Relatività per principianti

Fondamenti di Relatività Ristretta e Generale con breve Compendio di Fisica Classica

Francesco Borghero e Francesco Demontis

UNICApress/didattica

Questo libro è stato ideato per un corso che spieghi, nella maniera più semplice possibile, le Teorie della Relatività Ristretta e Generale a studenti che si accostano ad esse per la prima volta. I contenuti di queste due teorie fanno ormai parte del bagaglio culturale necessario ad ogni fisico e ad ogni matematico e, in particolare, esse costituiscono un insieme di conoscenze indispensabili per chiunque si occupi di fisica, astrofisica e cosmologia. Il libro è pertanto destinato soprattutto agli studenti delle lauree magistrali in Matematica, Fisica, Astronomia, Astrofisica e Ingegneria Aerospaziale e può essere utilizzato anche in corsi del 3° anno delle lauree triennali. Il testo può essere adattato anche per corsi monografici che trattino solo la Relatività Ristretta o solo la Relatività Generale. Infine, esso ambisce ad interessare un pubblico più vasto avente una buona preparazione matematica e fisica di base, quali possono essere i docenti della scuola secondaria.

UNICApress / didattica

Relatività per principianti
Fondamenti di Relatività Ristretta e Generale con breve Compendio di Fisica Classica

Francesco Borghero e Francesco Demontis

Cagliari
UNICApress
2021

Prima edizione

Immagine di copertina, impaginazione e grafica di Francesco Borghero e Francesco Demontis

© Francesco Borghero, Francesco Demontis e UNICApress, 2021

Licenza CC-BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0/)
Cagliari, UNICApress, 2021 (http://unicapress.unica.it)

ISBN 9788833120300 (versione cartacea)
9788833120317 (versione online)

DOI https://doi.org/10.13125/unicapress.978-88-3312-019-5

Francesco Borghero:

dedico questo libro alla memoria dei miei genitori Giannino e Angela per i grandi sacrifici compiuti per darmi la possibilità di studiare; a mia moglie Maria che ha sempre condiviso le mie aspirazioni: senza di lei non sarei io quel che sono; ai miei amatissimi figli Angela e Pierfrancesco e ai miei deliziosi nipotini Francesco e Giovanni Paolo.

Francesco Demontis:

dedico questo libro ai miei genitori Efisio e Maria Giovanna che, con tanti sacrifici, mi hanno prima di tutto consentito di studiare e, successivamente, sostenuto nelle mie scelte di vita; alla mia compagna Loredana che è sempre stata al mio fianco condividendo sia le gioie che i momenti tristi e al mio adorato figlio Nicolò che, con i suoi sorrisi, illumina le mie giornate. Desidero inoltre dedicare il libro alla memoria dell'altro mio figlio Lorenzo che, purtroppo, ci ha lasciato troppo presto generando, nell'animo dei genitori, un vuoto incolmabile.

Prefazione

Ceux qui aiment la science et qui ont trop
de raisons pour se dèfier de leurs facultés
d'invention, ont encore un rôle utile à jouer:
celui d'èlucider les recherches des autres et de les rèpandre.[1]

Quelli che amano la scienza e che hanno troppe
ragioni per diffidare delle loro capacità d'invenzione,
hanno ancora un ruolo utile da giocare:
quello di delucidare le ricerche degli altri e di diffonderle.
(Jules Tannery (1848-1910), matematico francese)

La Teoria della Relatività costituisce, assieme alla *Meccanica Quantistica*, una delle due grandi rivoluzioni scientifiche avvenute nell'ambito della Fisica del XX° secolo ed è senza dubbio uno dei pilastri della nostra conoscenza della Natura ed una delle strutture portanti della Fisica Moderna.
A differenza della *Meccanica Quantistica*, alla cui formulazione contribuirono i più grandi fisici del secolo scorso, *La Teoria della Relatività* è sostanzialmente opera di un solo uomo, Albert Einstein (1879-1955). Essa è articolata in due parti fondamentali:

1 - La Teoria della Relatività Ristretta (o Speciale o Particolare) (1905);

2 - La Teoria della Relatività Generale (1916).

La prima parte fu formulata da Einstein, alla giovane età di 26 anni, nel seguente articolo:
"Zur Elektrodynamic Bewegter Körper", Annalen der Physik, **7**, 1905, pagg. 891-921.
La seconda parte fu pubblicata da Einstein circa 11 anni dopo col titolo originale:
"Die Grundlage der Allgemeinen Relativitätstheorie", Annalen der Physik (4), **49**, 1916, pagg. 769-822.[2]
Queste due teorie costituiscono due pietre miliari nello sviluppo della conoscenza scientifica, in particolare la Teoria della Relatività Generale è forse uno dei vertici più alti raggiunti finora dall'ingegno umano.
Per la precisione storica è opportuno osservare che la formulazione matematica della Relatività Generale è stata resa possibile - oltre che ovviamente dal formidabile intuito fisico di Einstein - grazie a due fatti fondamentali:

[1] La citazione è riportata nella Prefazione di uno dei primi libri (pubblicati nel mondo) sulla Relatività, si tratta del libro di Roberto Marcolongo, *Relatività*, Casa Editrice G. Principato, 1921, 2^a edizione 1923.

[2] Di questi due lavori fondamentali abbiamo una traduzione italiana dovuta, per il primo di essi, a Paolo Straneo dal titolo *Sull'elettrodinamica dei corpi in moto*, per il secondo ad Aldo M. Pratelli dal titolo: *I fondamenti della teoria della Relatività Generale*. Queste traduzioni sono inserite nel volume *Cinquant'anni di Relatività* a cura di Mario Pantaleo, Editrice Universitaria, Firenze 1955, pagg. 479-504 e pagg. 509-559 (nuova edizione Giunti Barbera, Firenze 1980, pagg. 479-504). Le traduzioni di questi due lavori fondamentali sono state anche riprodotte nel volume: Albert Einstein, *Opere scelte* a cura di Enrico Bellone, Bollati Boringhieri, 1988, pagg. 148-177 e pagg. 282-343.

1) un'idea geniale del matematico lituano Hermann Minkowski (1908) (già professore di Einstein al Politecnico di Zurigo) consistente nel concetto di *"spaziotempo 4-dimensionale"*;

2) l'esistenza di una teoria matematica formulata, già da tempo (1901), da due grandi matematici italiani Gregorio Ricci Curbastro e dal suo allievo Tullio Levi-Civita.

Ciò è chiaramente messo in luce da Einstein stesso nella prima pagina del suo lavoro sui fondamenti della Relatività Generale, che qui riportiamo[3]:
"La generalizzazione della Teoria della Relatività è stata molto facilitata dalla forma data alla Teoria della Relatività Ristretta da Minkowski, il matematico che per primo ha compreso chiaramente l'equivalenza formale tra le coordinate spaziali e la coordinata temporale, rendendola applicabile alla teoria. I mezzi matematici necessari per la Teoria della Relatività Generale erano già pronti nel Calcolo Differenziale Assoluto, il quale si basa sulle ricerche di Gauss, Riemann e Christoffel sulle varietà non euclidee, ed è stato eretto a sistema da Ricci e Levi-Civita e da essi applicato a problemi della Fisica teorica."
La Relatività Speciale fornisce una nuova visione della natura dello spazio e del tempo e forse più di ogni altra teoria fisica, offre fin dall'inizio piacevoli sorprese intellettuali che non mancheranno d'interessare lo studente: essa, tra l'altro, è assolutamente necessaria per comprendere l'elettromagnetismo e le equazioni di Maxwell. Inoltre, lo sviluppo della fisica nucleare prima e della fisica delle particelle poi crearono le condizioni per una vasta verifica sperimentale e per molteplici applicazioni, e ne promosse le discussioni nell'ambito della vita quotidiana, questi concetti divennero di pubblico dominio, creando una nuova cultura.
La Relatività Generale, incorporando le implicazioni fondamentali di quella Speciale, ci consente una comprensione più profonda dello spazio, del tempo e della gravitazione, ed ha profondamente segnato l'intera nostra visione scientifica dell'universo.
" I fisici trovarono serie difficoltà dovendo confrontarsi con concetti e tecniche matematiche totalmente estranei alla cultura scientifica tradizionale. Inoltre **la più che modesta evidenza sperimentale** *per le conseguenze osservative della Relatività Generale, nei suoi primi cinquanta anni di esistenza, impedirono l'interazione fra questa teoria e il resto del mondo della fisica.*
In questo primo cinquantennio, la Relatività generale divenne il feudo di un numero ristretto di matematici e fisici eccelsi, che rimasero tutti incantati dalla profondità e bellezza di questa teoria. I loro nomi sono scolpiti nella cultura scientifica del novecento: Bargmann, Bergmann, Bondi, De Sitter, Dirac, Eddington, Eisenhart, Fermi, Fock, Froedmann, Hoffman, Infeld, Jordan, Gödel, Klein, Landau, Lemaitre, Levi-Civita, Lichnerowicz, Lifshitz, Møller, Pauli, Rosen, Shrödinger, Synge, Taub, Weyl, ecc. Il contatto fra la Relatività Generale ed il mondo reale degli esperimenti e delle osservazioni cresceva di ben poco in un'era dove il resto della fisica era in una fase a dir poco esplosiva, e non solo in senso figurato! Dal 1967, dopo la scoperta delle **pulsar**, *questa situazione è cambiata drasticamente. Da un lato, grazie al lavoro di fisici quali, Gupta, Feynmann, Thirring e Weinberg, venne lanciato un ponte per superare il baratro che separava la Relatività Generale dalle teorie dei campi tradizionali: sia l'idea di geometrizzazione implicita nella Relatività Generale, sia gli aspetti matematici delle non linearità delle equazioni di campo ricevettero la salutare iniezione di una vigorosa reinterpretazione ricca di contenuto fisico. D'altro lato la nascita dell'Astrofisica Relativistica creò una nuova realtà di attività sperimentali, in cui molti effetti relativistici vennero ad essere scoperti ed osservati. Diveniva chiaro che la stessa interpretazione di questi processi astrofisici non poteva essere fatta se non sulla base di una teoria che fondasse le sue radici sulla Relatività Generale. Questi concetti di geometrizzazione sono da allora divenuti di pubblico dominio e ancora una volta siamo stati testimoni della nascita di una nuova cultura."*[4]

[3]Il brano sopra citato si trova nella prima pagina del lavoro fondamentale di Einstein del 1916 (già citato nella nota 1 della pagina precedente) dal titolo: *I fondamenti della teoria della Relatività Generale*, inserito nel volume **Cinquant'anni di Relatività** a cura di Mario Pantaleo, Editrice Universitaria, Firenze 1955, pagg. 509-559. Tale lavoro è riprodotto anche nel volume: Albert Einstein, **Opere scelte** a cura di Enrico Bellone, Bollati Boringhieri, 1988, pagg. 282-343.

[4]Questo brano molto illuminante è preso a prestito dal bel libro di Hans C. Ohanian e Remo Ruffini, *Gravitazione e spazio-tempo*, Zanichelli Editore, Bologna, 1997, pagg. VII-VIII della Prefazione. Consigliamo allo studente la consul-

La Relatività Generale si è dunque brillantemente affermata, negli ultimi decenni, come la migliore teoria per spiegare numerosi oggetti scoperti dall'Astrofisica, **oltre alle pulsar: quasar, buchi neri, radiazione di fondo, lenti gravitazionali e onde gravitazionali,** queste ultime previste teoricamente da Einstein sono state rivelate sperimentalmente soltanto di recente, nel 2015, dando origine ad una nuovissima branca dell'Astronomia: la cosiddetta **Astronomia Multimessagera**, come vedremo in dettaglio nell'APPENDICE al Cap.12.
Inoltre la Relatività Generale ha consentito la nascita di una nuova Cosmologia e la scoperta di numerosi modelli di Universo impensabili nell'ambito della teoria newtoniana della gravitazione. Infine, non dobbiamo dimenticare due notevoli applicazioni di grandissima utilità pratica:

1) la definizione (travagliata) del cosiddetto **Tempo Coordinato Baricentrico** (TCB) fondamentale per descrivere il moto degli oggetti nel sistema solare ed oltre. Si tratta del miglior tempo scala che possa essere definito.
Esso rappresenta il **Tempo Proprio** *di un osservatore posto idealmente all'infinito ed in quiete rispetto al nostro sistema solare. Più precisamente è il* **Tempo Coordinato** *di un sistema di coordinate spazio-temporale che è centrato nel baricentro del nostro sistema solare, che si suppone non ruotante rispetto a galassie lontane e che tende asintoticamente al Tempo Proprio di un osservatore fisso rispetto a questo sistema di coordinate*;

2) la costruzione di uno strumento di grandissima utilità pratica come il GPS (Global Positioning System) che ormai tutti usano e conoscono.

