scena descritta nei riferimenti  $X$  e  $Y$ .

## Dilatazione dei tempi

Secondo i postulati newtoniani (N2-N3), un intervallo di tempo dovrebbe avere la stessa misura per tutti gli osservatori.

Consideriamo dunque due eventi  $A$  e  $B$  che per un osservatore  $X$  avvengono successivamente nello stesso punto dello spazio tridimensionale,  $x_A^i = x_B^i$ , e calcoliamo l'intervallo di tempo intercorrente misurato da un osservatore  $Y$ , connesso a  $X$  da un *boost* (2.9) lungo l'asse  $x$ . Poiché per ipotesi  $x_A^1 - x_B^1 = 0$ , si trova subito

$$c(t'_B - t'_A) = \gamma \cdot (ct_B + \beta x_B) - \gamma \cdot (ct_A + \beta x_A) = c\gamma \cdot (t_B - t_A) \quad (3.1)$$

ossia l'intervallo di tempo  $\Delta t' = t'_B - t'_A$  per l'osservatore  $Y$  risulta dilatato del fattore  $\gamma$  (2.10), che è sempre maggiore di 1 se gli osservatori sono in moto uno rispetto all'altro, rispetto all'intervallo di tempo  $\Delta t = t_B - t_A$  per l'osservatore  $X$ .

Naturalmente, l'intervallo di tempo fra (altri) due eventi (diciamo  $A$  e  $C$ ) che avvengono nello stesso punto dello spazio per l'osservatore  $Y$  risulterà a sua volta dilatato di un fattore  $\gamma$  se misurato dall'osservatore  $X$ .

È importante notare che gli eventi  $A$  e  $B$  sono descritti, nel piano a sinistra in Figura, da eventi di coordinate  $(ct_A, x_A)$  e  $(ct_B, x_B)$  rispetto agli assi di  $X$ , mentre sono descritti, nel piano a destra in Figura, da eventi di coordinate  $(ct'_A = \gamma(ct_A + \beta x_A), x'_A = \gamma(\beta ct_A + x_A))$  e  $(ct'_B = \gamma(ct_B + \beta x_B), x'_B = \gamma(\beta ct_B + x_B))$  rispetto agli assi di  $Y$ .

Nello stesso tempo,  $(ct'_A, x'_A)$  e  $(ct'_B, x'_B)$  sono anche le coordinate degli eventi nel grafico a sinistra rispetto agli assi di  $Y$ . Analogamente,  $(ct_A, x_A)$  e  $(ct_B, x_B)$  sono anche le coordinate degli eventi nel grafico a destra rispetto agli assi di  $X$ .

In altre parole, la trasformazione di Lorentz si può vedere sia come un cambiamento di riferimento nel medesimo piano (= spaziotempo di Minkowski) sia come la trasformazione che trasforma il grafico a sinistra in quello a destra.

In tutti i casi la distanza, calcolata con la metrica  $\eta$ , sui grafici a destra o a sinistra in figura di segmenti corrispondenti non cambia. Ad esempio, la lunghezza del segmento  $ct'_B - ct'_A$  non cambia sui grafici a destra e a sinistra, purché sia calcolato con la metrica  $\eta$  (che fornisce un risultato diverso dalla semplice misura euclidea).

Più in generale, dati due eventi  $A$  e  $B$  qualsiasi, un osservatore  $X$  definisce l'*intervallo di tempo intercorso tra  $A$  e  $B$*  come la quantità  $\Delta t = t_B - t_A$  riferendosi ai suoi assi (quindi, indifferentemente, a destra o a sinistra in Figura).

Infatti, anche se  $A$  e  $B$  non sono nello stesso punto spaziale per  $X$ , tutti gli eventi  $(t_A, x)$  sono per  $X$  contemporanei a  $A$ , qualunque sia la loro posizione nello spazio. Tra questi eventi ce ne sarà uno  $C$  che  $X$  giudica nello stesso punto spaziale di  $B$  e può usare il discorso precedente per calcolare  $\Delta t$  che è sia l'intervallo di tempo trascorso tra  $C$  e  $B$ , che l'intervallo di tempo intercorso tra  $A$  e  $B$  (sia l'intervallo di tempo trascorso tra qualunque  $(t_A, x)$  e  $B$ ) rispetto all'osservatore  $X$ .

A questo punto non stupisce troppo che due osservatori inerziali diversi, misurino tempi diversi tra i medesimi eventi  $A$  e  $B$ , visto che  $X$  e  $Y$  in pratica calcolano la lunghezza di segmenti diversi.

L'effetto sopra descritto corrisponde alla "dilatazione degli intervalli di tempo" introdotta da Lorentz, ma la sua interpretazione corretta nel modello einsteiniano non è quella di un fenomeno fisico per cui "un orologio in moto rallenta": questo presupporrebbe l'esistenza di un riferimento privilegiato (quello dell'etere, o di un metodo operativo di controllare a distanza la sincronizzazione di orologi), in contrasto con il principio di relatività *E1*. Viceversa, l'effetto è conseguenza del fatto che *la misura degli intervalli temporali non è invariante* rispetto a un

cambio di osservatore. L'osservatore che misura il minimo intervallo temporale fra i battiti successivi di un orologio è quello solidale all'orologio.

### Contrazione delle lunghezze

Anche per comprendere l'effetto di “contrazione dello spazio” si deve analizzare con cura in che cosa consista la *misura di una distanza spaziale*. Questo è quello che fa preliminarmente Einstein nel suo lavoro del 1905: egli suppone di compiere tale misura con l'invio di un segnale luminoso da un punto (nello spazio)  $a$  a un punto (nello spazio)  $b$ , ove il segnale viene riflesso indietro verso  $a$ : dal tempo totale di percorrenza, sapendo che la velocità del segnale è  $c$ , si ricava facilmente la distanza fra  $a$  e  $b$ . Questo modo di effettuare la misura può apparire inusuale, ma corrisponde esattamente a una misura interferometrica, come nell'esperimento di Michelson e Morley (o anche al funzionamento del radar).

In realtà tutti i modi per misurare distanze considerevoli, sulla superficie terrestre o fra i pianeti del Sistema Solare, prevedono l'uso di segnali elettromagnetici, eventualmente con una variante del procedimento, e cioè con una misura del *parallasse*, in cui si misura un angolo anziché un intervallo di tempo. Ai nostri giorni, comunque, in molti contesti “quotidiani” (ad esempio in edilizia) si usano dispositivi laser per effettuare misure di distanze anche di pochi metri.

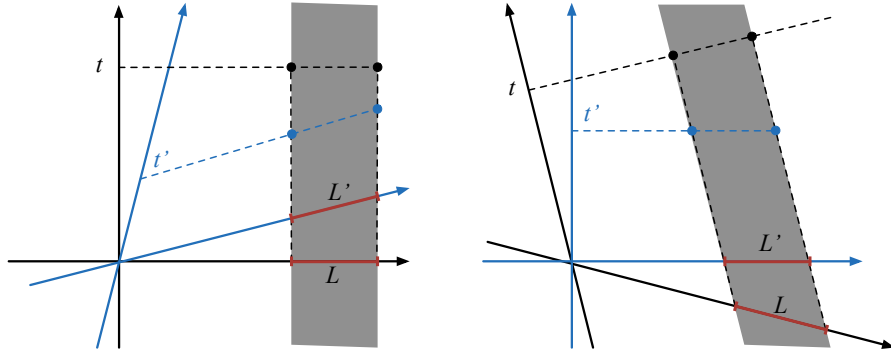


Figura 5: Contrazione delle lunghezze. La stessa scena descritta nei riferimenti  $X$  e  $Y$ .

Qui descriveremo invece la misura della distanza spaziale indipendentemente da come essa sia effettuata in pratica, basandoci sul concetto di simultaneità: questo potrà risultare più aderente all'idea di “misurare una lunghezza con un righello graduato” (ma vedremo più oltre che la nozione di simultaneità non ha affatto il significato empirico “immediato” con cui siamo abituati a identificarla).

Consideriamo due punti  $a$  e  $b$  nello spazio tridimensionale, fissi rispetto a un osservatore  $X$ . Entrambi i punti materiali tracciano una linea di universo nello spaziotempo di Minkowski. Per il punto  $a$  la linea di universo scritta nelle coordinate di  $X$  sarà  $\sigma_a : s \mapsto (s, x_a^i)$  con  $x_a^i$  costanti, visto che il punto  $a$  è fisso nello spazio per  $X$ . Analogamente,  $\sigma_b : s \mapsto (s, x_b^i)$ . Sia inoltre, ragionevolmente  $|x_b - x_a| = L$  la distanza tra  $a$  e  $b$  misurata da  $X$ .

Ciò che  $X$  misura come “distanza spaziale” fra  $a$  e  $b$  è la distanza euclidea fra i due eventi  $A$ , *posizione di  $a$  in un dato istante*  $x_A^0 = ct_0$ , e  $B$ , *posizione di  $b$  nello stesso istante*  $x_B^0 = ct_0$ .



Nello spaziotempo, questo corrisponde, scegliendo opportunamente la parametrizzazione e il riferimento in modo che i punti siano entrambi posti sull'asse  $x^1$ ), a considerare  $x_b^1 = x_a^1 + L$ . Le due linee di universo sono quindi:

$$\begin{cases} x_a^0(\tau) = \tau \\ x_a^1(\tau) = x_a^i \end{cases} \quad \begin{cases} x_b^0(\sigma) = \sigma \\ x_b^1(\sigma) = x_a^i + L \equiv x_b^i \end{cases} \quad (3.2)$$

e due punti simultanei sulle due linee di universo si ottengono, ovviamente, ponendo  $\sigma = \tau = ct_0$ , cioè

$$x_a^0(\tau) = x_b^0(\sigma) = ct_0 \quad (3.3)$$

e la distanza spaziale fra questi due eventi è  $x_b^1(\tau) - x_a^1(\tau) = L$ .

Per l'osservatore  $Y$ , le medesime linee di universo dei punti  $a$  e  $b$  sono rappresentate parametricamente da

$$\begin{cases} y_a^0(\tau) = \gamma x_a^0(\tau) + \beta \gamma x_a^1(\tau) = \gamma \tau \\ y_a^1(\tau) = \beta \gamma x_a^0(\tau) + \gamma x_a^1(\tau) = \beta \gamma \tau \end{cases} \quad \begin{cases} y_b^0(\sigma) = \gamma x_b^0(\sigma) + \beta \gamma x_b^1(\sigma) = \gamma \sigma + \beta \gamma L \\ y_b^1(\sigma) = \beta \gamma x_b^0(\sigma) + \gamma x_b^1(\sigma) = \beta \gamma \sigma + \gamma L \end{cases} \quad (3.4)$$

Ora  $Y$  deve scegliere su queste linee di universo due eventi simultanei, diciamo a  $y_b^0(\sigma) = y_a^0(\tau) = ct'_0$ . Da questa condizione che individua due eventi simultanei per  $Y$  si ottiene

$$y_a^0(\tau) = y_b^0(\sigma) \quad \Longleftrightarrow \quad \tau = \sigma + \beta L, \quad (3.5)$$

e la distanza fra le coordinate spaziali di questi due eventi è

$$L' = y_b^1(\sigma) - y_a^1(\sigma + \beta L) = \beta \gamma \sigma + \gamma L - \beta \gamma (\sigma + \beta L) = \gamma (1 - \beta^2) L = \frac{L}{\gamma}. \quad (3.6)$$

Quindi la distanza spaziale misurata da  $X$  è ridotta per il fattore  $\gamma^{-1}$ , che in effetti è sempre minore (o uguale quando  $\beta = 0$ ) di 1, l'inverso del fattore di dilatazione temporale.

Anche in questo caso, non si tratta di un fenomeno fisico per cui “i corpi in moto rispetto all'etere sono schiacciati nella direzione del moto”, come ipotizzato inizialmente da FitzGerald e da Lorentz, bensì della combinazione di due effetti del cambiamento di osservatore:

- i) le distanze spaziali fra due eventi non sono invarianti;
- ii) i due osservatori, in questo caso, non misurano la distanza fra gli stessi due eventi: le posizioni simultanee dei punti  $a$  e  $b$  per l'osservatore  $X$  non sono simultanee per l'osservatore  $Y$ , e viceversa.

Come nel caso della dilatazione dei tempi, sia la lunghezza del segmento  $L$  che quella del segmento  $L'$  possono essere misurate (usando la metrica  $\eta$ , non un righello) indifferentemente sul grafico di destra o di sinistra, e non sorprende che  $X$  e  $Y$  misurino lunghezze diverse visto che, per implementare la simultaneità, finiscono per misurare segmenti diversi nello spaziotempo.

## Effetto Doppler

Torniamo ancora all'effetto di dilatazione temporale. Nella letteratura divulgativa, il fenomeno è spesso descritto in termini come questi: “l'osservatore  $X$  vede l'orologio solidale a  $Y$  battere più lentamente del proprio orologio”. Dicendolo in questo modo, ad essere ottimisti, si rischia un fraintendimento.

L'osservatore  $X$ , inteso come “sperimentatore dotato di uno strumento di misura”, occupa fisicamente una posizione precisa nello spazio (diciamo l'origine del proprio riferimento spaziale), e ha una propria linea di universo (che coincide con l'asse dei tempi  $ct$ ). L'osservatore  $X$  *vede* un evento nello spazio tempo quando viene raggiunto da un segnale luminoso emesso nell'evento in questione. L'evento  $A$  e l'evento “un segnale luminoso emesso in  $A$  raggiunge la linea di universo dell'osservatore” sono due eventi distinti (a meno che l'evento  $A$  non si trovi sulla stessa linea di universo dell'osservatore).