Il libro è destinato soprattutto agli studenti delle lauree magistrali in Matematica, Fisica, Astronomia, Astrofisica e Ingegneria Aerospaziale. Può essere comunque utilizzato anche in corsi del 3° anno delle lauree triennali. *Esso è stato ideato soprattutto per un corso che spieghi nella maniera più semplice possibile, a studenti che si accostano per la prima volta alla materia, sia la Teoria della Relatività Ristretta sia la Teoria della Relatività Generale.* I contenuti di queste due teorie fanno ormai parte del bagaglio culturale necessario ad ogni fisico e ad ogni matematico ed in particolare esse costituiscono un insieme di conoscenze indispensabili per chiunque si occupi di fisica nucleare, di fisica delle particelle elementari, di fisica delle alte energie, di fisica dei plasmi, di astrofisica e di cosmologia. Tuttavia il testo può essere adattato anche per corsi monografici che trattino solo la Relatività Ristretta o solo la Relatività Generale.
Infine, il libro ambisce ad interessare un pubblico più vasto avente una buona preparazione matematica e fisica di base, quali possono essere i docenti della scuola secondaria.
Obbiettivi principali del corso sono:

1) essere il più possibile fedeli al pensiero originale di Einstein espresso nei suoi due lavori fondamentali citati in precedenza;

2) fornire i fondamenti della Relatività Ristretta e le loro implicazioni nella Dinamica e nell'Elettrodinamica quale base per svariate applicazioni alla Fisica Matematica ed alla Fisica Teorica;

3) badare soprattutto alla compiutezza delle idee ed alla loro precisa e rigorosa formulazione matematica nello spaziotempo 4-dimensionale;

4) mettere in luce il contenuto fisico della teoria, facendo ogni sforzo per trattare nel modo più completo i dubbi insidiosi ed i problemi paradossali che assillano tutti coloro che intraprendono per la prima volta lo studio della Relatività;

tazione di questo ottimo libro, per approfondire molte delle tematiche da noi trattate ad un livello necessariamente più elementare.

5) fornire i fondamenti matematici e concettuali della teoria della Relatività Generale fino alla formulazione delle equazioni gravitazionali di Einstein che ne costituiscono il nucleo centrale;

6) ritrovare, in prima approssimazione, la teoria newtoniana della gravitazione;

7) ricavare la soluzione esatta di Schwarzschild per una massa isolata a simmetria sferica (tipo il Sole o una qualunque stella) ed applicarla come introduzione allo studio dei buchi neri;

8) spiegare in maniera dettagliata ed esauriente le tre prove classiche:
   - avanzamento del perielio del pianeta Mercurio;
   - deflessione dei raggi luminosi in un campo gravitazionale;
   - spostamento delle righe spettrali (red-shift) in un campo gravitazionale;

9) fornire una semplice e chiara introduzione alla studio della moderna Cosmologia;

10) illustrare i più importanti modelli di Universo il cui studio fu iniziato, per la prima volta, dal matematico e cosmologo russo Alexander Alexandrovic Friedmann.

Questo libro è maturato attraverso l'insegnamento di vari corsi di Relatività, sia Speciale sia Generale che uno di noi (F. B.) ha tenuto negli ultimi 20 anni e l'altro (F. D.) ne è diventato recentemente titolare, per studenti di Matematica e Fisica, sia della laurea triennale sia della laurea specialistica o della magistrale, presso il Dipartimento di Matematica e Informatica della Facoltà di Scienze dell'Università di Cagliari.
Esso è frutto di una attenta e scrupolosa elaborazione, suggerita dall'esperienza didattica, di Appunti forniti, durante parecchi anni, agli studenti frequentanti i suddetti corsi e perciò sottoposti al loro giudizio.
Da questa sua origine didattica il libro ha desunto uno specifico carattere elementare, per cui esso non può (e non vuole) essere considerato un vero e proprio trattato.
Con molta modestia ed umiltà abbiamo fatte nostre le parole di Einstein (1879-1955)[5]:
"l'autore ha compiuto ogni sforzo nel tentativo di esporre le idee basilari nella forma più chiara e più semplice possibile,....Per raggiungere la massima chiarezza mi è parso inevitabile ripetermi di frequente, senza avere la minima cura per l'eleganza dell'esposizione; ho scrupolosamente seguito il precetto del geniale fisico teorico Ludwig Boltzmann, secondo cui i problemi dell'eleganza vanno lasciati al sarto e al calzolaio".
A più di cent'anni dalla nascita della Teoria della Relatività esistono innumerevoli ottime trattazioni alcune fatte da matematici e fisici eccelsi; ovviamente abbiamo un debito enorme verso tutti gli autori dai quali abbiamo attinto il materiale che è stato fonte di ispirazione per questa nostra modesta esposizione. Quelli che abbiamo avuto modo di consultare più frequentemente sono tutti citati in Bibliografia e brevemente commentati, mentre alcuni riferimenti specifici sono indicati, di volta in volta a piè di pagina, nei vari paragrafi. Consigliamo vivamente allo studente di prendere familiarità con almeno uno di essi di livello universitario.
Quando abbiamo cominciato a redigere il libro *" ci siamo prefissi di raccogliere, nei vari capitoli ed anche nei singoli paragrafi, quelle che a nostro giudizio sono le migliori impostazioni della materia di cui siamo a conoscenza"*.[6] Inoltre, il nostro lavoro è stato fatto *"con lo spirito di considerare ogni spiegazione che aiuti la comprensione della materia - anche se presa a prestito da altri autori - come parte delle conoscenze umane consolidate e quindi di pubblico dominio"*.[7]

[5] Albert Einstein, Prefazione alla *Relatività: esposizione divulgativa*, pag. 44, Boringhieri, Torino, 1967.

[6] A tal proposito, ci siamo attenuti alle considerazioni, che condividiamo pienamente, esposte nell'ottimo libro di Ray d'Inverno, *Introduzione alla Relatività di Einstein*, traduzione italiana a cura di Silvio Bergia, Cooperativa Libraria Universitaria Editrice Bologna (CLUEB), 2001, paragrafo 1.2 Ringraziamenti, pagg. 5-6.

[7] Ibidem, pag. 6.

In effetti molti degli approcci usati nel corso dell'esposizione sono stati presi in prestito da vari autori, ma in ogni caso da noi profondamente ripensati ed elaborati.
La nostra intenzione e preoccupazione costante è stata quella di organizzare il materiale in modo che fosse accessibile, nel modo più semplice, alla maggior parte degli studenti, seguendo in ciò un famoso detto del grande matematico italiano Tullio Levi-Civita: *Lo sforzo deve essere di chi scrive e non di chi legge!*
I 14 capitoli del libro potrebbero essere raggruppati (anche se ciò non è stato fatto materialmente) in cinque parti:
La Parte I^a comprende soltanto 2 Capitoli denominati rispettivamente Cap.0 e Cap.1.
Il Cap.0 consiste di 2 paragrafi, il primo dei quali contiene una Premessa Metodologica in cui si discute brevemente in che cosa consiste una *teoria matematica* e una *teoria fisica* e ciò che le differenzia l'una dall'altra: vi si tratteggia inoltre l'uso degli assiomi, la loro coerenza logica interna e l'ufficio *dell'evidenza sperimentale.* Il secondo paragrafo introduce il concetto di *realtà fisica o fenomenologica* secondo Bridgman, si accenna alla struttura di una teoria fisica e si introducono tre principi metodologici fondamentali: *Principio di oggettività degli osservatori, Principio di verifica sperimentale e Principio di antecedenza (o di causalità).*
Il Cap.1, intitolato *Assiomi della Fisica Classica*, contiene un compendio molto dettagliato di ciò che uno studente dovrebbe sapere della Fisica Classica, prima di intraprendere lo studio della Relatività. Vi si discute il concetto di spazio, tempo e movimento, riportando diversi brani tratti dalle opere di grandissimi scienziati: Newton, Bridgman, Born, Einstein, Landau e Lifschitz, Rindler. In particolare, vi si tratta dei concetti fondamentali della Meccanica di Galileo-Newton, dell'Elettrodinamica di Maxwell-Lorentz e della Teoria della Gravitazione di Newton.
Sui concetti di spazio e tempo ci siamo soffermati molto di più di quello che fanno la stragrande maggioranza degli autori nelle loro trattazioni. Ciò è stato fatto a ragion veduta ritenendo che sia estremamente utile ed istruttivo per gli studenti che, generalmente, non hanno né il tempo né la competenza per scegliere i libri più adeguati da consultare e metterne a confronto il contenuto.
Il lettore desideroso di arrivare rapidamente alla Relatività può, in un primo tempo, tralasciare la lettura di questi due capitoli introduttivi, senza pregiudizio per la comprensione di ciò che segue.
La Parte II^a include i capitoli dal Cap.2 al Cap.7 e tratta compiutamente della teoria della Relatività Ristretta, in particolare:
Il Cap.2 presenta, in maniera molto sobria, i motivi che hanno portato alla crisi della fisica classica.
Nel Cap.3 si introducono i due assiomi di Einstein che sono alla base della Teoria della Relatività Ristretta e da essi si deducono le trasformazioni di Lorentz.
Il Cap.4 presenta le conseguenze fondamentali delle trasformazioni di Lorentz e svolge tutta la Cinematica Relativistica incluse le conferme sperimentali.
Il Cap.5 è dedicato completamente allo studio dello spaziotempo di Minkowski introducendo anche il concetto di tempo proprio, di 4-velocità e di 4-accelerazione.
Il Cap.6 svolge in maniera semplice ma approfondita la dinamica relativistica di una particella arrivando a discutere l'equivalenza tra massa ed energia, illustrando compiutamente la famosa formula $E = mc^2$ emblematica di tutta la Relatività Speciale. Si considerano poi alcune applicazioni ai fenomeni d'urto fra particelle.
Il Cap.7 svolge, in modo sobrio, i fondamenti dell'elettrodinamica relativistica.
La Parte III^a, comprendente i Cap.8, 9, 10, presenta, nel modo più semplice, i concetti di base del *Calcolo Tensoriale su di una varietà differenziabile e della Geometria Riemanniana* che costituiscono il fondamento matematico della Relatività Generale. Nel trattare il calcolo tensoriale e la geometria riemanniana, pur usando il linguaggio ed il simbolismo della moderna geometria differenziale, ci siamo sforzati di essere il più vicino possibile al pensiero di Levi-Civita (uno dei

fondatori del Calcolo differenziale assoluto) specialmente nel definire i concetti fondamentali di **connessione** (*connessione di Levi- Civita*) e di **derivata covariante**. Ciò viene fatto utilizzando **la nozione di parallelismo**, prima su di una superficie dello spazio euclideo ordinario, ed estendendolo poi ad una varietà riemanniana. Sulla nozione di parallelismo, dovuta proprio a Levi-Civita (1917), ci soffermiamo più di quanto fanno molti autori, essendo un concetto veramente basilare e di una bellezza sorprendente. Il nostro scopo è quello di far sì che il pensiero di Levi- Civita non venga dimenticato con trattazioni astratte moderne e alla moda.
La Parte IV^a comprende i Cap.11, Cap.12 e Cap.13.
Il Cap.11 presenta i fondamenti della teoria einsteiniana della gravitazione, cioè la Relatività Generale, gli assiomi sui quali è fondata e la necessità della geometria riemanniana per la sua formulazione secondo Einstein. Vi si introduce il concetto di spaziotempo curvo, se ne discute la sua struttura, si introduce il concetto generale di riferimento fisico, si discutono i concetti di distanza ed intervallo temporale desunti dalla metrica e si discute il concetto, molto delicato, di simultaneità in Relatività Generale. Inoltre è studiato nei dettagli il moto di una particella liberamente gravitante e si determina il limite newtoniano per le equazioni delle geodetiche. Il capitolo culmina nella discussione delle equazioni gravitazionali o equazioni di campo di Einstein considerate da diversi punti di vista, incluso l'approccio variazionale di Hilbert. Si discutono le espressioni più importanti del tensore energia-impulso, si trova il limite newtoniano per le equazioni gravitazionali ed infine si fanno alcune utili considerazioni sulla costante cosmologica.
I Cap.12 e Cap.13 trattano, in modo abbastanza esauriente ed approfondito alcune fra le più notevoli applicazioni della Relatività Generale, e precisamente:
il Cap.12 tratta esaurientemente il problema esterno di Schwarzschild e la metrica a cui si perviene come soluzione esatta delle equazioni gravitazionali di Einstein. Vi si introduce il concetto di **"buco nero"** del quale si fa uno studio abbastanza approfondito non solo utilizzando la metrica di Schwarzschild ma introducendo anche le metriche di Eddington- Finkelstein e di Kruskal-Szekeres. Il capitolo termina con una sostanziosa APPENDICE dedicata ad una concisa trattazione di un argomento importantissimo ed attuale quello delle **onde gravitazionali** che furono predette teoricamente cento anni fa, nel 1916, da Einstein ma che sono state rivelate sperimentalmente soltanto pochi anni fa, nel 2015.
La metrica di Schwarzschild viene poi applicata, nel Cap.13, allo studio approfondito dei tre test classici: l'avanzamento del perielio del pianeta Mercurio, la deflessione dei raggi luminosi ed il red-shift delle linee spettrali in un campo gravitazionale.
La V^a ed ultima parte, costituita dal Cap.14, fornisce un'altra importante applicazione della Relatività Generale allo studio della Cosmologia e dei modelli di Universo.
Infine, ci corre l'obbligo di fare un'ultima considerazione che riguarda la limitatezza del numero degli esercizi risolti o proposti. Le motivazioni sono sostanzialmente tre:

a) la stragrande maggioranza dei corsi che abbiamo tenuto in passato e che continuiamo a tenere, erano e sono, ancora oggi, destinati a studenti dell'ultimo anno della laurea magistrale ed hanno sempre avuto tradizionalmente carattere prevalentemente teorico con enfasi soprattutto concettuale;

b) il numero esiguo delle ore di lezione non consentono di dedicare molto tempo agli esercizi a scapito della compiutezza della trattazione teorica dei concetti che nel nostro libro è l'obbiettivo e la caratteristica principale;

c) infine, riteniamo, che per gli esercizi, allo studente convenga ricorrere a libri specializzati alcuni dei quali sono citati nella Bibliografia alla fine del volume.

Si raccomanda comunque agli studenti di svolgere a casa un congruo numero di esercizi che vengono forniti, di volta in volta a lezione, da noi docenti.

Nonostante l'attenzione con la quale abbiamo redatto il libro, è possibile che ci siano ancora degli errori, speriamo non gravi e non concettuali, e di ciò ce ne scusiamo col lettore. Naturalmente ogni critica od osservazione da parte dei lettori che serva a migliorare questo nostro lavoro sarà, da parte nostra, ben gradita.
Infine, osserviamo che il testo è corredato da molte figure, generalmente a carattere naif, che hanno la sola funzione di rendere più chiaro il testo. Tali figure sono state ottenute dagli autori utilizzando i software Xfig e la versione 9 di Mathematica.

Per lo studente

La Teoria della Relatività, specialmente quella Generale, ebbe fin dalla sua nascita la reputazione di essere estremamente difficile. In parte, questa reputazione è dovuta ad un aneddoto attribuito al famoso astronomo inglese Sir Arthur Eddington, all'epoca direttore dell'osservatorio astronomico di Greenwich, ed uno dei primi studiosi autorevoli dell'opera di Einstein. Secondo questo aneddoto lo scienziato fu interrogato da un intervistatore che gli chiese *se ritenesse vero che solo tre persone al mondo capissero la Relatività. Dopo una breve riflessione, Eddington rispose: Chi è la terza?*
Nel 1987, quando la Teoria della Relatività aveva ormai più di ottant'anni, un collega e biografo di Einstein, Abraham Pais, aggiornò questo numero durante un suo intervento in occasione delle celebrazioni del 900-esimo anniversario della nascita dell'Università di Bologna. Disse Pais: *"secondo una stima ottimistica, oggi soltanto una persona su centomila capisce l'essenza della Teoria della Relatività di Einstein."*
A dire il vero, la Teoria della Relatività Ristretta non è difficile dal punto di vista matematico in quanto si può benissimo comprendere usando soltanto le quattro operazioni fondamentali dell'aritmetica e l'estrazione di radice quadrata. Per capire poi il concetto di *spaziotempo 4-dimensionale di Minkowski* basta un semplice calcolo tensoriale basato sulle trasformazioni lineari e le nozioni di base sugli spazi vettoriali *pseudoeuclidei.* La difficoltà è invece di natura *psicologica* in quanto gli effetti relativistici sono veramente sorprendenti e controintuitivi, soprattutto se si tiene conto che la nostra intuizione è basata sull'esperienza quotidiana maturata, fin dalla nascita, sull'osservazione di oggetti che si muovono a velocità estremamente basse rispetto alla velocità della luce nel vuoto.
Nel passare dalla Relatività Ristretta alla Relatività Generale il salto concettuale è senza dubbio uno dei più grandi nella storia del pensiero umano. Tuttavia la Relatività Generale ha compiuto ormai cent'anni ed è perciò stata studiata ed analizzata da diversi punti di vista e la sua comprensione è facilitata utilizzando una serie di passaggi logici semplici. Ciò rende la teoria alla portata di tutti gli studenti volenterosi, pazienti, curiosi e con una adeguata preparazione matematica.
A proposito della preparazione matematica è interessante sentire quel che dice a tal proposito il grande fisico teorico Paul Adrien Maurice Dirac:
" La Matematica è lo strumento particolarmente adatto per trattare concetti astratti di qualunque tipo, e non vi è alcun limite alle sue possibilità in questo campo. Per questa ragione un libro sulla nuova Fisica, a meno che non descriva semplicemente il lavoro sperimentale, deve essere essenzialmente matematico."[8]
Einstein, dal canto suo, aggiunge le seguenti considerazioni:
"Da studente non capivo molto bene che la possibilità di conoscere più profondamente i principi fondamentali della Fisica è legata ai metodi matematici più complicati. Me ne resi conto solo a poco a poco, dopo anni di lavoro scientifico indipendente."[9]
Perciò, caro studente, non scoraggiarti di fronte alle difficoltà matematiche che incontrerai nello studio della Relatività soprattutto di quella Generale: il libro che ti sta di fronte ha la caratteri-

[8] Paul A. M. Dirac, *I principi della Meccanica Quantistica*, estratto dalla Prefazione.
[9] Albert Einstein, *Autobiografia scientifica*, Boringhieri-Torino, 1979, pag.16.

stica di fornire le spiegazioni suddivise in piccoli passi, in modo tale che i calcoli, anche quelli più lunghi e difficili, sono tutti alla tua portata. Soprattutto ci siamo sforzati di evitare frasi del tipo "....si può vedere che...." o peggio ancora "....è facile vedere che....".

Prerequisiti: Le nozioni matematiche di base che occorre conoscere sono del tutto modeste: concetti fondamentali dell'Analisi Matematica per funzioni di una e più variabili, risoluzione di semplici equazioni differenziali, spazi vettoriali, spazi affini e spazi euclidei, più la matematica svolta nella parte III[a] del presente libro.
Per quanto riguarda la Fisica, ovviamente si devono conoscere i fondamenti della Meccanica Classica compreso il Principio di Relatività di Galileo ed i fondamenti della teoria della gravitazione di Newton. Inoltre è opportuno avere un po' di familiarità con l'Elettromagnetismo, le equazioni di Maxwell e la forza di Lorentz. Il Cap.1 è appositamente dedicato alla rivisitazione di tutto ciò.
A conclusione di queste considerazioni, caro studente, ti suggeriamo la lettura del seguente brano, scritto proprio per principianti:
"Una delle maggiori difficoltà che si presentano nell'insegnamento di una qualsiasi disciplina scientifica è costituita dal fatto che, proprio al principio del corso, si addensano le definizioni dei concetti fondamentali, sui quali la teoria poggia, e che richiedono da parte dell'allievo una capacità di astrazione alla quale egli non è ancora abituato. Né d'altronde, non volendo rinunciare ad una trattazione rigorosa della materia, tale difficoltà può essere facilmente eliminata. Per quanto piana ed elementare sia condotta l'esposizione, è impossibile che colui il quale per la prima volta si accosta ad un nuovo concetto, possa valutarne subito tutta la portata, ciò avverrà in un secondo tempo quando altri concetti verranno introdotti e si intrecceranno fra di loro dando corpo alla teoria la quale sarà sottoposta al vaglio delle diverse applicazioni per risolvere disparati problemi. Il giovane studioso, pertanto, man mano che si avvicina alle idee nuove del corso del quale sta per intraprendere lo studio, non sia scoraggiato se ai primi contatti con esse, pur avendo compreso sul piano logico il significato di una data definizione o di un dato concetto, non si rende subito conto della sua necessità e gli vien fatto di chiedersi - come accade agli studenti migliori - ma questo perché? Nel seguito, se avrà lavorato con diligenza, oltre che con intelligenza, i suoi dubbi troveranno un naturale chiarimento".[10]

Ringraziamenti

Francesco Borghero, anche se a distanza di tanti anni, desidera rivolgere un ricordo particolare ed affettuoso ed esprimere la propria riconoscenza alla memoria del Prof. Carlo Cattaneo (1911-1979), Maestro incomparabile, dal quale ha appreso, per la prima volta, il fascino e la bellezza della Teoria della Relatività.[11]
Inoltre, desidera ringraziare sentitamente il coautore Francesco Demontis per aver seguito con entusiasmo i suoi corsi di Relatività e soprattutto per averlo stimolato con insistenza e passione a trasformare, con il suo apporto decisivo, i suoi appunti, raccolti in tanti anni di insegnamento, in un libro organico per studenti. Qualunque possa essere il merito didattico del libro esso va comunque condiviso in parti uguali fra i due autori.
Francesco Demontis desidera ringraziare Francesco Borghero per avergli trasmesso la passione per la Relatività e l'arte della pazienza nella ricerca dei dettagli che permettono alle spiegazioni di diventare illuminanti.

[10]Questo brano, adattato al nostro scopo, è liberamente tratto dal libro di Mauro Picone e Gaetano Fichera, *Lezioni di Analisi Matematica*, vol.1, Dispense Universitarie Tumminelli - Città Universitaria, Roma (1962), pag. 3, sul quale uno di noi (F.B.) ha imparato, a suo tempo, i fondamenti dell'Analisi Matematica, come studente del 1° anno del corso di laurea in Matematica presso l'Università di Roma "La Sapienza".

[11]Ci si riferisce al bellissimo corso di *Teorie Relativistiche* tenuto dal Prof. Carlo Cattaneo nell'anno accademico 1969-70 e che F. B. frequentò come studente all'Istituto Matematico "Guido Castelnuovo" dell'Università di Roma La Sapienza.