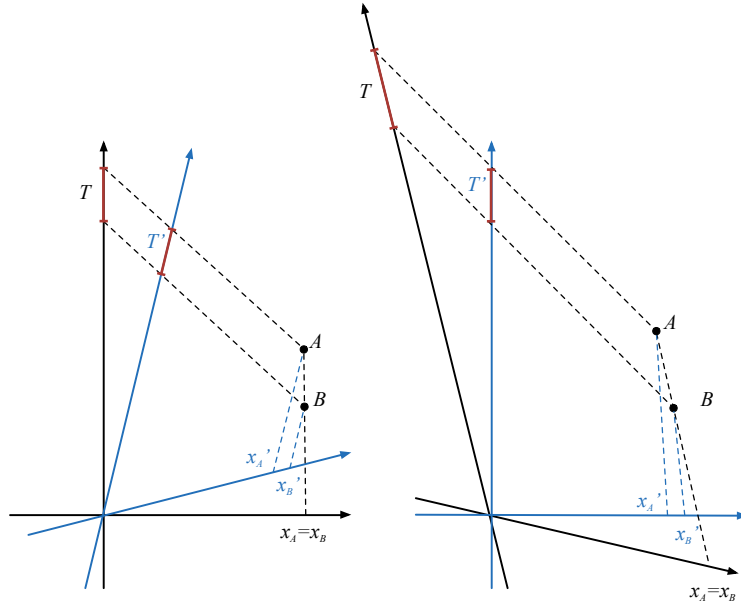


Figura 6: Effetto Doppler. La stessa scena descritta nei riferimenti  $X$  e  $Y$ .

Chiediamoci allora, nel caso di due segnali emessi da una sorgente in quiete rispetto a un osservatore  $X$ , con un intervallo temporale  $T$ , quale intervallo  $T'$  intercorre fra le *ricezioni* dei due segnali da parte dell'osservatore  $Y$ , rispetto a cui la sorgente è in moto.

Che l'intervallo di ricezione  $T'$  sia diverso da quello di emissione  $T$  non è una conseguenza peculiare della relatività: anche nella fisica newtoniana l'intervallo  $T'$  risulta minore di  $T$  se la sorgente si avvicina all'osservatore e maggiore di  $T$  se la sorgente si allontana. È il ben noto *effetto Doppler (newtoniano)*, che si verifica qualunque sia la velocità di propagazione del segnale (è esperienza comune sperimentarlo per segnali acustici emessi da sorgenti in movimento). Tuttavia, l'entità dell'effetto risulta diversa se si considerano trasformazioni di Lorentz anziché di Galileo.

Supponiamo dunque che negli eventi  $A$  e  $B$  (considerati più sopra per esemplificare la dilatazione temporale) siano emessi due segnali luminosi, diretti verso la linea di universo dell'osservatore  $X$ ; per comodità, supponiamo  $x_A^1 = x_B^1 > 0$ , cosicché entrambi i segnali devono propagarsi a velocità  $-c$  (rispetto a  $X$  e quindi pure rispetto a  $Y$ ). Le linee di universo dei due segnali, per l'osservatore  $X$ , sono rette parallele alla bisettrice  $x^1 = -x^0$ , e si possono descrivere come segue:

$$\begin{cases} x_A^0(\tau) = \tau + x_A^0 \\ x_A^1(\tau) = -\tau + x_A^1 \end{cases} \quad \begin{cases} x_B^0(\sigma) = \sigma + x_A^0 + T \\ x_B^1(\sigma) = -\sigma + x_A^1 \end{cases} \quad (3.7)$$

Le coordinate dei due eventi  $A'$  (il segnale emesso in  $A$  raggiunge l'origine di  $X$ ) e  $B'$  (il segnale emesso in  $B$  raggiunge l'origine di  $X$ ) si ottengono intersecando con l'asse  $x = 0$ , cioè si ottiene  $\tau = \sigma = x_A^1$ , e quindi

$$\begin{aligned} A' &= (x_A^1 + x_A^0, 0) & B' &= (x_A^1 + x_A^0 + T, 0) \\ \Rightarrow x_{B'}^0 - x_{A'}^0 &= x_A^1 + x_A^0 + T - x_A^1 - x_A^0 = T \end{aligned} \quad (3.8)$$

Per il segnale emesso da  $A$  abbiamo anche la descrizione di  $Y$

$$\begin{aligned} y_A^0(\tau) &= \gamma(x_A^0 + \beta x_A^1) = \gamma(\tau + x_A^0 - \beta\tau + \beta x_A^1) \\ y_A^1(\tau) &= \gamma(\beta x_A^0 + x_A^1) = \gamma(\beta\tau + \beta x_A^0 - \tau + x_A^1) \end{aligned} \quad (3.9)$$

e per il segnale emesso da  $B$

$$\begin{aligned} y_B^0(\sigma) &= \gamma(x_B^0 + \beta x_B^1) = \gamma(\sigma + x_A^0 + T - \beta\sigma + \beta x_A^1) \\ y_B^1(\sigma) &= \gamma(\beta x_B^0 + x_B^1) = \gamma(\beta\sigma + \beta x_A^0 + \beta T - \sigma + x_A^1) \end{aligned} \quad (3.10)$$

Le coordinate dei due eventi  $A''$  (il segnale emesso in  $A$  raggiunge l'origine di  $Y$ ) e  $B''$  (il segnale emesso in  $B$  raggiunge l'origine di  $Y$ ) si ottengono intersecando con l'asse  $y = 0$ , cioè

$$\begin{aligned} x_A^1(\tau) = 0 &\Rightarrow \tau = \frac{\beta x_A^0 + x_A^1}{1 - \beta} \Rightarrow y_{A''}^0 = \gamma(1 + \beta)(x_A^0 + x_A^1) \quad y_{A''}^1 = 0 \\ x_B^1(\sigma) = 0 &\Rightarrow \sigma = \frac{\beta x_A^0 + \beta T + x_A^1}{1 - \beta} \Rightarrow y_{B''}^0 = \gamma(1 + \beta)(x_A^0 + x_A^1 + T) \quad y_{B''}^1 = 0 \end{aligned} \quad (3.11)$$

Quindi l'intervallo temporale fra  $A''$  e  $B''$  è

$$T' = y_{B''}^0 - y_{A''}^0 = \gamma(1 + \beta)T = \sqrt{\frac{1+\beta}{1-\beta}} T \quad (3.12)$$

(nel caso considerato,  $\beta > 0$  se la sorgente si allontana dall'origine di  $Y$ ,  $\beta < 0$  se si avvicina). Non sorprende il fatto che il coefficiente dell'effetto Doppler sia esattamente  $\gamma$  volte il coefficiente classico.

Il coefficiente classico, che si deve ricalcolare ripetendo il conto in uno spaziotempo galileiano, si può ottenere sviluppando il risultato per piccoli  $\beta$ , in approssimazione lineare:

$$T' = (1 + \beta)T = \left(1 + \frac{v}{c}\right) T \quad (3.13)$$

Dovrebbe essere chiaro, a questo punto, che se  $Y$  osservasse un orologio solidale a  $X$  con un telescopio, lo vedrebbe rallentare con un fattore  $(1 + \beta)\gamma$ , e non  $\gamma$ : i valori di tempo rilevati al telescopio sarebbero infatti il risultato dell'effetto Doppler e non della sola dilatazione temporale.

In pratica, l'effetto Doppler è somma di due effetti. Il fatto che, essendo la sorgente in moto, ogni segnale percorre una distanza diversa, se si allontana la distanza percorsa è sempre maggiore e perciò l'intervallo tra due segnali aumenta.

Nel caso classico, questo rende conto di tutto l'effetto. La contrazione dei tempi, invece, è un effetto puramente relativistico, che non tiene affatto dell'eventuale tempo di volo del segnale. L'effetto Doppler è ottenuto combinando l'effetto relativistico  $T \rightarrow \gamma T = \hat{T}$  e l'effetto Doppler newtoniano  $\hat{T} \rightarrow (1 + \beta)\hat{T} = T'$ .

## 4. Simultaneità e causalità

La relatività della simultaneità (e quindi, come vedremo, dell'ordinamento temporale degli eventi) fu l'aspetto della Relatività Speciale (SR) che apparve subito maggiormente contrastante non solo con le convinzioni comuni e con i postulati della meccanica newtoniana, ma anche con il concetto di tempo fino ad allora considerato dai filosofi (in particolare da Immanuel Kant e, ai tempi di Einstein, da Henri Bergson). Converrà quindi chiarire bene che il concetto fisico–matematico di simultaneità in un sistema di riferimento, che individua i sottospazi  $x^0 = \text{costante}$ , non corrisponde affatto all'idea “intuitiva” di simultaneità che l'esperienza quotidiana ci suggerisce: quest'ultima è legata all'idea che siano simultanei gli eventi che “vediamo simultaneamente”.

Ad esempio, di fronte a una fotografia di un gruppo di persone, o di un panorama, tendiamo a pensare che l'immagine memorizzata rappresenti lo spazio fisico nell'istante in cui è stata scattata la foto. Quando osserviamo una foto astronomica che include corpi celesti a grande distanza dalla Terra, però, ci rendiamo ben conto che non è così. L'immagine di un pianeta del Sistema Solare, ad esempio, corrisponderà allo stato del pianeta qualche minuto (o ora) prima che fosse scattata la foto, mentre l'immagine di una stella potrà risalire ad anni, secoli o millenni prima, a seconda della sua distanza dalla Terra.

A quale insieme di eventi corrisponde quindi l'immagine dello spazio che vediamo in un certo istante?

Dato un generico evento  $(x_o^0, x_o^1, x_o^2, x_o^3)$  dello spaziotempo, consideriamo gli eventi di coordinate  $(x^\mu)$  descritti dall'equazione

$$\eta_{\mu\nu}(x^\mu - x_o^\mu)(x^\nu - x_o^\nu) = 0 : \quad (4.1)$$

Questi formano un'ipersuperficie tridimensionale detta *cono luce (nello spaziotempo)* con vertice nell'evento  $x_o^\mu$ . Poiché l'equazione (4.1) è invariante sotto trasformazioni di Poincaré, il cono luce basato in un dato punto dello spaziotempo è “lo stesso” per tutti gli osservatori.

Con questo si intende che il cono luce con vertice in un evento  $(x_i^0, x_i^1, x_i^2, x_i^3)$  è ottenuto semplicemente traslando quello in  $(x_o^0, x_o^1, x_o^2, x_o^3)$ .

Metà del cono luce si trova nel semispazio  $x^0 < x_o^0$  (*cono luce passato*) e metà nel semispazio  $x^0 > x_o^0$  (*cono luce futuro*).

I segnali luminosi si propagano nello spaziotempo lungo rette il cui vettore tangente soddisfa in ogni punto la condizione

$$\eta_{\mu\nu}u^\mu u^\nu = 0; \quad (4.2)$$

infatti, tenendo conto della (2.11), questa condizione diventa per un qualunque osservatore

$$-(u^0)^2 + \left(\frac{u^0}{c}\right)^2 |v|^2 = 0 \quad \Leftrightarrow \quad |v| = c. \quad (4.3)$$

Tutte le rette di questo tipo passanti per il punto  $(x_o^0, x_o^1, x_o^2, x_o^3)$  giacciono sul cono luce (4.1). Tutti i segnali luminosi ricevuti nell'evento  $(x_o^0, x_o^1, x_o^2, x_o^3)$  (cioè nel vertice del cono) devono dunque provenire da eventi appartenenti al cono luce passato, mentre i segnali emessi nello stesso vertice si propagano lungo il cono luce futuro.

Dunque, l'immagine che vediamo (o che, ad esempio, viene memorizzata da una fotocamera) in  $(x_o^0, x_o^1, x_o^2, x_o^3)$  rappresenta una porzione del cono luce passato, non dell'iperpiano di simultaneità

$x^0 = x_o^0$ . Per strano che possa sembrare, il cono luce è qualcosa che percepiamo concretamente, mentre lo spazio di simultaneità è una pura costruzione matematica: non il contrario.

Osservazione: per fissare le idee, consideriamo l'intersezione fra il cono luce nell'origine e il piano  $(x^0, x^1)$ : questa è data dalle due bisettrici,  $(x^1)^2 - (x^0)^2 = 0$ . Dobbiamo però ricordarci che stiamo usando coordinate in cui  $x^0 = ct$ : se misuriamo il tempo in secondi, quindi, la nostra unità di lunghezza sarà il secondo-luce, ossia circa  $3 \cdot 10^8 m$ . Se disegnassimo lo stesso diagramma spaziotemporale usando come coordinate  $(t, x)$  e come unità di misura il secondo e il metro, allora le due rette che rappresentano il cono luce formerebbero con l'asse  $t$  un angolo uguale a  $\arctan(3 \cdot 10^8) \approx \frac{\pi}{2} - 3.3 \cdot 10^{-9}$ , pressoché indistinguibile dall'angolo retto (corrispondente all'asse  $x$ , cioè alla retta di simultaneità). Ecco perché alle scale di tempo e distanza tipiche della nostra esperienza quotidiana non percepiamo la differenza fra cono luce e iperpiano di simultaneità.

È importante rendersi conto del fatto che la condizione di simultaneità di due eventi per un osservatore è fisicamente e matematicamente ben definita, ma non è sperimentabile in modo diretto: il *sistema di coordinate cartesiane* associato a un osservatore si estende, per definizione, all'intero spaziotempo, ma è un puro costrutto matematico: qualunque *apparato di misura* è localizzato nello spazio. Un osservatore sulla Terra, ad esempio, può a buon diritto affermare che un certo evento  $A$  sulla superficie del Sole è simultaneo a un evento  $B$  avvenuto sulla Terra circa 8 minuti prima rispetto al momento in cui l'immagine dell'evento  $A$  giunge alla Terra. Ma può farlo solo perché sono stati preventivamente misurati il valore di  $c$  e la distanza Terra/Sole: non avrebbe alcun modo di *misurare direttamente* la simultaneità fra  $A$  e  $B$  né, più in generale, l'intervallo di tempo che intercorre fra eventi distanti nello spazio.

Si ripensi a quanto abbiamo detto nel capitolo precedente a proposito dell'effetto Doppler: dalla misura dell'intervallo  $T'$  l'osservatore  $X$  può risalire alla distanza temporale fra gli eventi  $A$  e  $B$  se conosce la velocità  $\beta$  della sorgente, oppure se conosce l'intervallo  $T$  misurato nel sistema di riferimento della sorgente, ma non se ignora entrambi questi dati.