Gli autori sono in debito anche con molti colleghi (non solo dell'Università di Cagliari) con cui hanno discusso varie parti del libro. Non volendo fare un lunghissimo elenco, si limitano a dire grazie ai colleghi con i quali l'interazione è stata più intensa. Quindi tante grazie: al Prof. Andrea Loi (UniCa) per aver revisionato alcune parti di geometria differenziale, al Prof. Sebastiano Pennisi (UniCa) per averci aiutato a impostare meglio il capitolo della meccanica classica e dato utili suggerimenti riguardo al "paradosso dell'asta nel ripostiglio", al Prof. Giovanni Ortenzi (Università di Milano Bicocca), per aver suggerito importanti modifiche nella trattazione dell'effetto Doppler, al Prof. Giuseppe Rodriguez (UniCa) e alla Prof.ssa Caterina Fenu (UniCa) per averci suggerito alcuni utili accorgimenti nell'uso del programma *LaTeX* con il quale è stato scritto il libro, al Prof. Stefano Montaldo, alla Prof.ssa Maria Polo e al Prof. Cornelis Van der Mee (UniCa) per l'incoraggiamento in questi anni a portare avanti questo lavoro. Un particolare ringraziamento lo dobbiamo a Pierfrancesco Borghero per la cura impiegata nel produrre l'illustrazione di copertina. Un sentito ringraziamento va inoltre agli anonimi Revisori per il loro lavoro che si è concretizzato in molti utili suggerimenti che ci hanno consentito di migliorare il testo e di renderlo più completo.
Infine gli autori desiderano rivolgere un affettuoso ringraziamento agli studenti dei vari corsi da loro tenuti presso il Dipartimento di Matematica e Informatica dell'Università di Cagliari per il loro interesse, per la loro attenzione, per i loro utili commenti e per le loro stimolanti osservazioni: sono loro che talvolta ci hanno fatto capire "ciò che andava" e "ciò che non andava" nelle nostre lezioni.

Cagliari, 11 - 02 - 2021

Francesco Borghero e Francesco Demontis
Dipartimento di Matematica e Informatica
Università degli Studi di Cagliari

Indice

0 PREMESSA METODOLOGICA **3**
0.1 Teoria matematica e teoria fisica . . . 3
0.2 Alcune considerazioni introduttive sui fondamenti della fisica . . . 5

1 ASSIOMI DELLA FISICA CLASSICA **9**
1.1 Spazio e tempo nella Fisica Classica . . . 9
1.1.1 Spazio, tempo e movimento secondo Newton . . . 10
1.2 I° Gruppo - Assiomi generali . . . 11
1.2.1 Lo spazio è relativo . . . 12
1.2.2 Il tempo è assoluto . . . 13
1.3 II° Gruppo - Assiomi della Meccanica Newtoniana . . . 14
1.3.1 II° Gruppo (A) - Assiomi della Meccanica Newtoniana in assenza di Gravitazione . . . 15
1.3.2 Altre proprietà dello spazio e del tempo: omogeneità ed isotropia . . . 17
1.3.3 Ulteriori considerazioni su omogeneità ed isotropia e loro legame con i riferimenti inerziali ed il Principio di Relatività . . . 19
1.3.4 Relazioni tra coordinate di un medesimo evento in riferimenti galileiani diversi . . . 20
1.3.5 Trasformazioni di Galileo . . . 24
1.3.6 Conseguenze delle trasformazioni di Galileo . . . 28
1.3.7 Principio di Relatività e invarianza della 2^a legge della dinamica per trasformazioni di Galileo . . . 29
1.3.8 Il Principio di Relatività come Principio di Simmetria . . . 30
1.3.9 II° Gruppo (B) - Assiomi della Meccanica Newtoniana in presenza di Gravitazione . . . 34
1.3.10 La massa nella teoria newtoniana . . . 35
1.3.11 Digressione sulla Teoria del Potenziale Newtoniano . . . 38
1.4 III° Gruppo - Assiomi dell'elettromagnetismo . . . 42
1.4.1 Carica elettrica, campo elettrico, campo magnetico, densità di carica e densità di corrente . . . 42
1.4.2 Equazioni di Maxwell nel sistema di Gauss . . . 43
1.4.3 Forza di Lorentz, conservazione della carica, energia elettromagnetica e vettore di Poynting . . . 44
1.4.4 Equazioni di Maxwell in mezzi omogenei non conduttori e legge di propagazione delle onde elettromagnetiche. Equazioni di Maxwell nel vuoto . . . 45
1.5 APPENDICE: Letture su spazio e tempo nella Fisica Classica . . . 47
1.5.1 Spazio e tempo classici commentati da Einstein . . . 47
1.5.2 Spazio e tempo secondo l'operativismo di Bridgman . . . 48

2 CRISI DELLA FISICA CLASSICA **49**
2.1 La ricerca dell'etere cosmico . . . 49
2.2 Non-invarianza delle equazioni di Maxwell per trasformazioni di Galileo . . . 51
2.3 Non-invarianza della legge di propagazione delle onde elettromagnetiche . . . 52
2.4 L'aberrazione annua delle stelle: spiegazione classica . . . 54
2.5 Effetto Doppler ottico: spiegazione classica . . . 61
2.6 Un esercizio istruttivo di Cinematica Classica: regata su un fiume . . . 64
2.7 L'esperimento di Michelson-Morley . . . 68
2.8 Conseguenze dell'esperimento di Michelson-Morley . . . 71
2.9 APPENDICE 1: Alcune nozioni utili di Astronomia Sferica . . . 73
2.10 APPENDICE 2: Richiami sulla propagazione delle onde . . . 76

3 FONDAMENTI DELLA RELATIVITÀ RISTRETTA **81**
3.1 Concezione unitaria della Fisica: I nuovi Assiomi . . . 81
3.2 I Postulati della Teoria della Relatività Ristretta . . . 82
3.2.1 Il tempo è relativo . . . 84
3.2.2 Il concetto einsteiniano di tempo ed il concetto di simultaneità . . . 85
3.2.3 Ulteriori riflessioni sulla sincronizzazione degli orologi . . . 87
3.2.4 Ancora alcune riflessioni sul concetto einsteiniano di tempo e di simultaneità 90
3.2.5 Una terza definizione di simultaneità più aderente al senso comune . . . 91
3.2.6 La relatività della simultaneità . . . 93
3.2.7 Formulazione matematica dei postulati della relatività einsteiniana . . . 96
3.3 Trasformazioni speciali di Lorentz . . . 98
3.3.1 Trasformazioni speciali di Lorentz come rotazioni nello spazio di Minkowski 103
3.4 Trasformazioni generali di Lorentz . . . 104
3.5 APPENDICE: Un teorema sulle forme quadratiche reali in 4 variabili . . . 107

4 CINEMATICA RELATIVISTICA **109**
4.1 Conseguenze delle trasformazioni di Lorentz . . . 109
4.1.1 Osservazioni fondamentali . . . 109
4.1.2 Simmetria nelle variabili y e z e nelle variabili x e $T = ct$ . . . 110
4.1.3 Limite Newtoniano . . . 110
4.1.4 Proprietà del fattore di Lorentz . . . 110
4.1.5 Esistenza di una sola velocità invariante . . . 112
4.1.6 Relatività del concetto di simultaneità . . . 112
4.1.7 Relatività delle durate temporali (dilatazione dei tempi) . . . 113
4.1.8 Relatività dell'ordine temporale . . . 115
4.1.9 Relatività delle lunghezze di regoli rigidi (contrazione di Lorentz-Fitzgerald) 117
4.1.10 Legge relativistica di composizione delle velocità . . . 120
4.1.11 Legge di trasformazione del fattore di Lorentz . . . 123
4.1.12 Insuperabilità della velocità della luce nel vuoto . . . 123
4.1.13 Invarianza della legge di propagazione delle onde elettromagnetiche per trasformazioni di Lorentz . . . 125
4.2 Conferme sperimentali della Cinematica Relativistica . . . 127
4.2.1 L'aberrazione della luce stellare: spiegazione relativistica . . . 127
4.2.2 Effetto Doppler relativistico longitudinale e trasversale . . . 129
4.2.3 L'esperimento di Michelson-Morley: spiegazione relativistica . . . 133
4.2.4 L'allungamento della vita media del mesone μ causata dalla dilatazione del tempo . . . 135

5 SPAZIOTEMPO DI MINKOWSKI **137**
5.1 Invariante fondamentale della Relatività Ristretta . . . 137
5.1.1 Cenno alle trasformazioni generali di Lorentz in notazione tensoriale . . . 140
5.1.2 Digressione sugli spazi affini, euclidei e pseudo-euclidei . . . 142
5.2 Lo spaziotempo di Minkowski come modello di Universo . . . 143
5.3 Classificazione degli intervalli spaziotemporali . . . 144
5.3.1 Divario spazio-temporale del genere tempo . . . 145
5.3.2 Divario spazio-temporale del genere spazio . . . 145
5.3.3 Divario spazio-temporale del genere luce . . . 146
5.4 Quadrivettori dello spazio di Minkowski . . . 146
5.4.1 Proprietà geometriche di $\mathcal{M}^4$ . . . 146
5.4.2 Prodotto scalare, norma e modulo dei vettori di $\mathcal{M}^4$ . . . 147
5.5 Proprietà dei quadrivettori . . . 148
5.5.1 Classificazione dei quadrivettori e definizione del cono di luce . . . 148
5.5.2 Diagrammi spazio-temporali e linee di universo in $\mathcal{M}^2$. . . . 151
5.5.3 Interpretazione geometrica della trasformazione di Lorentz in $\mathcal{M}^2$ . . . 153
5.5.4 Deduzione geometrica della contrazione di Lorentz in $\mathcal{M}^2$ . . . 156
5.6 Grandezze cinematiche assolute . . . 158
5.6.1 Tempo proprio . . . 158
5.6.2 Quadrivelocità . . . 160
5.6.3 Quadriaccelerazione . . . 161
5.7 Cenno al Paradosso dei Gemelli o Paradosso degli orologi . . . 163

6 MECCANICA RELATIVISTICA DI UNA PARTICELLA **167**
6.1 Introduzione . . . 167
6.2 Grandezze dinamiche assolute e relative . . . 169
6.2.1 Massa propria di una particella . . . 169
6.2.2 Quadri-quantità di moto, massa relativistica e tri-quantità di moto . . . 170
6.2.3 La conservazione del 4-impulso come base della Meccanica Relativistica . 173
6.2.4 Energia totale, energia di riposo ed energia cinetica di una particella . . . 177
6.3 Le leggi della dinamica relativistica . . . 180
6.3.1 Quadriforza e leggi assolute del moto per una particella di massa propria costante . . . 180
6.3.2 Formulazione relativa 3-dimensionale della dinamica di una particella . . 182
6.3.3 Teorema dell'energia . . . 184
6.3.4 Energia cinetica relativistica ed energia totale dedotte dal teorema dell'energia . . . 184
6.3.5 Relazione tra energia e momento: relazione di mass shell . . . 188
6.3.6 Particelle di massa propria nulla . . . 189
6.3.7 Relazione tra forza, accelerazione e velocità di una particella nello spazio fisico . . . 191
6.3.8 Trasformazione di Lorentz della 3-quantità di moto, dell'energia e della massa . . . 192
6.3.9 Trasformazione di Lorentz della 3-forza . . . 195
6.4 Equivalenza tra massa ed energia . . . 197
6.4.1 Leggi di conservazione nei fenomeni d'urto . . . 197
6.4.2 Equivalenza tra massa ed energia . . . 199
6.4.3 Moto di una particella carica in un campo magnetico uniforme . . . 201
6.4.4 Cenno alle particelle di massa propria variabile . . . 204
6.5 APPENDICE . . . 206