Quindi, il fatto che la simultaneità sia relativa all'osservatore non è in contrasto con alcuna evidenza empirica. Per contro, il contenuto di un'immagine fotografica non dipende dall'osservatore, *proprio perché non rappresenta una porzione di un iperpiano di simultaneità ma una porzione del cono luce, che è lo stesso per tutti gli osservatori*.

Chiediamoci ora: dati due eventi distinti  $A$  e  $B$  nello spaziotempo, esisterà sempre un osservatore per cui tali eventi risultano simultanei? Proviamo a costruire questo ipotetico osservatore  $X$ . Con una rotazione spaziale e un'opportuna scelta dell'origine possiamo quindi sempre fare in modo che i due eventi  $A$  e  $B$  giacciono entrambi sul piano  $(x^0, x^1)$ . A questo punto si tratta di trovare un  $\beta = \beta_0$  tale che attraverso il *boost* (2.9) si abbia  $y_A^0 = y_B^0$ . La condizione è quindi semplicemente

$$0 = y_A^0 - y_B^0 = \gamma(x_A^0 - x_B^0) + \beta\gamma(x_A^1 - x_B^1) \quad \Rightarrow \quad \beta_0 = -\frac{x_A^0 - x_B^0}{x_A^1 - x_B^1} \quad (4.4)$$

ma  $\beta = v_{tr}/c = \tanh(\theta)$ , come sappiamo, quindi  $\beta$  può assumere solo valori compresi nell'intervallo  $(-1, 1)$ . Pertanto, la trasformazione di Lorentz cercata esiste solo se  $|x_A^0 - x_B^0| < |x_A^1 - x_B^1|$ , ossia

$$\eta_{\mu\nu}(x_A^\mu - x_B^\mu)(x_A^\nu - x_B^\nu) > 0 \quad (4.5)$$

La quantità  $\eta_{\mu\nu}(x_A^\mu - x_B^\mu)(x_A^\nu - x_B^\nu)$ , che come sappiamo è indipendente dall'osservatore, è detta *separazione spaziotemporale*. Tale quantità è l'analogo della norma (quadratica) del vettore spostamento in uno spazio affine euclideo, ma con una metrica pseudoeuclidea può assumere anche valori negativi, e può annullarsi anche per spostamenti non nulli.

Due eventi sono simultanei per alcuni osservatori, quindi, se e solo se la loro separazione spaziotemporale è positiva. Diciamo quindi che  $A$  e  $B$  hanno separazione *di tipo spazio*.

In questo caso, inoltre, dato un osservatore  $X$  per cui  $x_B^0 - x_A^0 > 0$  (ossia  $A$  precede  $B$  nel tempo) è possibile trovare un secondo osservatore  $X'$  tale che  $x_B'^0 - x_A'^0 < 0$ , cioè tale che la successione temporale dei due eventi sia invertita: a questo scopo basta prendere un  $\beta$  con lo stesso segno di  $\beta_0$  ma tale che  $|\beta_0| < |\beta| < 1$ .

Questo è impossibile, invece, se  $\left| \frac{x_A^0 - x_B^0}{x_A^1 - x_B^1} \right| \geq 1$ , ossia se

$$\eta_{\mu\nu}(x_A^\mu - x_B^\mu)(x_A^\nu - x_B^\nu) \leq 0 \quad (4.6)$$

Se la separazione spaziotemporale fra  $A$  e  $B$  è negativa o nulla, allora l'ordine temporale fra i due eventi è necessariamente lo stesso per tutti gli osservatori e si dice che  $A$  e  $B$  hanno una separazione spaziotemporale *di tipo tempo* (negativa) o *di tipo luce* (nulla). In compenso, in questo caso di separazione spaziotemporale di tipo tempo, possiamo sempre trovare un osservatore che vede  $A$  e  $B$  accadere nello stesso punto spaziale, cioè tali che

$$0 = y_A^1 - y_B^1 = \gamma\beta(x_A^0 - x_B^0) + \gamma(x_A^1 - x_B^1) \quad \Rightarrow \quad \beta_0 = -\frac{x_A^1 - x_B^1}{x_A^0 - x_B^0} \quad (4.7)$$

Questo implica  $|x_A^1 - x_B^1| < |x_A^0 - x_B^0|$ , e

$$\eta_{\mu\nu}(x_A^\mu - x_B^\mu)(x_A^\nu - x_B^\nu) < 0 \quad (4.8)$$

cioè  $A$  e  $B$  devono avere separazione temporale di tipo tempo. Tale osservatore che vede  $A$  e  $B$  nello stesso punto spaziale si chiama anche *osservatore comovente*.

Nei testi che utilizzano la convenzione opposta (+ ---) sulla segnatura della metrica di Minkowski le disuguaglianze nelle condizioni (4.5), (4.6) e (4.8) risultano invertite. Per questa ragione è conveniente parlare di “separazione di tipo spazio” e “separazione di tipo tempo”, rispettivamente, anziché di “separazione positiva” e “separazione negativa”.

Queste osservazioni sono cruciali in relazione al *principio di causalità*. Qualsiasi interazione di un sistema fisico con un agente esterno può influire sullo stato del sistema solo a partire dal momento dell'interazione, non precedentemente. Quindi, se l'evento  $A$  precede temporalmente l'evento  $B$ , allora  $A$  può avere effetto su  $B$ , mentre  $B$  non può avere un effetto “retroattivo” su  $A$ .

Ma nel modello einsteiniano l'ordinamento temporale fra due eventi, come abbiamo appena visto, può dipendere dall'osservatore.

Il concetto di separazione spaziotemporale permette di risolvere l'apparente paradosso: *un evento  $A$  può condizionare un evento  $B$  solo se*

- (i) *la separazione spaziotemporale fra i due eventi è di tipo tempo o nulla e*
- (ii)  *$A$  precede  $B$*

Si noti che se è vera per un osservatore la condizione (ii), sotto la condizione (i), è vera per tutti.

Questo implica che nessuna particella può muoversi a velocità maggiore di  $c$ , e nessun tipo di interazione (o informazione) può trasmettersi a una velocità superiore a  $c$ . Infatti, se una particella o un segnale avessero velocità maggiore  $c$  in un dato istante  $x_o^0 = x^0(\tau_o)$ , allora in quell'evento il prodotto scalare (pseudoeuclideo) del vettore quadrivelocità con se stesso sarebbe maggiore di zero:

$$\eta_{\mu\nu}u^\mu u^\nu > 0 \quad \Leftrightarrow \quad |v| > c. \quad (4.9)$$

Ma in questo caso gli eventi lungo la linea di universo avrebbero separazione di tipo spazio per valori del parametro sufficientemente vicini a  $\tau_o$ ,

$$\eta_{\mu\nu}(x^\mu - x_o^\mu)(x^\nu - x_o^\nu) \approx \eta_{\mu\nu}u_o^\mu u_o^\nu (\tau - \tau_o)^2 > 0, \quad (4.10)$$

e il loro ordinamento temporale dipenderebbe dall'osservatore. Questo violerebbe il principio di causalità, poiché la relazione causale fra gli stati di un sistema fisico in istanti diversi corrisponde al loro ordinamento temporale.

In uno spaziotempo newtoniano, fissato un evento  $A$  e dato un evento  $B$ , o  $B$  è contemporaneo (cioè simultaneo) ad  $A$  o è precedente (cioè nel passato) rispetto ad  $A$ , o è successivo (cioè nel futuro) rispetto ad  $A$ .

Nello spaziotempo einsteiniano, dato  $A$  ci sono eventi futuri (quelli che hanno separazione di tipo tempo o luce e  $t_B > t_A$ ) ed eventi passati (quelli che hanno separazione di tipo tempo o luce e  $t_B < t_A$ ).

Gli eventi con separazione di tipo spazio, invece non sono necessariamente simultanei. Gli eventi con separazione di tipo spazio sono *contemporanei* nel senso che non può esserci influenza causale reciproca, ma contengono gli eventi giudicati simultanei da ogni osservatore, ma ogni osservatore ritiene simultanei eventi diversi. In ogni caso gli eventi contemporanei formano un aperto dello spaziotempo di Minkowski, mentre sono un chiuso (una ipersuperficie) nello spaziotempo newtoniano.

Questo “*presente spesso*” è una delle principali differenze dello spaziotempo relativistico rispetto alla fisica newtoniana.

Infine va detto che siccome lo spazio di Minkowski è affine, dati due eventi  $A$  e  $B$ , allora  $A - B$  è un vettore tangente, in particolare è il vettore tangente la cui norma  $\eta(A - B, A - B)$  definisce se la separazione è di tipo tempo, spazio o luce. Quindi possiamo definire anche il tipo di un vettore  $u$  tangente allo spaziotempo di Minkowski. Diciamo che  $u$  è di tipo tempo (luce, spazio) se  $\eta(u, u)$  è negativo (nullo, positivo).

Una linea di universo è di tipo tempo (luce, spazio) se la sua quadrivelocità in ogni suo punto è di tipo tempo (luce, spazio). Ad esempio, le linee di universo di particelle (raggi di luce) sono curve di tipo tempo (luce).

Data una superficie  $S \subset M$ , essa è di tipo spazio se *ogni* vettore tangente a  $S$  è di tipo spazio. È di tipo tempo (luce) se *esiste* un vettore tangente a  $S$  che sia di tipo tempo (luce). Ad esempio, le superfici di contemporaneità sono di tipo spazio.

Ovviamente esistono curve e superfici che non sono né di tipo tempo, né di tipo luce, né di tipo spazio, ad esempio perché la quadrivelocità in un punto è di tipo tempo mentre è di tipo spazio in un altro punto.

Riassumiamo in termini geometrici quanto discusso in questa sezione. Denotiamo con  $\eta(u, u) = \eta_{\mu\nu}u^\mu u^\nu$  il prodotto scalare (pseudoeuclideo) di un quadrivettore  $u$  con se stesso.

1. *Coni luce nello spaziotempo*: a ogni punto dello spaziotempo è associato un *cono luce*, indipendente dall'osservatore, con equazione (4.1). Tutti i punti del cono luce con vertice in  $A$  hanno separazione spaziotemporale nulla rispetto ad  $A$  (ma non necessariamente fra di loro!). Il cono luce con vertice in  $A$  separa tre regioni disgiunte dello spazio tempo:
  - gli eventi che hanno separazione di tipo spazio (4.5) rispetto ad  $A$ , il cui ordinamento cronologico rispetto ad  $A$  dipende dall'osservatore;

- gli eventi che hanno separazione nulla o di tipo tempo (4.6) da  $A$  e che (per tutti gli osservatori) precedono  $A$  nel tempo (il *passato* di  $A$ );
  - gli eventi che hanno separazione nulla o di tipo tempo da  $A$  e che seguono  $A$  nel tempo (il *futuro* di  $A$ ).
2. *Coni luce nello spazio tangente*: nello spazio tangente a ciascun evento  $A$  è definito il *cono luce (tangente)* con equazione  $\eta(u, u) = 0$  (4.2). I quadrivettori appartenenti al cono luce tangente sono detti *di tipo luce*

Ci si potrebbe aspettare che i vettori tangenti a un cono luce nello spaziotempo debbano essere sempre di tipo luce. E invece no! La condizione di tangenza a una sottovarietà è un'equazione lineare nello spazio tangente (e questo vale anche per il cono luce, in ogni punto diverso dal vertice che è un punto singolare), mentre l'equazione del cono luce nello spazio tangente è quadratica. Solo nel vertice del cono luce (4.1) i vettori tangenti al cono sono effettivamente tutti e soli quelli appartenenti al cono luce tangente (4.2); in ogni altro punto, i vettori tangenti al cono possono essere di tipo luce o di tipo spazio (esercizio: spiegare perché).

3. *Quadrivettori e linee di universo*: un quadrivettore  $u$  è detto *di tipo spazio* se  $\eta(u, u) > 0$  e *di tipo tempo* se  $\eta(u, u) < 0$ .

La ragione di questi nomi è che per il sistema di coordinate ortonormali associate a un qualsiasi osservatore interziale i vettori  $\frac{\partial}{\partial x^i}$  sono sempre di tipo spazio, mentre il vettore  $\frac{\partial}{\partial x^0}$  è sempre di tipo tempo.

Per il principio di causalità, il vettore tangente alla linea di universo di una particella o di un segnale ha necessariamente  $\eta(u, u) \leq 0$ . In questo modo, ogni evento  $A$  sulla linea di universo la divide in una parte interamente contenuta nel passato di  $A$  e una parte interamente contenuta nel futuro di  $A$ .

## 5. Parametrizzazione delle linee di universo

Una linea di universo è una traiettoria nello spaziotempo: la sua parametrizzazione non ha significato fisico, nel senso che differenti parametrizzazioni della medesima traiettoria descrivono il medesimo moto nello spazio fisico.

Come abbiamo già notato, a parametrizzazioni diverse corrispondono quadrivelocità diverse: a seguito di un cambiamento di parametro,  $\tau \mapsto \sigma(\tau)$  avremo  $u^\mu = \frac{dx^\mu}{d\tau} \mapsto \frac{dx^\mu}{d\sigma} = \frac{dx^\mu}{d\tau} \left(\frac{d\sigma}{d\tau}\right)^{-1}$ . Quindi le quadrivelocità corrispondenti a due parametrizzazioni diverse di una stessa linea di universo hanno, in ciascun punto, la stessa direzione: cambia solo il modulo del quadrivettore  $u^\mu$ . Fissare il valore di  $\eta(u, u) \equiv \eta_{\mu\nu} u^\mu u^\nu$  determina completamente la parametrizzazione di una linea di universo, a meno di una costante additiva.

Anche se la parametrizzazione delle linee di universo e il modulo della quadrivelocità non sono proprietà fisiche, nella descrizione del moto di un punto materiale vi sono due possibili parametrizzazioni che risultano particolarmente utili: il *tempo relativo* e il *tempo proprio*.

### Tempo relativo

Al fine di descrivere il moto visto da uno specifico osservatore, può essere conveniente parame-



trizzare la linea di universo usando il *tempo relativo all'osservatore*:

$$\tau = t \quad \Rightarrow \quad u^0 = \frac{dx^0}{d\tau} = c \quad \Rightarrow \quad u^i = v^i \quad (5.1)$$

Con questa parametrizzazione,  $\eta_{\mu\nu}u^\mu u^\nu = |v|^2 - c^2$ .