7 ELETTRODINAMICA RELATIVISTICA NEL VUOTO **211**
7.1 Introduzione . . . 211
7.2 Densità di carica e densità di corrente . . . 212
7.3 Covarianza delle equazioni di Maxwell . . . 214
7.4 Leggi di trasformazione del campo elettrico **E** e magnetico **H** . . . 220
7.5 Digressione sui 4-tensori doppi dello spazio di Minkowski . . . 224
7.6 Tensore del campo elettromagnetico . . . 226
7.7 Forma tensoriale delle equazioni di Maxwell . . . 229
7.8 Formulazione potenziale delle equazioni di Maxwell . . . 233

8 ALGEBRA TENSORIALE **237**
8.1 Introduzione . . . 237
8.2 Varietà differenziabili e sistemi di coordinate . . . 238
8.3 Esempi di varietà differenziabili . . . 243
8.4 Curve e superfici di una varietà differenziabile . . . 245
8.5 Trasformazioni di coordinate su di una varietà differenziabile . . . 246
8.6 Digressione sulla convenzione di sommazione di Einstein . . . 248
8.7 Spazio vettoriale tangente $T_P(M)$ ad una varietà differenziabile . . . 250
8.8 Dimensione dello spazio tangente. Base naturale. Cenno allo spazio cotangente $T_P^*(M)$ . . . 251
8.9 Tensori controvarianti di rango 1 . . . 252
8.10 Tensori covarianti di rango 1 . . . 254
8.11 Invarianti . . . 256
8.12 Tensori del secondo ordine e di ordine superiore . . . 257
8.13 Tensori simmetrici e tensori antisimmetrici . . . 260
8.14 Operazioni algebriche sui tensori . . . 262
8.15 Criteri di tensorialità e seconda definizione di tensore . . . 263
8.16 Varietà riemanniane . . . 269
8.17 Varietà euclidee e varietà globalmente euclidee . . . 270
8.18 Il tensore coniugato del tensore metrico . . . 271
8.19 Abbassamento ed innalzamento degli indici mediante il tensore metrico . . . 272

9 PARALLELISMO DI LEVI-CIVITA **273**
9.1 Parallelismo di Levi-Civita . . . 273
9.1.1 Osservazione fondamentale di Gauss circa la geometria intrinseca di una superficie . . . 274
9.1.2 Superfici sviluppabili . . . 275
9.1.3 Parallelismo superficiale: Definizione geometrica . . . 277
9.1.4 Prime conseguenze. Equipollenza superficiale di vettori . . . 279
9.1.5 Trasporto di un vettore lungo una linea chiusa. Angolo di parallelismo . . . 281
9.1.6 Trasporto infinitesimo. Forma differenziale della legge di parallelismo . . . 282
9.1.7 Equazione simbolica ed equazioni intrinseche di parallelismo . . . 285
9.1.8 Proprietà dei simboli di Christoffel. Identità di Ricci. 1^a e 2^a formula di Christoffel . . . 287
9.1.9 Equazioni di parallelismo nelle componenti covarianti . . . 288
9.1.10 Alcune conferme analitiche . . . 290
9.1.11 Trasporto parallelo su di una varietà riemanniana . . . 291
9.1.12 Ancora su alcune proprietà dei simboli di Christoffel e loro formule di trasformazione . . . 293

10 GEOMETRIA RIEMANNIANA E DERIVAZIONE COVARIANTE **297**
10.1 Derivazione ordinaria di un campo tensoriale . . . 297
10.2 Derivazione covariante di un campo vettoriale . . . 300
10.3 Derivazione covariante di un tensore misto del 2° ordine . . . 302
10.4 Derivazione covariante di un tensore qualunque . . . 304
10.5 Proprietà formali della derivazione covariante e teorema di Ricci . . . 304
10.6 Derivazione controvariante . . . 306
10.7 Differenziale assoluto e derivata assoluta . . . 306
10.7.1 Cenno sulle varietà a connessione affine . . . 308
10.7.2 Connessione di Levi-Civita o riemanniana . . . 309
10.7.3 L'operatore divergenza . . . 311
10.8 Cenni di Calcolo delle variazioni . . . 312
10.8.1 Il concetto di funzionale . . . 312
10.8.2 Lemma fondamentale del Calcolo delle variazioni . . . 313
10.8.3 Variazione prima di un integrale e metodo delle variazioni . . . 314
10.8.4 Equazione di Eulero associata ad un principio variazionale . . . 316
10.8.5 Equazioni di Eulero nel caso di un funzionale dipendente da più argomenti 317
10.8.6 Digressione sul Principio di Hamilton . . . 318
10.9 Le geodetiche di una varietà riemanniana . . . 321
10.9.1 Ulteriori proprietà delle geodetiche . . . 325
10.9.2 Curve nulle e geodetiche nulle . . . 326
10.10 Coordinate geodetiche . . . 329
10.11 La curvatura . . . 332
10.11.1 Non commutabilità di due derivate covarianti successive: Tensore di Riemann . . . 333
10.11.2 Forma totalmente covariante del tensore di Riemann . . . 335
10.11.3 Proprietà di simmetria ed antisimmetria del tensore di Riemann . . . 337
10.11.4 Identità di Bianchi . . . 338
10.11.5 Tensore di Ricci e scalare di curvatura . . . 339
10.11.6 Tensore di Einstein e sua divergenza . . . 341
10.11.7 Relazione tra tensore di Riemann e parallelismo . . . 343
10.12 APPENDICE 1: Un approccio alternativo alla derivazione covariante . . . 344
10.12.1 Derivazione covariante di un campo tensoriale in $\mathbb{R}^n$ . . . 344
10.12.2 Carattere non tensoriale dei simboli di Christoffel di seconda specie . . . 346
10.12.3 Derivazione covariante di un campo tensoriale su di una varietà . . . 348
10.13 APPENDICE 2: Derivata di un determinante . . . 349

11 FONDAMENTI DELLA TEORIA EINSTEINIANA DELLA GRAVITAZIONE **353**
11.1 Introduzione . . . 354
11.2 Programma della teoria einsteiniana . . . 355
11.2.1 Le critiche di Einstein alla teoria newtoniana e programma della teoria einsteiniana . . . 355
11.2.2 Gli assiomi della Relatività Generale . . . 358
11.3 Necessità della geometria riemanniana . . . 361
11.4 Lo spaziotempo $\mathcal{V}^4$ della Relatività generale . . . 363
11.5 Influenza della struttura di $\mathcal{V}^4$ sui fenomeni fisici . . . 365
11.6 Influenza della materia e del suo moto sulla struttura di $\mathcal{V}^4$. . . . 368
11.7 Grandezze assolute e grandezze relative. Concetto di riferimento fisico. . . . 369
11.8 Distanze, intervalli di tempo e metrica spaziale . . . 374
11.9 Direzione temporale e piattaforma spaziale associata. . . . 379
11.10 Grandezze relative standard . . . 382

11.11 Concetto di simultaneità in Relatività Generale . . . 386
11.12 Velocità della luce e velocità di una particella in Relatività Generale . . . 390
11.13 Moto di una particella liberamente gravitante . . . 392
11.14 Gravità come fenomeno metrico. Limite newtoniano per le geodetiche . . . 395
11.15 Le equazioni gravitazionali di Einstein . . . 400
11.15.1 Genesi euristica delle equazioni gravitazionali . . . 400
11.15.2 Il teorema di Cartan . . . 402
11.15.3 Una forma alternativa dell'equazione gravitazionale . . . 403
11.15.4 Digressione: Cenno sulle densità scalari e densità tensoriali . . . 404
11.16 Deduzione variazionale delle equazioni gravitazionali . . . 407
11.17 Tensore energia-impulso in RR ed in RG . . . 412
11.18 Limite newtoniano delle equazioni gravitazionali . . . 419
11.19 Osservazioni sulla costante cosmologica Λ . . . 424
11.20 APPENDICE: Osservazioni su alcune convenzioni fondamentali . . . 426

12 SOLUZIONE DI SCHWARZSCHILD E SUE CONSEGUENZE 427
12.1 Metrica indipendente dal tempo, a simmetria radiale . . . 428
12.2 Simboli di Christoffel dedotti dalle equazioni delle geodetiche . . . 431
12.3 Soluzione esatta delle equazioni di campo: metrica di Schwarzschild . . . 434
12.4 Alcune considerazioni sulla metrica di Schwarzschild . . . 437
12.5 L'idea di "buco nero" nella teoria newtoniana . . . 439
12.6 Singolarità della metrica di Schwarzschild: buchi neri . . . 440
12.7 Diagrammi spazio-temporali in coordinate di Schwarzschild . . . 443
12.8 Caduta libera radiale di una particella in un buco nero . . . 446
12.9 Coordinate e metrica di Eddington-Finkelstein . . . 448
12.10 Coordinate di Eddington-Finkelstein avanzate e ritardate . . . 451
12.11 Relazioni fra le coordinate di Schwarzschild e quelle di Kruskal . . . 454
12.12 Diagramma spazio-temporale e metrica di Kruskal . . . 460
12.13 APPENDICE: Cenno sulle onde gravitazionali . . . 465
12.13.1 Introduzione . . . 465
12.13.2 Alcune analogie fra campo elettromagnetico e campo gravitazionale . . . 465
12.13.3 Condizioni sulle coordinate: invarianza di gauge delle equazioni di Einstein 467
12.13.4 Soluzioni di campo debole . . . 469
12.13.5 Linearizzazione delle equazioni di Einstein. Onde gravitazionali . . . 471
12.13.6 Soluzione in termini di potenziali ritardati delle equazioni linearizzate di Einstein. . . . 475
12.13.7 Cenno sulle onde gravitazionali piane . . . 476
12.13.8 Scoperta sperimentale delle onde gravitazionali . . . 477

13 TEST CLASSICI DELLA RELATIVITÀ GENERALE 479
13.1 Introduzione . . . 479
13.2 Avanzamento del perielio dei pianeti . . . 480
13.2.1 Moto ellittico: perielio ed afelio . . . 480
13.2.2 Equazioni differenziali delle geodetiche del $\mathbf{ds}^2$ esterno di Schwarzschild 482
13.2.3 Equazione differenziale dell'orbita ed avanzamento del perielio . . . 484
13.2.4 Calcolo numerico dell'avanzamento del perielio di Mercurio . . . 489
13.3 Deflessione dei raggi luminosi . . . 490
13.3.1 Raggi luminosi in quanto geodetiche di lunghezza nulla . . . 491
13.3.2 Deflessione dei raggi luminosi previsto dalla teoria newtoniana . . . 495
13.4 Redshift delle righe spettrali in un campo gravitazionale . . . 497

14 INTRODUZIONE ALLA COSMOLOGIA **501**
14.1 Introduzione 501
14.2 L'evoluzione delle nostre idee sull'Universo 502
14.3 Alcuni dati numerici su stelle e galassie 505
14.4 Principio Cosmologico 507
14.5 Difficoltà nel modello cosmologico newtoniano: paradosso di Olbers 508
14.6 Postulato di Weyl 510
14.7 Definizione generale di modello cosmologico relativistico 511
14.8 Metrica di Friedmann-Robertson-Walker (FRW) 513
14.9 Alcune proprietà cinematiche della metrica FRW 515
14.10 Condizioni dinamiche. Equazioni di Friedmann 516
14.11 F-modelli statici: Modello di Einstein 520
14.12 Osservazioni sull'equazione di Friedmann. Soluzioni normalizzate 521
14.13 F-modelli vuoti 522
14.14 F-modelli senza azione cosmologica: i tre modelli standard 524
14.15 Alcune considerazioni sulla costante cosmologica 528
14.16 F-modelli generali con azione cosmologica 530
14.17 Densità attuale e critica, costante di Hubble e parametro di decelerazione 538
14.18 Considerazioni conclusive 540
14.19 APPENDICE AL CAP. 14 542
14.20 Epilogo 547

15 Bibliografia Commentata **549**
15.1 Testi divulgativi 549
15.2 Relatività Ristretta o Speciale 549
15.3 Calcolo Tensoriale e Geometria Riemanniana 551
15.4 Relatività Generale 552
15.5 Autobiografia, Opere scelte e Biografie di Einstein 554
15.6 Cosmologia 554
15.7 Libri di Esercizi 555

Ho progettato di scrivere in modo che
lo studioso possa sempre vedere l'essenza
delle cose che impara, anche in modo
che possa apparire la sorgente delle invenzioni,
e quindi in modo tale che lo studioso possa
capire ogni cosa come se l'avesse inventata egli stesso.
(G. W. Von Leibniz, *Matematische Schriften*, edito da Gerhard, vol.VII, pag.9.)