La parametrizzazione con il tempo relativo è conveniente perché le componenti della quadrivelocità coincidono con quantità fisicamente misurabili, ancorché relative all'osservatore. Se si cambia osservatore, tuttavia, si deve anche cambiare la parametrizzazione.

### *Tempo proprio*

Un'altra parametrizzazione utile, questa volta indipendente dall'osservatore, è quella del *tempo proprio*. Su una generica linea di universo di tipo tempo si può calcolare la *lunghezza d'arco pseudoeuclidea* fra due eventi

$$s(\tau_0, \tau_1) = \int_{\tau_0}^{\tau_1} \sqrt{-\eta_{\mu\nu}u^\mu(\tau)u^\nu(\tau)} d\tau;$$

Il valore di questo integrale non dipende dalla parametrizzazione usata, né dal sistema di coordinate. Il tempo proprio è definito fissando arbitrariamente l'istante iniziale  $\tau_0$ , e usando

$$\tilde{\tau} = \frac{s(\tau_0, \tau)}{c}$$

come parametro sulla curva. Rispetto a una parametrizzazione generica  $\tau$ , derivando si ottiene

$$\frac{d\tilde{\tau}}{d\tau} = \frac{1}{c} \sqrt{-\eta_{\mu\nu}u^\mu(\tau)u^\nu(\tau)}; \quad (5.2)$$

quindi se la parametrizzazione coincide con il tempo proprio,  $\tau = \tilde{\tau}$ , si ha

$$\eta_{\mu\nu}u^\mu u^\nu = -c^2 \quad (5.3)$$

Questa condizione sul vettore quadrivelocità è invariante sotto trasformazioni di Lorentz e caratterizza il tempo proprio (a meno di una costante additiva arbitraria).

Perché  $\tilde{\tau}$  si chiama “tempo proprio”? Supponiamo che la linea di universo in questione sia una retta. In questo caso esiste un osservatore inerziale  $Y$  per il quale il punto materiale descritto dalla linea di universo è in quiete. Nelle coordinate di questo osservatore, usando come parametro il tempo relativo, il vettore quadrivelocità ha componenti  $(c, 0, 0, 0)$ , pertanto  $\tilde{\tau} = t + \text{cost.}$  (dove  $t$  è il tempo relativo all'osservatore  $Y$  solidale con il punto materiale). Per un osservatore solidale al punto materiale, dunque, tempo relativo e tempo proprio coincidono.

Per un generico osservatore  $X$  che vede il punto materiale in moto a velocità  $v$ , con la parametrizzazione del tempo relativo il vettore quadrivelocità ha componenti  $(c, v^1, v^2, v^3)$ ; quali sono le componenti della quadrivelocità corrispondenti al tempo proprio? Per ottenerle, basta osservare che il rapporto fra i moduli dei due vettori è uguale a  $\sqrt{\frac{-c^2}{|v|^2 - c^2}} = \sqrt{\frac{1}{1 - |v|^2/c^2}} = \gamma$ .

Dunque con la parametrizzazione del tempo proprio il vettore quadrivelocità ha componenti  $(\gamma c, \gamma v^1, \gamma v^2, \gamma v^3)$ . La quantità  $\gamma$  corrisponde al fattore di Lorentz per la trasformazione che connette l'osservatore  $X$  con un osservatore solidale al punto materiale: è per l'appunto il fattore di dilatazione fra il tempo relativo e il tempo proprio.

Più in generale, se la linea di universo non è una retta, possiamo definire *in un dato istante* un *boost* che connetta l'osservatore  $X$  con un secondo osservatore  $Y$  (diverso in ciascun istante) per il quale il punto materiale ha, in quell'istante, velocità nulla. A questo fine è sufficiente che la velocità di trascinamento  $v_{tr}$  fra i due sistemi sia uguale alla velocità  $v$  del punto materiale vista da  $X$  nell'istante dato. È facile verificare che nel medesimo istante si ha  $\frac{d\tilde{\tau}}{dt} = 1$ , dove  $t$  è il tempo relativo a  $Y$ . In questo senso, anche se per un moto che non sia rettilineo uniforme non esiste un osservatore inerziale solidale con il punto materiale, il tempo proprio definito dall'ascissa pseudoeuclidea coincide in ogni istante con il tempo relativo di un osservatore per il quale il punto materiale ha velocità nulla. In accordo con questo è uso comune, in riferimento a un punto materiale in moto a velocità  $v$ , impiegare i simboli  $\beta$  e  $\gamma$  per denotare, rispettivamente, il rapporto  $\frac{v}{c}$  e la quantità  $\frac{1}{\sqrt{1-\beta^2}}$ : anche noi adotteremo nel seguito questa notazione.

Una parametrizzazione è detta *uniforme* se si ha  $\eta(u, u) = \text{cost}$ . La parametrizzazione con il tempo proprio è sempre uniforme (5.3). La parametrizzazione con il tempo relativo a un osservatore, invece, in generale non è uniforme, poiché come abbiamo visto  $\eta(u, u) = |v|^2 - c^2$ : è uniforme solo se la velocità scalare  $|v|$  del punto materiale è costante.

Se si considerano più linee di universo, corrispondenti a punti materiali distinti, è possibile parametrizzarle tutte con il tempo relativo all'osservatore. Per contro, il tempo proprio sarà definito diversamente su ciascuna linea di universo.

Riassumiamo le formule relative alle due parametrizzazioni, che ci serviranno in seguito:

$$\text{Tempo relativo:} \quad \tau = t, \quad u = (c, v^1, v^2, v^3), \quad \eta_{\mu\nu} u^\mu u^\nu = |v|^2 - c^2$$

$$\text{Tempo proprio:} \quad \frac{dt}{d\tau} = \gamma = \frac{1}{\sqrt{1-\frac{|v|^2}{c^2}}}, \quad u = (\gamma c, \gamma v^1, \gamma v^2, \gamma v^3), \quad \eta_{\mu\nu} u^\mu u^\nu = -c^2$$

## 6. Il paradosso dei gemelli

Due gemelli, Aldo ( $A$ ) e Barbara ( $B$ ) vengono separati alle 00:00 del 1 gennaio 2100. Aldo resta sulla terra e continua la sua vita con al collo un orologio atomico che segna, con buona approssimazione, il tempo che passa per lui.

Barbara invece, anche lei con al collo un orologio atomico, identico a quello di Aldo, collaudato per tenere lo stesso tempo con un'approssimazione, diciamo, di poche parti per miliardo (cioè pochi centesimi di secondo per anno), viene messa su una astronave e spedita a una velocità  $v = \beta c$  ( $\beta = \frac{3}{4}$ ) verso Alpha Centauri. Per mezzo di una tecnologia oggi sconosciuta (o per brutale approssimazione) il viaggio di Barbara avviene a velocità costante. Dopo un mese (misurato dal suo— di  $B$ —orologio, in  $2.628 \cdot 10^6 s$ ) inverte la direzione (istantaneamente senza sfraccellarsi) e torna sulla terra dopo ovviamente un altro mese misurato da lei. All'arrivo il suo orologio viene confrontato con quello di  $B$ .

A questo punto, uno non dovrebbe più stupirsi che i due orologi possano segnare tempi diversi, se non perché la situazione sembra simmetrica per  $A$  e  $B$ . Dal punto di vista di  $B$  è  $A$  ad essersi allontanato a velocità  $v = \beta c$  da lei, e poi ad essersi riavvicinato. Al massimo cambia la durata del viaggio che cambia, visto che 2 mesi per  $B$  non necessariamente corrisponde a 2 mesi per  $A$ , che sarebbe la situazione simmetrica in cui chi viaggia misura la durata.

Ma comunque è difficile pensare, anche prima di iniziare il conto, a una ragione per cui l'orologio di  $A$  debba essere avanti (o indietro) rispetto a quello di  $B$  quando si reincontrano. Almeno

sembra difficile propendere per un segno piuttosto che l'altro in virtù della simmetria tra  $A$  e  $B$ . Quindi in genere uno si aspetterebbe, per simmetria, nessuna differenza, visto che è difficile stabilire quale orologio ci debba aspettare sia avanti rispetto all'altro. Comunque possiamo calcolare tutto e poi discutere.

Un paradosso è definito come:

*Proposizione formulata in apparente contraddizione con l'esperienza comune ( i p. degli stoici ) o con i principi elementari della logica, ma che all'esame critico si dimostra valida.*

L'enfasi per noi qui è su *apparente*. Non stiamo mostrando che la relatività speciale non è corretta, men che meno per le sue previsioni controintuitive. Stiamo cercando di mostrare che il paradosso dei gemelli è un effetto reale previsto dalla SR. Esempi sono chiaramente osservati in laboratorio. Ad esempio, i muoni prodotti dai raggi cosmici nella stratosfera dovrebbero decadere prima di raggiungere il suolo se si tiene conto della loro vita media e della loro velocità. Invece si osserva il contrario che ad esempio, è perfettamente spiegato dal fatto che la loro vita media in moto sia effettivamente, misurabilmente più lunga se misurata dall'osservatore a terra.

Cominciamo a rappresentare graficamente la situazione vista da  $B$ , per la quale abbiamo più dati. La partenza di  $B$  la scegliamo come origine  $O$ , durante il viaggio di andata la linea di universo di  $A$  è  $OP$ , durante il viaggio di ritorno è  $PR$ .

Il tempo trascorso per  $B$  durante l'andata è la lunghezza del segmento  $OQ$ , la distanza percorsa è la lunghezza del segmento  $PQ$ . Sappiamo che la velocità del viaggio è  $\beta$ , quindi la linea  $OP$  ha equazione  $-\beta ct = x$ . Il viaggio arriva alla sua massima distanza quanto, in  $P$ ,  $t_P = 2.628 \cdot 10^6 s$ , cioè a distanza  $PQ$  pari a  $\Delta x_B = 5.913 \cdot 10^{14} m$  misurata da  $B$  ( $x_P = -5.913 \cdot 10^{14} m$ ). Sono qualche centinaio di volte oltre la distanza tra sole e Urano, qualche decina di volte più lontano della sonda Pioneer (nel 2019).

A questo punto,  $B$  inverte il suo moto e torna a terra nel tempo corrispondente alla lunghezza del segmento  $QR = QO = c \cdot 2.628 \cdot 10^6 s$ . La retta su cui giace il segmento  $PR$  ha equazione  $x = x_P + \beta c(t - t_P)$ . Il punto  $R$ , per  $B$  ha coordinate  $x_R = 0$  e  $t_R = 5.256 \cdot 10^6 s$ . Il punto  $Q$  sempre rispetto a  $B$ , ha coordinate  $x_Q = 0$  e  $t_Q = 2.628 \cdot 10^6 s$ . In particolare, per  $B$  il viaggio dura ( $OR$ )  $\Delta T_B = 5.256 \cdot 10^6 s$ , cioè 2 mesi.

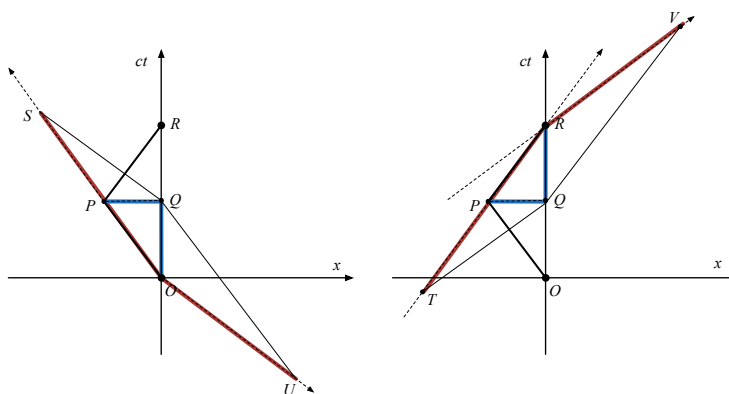


Figura 7: L'esperimento dei gemelli. a) andata, b) ritorno. Il rosso visto da  $A$ , il blue visto da  $B$ .

È chiaro che possiamo calcolare tutta la scena riferita a  $B$  facilmente. Per quanto riguarda la

scena riferita a  $A$ , possiamo semplicemente applicare una trasformazione di Poincaré opportuna, in realtà due, una per l'andata e una per il ritorno, in cui  $A$  diventa solidale con sistema di riferimento. Ma se siamo interessati a misurare quello che può misurare  $A$ , è sufficiente individuare i segmenti che misurerebbe e calcolarne la lunghezza nel piano riferito a  $B$ , visto che la separazione spaziotemporale è invariante per trasformazioni di Poincaré.

Per l'andata,  $P$  “vede” invertire il moto nell'evento  $Q$  (anche se in realtà lo vedrà solo quando la luce emessa in  $Q$  lo raggiungerà). Ma la distanza di  $Q$  misurata da  $A$  corrisponde alla lunghezza del segmento  $UO$  (che in effetti sono contemporanei per  $A$ ) e il tempo a cui  $Q$  avviene (per  $B$ ) corrisponde alla lunghezza di  $SO$ . La retta per  $SO$  è già nota ( $x = -\beta ct$ ) mentre le rette di contemporaneità per  $A$  sono parallele a  $-\beta x = ct$ . Quindi bisogna solo proiettare  $Q$  su questi assi.

Basta risolvere i sistemi

$$\begin{cases} (SO): & x = -\beta ct \\ (QS): & -\beta(x - x_Q) = c(t - t_Q) \end{cases} \quad \begin{cases} (UQ): & (x - x_Q) = -\beta c(t - t_Q) \\ (UO): & -\beta x = ct \end{cases} \quad (6.1)$$

per ottenere rispettivamente  $S$  e  $U$ . Si ottiene

$$S = (ct_S = c\gamma^2 \Delta T_B, x_S = -\beta\gamma^2 c \Delta T_B) \quad U = (ct_U = -\beta^2 \gamma^2 c \Delta T_B, x_U = \beta\gamma^2 c \Delta T_B)$$

Si ottengono le distanze dall'origine calcolate con la metrica  $\eta$

$$|SO| = c\gamma \Delta T_B \quad |UO| = \gamma\beta c \Delta T_B \quad (6.2)$$

Fin qui nulla di nuovo, visto che abbiamo solo ripetuto il conto per contrazione delle distanze e dilatazione dei tempi (l'unica differenza è che stiamo considerando un orologio solidale a  $B$  e un regolo solidale ad  $A$ , come si evince dalla figura 7.a).