Capitolo 0

PREMESSA METODOLOGICA

Einstein ha introdotto nella Fisica un metodo che deve
costituire la base di tutti i nostri sforzi critici.
Questo metodo si fonda sulla convinzione che l'aspetto più importante
di una teoria è quello che questa teoria effettivamente fa, non quello
che dice di fare né quello che il suo autore ritiene essa faccia,
tutte cose che spesso sono parecchio diverse.
(Percy Williams Bridgman (1882-1961))[1].

In questo primo capitolo introduttivo discuteremo brevemente alcune questioni di natura metodologica[2]: che cosa si intende per teoria matematica e per teoria fisica, le loro differenze fondamentali ed i loro rapporti. Discuteremo anche succintamente il concetto di realtà fisica (o fenomenologica) e tre principi fondamentali di uso corrente in fisica dei quali però, generalmente, non si fa menzione esplicita: *il principio di oggettività degli osservatori, il principio di verifica sperimentale ed il principio di antecedenza (o di causalità)*. Infine accenneremo alla gerarchia della conoscenza fisica ed al rapporto fra le leggi fisiche ed i cosiddetti principi di simmetria.

0.1 Teoria matematica e teoria fisica

È ben noto quel che si intende per **teoria matematica**. Essa è costituita da un certo numero di oggetti $p_1, p_2, p_3, ...$ assunti senza definizione (**enti primitivi**) e legati tra loro da relazioni $A_1, A_2, A_3, ...$ non dimostrate ma assunte a priori (**assiomi o postulati**). Agli assiomi non è richiesto altro che di essere compatibili ed indipendenti, quest'ultima proprietà non essendo veramente essenziale. Da questi costituenti essenziali possono derivarsi, mediante definizioni, altri oggetti $d_1, d_2, d_3, ...$ (**enti derivati**) ed altre relazioni $T_1, T_2, T_3, ...$ tra i vari oggetti, primitivi e derivati, possono essere dedotte per via logica (**teoremi**). In una teoria puramente matematica la scelta degli enti primitivi e degli assiomi è, in linea di principio, completamente libera, salvo il requisito della loro compatibilità logica. In pratica gli enti primitivi e gli assiomi si scelgono sempre con criteri di semplicità o di evidenza sperimentale.
Agli oggetti primitivi, e pertanto anche a quelli derivati, la teoria matematica può non attribuire alcuna interpretazione concreta. **Vera essenza della teoria matematica sono soltanto le relazioni tra quegli oggetti, i quali non hanno quindi che un significato nominale**[3].

[1] P. W. Bridgman, *La natura della teoria fisica*, La Nuova Italia Editrice, Firenze (1965), pagg. 6-7.

[2] Per la stesura di questa prima parte del libro abbiamo tratto ispirazione dalle Lezioni sulla Teoria della Relatività professate dal Prof. Carlo Cattaneo all'Università di Roma, "La Sapienza", nell'Anno Accademico 1968-1969.

[3] Questa completa astrattezza chiarisce e delimita il compito assolto nella scienza dalla matematica, che è quello di studiare le possibili forme razionali del pensiero umano e di fornire sempre nuovi schemi atti ad esprimere quanto vi è di razionalizzabile nel mondo che ci circonda. Questa delimitazione di compiti nulla ha a che vedere con il

Una **teoria fisica** richiede, per sua destinazione, qualche cosa di più. Essa è bensì una teoria matematica ma ad essa si domanda che gli oggetti che la costituiscono, primitivi o derivati (o almeno una loro parte) trovino i loro corrispondenti in altrettanti oggetti ben individuabili della realtà fisica; e che **ogni relazione osservata** tra questi ultimi si trovi in accordo con le relazioni, postulate o dedotte, della teoria. In altri termini ad una teoria fisica oltre alla coerenza logica interna si domanda anche che tutte le relazioni fra enti primitivi e derivati che vi figurano siano confermate dall'osservazione sperimentale ("evidenza sperimentale").
Come nasce e come si modifica una teoria fisica? Se è ben probabile che la sua prima origine stia nell'osservazione della Natura è tuttavia certo che gli assiomi fisici non sono mai imposti in modo immediato e definitivo dalla osservazione stessa. In verità il mondo fisico non si lascia interrogare e i fenomeni non si lasciano descrivere se non alla luce di una teoria di tentativo già costituitasi, magari inconsapevolmente, nella mente di chi osserva ("le sensate esperienze" di Galileo). Questa teoria, nata da un atto di intuizione, e le sue conseguenze logiche osservabili vengono a mano a mano sottoposte al controllo sperimentale. Se ad un certo momento l'esperienza rivela un disaccordo con le deduzioni teoriche, un ulteriore atto di intuizione suggerisce modificazioni agli assiomi precedenti, la teoria modificata viene sottoposta ad ulteriori controlli e così via[4].
Una teoria fisica concernente una certa classe di fenomeni è fondata su di un certo numero di enti primitivi e di assiomi. Alcuni sono specifici dei fenomeni che la teoria ha di mira, altri sono propri anche di altre classi di fenomeni; altri infine sono del tutto generali e sono comuni a tutti i fenomeni fisici. Spesso alcuni tra questi ultimi assiomi non vengono neppure formulati esplicitamente, tanto appaiono ovvi. Allorchè un disaccordo con l'esperienza mette in dubbio una teoria è agli assiomi più specifici che si rivolgono i primi tentativi di modificazione. Ma se reiterati tentativi limitati a quest'ambito risultano inefficaci, il dubbio comincia ad interessare assiomi più generali la cui modificazione diviene tuttavia più difficile non dovendosi turbare il preesistente accordo degli stessi postulati generali con le altre classi di fenomeni fisici.
Nella storia della fisica moderna si sono avuti due momenti critici che hanno condotto a modificare assiomi generali. Il primo, originato dal disaccordo tra previsioni teoriche e risultati sperimentali relativi allo spettro di emissione del corpo nero, fu drasticamente risolto con la rinuncia alla continuità degli scambi energetici tra materia e campo elettromagnetico (Planck, 1900) e diede origine alla **meccanica quantistica (Einstein 1905, Bohr 1913, Schrödinger 1925, Heisenberg 1925, Dirac 1925).** L'altro momento critico, originato anch'esso dal disaccordo tra previsioni teoriche ed i risultati di alcune esperienze ottiche (Michelson 1881, Michelson-Morley 1887-1891), ma anche tra l'apparente incompatibilità tra il fenomeno della propagazione della luce ed il Principio di Relatività già introdotto nella fisica da Galileo, condusse alla rinuncia - ben più radicale della precedente - delle idee tradizionali sullo spazio e sul tempo e diede origine alla **Teoria della Relatività (Ristretta e Generale)(Einstein 1905-1915).**
Le considerazioni precedenti mostrano quanto sia opportuno, per l'esame critico di una teoria scientifica, formulare esplicitamente in modo preciso e completo, gli assiomi in essa adottati. La tendenza, sempre più diffusa, a formulare molte teorie fisiche (oltre che matematiche) in veste astratta puramente assiomatica, non è che la manifestazione più spinta di tale esigenza. **Una precisa enunciazione degli assiomi non è però incompatibile con un'esauriente illustrazione del significato fisico dei concetti**.
Un'ultima avvertenza: l'esame critico di una teoria fisica deve fondarsi sul controllo di due esigenze fondamentali:

valore di una particolare teoria matematica nata dall'intuizione della mente umana, a giudicare il quale possono intervenire o criteri di pura estetica, o criteri di utilità applicativa, o comunque considerazioni che vanno al di fuori della teoria stessa.

[4]Questo continuo alternarsi di atteggiamenti mentali diversi, **osservazione ed intuizione, deduzione logica e controllo sperimentale**, spiega perchè matematica e fisica hanno sempre avuto tanta reciproca influenza.

1) **coerenza logica interna;**

2) **accordo della teoria con la realtà fisica (verifica sperimentale)**.

Occorre evitare l'errore, piuttosto frequente, di giudicare una nuova teoria fisica sulla base di idee rese in noi abituali da una teoria precedente e facilmente scambiabili, tanto sono in noi radicate, per verità a priori (preconcetti).
Aggiungiamo infine un'osservazione tanto utile quanto importante: Affinché una nuova teoria sia **"una buona teoria"** deve verificare almeno tre condizioni essenziali:

1) **prevedere nuovi fenomeni;**

2) **contenere come caso limite la teoria precedente**;

3) **fornire risultati verificabili sperimentalmente** (*cioè, come spesso si dice, soddisfare al criterio dell'evidenza sperimentale*).

Vedremo che le teorie della Relatività sia Ristretta sia Generale soddisfano pienamente a tutte e tre le precedenti condizioni.

0.2 Alcune considerazioni introduttive sui fondamenti della fisica

Riportiamo qui di seguito alcune citazioni - che anche se risalenti al 1927 sono ancora molto utili ed attuali - del grande fisico sperimentale Percy Williams Bridgman (premio Nobel per la fisica nel 1947) desunte dal suo interessantissimo libro **La logica della fisica moderna**[5]
"Qualunque sia la nostra opinione circa l'accettazione definitiva dei dettagli analitici delle teorie Ristretta e Generale della Relatività di Einstein, non v'è dubbio che grazie a queste teorie la Fisica ha subito un cambiamento permanente. Lo scoprire che i nostri concetti classici, ammessi senza discussione, erano inadatti ad affrontare l'effettiva situazione, ha costituito una grande sorpresa, da cui è nato un atteggiamento critico, nei riguardi di tutta la nostra struttura concettuale, che almeno in parte permarrà. Un esame della situazione dopo questa sorpresa mostra che non avrebbero dovuto occorrere i fatti sperimentali che hanno portato alla Relatività, per convincerci dell'insufficienza dei nostri vecchi concetti, e che invece una analisi sufficientemente profonda avrebbe dovuto perlomeno prepararci alla possibilità di ciò che Einstein fece".
"Dunque il nostro problema è duplice: da un lato capire cosa stiamo cercando di fare e quali dovrebbero essere le nostre mete in Fisica, dall'altro capire la natura della struttura della Fisica così come si presenta oggi"[6].
"Accetteremo come dotato di significato il giudizio di senso comune, secondo cui **vi è un mondo a noi esterno**, *e limiteremo per quanto è possibile la nostra indagine al comportamento e all'interpretazione di questo mondo esterno"*[7].
"Nel riconoscere l'essenziale impossibilità di estrapolare il significato degli esperimenti al di là dei loro limiti attuali, il fisico deve, per non essere costretto a rivedere continuamente il proprio atteggiamento, usare nella descrizione e nel collegamento reciproco dei fenomeni naturali concetti di un genere tale, che la nostra esperienza del momento non ponga ipoteche sul futuro. In ciò mi sembra consistere il contributo maggiore di Einstein"[8].
"In generale per **concetto** *noi non intendiamo altro che un gruppo di operazioni; il concetto è sinonimo del corrispondente gruppo di operazioni necessarie per definirlo".*
"Dobbiamo esigere che il gruppo di operazioni equivalente a un dato concetto sia un gruppo unico, altrimenti sorgono inammissibili possibilità di ambiguità nelle applicazioni pratiche..... È evidente che se noi

[5] P. W. Bridgman, *La logica della fisica moderna*, Ediz. Scient. Einaudi, Torino (1957), pag. 21.
[6] Ibidem, pag. 17.
[7] Ibidem, pag. 18.
[8] Ibidem, pag. 23.