Durante il viaggio di ritorno, fissiamo il riferimento di  $A$  all'arrivo nell'origine  $R$ . Le rette che fungono da assi sono  $x - x_R = \beta c(t - t_R)$  per l'asse dei tempi,  $\beta(x - x_R) = c(t - t_R)$  per la retta di contemporaneità. Rispetto a questo sistema riproiettiamo l'evento  $Q$ . Il segmento  $TR$  sta per il tempo di ritorno misurato da  $A$ , il segmento  $RV$  per la distanza percorsa durante il ritorno.

Basta risolvere i sistemi

$$\begin{cases} (TR): & (x - x_R) = \beta c(t - t_R) \\ (QT): & \beta(x - x_Q) = c(t - t_Q) \end{cases} \quad \begin{cases} (VQ): & (x - x_Q) = \beta c(t - t_Q) \\ (VR): & \beta(x - x_R) = c(t - t_R) \end{cases} \quad (6.3)$$

per ottenere rispettivamente  $T$  e  $V$ .

Si ottiene

$$TR = (ct_T = -\gamma^2 c \Delta T_B, -\gamma^2 \beta c \Delta T_B) \quad VR = (\beta^2 \gamma^2 c \Delta T_B, \beta\gamma^2 c \Delta T_B)$$

Si ottengono le distanze dall'origine calcolate con la metrica  $\eta$

$$|TR| = c\gamma \Delta T_B \quad |VR| = \gamma\beta c \Delta T_B \quad (6.4)$$

**Esercizio:** produrre i grafici nel sistema di riferimento di  $A$  applicando le opportune trasformazioni di Poincaré.

Per riassumere, per  $B$  il viaggio dura  $T_B = 2\Delta T_B$ , lungo una distanza  $\Delta_B = \beta c\Delta T_B$  (sia all'andata che al ritorno). Per  $A$  il viaggio dura  $cT_A = |SO| + |TR| = \gamma c\Delta T_B + \gamma c\Delta T_B = 2\gamma c\Delta T_B$ . Notare che l'andata e il ritorno durano lo stesso tempo anche per  $A$  (ma dilatati di un fattore  $\gamma$  rispetto a  $B$ ) e la distanza percorsa è  $\Delta_A = |OU| = |VR| = \gamma\beta c\Delta T_B$ , anch'essi uguali e dilatati dello stesso fattore  $\gamma$  rispetto alla misura di  $B$ .

Quindi al loro ritorno,  $A$  è più vecchio di  $B$  di un fattore  $\gamma = \frac{4}{7}\sqrt{7} \simeq 1.512$ , cioè di circa un mese. L'apparente simmetria tra  $A$  e  $B$  è solo apparente. Quando  $B$  inverte il suo moto lei percepisce una accelerazione,  $A$  al contrario no.

## Appendix A. L'esperimento di Michelson e Morley

Nel 1887 Albert A. Michelson e Edward W. Morley cercarono di confrontare le velocità di propagazione della luce lungo direzioni diverse dello spazio. La radiazione elettromagnetica, in quegli anni, era correntemente interpretata come un'onda di deformazione in un mezzo elastico detto *etere luminifero*, che si supponeva permeasse tutto lo spazio fisico.

In conseguenza delle equazioni di Maxwell, la propagazione delle onde elettromagnetiche nel vuoto è descritta dall'*equazione d'onda di d'Alembert*,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad (A.1)$$

dove  $u$  è una qualsiasi componente del vettore campo elettrico  $E$  o del campo magnetico  $B$ , e la costante  $c$  è legata a due costanti fisiche che compaiono nelle equazioni di Maxwell, la *costante dielettrica del vuoto*  $\varepsilon_0$  e la *permeabilità magnetica del vuoto*  $\mu_0$ , attraverso la formula

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \quad (A.2)$$

La costante  $c$  corrisponde alla velocità scalare con cui si propaga un fronte d'onda. Nel caso di onde in un mezzo elastico, questa velocità è quella relativa a un osservatore solidale con il mezzo in cui l'onda si propaga.

Misurando la velocità della luce relativamente a un laboratorio terrestre, quindi, ci si attendeva di poter misurare il moto della Terra rispetto all'etere luminifero. Infatti, la velocità osservata sarebbe dovuta essere la somma (vettoriale) della velocità del fronte d'onda luminoso nell'etere (con modulo  $c$ ) e della velocità di trascinamento dell'osservatore terrestre rispetto al sistema solidale con l'etere: pertanto, sarebbe dovuta risultare massima (o minima, a seconda del verso di propagazione) nella direzione collineare a  $v_{tr}$ .

L'apparato sperimentale ideato da Michelson consiste in un interferometro con due bracci complanari e perpendicolari  $OA$  e  $OB$ . Alle estremità  $A$  e  $B$  dei due bracci sono posti due specchi, che riflettono la luce emessa da una singola sorgente posta in  $O$ . Agendo sulle lunghezze dei bracci e osservando le figure di interferenza che si formano in  $O$  si può garantire che il tempo in cui due raggi luminosi percorrono avanti e indietro i due bracci sia lo stesso. Se si suppone che la luce abbia la stessa velocità nelle due direzioni ortogonali, la lunghezza dei due bracci sarà allora uguale.

Supponiamo ora che l'interferometro sia in moto rispetto al mezzo in cui si propaga la luce, e che uno dei bracci (diciamo  $OA$ ) sia collineare alla velocità di trascinamento: tenendo conto

di (1.17), detta  $L_A$  la lunghezza del braccio  $OA$  dell'interferometro, il tempo  $T_A$  in cui il fronte d'onda luminoso percorre avanti e indietro la distanza  $L_A$  sarà

$$T_A = \frac{L_A}{c + v_{tr}} + \frac{L_A}{c - v_{tr}} = \frac{2L_A}{c(1 - v_{tr}^2/c^2)} \quad (A.3)$$

Per contro, consideriamo nel sistema di riferimento dell'etere la luce che percorre avanti e indietro a velocità  $c$  l'altro braccio  $OB$ , di lunghezza  $L_B$ , mettendoci un tempo  $T_B$ . Rispetto all'etere, la luce percorre una distanza pari a  $2\sqrt{L_B^2 + v_{tr}^2 T_B^2}$ . Questa distanza, divisa per  $c$ , dà a sua volta il tempo di percorrenza  $T_B$ : risolvendo, si trova

$$T_B = \frac{2L_B}{c\sqrt{1 - v_{tr}^2/c^2}} \quad (A.4)$$

In una prima analisi dell'esperimento Michelson aveva erroneamente supposto che lungo il braccio  $OB$  la luce percorresse la lunghezza  $L_B$  a velocità  $c$ , ma questo errore fu poi corretto.

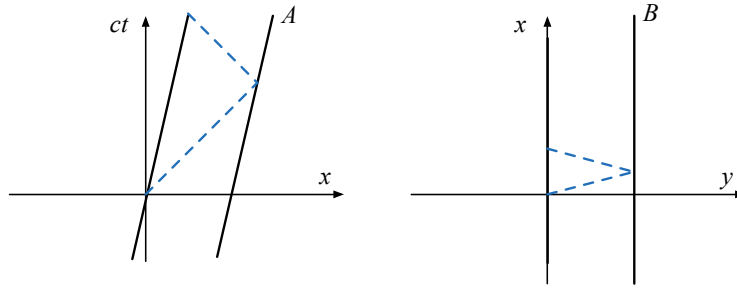


Figura 7: esperimento di Michelson e Morley nello spaziotempo newtoniano.  
a) Lo specchio  $A$  nella direzione del boost. b) Lo specchio  $B$  ortogonale al boost.

Quindi, se lo sperimentatore regola la distanza  $L_B$  in modo che  $T_A = T_B$ , cosicché le frange di interferenza in  $O$  scompaiano, si avrà

$$\frac{L_B}{L_A} = \frac{1}{\sqrt{1 - v_{tr}^2/c^2}}. \quad (A.5)$$

A questo punto, lo sperimentatore ruota l'interferometro di  $\frac{\pi}{2}$ , senza variare la lunghezza dei bracci.

A questo scopo, lo strumento di Michelson e Morley era montato su una base estremamente rigida: una lastra di pietra, galleggiante in una vasca piena di mercurio, in cui la lastra poteva ruotare praticamente senza attrito e senza subire deformazioni. Per amplificare l'effetto, gli specchi erano disposti in modo da riflettere la luce otto volte lungo ciascun braccio prima di giungere a formare le figure di interferenza: le distanze  $L_A$  e  $L_B$  così risultanti erano intorno agli  $11m$ .

La direzione dei bracci rispetto al moto dell'etere risulta ora invertita, quindi i due tempi di percorrenza saranno diversi e si produrranno delle diverse frange di interferenza. Nella pratica, inizialmente la direzione del moto rispetto all'etere non è nota, per cui l'esperimento viene fatto ruotando progressivamente l'interferometro: ci si aspetta di osservare uno spostamento delle frange di interferenza, la cui entità dovrebbe variare sinusoidalmente in funzione dell'angolo

di rotazione dell'apparecchio. Sotto queste ipotesi, dallo spostamento massimo delle frange è possibile risalire alla velocità  $v_{tr}$ . Inoltre, tenendo conto che la Terra ruota su se stessa e si sposta lungo la sua orbita, ripetendo l'esperimento in momenti diversi ci si aspetta di osservare valori differenti di  $v_{tr}$ .

Nulla di tutto questo fu osservato da Michelson e Morley. Le frange di interferenza, al netto delle fluttuazioni dovute alle variazioni della temperatura ambientale, restarono fisse indipendentemente dall'angolo di rotazione dell'interferometro.

Si sarebbe quindi dovuto concludere che l'etere è immobile rispetto alla Terra; questo non solo appariva inverosimile, ma avrebbe dovuto comportare dei precisi effetti sull'aberrazione della luce proveniente dalle stelle, effetti che invece non si osservavano.

In astronomia si definisce *aberrazione* lo spostamento della posizione apparente dei corpi celesti per un osservatore terrestre, dovuto al moto della Terra.

Un primo tentativo di spiegazione del mancato spostamento delle frange di interferenza fu avanzata da George F. FitzGerald nel 1889 e successivamente da Hendrik A. Lorentz nel 1892: essi supposero che un fenomeno fisico di natura elettromagnetica determinasse uno schiacciamento dei corpi in moto rispetto all'etere, nella direzione della velocità relativa all'etere, con un fattore  $\sqrt{1 - (v_{tr}/c)^2}$ . In questo modo, qualunque sia l'angolo che i bracci dell'interferometro formano rispetto alla velocità di trascinamento dell'etere, la diversa velocità di propagazione nella direzione di  $v_{tr}$  sarebbe esattamente compensata dalla contrazione della lunghezza del braccio in quella direzione.

Tuttavia, la sola contrazione delle lunghezze non basterebbe a spiegare l'assenza di spostamento delle frange al variare di  $v_{tr}$  quando si ripete l'esperimento in momenti diversi della rotazione terrestre. Infatti, anche se il tempo di percorrenza è lo stesso per i due bracci ( $T_A = T_B$ ) qualunque sia la loro direzione, tuttavia per valori diversi di  $v_{tr}$  tale tempo di percorrenza dovrebbe variare. Dato che nemmeno questa variazione si osserva, si deve supporre non solo che le lunghezze si contraggano nella direzione del moto relativo all'etere, ma anche che gli intervalli temporali risultino dilatati di un fattore  $\frac{1}{\sqrt{1 - (v_{tr}/c)^2}}$ : questo effetto appare ancora più misterioso della contrazione delle lunghezze.

In un contesto relativistico, sia la contrazione delle lunghezze che la dilatazione dei tempi ipotizzate da Lorentz e FitzGerald sono perfettamente spiegate e si può dire che per l'esperimento di Michelson e Morley, in SR non si prevede nessuna frangia di interferenza, come osservato.

## 8. Dinamica della particella relativistica

Sotto una trasformazione di Lorentz, l'accelerazione di un punto materiale non resta invariata. Pertanto l'equazione fondamentale della dinamica newtoniana,  $\vec{F} = m\vec{a}$ , non è compatibile con la Relatività Ristretta: la forza agente su un punto materiale in un dato istante, infatti, non risulta più la stessa per tutti gli osservatori inerziali.

Esistono approcci diversi al problema di costruire una dinamica del punto materiale compatibile con la Relatività Ristretta. Qui ci baseremo sull'assunzione che la dinamica debba essere descritta da un principio variazionale, ossia da una lagrangiana. Altri approcci, basati su diversi presupposti (o strategie euristiche) conducono comunque al medesimo risultato.

Si tratta innanzitutto di stabilire se il principio variazionale che cerchiamo debba selezionare

i moti (nello spazio fisico tridimensionale) visti da un generico osservatore inerziale, oppure debba selezionare le linee di universo nello spaziotempo che corrispondono ai moti fisici. Entrambe le strade sono percorribili. Il primo passo è comunque la costruzione di un principio variazionale che descriva correttamente il moto di un punto materiale *in assenza di interazioni* (particella libera): questo deve essere un moto rettilineo uniforme per qualunque osservatore inerziale, e deve essere rappresentato (in qualunque sistema di coordinate inerziali) da una retta nello spaziotempo. La funzione di Lagrange per la particella libera, quindi, dovrà essere invariante sotto trasformazioni di Lorentz (anche non omogenee, ossia anche con una possibile traslazione dell'origine delle coordinate nello spaziotempo). Costruire una lagrangiana Lorentz-invariante che sia funzione delle componenti della velocità tridimensionale relativa a un osservatore,  $v^i$ , è possibile ma non è immediato perché queste componenti, come abbiamo visto, si trasformano in modo nonlineare. Quindi proponiamoci inizialmente di costruire una lagrangiana Lorentz-invariante che dipenda dalle componenti  $u^\mu$  della quadrivelocità. Una funzione Lorentz-invariante è ovviamente la norma pseudoeuclidea  $\eta_{\mu\nu}u^\mu u^\nu$ . Proviamo quindi a scegliere come lagrangiana questa funzione (con una costante moltiplicativa, che diventerà rilevante quando dovremo aggiungere un termine di interazione):

$$L = \frac{m}{2} \eta_{\mu\nu} u^\mu u^\nu \quad (8.1)$$

Poiché nelle coordinate che usiamo la metrica di Minkowski  $\eta_{\mu\nu}$  ha componenti costanti, le equazioni di Eulero-Lagrange sono semplicemente

$$\frac{d}{d\tau}(m\eta_{\mu\nu}u^\mu) = m\eta_{\mu\nu}\frac{du^\mu}{d\tau} = 0 \quad (8.2)$$

le cui soluzioni sono rette parametrizzate uniformemente,  $x^\mu(\tau) = u_0^\mu \tau + x_0^\mu$ .