adottiamo nei riguardi dei concetti questo punto di vista, cioè che la definizione propria di un concetto va data non in termini di proprietà, ma in termini di operazioni effettive, evitiamo il pericolo di dover rivedere il nostro atteggiamento verso la Natura"[9].
Possiamo perciò dire che Einstein ha introdotto nella Fisica un metodo che deve costituire la base di tutti i nostri sforzi critici. *"Questo metodo si fonda sulla convinzione che l'aspetto più importante di una teoria è quello che questa teoria effettivamente fa, non quello che dice di fare nè quello che il suo autore ritiene essa faccia, tutte cose che spesso sono parecchio diverse. Dal punto di vista critico la tecnica più importante consiste nel capire chiaramente e nel riferire altrettanto chiaramente quello che uno effettivamente fa e quello che effettivamente avviene in una determinata situazione, e si tratta di una tecnica che non si acquista tanto facilmente e in cui si diventa esperti solo praticandola di continuo"*[10].
"Diciamo **avvenimento** *un qualunque fenomeno fisico il cui elemento oggettivo più semplice è il fatto che esso* **accade***; inoltre attribuiamo ad esso un'esistenza oggettiva indipendente dall'osservatore".*
Spesso il termine avvenimento è sostituito da **evento**, anche se forse sarebbe più corretto riservare il termine **evento** alla rappresentazione matematica di un dato avvenimento mediante le sue coordinate spazio-temporali. Non daremo alcuna definizione astratta di "**osservatore**", come fanno oggi alcuni autori, ma lo concepiremo semplicemente come uno **sperimentatore** (il fisico o lo scienziato in generale), dotato di strumenti per eseguire misure ed effettuare esperimenti sugli avvenimenti del mondo esterno. Tenuto conto del precedente "**operativismo**" di Bridgman ammetteremo perciò esplicitamente, quale premessa indispensabile di ogni teoria scientifica, i seguenti principi metodologici:

1) **Principio di oggettività degli osservatori** (Bridgman)[11]
   Osservatori diversi devono inequivocabilmente accordarsi sull'identità di un determinato "avvenimento" e "sull'identità delle circostanze locali", inoltre, posti nelle stesse condizioni, devono comportarsi nello stesso modo.
   Tali osservatori saranno chiamati **osservatori simili**. Supporremo che osservatori simili eseguano esperimenti simili. Per far ciò si può, per esempio, supporre che ogni osservatore esegua esperimenti in base alle istruzioni contenute in un manuale di *Guida all'esecuzione degli esperimenti*.

2) **Principio di verifica sperimentale** (Mach-Einstein)[12]
   Nessun concetto e nessuna affermazione che non siano suscettibili di verifica sperimentale devono trovar posto in una teoria fisica.

3) **Principio di antecedenza (o di causalità)**[13] *Se due eventi A, B, sono tali che B (effetto) sia provocato da A (causa), allora A deve precedere B per qualunque osservatore.*
   In altre parole possiamo dire che: *L'ordine temporale degli "eventi causalmente connessi" è lo stesso in tutti i sistemi di riferimento.* Ovvero più concisamente: *Il presente determina causalmente il futuro, in qualunque sistema di riferimento.* O ancora più semplicemente: *L'effetto non può mai precedere la causa.*

Infine vogliamo accennare molto brevemente alla *gerarchia della conoscenza del mondo fisico*. Nella nostra investigazione sulla Natura abbiamo sostanzialmente a che fare con *eventi, dati sperimentali, grandezze fisiche, leggi fisiche ed una sorta di "super-leggi" o "meta-leggi"*, queste ultime sono spesso chiamate **principi di simmetria**. Possiamo grosso modo dire che **le leggi organizzano ed ordinano gli eventi, i dati e le grandezze, mentre i principi di simmetria regolano le**

[9] Ibidem, pag. 24-26.
[10] P. W. Bridgman, *La natura della teoria fisica*, La Nuova Italia Editrice, Firenze (1965), pagg. 6-7.
[11] Ibidem, pagg. 88-90. Cfr. anche T. Levi-Civita e U. Amaldi, *Lezioni di Meccanica Razionale*, vol. 1, Zanichelli, Bologna (1974), pag. 86.
[12] Max Born, *La sintesi einsteiniana* , Boringhieri, Torino (1969), pag. 17.
[13] Max Born, *Filosofia naturale della causalità e del caso*, Boringhieri, Torino (1962), pag. 19.

leggi fisiche.[14] Daremo qualche esempio di principio di simmetria, nel seguito, a proposito del Principio di Relatività, introdotto per la prima volta da Galileo, nell'ambito della Meccanica, ed esteso poi da Einstein a tutta la Fisica.

[14]Sui Principi di simmetria si veda l'ottimo libro di Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), pagg. 9-10. Sulla gerarchia della conoscenza fisica e sul rapporto tra principi di simmetria e leggi si possono consultare anche le seguenti pubblicazioni: 1) Carlo Bernardini, *Le leggi fisiche*, Editori Riuniti, Roma 1983; 2) E. P. Wigner, *Symmetries and Reflections*, Indiana University Press, 1967; 3) M. Bunge, in "Am. J. Phys.", **29**, 1961, pag. 519; 4) R.M.F. Houtappel, H. Van Dam, e E. P. Wigner, in "Rev. Mod. Phys.", **37**, 1965, pag. 595; quest'ultimo articolo ha carattere più tecnico dei precedenti.

Capitolo 1

ASSIOMI DELLA FISICA CLASSICA

Se ho visto più lontano, è perché stavo sulle spalle di giganti. [1]
(Isaac Newton 1643-1727)

Io non so come possa apparire al mondo, ma, per quel che mi riguarda, mi sembra di essere stato soltanto come un bambino che gioca sulla spiaggia ed è contento se di quando in quando trova un ciottolo più levigato o una conchiglia più bella del solito, mentre un grande oceano di verità gli sta dinanzi ancora da scoprire.[2]
(Isaac Newton 1643-1727)

1.1 Spazio e tempo nella Fisica Classica

Non abbiamo certamente la pretesa di svolgere un'indagine storica e critica esauriente ed approfondita dell'evoluzione subita dai concetti fondamentali di **spazio, tempo e movimento (o moto)** da Galileo e Newton fino ad Einstein che sarebbe, per i nostri scopi, fuori luogo; per questo ci sono moltissimi trattati interamente dedicati a tali problematiche.[3] Il nostro scopo è quello di mettere in luce alcune idee fondamentali riguardo a tali concetti che possano interessare soprattutto gli studenti di Fisica e di Matematica. In particolare, si intende far acquisire loro coscienza e consapevolezza che proprio sui concetti basilari riguardanti i fondamenti dell'intera fisica ci sono stati nel passato, e ci sono ancora a tutt'oggi, dibattiti accesi su questioni concettuali e su problemi, non del tutto risolti, che investono non solo la fisica, ma anche la matematica, la filosofia della scienza e l'intero pensiero umano. E lo faremo attraverso molteplici citazioni tratte da alcuni tra i fondatori della fisica moderna a partire da Newton (alcune di queste citazioni sono distribuite nei vari paragrafi di questo capitolo, altre sono riportate nell'APPENDICE (1.5) alla fine di questo stesso capitolo.
Spesso i trattati di fisica, anche i migliori, non si soffermano quasi mai su quello che, con Bridgman, si può chiamare *"l'aspetto interpretativo della fisica"*, definendo tali speculazioni *"metafisiche"*. Per esempio, riguardo ai concetti di cui ci stiamo occupando, molti autori dichiarano che **nella Meccanica Newtoniana "lo spazio ed il tempo sono assoluti"** ma non spiegano

[1] Questa bella espressione di Newton è riportata dal famoso cosmologo austriaco (naturalizzato inglese) Hermann Bondi nel suo libro *La Relatività e il senso comune*, Zanichelli, Bologna, 1975, pag. 9.

[2] Questo famosissimo passo di Newton è riportato da Arthur Stanley Eddington nel suo libro *Spazio, Tempo e Gravitazione*, Universale Scientifica Boringhieri, 1971, pag. 122.

[3] Per esempio, oltre i libri già citati di Bridgman, possiamo tener presente: Max Jammer, *Storia del concetto di spazio*, Feltrinelli Editore, Milano (1963); Giuliano Toraldo di Francia, *L'indagine del mondo fisico*, Giulio Einaudi Editore, Torino (1976), Steven Hawking, *Dal Big Bang ai buchi neri. Breve storia del tempo*, Rizzoli, Milano (1988); Paul Davies, *I misteri del tempo*, Oscar Saggi Mondadori, Milano (1997); Ilya Prigogine, *La fine delle certezze. Il tempo, il caos e le leggi della natura*, Bollati Boringhieri (1997).

quasi mai dettagliatamente che cosa si deve intendere con tale affermazione e meno che mai si soffermano sull'evoluzione storica e critica di tali concetti fondamentali.

1.1.1 Spazio, tempo e movimento secondo Newton

E' esperienza comune che ogni fenomeno fisico si svolge nello "spazio" e nel "tempo", perciò **"spazio" e "tempo" sono gli ingredienti fondamentali di ogni teoria fisica**. Da questa semplice constatazione ne segue che ogni legge fisica comporta esplicitamente o implicitamente delle **"relazioni spazio-temporali"** cioè delle distanze e degli intervalli di tempo.
Sentiamo, a tal proposito, che cosa dice Newton, del tempo e dello spazio nella prima parte della sua opera immortale **"Philosophiae Naturalis Principia Mathematica"** (Londra, 1687; trad. italiana a cura di Alberto Pala, Utet, Torino, 1965).
Definizioni - Scolio, pagg.101-102:
"Fin qui è stato indicato in quale senso siano da intendere, nel seguito, parole non comunemente note. Non definisco, invece, **tempo, spazio, luogo e moto**, *in quanto notissimi a tutti. Va notato, tuttavia, come comunemente non si concepiscano queste quantità che in relazione a cose sensibili. Di qui nascono i vari pregiudizi, per eliminare i quali bisogna distinguere le medesime quantità in assolute e relative, vere e apparenti, matematiche e volgari.*
I. *Il tempo assoluto, vero, matematico, in sé e per sua natura senza relazione ad alcunché di esterno, scorre uniformemente, e con altro nome è chiamato durata: quello relativo, apparente e volgare, è una misura (esatta o inesatta) sensibile ed esterna della durata per mezzo del moto, che comunemente viene impiegata al posto del vero tempo: tali sono l'ora, il giorno, l'anno.*
Pagg. 102-103:
II. *Lo spazio assoluto, per sua natura senza relazione ad alcunché di esterno, rimane sempre uguale ed immobile; lo spazio relativo è una misura o dimensione mobile dello spazio assoluto, che i nostri sensi definiscono in relazione alla sua posizione rispetto ai corpi, ed è comunemente preso al posto dello spazio immobile;....*
Pagg. 105-106:
IV. *....I giorni naturali, che di consueto sono ritenuti uguali, e sono usati come misura del tempo, sono ineguali. Gli astronomi correggono questa ineguaglianza affinché, con un tempo più vero, possano misurare i moti celesti. E' possibile che non vi sia movimento talmente uniforme per mezzo del quale si possa misurare accuratamente il tempo. Tutti i movimenti possono essere accelerati e ritardati, ma il flusso del tempo assoluto non può essere mutato. Identica è la durata o la persistenza delle cose, sia che i moti vengano accelerati, sia che vengano ritardati, sia che vengano annullati;....*
Pagg. 106-107
...Così, invece dei luoghi e dei moti assoluti usiamo i relativi; né ciò riesce scomodo nelle cose umane: ma nella Filosofia occorre astrarre dai sensi. Potrebbe anche darsi che non vi sia alcun corpo in quiete al quale possano venire riferiti sia i luoghi che i moti".