Come sappiamo, le componenti della quadrivelocità non sono fisicamente osservabili: esse dipendono dalla parametrizzazione. Le equazioni (8.2) ammettono una e una sola soluzione per ogni dato iniziale  $(x_0^\mu, u_0^\mu)$ : ma due quadrivelocità iniziali che appartengano alla stessa retta corrispondono allo stesso stato fisico del sistema, in quanto le componenti della velocità “fisica”  $v^i$  sono le stesse. D'altra parte, lungo ogni soluzione la norma della quadrivelocità deve restare costante,

$$\frac{d}{d\tau}(\eta_{\mu\nu}u^\mu u^\nu) = 2u^\nu \eta_{\mu\nu} \frac{du^\mu}{d\tau} = 0 \quad (8.3)$$

pertanto, se imponiamo la condizione iniziale

$$\eta_{\mu\nu}u_0^\mu u_0^\nu = -c^2$$

questa si conserverà lungo le soluzioni, che risulteranno quindi rette parametrizzate con il tempo proprio. In questo modo, per ogni stato fisico della particella otteniamo una sola condizione iniziale e un'unica soluzione. La scelta della (8.1), dunque, corrisponde a tutte le nostre necessità.

Possiamo ora chiederci: se scegliessimo una lagrangiana diversa dalla (8.1) che cosa cambierebbe? Ad esempio, qualunque funzione della norma pseudoeuclidea della quadrivelocità sarebbe invariante sotto trasformazioni di Lorentz. Una tale lagrangiana sarebbe allora una funzione della lagrangiana (8.1):  $L' = f(L)$ . Sappiamo inoltre che per  $L$  il modulo della quadrivelocità (e quindi  $L$  stessa) si conserva lungo le curve di moto. Abbiamo già discusso approfonditamente questa situazione quando abbiamo introdotto la formulazione variazionale



della meccanica lagrangiana: abbiamo osservato che se si sceglie  $f(L)$  in modo tale che anche  $L'$  risulti regolare, allora le soluzioni delle equazioni di Eulero-Lagrange per  $L$  e per  $L'$  sono le stesse. Se però consideriamo

$$L' = -mc\sqrt{-\eta_{\mu\nu}u^\mu u^\nu} \quad (8.4)$$

(il corrispondente integrale d'azione  $A = -mc \int \sqrt{-\eta_{\mu\nu}u^\mu u^\nu} d\tau$  è invariante per riparametrizzazioni e coincide con la lunghezza d'arco pseudoeuclidea, a meno di una costante moltiplicativa), le curve estremali sono ancora rette nello spaziotempo, ma con tutte le parametrizzazioni possibili. In effetti non vale l'unicità della soluzioni per ogni dato iniziale: a parità di dato iniziale si può riparametrizzare arbitrariamente la traiettoria (negli istanti successivi) ottenendo una nuova soluzione.

Le equazioni del moto per la lagrangiana  $L'$  sono

$$\frac{d}{d\tau} \left( \frac{mc \eta_{\mu\nu} u^\mu}{\sqrt{-\eta_{\alpha\beta} u^\alpha u^\beta}} \right) = 0 \quad (8.5)$$

Anche da qui si vede facilmente che tutte le soluzioni dell'eq. (8.2), per le quali vale (8.3), sono soluzioni anche per la lagrangiana  $L'$  (ma non vale il viceversa).

Esercizio: provare a definire l'hamiltoniana della lagrangiana (8.4).

Nei testi di relatività compare prevalentemente una diversa formulazione lagrangiana, nello spazio fisico tridimensionale. Sulla base di considerazioni fisiche diverse da quelle viste qui, si sceglie come lagrangiana la funzione

$$L'' = mc\sqrt{c^2 - |v|^2} = \frac{mc^2}{\gamma} \quad (8.6)$$

(si veda ad es. [Regge], pag. 844). L'integrale d'azione per questa lagrangiana coincide ancora –a meno di una costante moltiplicativa– con il tempo proprio (questo, tra l'altro, mostra immediatamente che l'azione è invariante sotto trasformazioni di Lorentz), ma nella (8.6) le curve sono per definizione parametrizzate con  $t$ , ossia con il tempo relativo all'osservatore.

Osservazione: la lagrangiana (8.6) si ottiene direttamente dalla (8.4) sostituendo in  $L'$  la condizione (5.1). In generale, se si impone una restrizione sulle variabili dinamiche che compaiono in una lagrangiana (ad esempio un vincolo, o come in questo caso una condizione sulle componenti della velocità) e poi si ricavano le equazioni di Eulero-Lagrange, *non si trovano* le stesse equazioni che si troverebbero se la condizione fosse imposta a posteriori sulle equazioni del moto. Ad esempio, se invece della condizione che caratterizza la parametrizzazione con il tempo relativo si cercasse di sostituire nella  $L'$  la condizione corrispondente alla scelta del tempo proprio, e cioè  $\eta_{\mu\nu}u^\mu u^\nu = -c^2$ , la lagrangiana diverrebbe indenticamente costante (uguale a  $mc^2$ ) e quindi non produrrebbe alcuna equazione del moto. Invece, scegliendo la condizione di parametrizzazione con il tempo relativo, sorprendentemente si trova che le soluzioni delle equazioni di Lagrange per  $L''$ ,

$$\frac{d}{dt} \frac{\partial L''}{\partial v^i} = \frac{d}{dt} \frac{mc v^i}{\sqrt{c^2 - |v|^2}} = 0 \quad (8.7)$$

coincidono esattamente con le particolari soluzioni di (8.5) per le quali si sia imposto  $u^0 \equiv c$ . Dunque la scelta del tempo proprio, eccezionalmente, si può imporre equivalentemente sulla lagrangiana  $L'$  o sulle curve di moto.

Abbiamo quindi tre diverse formulazioni lagrangiane per la particella libera: la (8.1), adatta a descrivere le linee di universo nello spaziotempo con la parametrizzazione del tempo proprio; la

(8.4) che descrive le stesse linee di universo ma è compatibile con qualunque parametrizzazione; infine, scegliendo la parametrizzazione con il tempo relativo si ottiene la lagrangiana (8.6). La relazione matematica fra le soluzioni per (8.1) e (8.4) è completamente spiegata dalle considerazioni fatte sopra, mentre il fatto che anche  $L''$  produca le stesse traiettorie, parametrizzate con il tempo relativo, non è affatto una conseguenza banale del fatto di aver imposto la condizione  $u^0 \equiv c$  nella (8.4). Per verificare che le traiettorie sono le stesse è sufficiente confrontare le equazioni del moto, ma per spiegare *perché* questo avvenga, dal punto di vista matematico, occorrono considerazioni geometriche (il metodo della riduzione simplettica su varietà coisotrope nello spazio delle fasi) che esulano dai nostri obiettivi.

### Particella in interazione

Passiamo ora a costruire un termine di interazione che sia compatibile con questo quadro. Se supponiamo di aggiungere un termine di potenziale alla lagrangiana libera, come siamo abituati a fare in meccanica newtoniana, ad esempio

$$L = \frac{m}{2} \eta_{\mu\nu} u^\mu u^\nu + U(x^\lambda) \quad (8.8)$$

oppure

$$L' = -mc \sqrt{-\eta_{\mu\nu} u^\mu u^\nu} + U(x^\lambda), \quad (8.9)$$

nel primo caso le equazioni del moto risultanti non conservano la parametrizzazione uniforme, mentre nel secondo caso non sono invarianti per riparametrizzazioni arbitrarie. Queste lagrangiane, quindi, determinerebbero un effetto non fisico (determinare la parametrizzazione della linea di universo come conseguenza dell'interazione). L'effetto dell'interazione sulla parametrizzazione potrebbe essere sempre eliminato a posteriori con una riparametrizzazione opportuna, tuttavia questa non potrebbe essere stabilita a priori: concettualmente, questo rappresenterebbe un'incongruenza del modello.

Sappiamo, infatti, che l'informazione fisica sul moto della particella è interamente contenuta nella traiettoria: moti corrispondenti a una medesima traiettoria nello spaziotempo, indipendentemente dalla parametrizzazione, sono fisicamente equivalenti. Se è così, dovrebbe essere sempre possibile fissare *in modo indipendente da qualunque interazione* un rappresentante in ciascuna classe di moti equivalenti, al fine di ottenere una descrizione unica. Con l'introduzione di un potenziale dipendente solo dalla posizione del punto materiale, come si è visto, questo non è possibile. Dobbiamo quindi indagare se si possa fare con un potenziale generalizzato, ossia che dipenda anche dalla quadrivelocità.

Un generico potenziale dipendente dalle velocità produrrebbe nelle equazioni del moto un termine di interazione che contiene non solo le componenti della velocità, ma anche le loro derivate  $\frac{du^\lambda}{d\tau}$ . Per evitare questo, occorre che il potenziale dipenda dalle  $u^\lambda$  in modo lineare. Consideriamo quindi lagrangiane della forma

$$L = \frac{m}{2} \eta_{\mu\nu} u^\mu u^\nu - e A_\mu(x^\lambda) u^\mu \quad (8.10)$$

oppure

$$L' = -mc \sqrt{-\eta_{\mu\nu} u^\mu u^\nu} - e A_\mu(x^\lambda) u^\mu \quad (8.11),$$

dove è stata introdotta una costante di accoppiamento  $e$  di cui discuteremo più oltre il significato. Le corrispondenti equazioni di Lagrange sono rispettivamente

$$m\eta_{\mu\nu}\frac{du^\mu}{d\tau} = e\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)u^\mu \quad (8.12)$$

e

$$\frac{d}{d\tau}\left(\frac{mc\eta_{\mu\nu}u^\mu}{\sqrt{-\eta_{\alpha\beta}u^\alpha u^\beta}}\right) = e\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)u^\mu \quad (8.13)$$

Le equazioni (8.13) si riducono alle (8.12) nel caso della parametrizzazione col tempo proprio ( $\eta_{\alpha\beta}u^\alpha u^\beta = -c^2$ ).

Osservazione: le lagrangiane (8.10) e (8.11) non sono più, in generale, invarianti sotto trasformazioni di Lorentz: l'invarianza è rotta dalla presenza del quadripotenziale  $A_\lambda$ . Questo è del tutto corretto, dal punto di vista fisico: anche nella meccanica newtoniana la lagrangiana di una particella libera è invariante sotto trasformazioni di Galileo, ma se si aggiunge un potenziale allora le uniche simmetrie rimanenti sono quelle che lasciano invariato anche il potenziale (e potrebbero non essercene affatto).

Questo significa che la trasformazione di Lorentz  $x'^\nu = \Lambda^\nu_\mu x^\mu$  manda una soluzione per un certo  $A_\mu$ , non in soluzione delle equazioni del sistema per la *stessa*  $A_\mu$  ma in una soluzione del sistema con  $A'_\mu = \bar{\Lambda}^\nu_\mu A_\nu$ . Quindi il campo  $A = A_\mu(x)dx^\mu$  è una 1-forma sullo spaziotempo.

La combinazione di derivate di  $A_\lambda$  che compare nelle equazioni ha un significato matematico evidente. Le  $A_\lambda$ , che sono i coefficienti di una forma lineare applicata alla quadri-velocità, sono le componenti di una 1-forma  $A = A_\lambda(x^\mu)dx^\lambda$ ; detta  $F = F_{\mu\nu}dx^\mu \wedge dx^\nu$  la 2-forma definita dal differenziale esterno di  $A$ ,  $F = dA$ , le componenti  $F_{\mu\nu}$  corrispondono esattamente alle quantità fra parentesi in (8.12) e (8.13), che si possono quindi riscrivere così:

$$m\eta_{\mu\nu}\frac{du^\mu}{d\tau} = eF_{\mu\nu}u^\mu \quad (8.14)$$

e

$$\frac{d}{d\tau}\left(\frac{mc\eta_{\mu\nu}u^\mu}{\sqrt{-\eta_{\alpha\beta}u^\alpha u^\beta}}\right) = eF_{\mu\nu}u^\mu \quad (8.15)$$

Osservazione: è importante notare che nelle equazioni del moto compaiono le componenti del tensore  $F$  e non quelle del quadripotenziale  $A$ . Di conseguenza, due quadripotenziali che differiscano per una 1-forma chiusa (ad esempio  $A$  e  $A' = A + df$ , per una qualsiasi funzione  $f$  sullo spaziotempo) generano gli stessi effetti sul moto delle particelle e sono fisicamente indistinguibili. Il valore numerico delle componenti  $A_\lambda$  in un dato punto dello spazio tempo non ha quindi alcun significato fisico, dato che può essere cambiato a piacimento aggiungendo le derivate parziali di una qualsiasi funzione definita nell'intorno del punto: pertanto la 1-forma  $A$  non rappresenta una grandezza fisica misurabile. Ad essere misurabili sono invece le componenti del tensore  $F$ . La trasformazione  $A \mapsto A' = A + df$ , che lascia invariate le equazioni del moto, è detta *trasformazione di gauge*.

Sotto una trasformazione di gauge, alle lagrangiane (8.10) e (8.11) viene a sommarsi il termine  $\partial_\mu f u^\mu$ . Questo termine, tuttavia, coincide con  $\frac{df}{d\tau}$  su qualsiasi curva nello spaziotempo, e pertanto la sua variazione è identicamente nulla: questo spiega perché la trasformazione di gauge non ha effetto sulle equazioni del moto. Cfr. [wiki](#).

Osserviamo che le equazioni (8.14) conservano la parametrizzazione uniforme, come desiderato. Infatti in questo caso si ha

$$\frac{d}{d\tau}(\eta_{\mu\nu}u^\mu u^\nu) = \frac{e}{m}F_{\mu\nu}u^\mu u^\nu \equiv 0 \quad (8.16)$$

il termine a destra dell'uguale si annulla identicamente poiché la 2-forma  $F$  è antisimmetrica. Per quanto riguarda le (8.15), sono invarianti per riparametrizzazione (questa, infatti, moltiplica il membro destro e quello sinistro per uno stesso fattore), il che corrisponde al fatto che la lagrangiana (8.11), essendo ancora omogenea di grado 1 nelle  $u^\lambda$ , produce un funzionale d'azione che non dipende dalla parametrizzazione.

Le equazioni (8.15), oltre a ridursi come si è già detto alle (8.14) se si sceglie di parametrizzare con il tempo proprio, con la parametrizzazione del tempo relativo si riducono invece (anche questa volta “miracolosamente”) alle equazioni che si otterrebbero variando la lagrangiana

$$L'' = mc\sqrt{c^2 - |v|^2} - e(cA_0 + A_i v^i) \quad (8.17)$$

che è quella che si trova nella maggior parte dei testi di Fisica relativistica. In questa forma si osserva direttamente che il termine omogeneo  $A_\mu(x^\lambda)u^\mu$ , dal punto di vista della descrizione tridimensionale, corrisponde in realtà alla somma di termine di interazione lineare nelle velocità  $v^i$  e un termine, proporzionale a  $A_0(x^\lambda)$ , non dipendente dalle velocità. Quando si considera la parametrizzazione con il tempo proprio, quindi, la componente temporale  $A_0$  del quadripotenziale ha l'effetto di un potenziale “tradizionale” del tipo usualmente considerato in meccanica newtoniana.

Vediamo ora quale significato fisico può avere un'interazione come quella descritta. A questo fine consideriamo le equazioni del moto nello spazio tridimensionale, che a questo punto possiamo ottenere in tre modi:

- dalle equazioni (8.14), considerando che con la parametrizzazione del tempo proprio, come si è già detto, si ha

$$u^0 = c\gamma, u^i = \gamma v^i \quad \Rightarrow \quad \frac{dt}{d\tau} = \gamma \quad \Rightarrow \quad \frac{d}{d\tau} = \gamma \frac{d}{dt} \quad (8.18)$$

quindi le equazioni con indice spaziale  $i$  (ricordando che la metrica  $\eta$  è diagonale) diventano

$$\gamma \frac{d}{d\tau} (m\gamma v^i) = e\gamma (cF_{0i} + F_{ji}v^j) \quad (8.19)$$

mentre l'equazione con indice temporale diventa

$$-\gamma \frac{d}{d\tau} (mc\gamma) = e\gamma F_{j0}v^j \quad (8.20)$$

- dalle equazioni (8.15), imponendo la parametrizzazione con il tempo relativo  $\tau = t$ ,  $u = (c, v^1, v^2, v^3)$ : si ottengono le equazioni

$$\frac{d}{dt} \frac{mc v^i}{\sqrt{c^2 - |v|^2}} = \frac{d}{dt} m\gamma v^i = e(cF_{0i} + F_{ji}v^j), \quad (8.21)$$

che sono manifestamente equivalenti alle (8.19)

- usando la lagrangiana (8.17), che genera direttamente le equazioni (8.21).

A questo punto evidenziamo il diverso ruolo delle componenti di  $F$  nelle equazioni. Ribattezziamo queste componenti nel modo seguente:

$$F_{\alpha\beta} = \begin{pmatrix} 0 & E^1/c & E^2/c & E^3/c \\ -E^1/c & 0 & -B^3 & B^2 \\ -E^2/c & B^3 & 0 & -B^1 \\ -E^3/c & -B^2 & B^1 & 0 \end{pmatrix} \quad (8.22)$$

Da notare che le terne  $E = (E^1, E^2, E^3)$  e  $B = (B^1, B^2, B^3)$ , quando si considera una rotazione degli assi spaziali, si comportano effettivamente come due vettori di  $\mathbb{R}^3$ . Quando invece si passa da un osservatore  $X$  a un osservatore  $Y$ , in moto rispetto al primo, allora le componenti  $E^i$  e  $B^i$  si combinano fra loro.

In questo modo l'equazione del moto diventa

$$\frac{d}{dt} m\gamma v = e(E + v \times B), \quad (8.23)$$

e la restante equazione (8.20) diventa

$$\frac{d}{dt} (m\gamma c^2) = eE \cdot v. \quad (8.24)$$

Le equazioni che abbiamo ottenuto, quindi, nel limite  $\gamma \rightarrow 1$  (ossia per velocità piccole rispetto a  $c$ ) sono esattamente quelle della forza di Lorentz esercitata dal campo elettromagnetico su una particella con carica elettrica  $e$ . Con questa identificazione, la componente temporale del quadripotenziale coincide (a meno del segno, a seconda delle convenzioni) con il potenziale elettrico, mentre le componenti spaziali coincidono con quelle del potenziale vettore magnetico.

Osservazione: a questo punto dovrebbe risultare chiaro il motivo per cui abbiamo introdotto due costanti,  $m$  ed  $e$ , nella costruzione della lagrangiana. L'accelerazione della particella in uno stato assegnato, infatti, è determinata da tre elementi: la *massa*, il *campo elettromagnetico* presente nello spaziotempo, e infine la *carica elettrica*, che determina in che misura la particella stessa interagisce con il campo elettromagnetico. Questi tre elementi sono del tutto indipendenti fra loro: se si considera un sistema di più particelle, il potenziale elettromagnetico è lo stesso per tutte, ma la carica di ciascuna potrà essere diversa, e così la massa.

Rispetto all'elettrodinamica classica, però, le equazioni che abbiamo trovato introducono una correzione relativistica: nelle equazioni (8.23) e (8.24) l'inerzia della particella è rappresentata dal fattore  $m\gamma$ , che dipende dalla velocità della particella stessa. Nel “limite nonrelativistico”  $\gamma \rightarrow 1$  questo fattore tende a  $m$  (massa a riposo). La “massa relativistica”  $m\gamma$ , però, tende a infinito quando  $|v| \rightarrow c$ . Nel caso di un campo elettrico costante, ad esempio, il moto della particella relativistica non è uniformemente accelerato: più in generale, non è mai possibile accelerare una particella massiva fino a farle raggiungere la velocità della luce.

L'equazione (8.24) ci dice che la potenza istantanea della forza di Lorentz è uguale alla derivata temporale della quantità  $\mathcal{E} = mc^2\gamma$ , che quindi dovrebbe corrispondere all'energia cinetica. In effetti, questa è precisamente la grandezza che viene conservata, nel caso della lagrangiana libera, come conseguenza dell'invarianza sotto le traslazioni temporali generate nello spaziotempo dal campo vettoriale  $\frac{\partial}{\partial t} = c \frac{\partial}{\partial x^0}$ :

$$\mathcal{E} = -c \frac{\partial L'}{\partial u^0} = cp^0 \quad (8.25)$$

Sviluppando  $\gamma = (1 - \beta^2)^{-1/2}$  in potenze di  $\beta$  si ottiene

$$\mathcal{E} = mc^2 + \frac{1}{2}m|v|^2 + O(\beta^4) \quad (8.26)$$

Quindi l'energia cinetica relativistica si riduce a quella classica,  $T = \frac{1}{2}m|v|^2$ , quando  $\beta = \frac{|v|}{c} \ll 1$ , tranne per il termine costante  $\mathcal{E}_0 = mc^2$ .

Questo termine fu interpretato da Einstein come un'energia presente anche in una particella in quiete, per il fatto di essere dotata di massa: può essere osservato solo se, a seguito di un

fenomeno fisico che modifica la massa di una particella, questa *energia a riposo* si converte parzialmente in altre forme di energia, e questo è ciò che effettivamente avviene nei fenomeni nucleari e subnucleari.

Notiamo anche che  $F = dA$  implica  $dF \equiv 0$ . Se si scrivono le equazioni corrispondenti per le componenti  $E^i$  e  $B^i$ , si ottengono esattamente le equazioni di Maxwell omogenee:

$$\operatorname{div}(B) = 0 \quad \operatorname{rot}(E) + \frac{\partial B}{\partial t} = 0 \quad (8.27)$$

ossia metà delle equazioni del campo elettromagnetico. Quindi abbiamo visto che la costruzione di lagrangiane sulla base delle assunzioni che abbiamo formulato (invarianza di Lorentz nel caso di assenza di interazioni; indipendenza dalle interazioni della parametrizzazione della linea di universo) ci ha condotto a una teoria che ha tutte le proprietà dell'interazione delle particelle cariche con il campo elettromagnetico.

L'identificazione di  $F$  con il tensore elettromagnetico (o *tensore di Faraday*) richiederebbe anche che valgano le restanti due equazioni di Maxwell,

$$\operatorname{div}(E) = \frac{\rho}{\varepsilon_0}, \quad \operatorname{rot}(B) - \frac{1}{c^2} \frac{\partial E}{\partial t} = \mu_0 J, \quad (8.28)$$

che descrivono l'effetto della densità di carica elettrica  $\rho$  e della densità di corrente elettrica  $J$  sul campo elettromagnetico. Queste due equazioni si possono ottenere da una lagrangiana per il campo elettromagnetico stesso, il che esula dall'ambito della nostra trattazione.

In linea di principio, come abbiamo visto, si potrebbe rappresentare in questo contesto anche un campo di forze posizionali (eventualmente dipendenti dal tempo) del tipo usuale in meccanica newtoniana: basterebbe supporre  $A = (U(x^\lambda), 0, 0, 0)$ . Tuttavia, questo varrebbe solo per un dato osservatore inerziale: un secondo osservatore (inerziale!) in moto rispetto al precedente vedrebbe necessariamente comparire dei termini di forza dipendenti dalla velocità.

Vi sono altre situazioni tipiche della meccanica newtoniana che non sono più compatibili con il modello relativistico. Ad esempio, non possiamo pensare di descrivere un sistema di *due o più* punti materiali soggetti a forze interne (azioni a distanza). Dovremmo infatti immaginare un potenziale dipendente non solo dalle coordinate spaziotemporali di una particella in un dato istante, ma da quelle di *tutte* le particelle interagenti, *in un medesimo istante*. È evidente che questo sarebbe in contraddizione con la Relatività Ristretta, in quanto la simultaneità è relativa all'osservatore (e quindi l'interazione fra particelle dipenderebbe essa stessa dall'osservatore). Similmente non possiamo ammettere equazioni di vincolo relative, ad esempio, alla distanza spaziale fra punti del sistema (come nel caso dei vincoli di rigidità), dato che questa distanza dipende dall'osservatore.

Viene dunque da chiedersi se di tutti i tipi di forze che incontravamo nella meccanica newtoniana sopravviva, nel modello relativistico, solo l'interazione elettromagnetica. Si noti che anche questa non può più essere vista come azione a distanza fra cariche elettriche. Il campo elettromagnetico ha acquistato il ruolo di un ente fisico autonomo, che infatti può esistere (e propagarsi, come nel caso delle onde elettromagnetiche) in assenza di cariche e correnti. È proprio il campo elettromagnetico che determina l'effetto di interazione fra particelle cariche: ognuna di esse, infatti, interagisce con il campo elettromagnetico, e l'effetto di tale interazione si propaga (a velocità  $c$ ) da una particella all'altra, in modo perfettamente compatibile con le trasformazioni fra osservatori inerziali e con il principio di causalità.

Di fatto, dal punto di vista fisico tutte le forze che si incontrano in meccanica classica sono in realtà la manifestazione (macroscopica) di due sole interazioni fondamentali: quella elettromagnetica e quella gravitazionale.

La fisica moderna conosce altre due interazioni fondamentali, che però agiscono solo a livello nucleare e subnucleare.

Una forza elastica, ad esempio, non è altro che il risultato della deformazione di un reticolo solido tenuto insieme, a livello microscopico, da legami di natura elettrica.

L'interazione elettromagnetica, dunque, l'abbiamo descritta. E l'interazione con la gravità? Einstein ha mostrato, con la teoria della Relatività Generale, che l'interazione gravitazionale può essere descritta da una diversa modifica delle lagrangiane relativistiche che abbiamo qui considerato, e precisamente generalizzando la metrica dello spaziotempo. Lo spaziotempo, anziché uno spazio affine, diventa una varietà differenziabile pseudoriemanniana, in cui la metrica (di segnatura  $(3, 1)$ ) può avere curvatura non nulla – in presenza di materia ed energia – e pertanto non ammettere coordinate ortonormali. Come già sappiamo, se la metrica  $g_{\mu\nu}$  ha coefficienti non costanti, le equazioni del moto per una lagrangiana della forma

$$L = \frac{m}{2} g_{\mu\nu}(x^\lambda) u^\mu u^\nu + e A_\mu(x^\lambda) u^\mu \quad (8.29)$$

includono dei termini quadratici nelle velocità che hanno come coefficienti i simboli di Christoffel: questi, come ha mostrato Einstein, rappresentano correttamente l'effetto della gravità sul moto di una particella massiva, visto da un generico osservatore (non necessariamente inerziale); in particolare, in assenza di interazione elettromagnetica la linea di universo di una particella sarà una geodetica dello spaziotempo.

Ma questa è un'altra storia, e si dovrà raccontare un'altra volta.

## Appendix B. Il quadrimpulso

In meccanica classica, se si descrive la dinamica di un punto materiale usando coordinate cartesiane ortonormali, le derivate della lagrangiana libera (coincidente con l'energia cinetica) rispetto alle componenti della velocità forniscono le componenti della quantità di moto. Un teorema fondamentale della dinamica afferma che la quantità di moto totale, in un sistema di punti materiali, si conserva se il sistema non è soggetto a forze esterne (in generale, la derivata della quantità di moto nel tempo è uguale alla risultante delle forze, e le forze interne fra i punti materiali del sistema hanno risultante nulla per il principio di azione e reazione).

Classicamente, il baricentro di un sistema soggetto solo a forze interne si muove di moto rettilineo uniforme, come si muoverebbe un punto materiale non soggetto a forze e dotato della quantità di moto totale del sistema. Nella dinamica relativistica, invece, non si può parlare di baricentro di un sistema di particelle, se non relativamente a un osservatore. Infatti le coordinate del baricentro in un dato istante sono date dalla media delle posizioni in quello stesso istante dei punti materiali del sistema, pesate con le rispettive masse; ma sappiamo che per osservatori in moto l'uno rispetto all'altro gli spazi di simultaneità sono diversi, e quindi il risultato dipende dall'osservatore.

Definiamo innanzitutto il concetto relativistico di *quadrimpulso* di una particella. Le componenti del quadrimpulso sono ancora definite come le derivate della lagrangiana (per la particella

libera) rispetto alle componenti della quadrivelocità (usando coordinate cartesiane ortonormali):

$$p_\mu = \frac{\partial L}{\partial u^\mu} \quad (B.1)$$

Dato che abbiamo considerato diverse possibili lagrangiane per la particella libera, ci si aspetterebbe che ad ognuna di queste corrisponda una diversa definizione del quadrimpulso. In realtà, la relazione fra le componenti del quadrimpulso e la velocità osservata è sempre la stessa. Se si considera la lagrangiana quadratica (8.1), il risultato è

$$p_\mu = m\eta_{\mu\nu}u^\nu; \quad (B.2)$$

sappiamo che con la parametrizzazione del tempo proprio si ha  $u^0 = c\gamma$ ,  $u^i = \gamma v^i$  e quindi le componenti (covarianti) del quadrimpulso (in coordinate ortonormali) risultano essere  $p_0 = -mc\gamma$ ,  $p_i = m\gamma v^i$ . In genere ci si riferisce però al quadrimpulso come al *vettore*  $p$  ottenuto alzando gli indici con la metrica  $\eta$ : questo ha componenti

$$p = (mc\gamma, m\gamma v^1, m\gamma v^2, m\gamma v^3). \quad (B.3)$$

Se si considera invece la lagrangiana (8.4), allora si trova

$$p_\mu = \frac{\partial L'}{\partial u^\mu} = \frac{mc\eta_{\mu\nu}u^\nu}{\sqrt{-\eta_{\alpha\beta}u^\alpha u^\beta}} \quad (B.4)$$

Queste componenti non dipendono dalla parametrizzazione della linea di universo (sotto una qualsiasi riparametrizzazione, numeratore e denominatore risultano moltiplicati per lo stesso fattore); se si usa il tempo proprio, il denominatore è uguale a  $c$  e si riproduce esattamente l'espressione (B.2). Si può verificare per esercizio che anche se si utilizza il tempo relativo, sostituendo le relazioni ( $u^0 = c$ ,  $u^i = v^i$ ) che valgono in questo caso, si ottiene ancora l'espressione (B.3).

La definizione (B.4) del quadrimpulso, essendo indipendente dalla parametrizzazione, fornisce una grandezza fisica osservabile (diversamente dalla quadrivelocità). Per essa, come si vede facilmente, vale l'identità

$$p_\mu p^\mu = \eta^{\mu\nu} p_\mu p_\nu = -m^2 c^2. \quad (B.5)$$

Ciò significa che non c'è corrispondenza biunivoca fra quadrivelocità e quadrimpulso: per qualsiasi quadrivelocità  $u \in \mathbb{R}^4$  il quadrimpulso definito da (B.4) appartiene necessariamente alla sottovarietà tridimensionale di equazione (B.5), detta *iperboloide di massa* o *mass shell*. Dal punto di vista matematico, questa proprietà è un riflesso del fatto che il quadrimpulso, essendo invariante per riparametrizzazioni, è lo stesso per due quadrivelocità che appartengano alla stessa retta (queste, infatti, corrispondono a parametrizzazioni diverse della stessa linea di universo).

In meccanica classica, quando si devono risolvere problemi di *urti* fra punti materiali, si suppone che le particelle coinvolte si comportino, prima e dopo l'urto, come particelle libere (non soggette a forze), e si cerca di ricavare la relazione che lega le loro quantità di moto prima dell'urto con le quantità corrispondenti dopo l'urto, prescindendo dalla forma specifica del potenziale di interazione.



Se si considerano “urti istantanei” si deve supporre una discontinuità nella velocità dei punti, e in questo caso non è applicabile la descrizione usuale del moto come soluzione di un’equazione differenziale del secondo ordine.

A questo fine, nell’ipotesi che siano presenti solo forze interne, si usa la proprietà di conservazione della quantità di moto totale del sistema e –nel caso di urti elastici– la conservazione dell’energia totale. In questo tipo di problemi è spesso vantaggioso riferirsi a un osservatore inerziale per il quale la quantità di moto totale del sistema risulti nulla: questo è detto *riferimento del centro di massa*, dato che per esso il baricentro del sistema è in quiete (la velocità del baricentro, infatti, è sempre uguale alla quantità di moto totale divisa per la massa totale del sistema).

Nella dinamica relativistica, lo stesso tipo di problemi si risolve utilizzando la conservazione del quadrimpulso totale del sistema,  $P = \sum_{A=1}^N p_A$ , ottenuto sommando vettorialmente i quadripulsi di tutte le  $N$  particelle del sistema. Il modulo del quadrimpulso totale è invariante sotto trasformazioni di Lorentz e la sua radice quadrata è detta *massa invariante (totale)* del sistema.

L’equivalente relativistico del *riferimento del centro di massa* è il riferimento inerziale in cui il quadrimpulso totale del sistema ha *tutte le componenti spaziali nulle*, ossia coincide con il quadrimpulso di una singola particella in quiete dotata della massa totale del sistema. Per ottenerlo, a partire da un osservatore dato, si applica il *boost* nella direzione individuata dal quadrimpulso totale  $P$ , con parametro  $\beta = \frac{\sqrt{(P^1)^2 + (P^2)^2 + (P^3)^2}}{P^0}$ .

## Appendix C. Effetto Compton

La conservazione del quadrimpulso totale permette di descrivere anche l’urto fra un elettrone e un fotone (trascurando la loro natura quantistica, che non consentirebbe in generale di trattarle come punti materiali classici).

A questo scopo, però, si deve assegnare un quadrimpulso al fotone: per una particella di massa nulla non si può usare l’espressione (B.3). I lavori di Einstein sull’effetto fotoelettrico e quelli di A. Compton hanno consentito di identificare le proprietà del quadrimpulso che valgono nel caso di un fotone.

Sia Einstein che Compton ricevettero il premio Nobel proprio per questi studi, rispettivamente nel 1921 e nel 1927

In primo luogo, se consideriamo il limite per  $m \rightarrow 0$  dell’equazione di *mass shell* (B.5), ricaviamo che il quadrimpulso di una particella di massa nulla deve necessariamente essere di tipo luce,  $p_\mu p^\mu = 0$ . L’altro elemento rilevante per identificare il quadrimpulso di un fotone è la relazione (8.25) fra l’energia cinetica relativistica e la componente temporale del quadrimpulso. La teoria del fotone ci dice che l’energia trasportata da un fotone di frequenza  $\nu$  è data dal prodotto di quest’ultima per la *costante di Planck*  $h$ :  $\mathcal{E} = h\nu$ . Quindi possiamo assumere che per un fotone si abbia

$$p^0 = \frac{h\nu}{c} \quad (C.1)$$

Questa identità e l’equazione  $p_\mu p^\mu = 0$  determinano completamente le componenti del quadrimpulso: per un fotone la cui velocità abbia le componenti  $(v^1, v^2, v^3)$ , con  $|v| = c$ , si avrà

$$p = \left( \frac{h\nu}{c}, \frac{h\nu}{c^2} v^1, \frac{h\nu}{c^2} v^2, \frac{h\nu}{c^2} v^3 \right) \quad (C.2)$$

A questo punto possiamo affrontare l'effetto Compton. Supponiamo che un fotone di frequenza  $\nu$  urti un elettrone che è inizialmente in quiete nel sistema di riferimento considerato. Dopo l'urto, il fotone avrà una frequenza  $\nu'$  e avrà una diversa direzione di propagazione. Scegliamo gli assi coordinati in modo che prima dell'urto il fotone si incidente propaghi lungo l'asse  $x^1$ , e il piano  $(x^1, x^2)$  sia quello in cui si propaga il fotone dopo l'urto (fotone diffuso). Prima dell'urto, quindi, il quadrimpulso dell'elettrone e del fotone sono rispettivamente

$$\begin{aligned} p_e &= (m_e c, 0, 0, 0) \\ p_\gamma &= \left(\frac{h\nu}{c}, \frac{h\nu}{c}, 0, 0\right) \end{aligned} \quad (C.3)$$

designiamo il fotone con la lettera  $\gamma$ , come d'uso nella letteratura fisica: da non confondere con il fattore  $\gamma$  che compare nelle equazioni che seguono, che invece corrisponde alla velocità dell'elettrone!

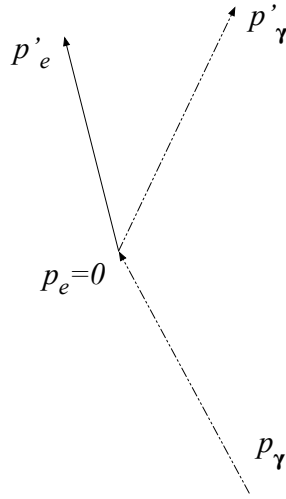


Figura 8: effetto Compton.

Dopo l'urto, l'elettrone si muove con velocità  $v_e$  e il fotone si muove con velocità  $v_\gamma$ : quest'ultima ha componenti  $(c \cos\theta, c \sin\theta, 0)$ , dove  $\theta$  è l'angolo rispetto all'asse  $x^1$ , cioè l'angolo di cui è stato deviato il fotone a seguito dell'urto (*angolo di diffusione* del fotone). Posto  $\gamma = \frac{1}{\sqrt{1-(|v_e|/c)^2}}$ , i quadrimpulsi delle due particelle dopo l'urto diventano

$$\begin{aligned} p'_e &= (m_e \gamma c, m_e \gamma v_e^1, m_e \gamma v_e^2, m_e \gamma v_e^3) \\ p'_\gamma &= \left(\frac{h\nu'}{c}, \frac{h\nu'}{c} \cos\theta, \frac{h\nu'}{c} \sin\theta, 0\right) \end{aligned} \quad (C.4)$$

Imponendo ora la conservazione del quadrimpulso totale,  $p_e + p_\gamma = p'_e + p'_\gamma$ , si trova innanzitutto che deve essere  $v_e^3 = 0$ : la traiettoria del fotone incidente, quella del fotone diffuso e quella dell'elettrone dopo l'urto sono necessariamente complanari. Le equazioni che si ottengono uguagliando le altre componenti di  $P$  e  $P'$  sono (moltiplicando tutto per  $c$  per comodità)

$$\begin{cases} m_e c^2 + h\nu = m_e \gamma c^2 + h\nu' \\ h\nu = m_e \gamma c v_e^1 + h\nu' \cos\theta \\ 0 = m_e \gamma c v_e^2 + h\nu' \sin\theta \\ 0 = m_e \gamma c v_e^3 \end{cases} \quad (C.5)$$

Nella pratica, le grandezze che si osservano sono l'angolo di diffusione  $\theta$  e le frequenze  $\nu$  e  $\nu'$ : vogliamo quindi ottenere una relazione fra queste, eliminando le componenti della velocità finale dell'elettrone,  $(v_e^1, v_e^2)$ , che non vengono misurate. A questo scopo ricaviamo innanzitutto, dalla seconda e terza equazione,

$$m_e^2 c^2 \gamma^2 |v_e|^2 = (h\nu - h\nu' \cos\theta)^2 + (h\nu' \sin\theta)^2 = (h\nu)^2 + (h\nu')^2 - 2h^2 \nu \nu' \cos\theta. \quad (C.6)$$

e riscriviamo la prima come

$$m_e c^2 (\gamma - 1) = h(\nu - \nu') \quad \Rightarrow \quad m_e^2 c^4 (\gamma - 1)^2 = (h\nu)^2 + (h\nu')^2 - 2h^2 \nu \nu'. \quad (C.7)$$

Sottraendo quest'ultima equazione da (C.6) si trova

$$2h^2 \nu \nu' (1 - \cos\theta) = m_e^2 c^4 (\gamma - 1)^2 - m_e^2 c^2 \gamma^2 |v_e|^2 = 2m_e^2 c^4 (1 - \gamma) \quad (C.8)$$

(nell'ultimo passaggio si è usato il fatto che  $\gamma^2 (c^2 - |v_e|^2) = c^2$ ). Usando nuovamente la prima equazione abbiamo

$$2m_e^2 c^4 (\gamma - 1) = 2m_e c^2 h(\nu - \nu') \quad (C.9)$$

e quindi

$$h\nu\nu'(1 - \cos\theta) = m_e c^2 (\nu - \nu') \quad \Rightarrow \quad c \frac{(\nu - \nu')}{\nu\nu'} = \lambda' - \lambda = \frac{h}{m_e c} (1 - \cos\theta) \quad (C.10)$$

In questo modo abbiamo eliminato le variabili associate al moto dell'elettrone, come volevamo:  $\lambda = \frac{c}{\nu}$  e  $\lambda' = \frac{c}{\nu'}$  sono, rispettivamente, le lunghezze d'onda del fotone incidente e del fotone diffuso. Il fattore costante  $\lambda_c = \frac{h}{m_e c}$  è detto *lunghezza d'onda Compton dell'elettrone*. Si noti che l'angolo di diffusione del fotone determina la differenza fra le due lunghezze d'onda  $\lambda$  e  $\lambda'$ , quale che sia la lunghezza d'onda del fotone incidente.