Riguardo le precedenti citazioni dai Principia di Newton, osserva giustamente il Bridgman:[4]
"... "Ora nulla ci garantisce che in Natura esista qualcosa avente proprietà simili a quelle indicate nelle precedenti definizioni newtoniane, e la fisica, ridotta a concetti di questo genere, diventa una scienza astratta e lontana dalla realtà quanto la geometria astratta dei matematici, costruita su postulati. E' compito dell'esperienza stabilire se concetti così definiti corrispondono a qualcosa di esistente in Natura; noi dobbiamo essere sempre preparati a constatare che questi concetti non corrispondono a nulla, o solo in parte corrispondono a qualcosa. In particolare, se esaminiamo la definizione di **tempo assoluto** *alla luce dell'esperienza, non troviamo in Natura nulla che abbia tali proprietà".*

Noi aggiungiamo che la stessa osservazione vale per il concetto di **spazio assoluto**. **Non**

[4] P.W. Bridgman, *La logica della fisica moderna*, Edizioni scientifiche Einaudi, Torino (1957), pag. 24.

ha senso parlare della posizione o del moto di oggetti nello spazio assoluto, o di proprietà dello spazio assoluto, ma ci si deve sempre riferire ad un ben determinato sistema di coordinate legato ad un corpo rigido, come è chiarito in un meraviglioso brano di Albert Einstein che abbiamo riportato, come lettura, nel paragrafo (1.5.1) dell' APPENDICE, e che consigliamo vivamente allo studente di leggere.

Ed il grande fisico teorico Max Born osserva:[5] *"L'affermazione in base alla quale lo spazio e il tempo assoluti esistono "indipendentemente da ogni oggetto esterno" sembra strana da parte di Newton, in quanto egli spesso dà risalto al fatto che intende indagare soltanto su ciò che è reale, che può cioè essere rilevato dall'osservazione. Il suo motto, definito e conciso, è "hypotheses non fingo" (non formulo ipotesi). Tuttavia, ciò che esiste "indipendentemente da ogni oggetto esterno" non è osservabile, e non costituisce un fatto reale. Ci troviamo qui evidentemente di fronte ad un caso in cui sono state applicate, al mondo oggettivo, idee preconcette sulla cui veridicità non era stato fatto alcun esame approfondito".*

Infine i fisici teorici Landau e Lifschitz - contrariamente a ciò che molti autori affermano, ripetendo con Newton, che lo spazio è assoluto - dicono chiaramente che[6]: *"già in Meccanica Classica lo spazio è relativo, cioè le relazioni spaziali tra differenti eventi dipendono soltanto dal sistema di riferimento nel quale vengono descritti. L'asserzione che due eventi avvengono ad istanti differenti in uno stesso punto dello spazio o, in generale, ad una determinata distanza l'uno dall'altro, acquista un senso soltanto se è indicato il sistema di riferimento al quale questa asserzione si riferisce".*
"Il tempo in Meccanica Classica è invece assoluto"; in altri termini, *"si suppone che le proprietà del tempo non dipendano dal sistema di riferimento; il tempo è identico per tutti i sistemi di riferimento"*. Ciò significa che *"se due eventi arbitrari sono simultanei[7] per un osservatore, essi sono ugualmente simultanei per ogni altro osservatore"*. In generale, *"l'intervallo di tempo tra due eventi dati deve essere identico in tutti i sistemi di riferimento".*

1.2 I° Gruppo - Assiomi generali

Tenuto conto di tutte le considerazioni precedenti, per renderci conto di quali fossero le difficoltà che si presentarono nella Fisica Classica verso la fine del XIX° secolo (cioè fine 1800), converrà richiamare esplicitamente gli assiomi fondamentali sui quali essa si fondava. In essi possiamo distinguere **un gruppo di assiomi generali**, implicanti concetti comuni a tutti i fenomeni fisici, cioè **i concetti di spazio e tempo**, e **due gruppi di assiomi più specifici attinenti l'uno ai fenomeni meccanici, l'altro ai fenomeni elettromagnetici.** Nell'una o nell'altra di queste due classi si pensava allora potersi inquadrare qualsiasi fenomeno fisico.

Gli assiomi generali introducono come enti primitivi lo spazio ed il tempo come vennero sostanzialmente concepiti dai fondatori della Meccanica, Galileo e Newton, con alcuni necessari ritocchi che tengono conto delle critiche fatte dagli scienziati autorevoli citati nel paragrafo precedente, e con l'aggiunta di alcune precisazioni maturate successivamente nel corso dei secoli. L'esposizione che segue degli assiomi fondamentali non pretende minimamente di essere rigorosa ed esaustiva: si raccolgono semplicemente le proprietà dello spazio e del tempo che vengono effettivamente usate nella Fisica Classica; ciò è stato fatto esclusivamente per comodità dello studente che non ha, generalmente, il tempo né la voglia di reperirle sui vari libri e che però, noi riteniamo, debba sempre tenere presenti nel proseguimento dei suoi studi.

[5] Max Born, *La sintesi Einsteiniana*, Boringhieri, Torino (1969), pag. 80.
[6] L. D. Landau e E. M. Lifschitz, *Teoria dei campi*, Editori Riuniti, Roma (1976), pag. 15.
[7] In Meccanica Classica si intende tacitamente che **due eventi sono simultanei se accadono nello stesso istante**, senza però dare nessuna definizione operativa di simultaneità e senza alcuna verifica sperimentale. Perciò tale nozione viene assunta come un **pregiudizio** ed in ciò si deve ravvisare la sua intrinseca debolezza.

1.2.1 Lo spazio è relativo

ASSIOMA I:

1) *Esiste in Natura uno* **spazio fisico** Σ *i cui elementi P diconsi* **punti (o luoghi)**, *contenente oggetti materiali di svariata natura, sede di tutti i possibili fenomeni fisici (ma anche chimici, biologici, ecc.), al quale sono perciò subordinate, non solo tutte le leggi fisiche, ma anche tutte le possibili leggi della Natura.*

2) *In tale spazio esiste una classe particolare di oggetti chiamati* **solidi** *dai quali, per astrazione, si deduce il concetto di* **corpo rigido.**[8] (vedi Fig.1.1).

Figura 1.1: Necessità di un sistema di riferimento, legato ad un corpo rigido, per la descrizione del moto.

3) **Lo spazio fisico Σ è euclideo tridimensionale (omeomorfo ad $\mathbb{R}^3$)**
Con ciò si vuol dire che esso è rappresentabile con uno spazio euclideo a 3 dimesioni E^3, vale cioè in esso l'ordinaria Geometria Euclidea.
In Σ esistono ∞^6 terne cartesiane ortonormali distinte.[9]

4) **Proprietà d'invarianza degli intervalli spaziali**
In sistemi di coordinate arbitrari e distinti $Oxyz$, $O'x'y'z'$ in quiete fra di loro, o in moto relativo - con la convenzione che in essi venga scelta la stessa unità di misura per le lunghezze - la distanza tra

[8]Per corpo rigido possiamo adottare la seguente definizione: *Dicesi* **corpo rigido** *un corpo al quale si attribuiscono in maniera completa le proprietà sensibilmente possedute dai diversi solidi.*
Questo concetto è necessario - almeno nell'impostazione tradizionale della Fisica Classica - per definire le lunghezze e le distanze fra i punti di Σ

[9]Dobbiamo osservare che, nell'ambito della Meccanica Razionale, per individuare la **posizione di una particella** $\mathcal{P}$ o studiarne **il moto** è necessario introdurre un **corpo rigido**, che spesso, per semplificare, chiameremo **solido rigido** S_ρ a cui riferire la quiete o il moto di $\mathcal{P}$. In genere, c'è un po' l'abitudine di identificare i **Riferimenti fisici** con i **Sistemi di coordinate**. In realtà il concetto di Riferimento fisico è più generale, e si collega ai *corpi naturali indeformabili (corpi rigidi o più brevemente solidi).* **Riferimento fisico** è un qualunque corpo rigido sul quale si possa innestare, in modo solidale uno o più sistemi di coordinate cartesiane. Ed è per questo che si parla di **solido di riferimento** o talvolta di **spazio di riferimento** come l'insieme S_ρ dei punti dello spazio fisico Σ **solidali** al corpo rigido prefissato. Entro S_ρ è però possibile scegliere, in modo arbitrario, una **terna cartesiana di riferimento** $\mathcal{T} \equiv (Oxyz)$, ed è proprio rispetto a quest'ultima che si studia il moto di $\mathcal{P}$. Ci si servirà sistematicamente di terne ortogonali, monometriche, destre (*terne ortonormali destre*); l'origine O può essere scelta tra gli ∞^3 punti dello **spazio di riferimento** S_ρ; l'orientamento di un asse tra le ∞^2 rette della stella di centro O; l'orientamento di un altro asse tra le ∞^1 rette del fascio di centro O nel piano ortogonale al primo asse. Risulta così individuata la terna e **si hanno in complesso ∞^6 terne distinte**, tutte associate ad una medesima unità di lunghezza. Osserviamo che la locuzione **spazio di riferimento** è molto usata dai matematici col significato di insieme di punti aventi una determinata proprietà; si parla anche di spazi sovrapposti ed in moto l'uno rispetto all'altro. I fisici, in generale, ritengono che abbia poco senso parlare si spazi sovrapposti e preferiscono usare sempre il concetto di **terna cartesiana di riferimento**. **(sistemi di coordinate o riferimenti cartesiani)** $\mathcal{T} = Oxyz$, solidali a corpi rigidi, **rispetto alle quali è possibile reperire i punti di** Σ, munite di un osservatore $\mathcal{O}$, che supponiamo essere collocato nell'origine della terna cartesiana di riferimento, e di un regolo rigido (**regolo campione**) per la misura delle distanze spaziali ovvero per la determinazione delle coordinate x, y, z dei punti di Σ.

due punti ben determinati (o, equivalentemente, la lunghezza di un corpo lineare rigido) misurata in un riferimento cartesiano è la stessa misurata in qualsiasi altro sistema (pur avendo i due punti coordinate diverse rispetto ai diversi sistemi di coordinate).

5) **Lo spazio fisico Σ è relativo**
Con ciò - contrariamente a quanto spesso si afferma dicendo che **lo spazio è assoluto** - si vuole intendere, seguendo Landau e Lifschitz, che: le relazioni spaziali tra differenti eventi dipendono soltanto dal sistema di riferimento nel quale vengono descritti. L'asserzione che due eventi avvengono ad istanti differenti in uno stesso punto dello spazio (**eventi equilocati**) o, in generale, ad una determinata distanza l'uno dall'altro, acquista un senso soltanto se è indicato il sistema di riferimento al quale questa asserzione si riferisce.

La fisica classica ha ritenuto di poter, almeno approssimativamente, identificare lo spazio fisico Σ con lo spazio Σ^* solidale alle cosiddette **stelle fisse**, pur senza assumere questa identificazione come definitiva. In altre parole, per reperire i punti di Σ, ci si può avvalere di una terna cartesiana $\mathcal{T}^* = (O^*x^*y^*z^*)$ fissandone l'origine O^* su di una arbitraria stella fissa e gli assi x^*, y^*, z^* in direzione di altre tre stelle fisse. La fisica classica permette anche di definire altri sistemi di coordinate nello spazio Σ, mobili l'uno rispetto all'altro di moto rigido.

1.2.2 Il tempo è assoluto

ASSIOMA II:

1) *Si postula l'esistenza di un* **tempo assoluto (pantopico)** Θ, *i cui elementi t diconsi* **istanti**, *del tutto indipendente dalla nozione di spazio, cioè indipendente sia dalla posizione dell'osservatore sia dalle sue condizioni di quiete o di moto, rispetto ad una data terna cartesiana di riferimento* $\mathcal{T} = Oxyz$. *Inoltre, deve potersi dire senza incertezza se due avvenimenti sono simultanei* **(nozione di simultaneità o contemporaneità)** *ovvero se l'uno precede l'altro* (**nozione di ordine temporale**).
La valutazione numerica di una durata temporale è naturalmente subordinata alla scelta di una **unità di misura temporale**, completamente disponibile, così come deve essere disponibile **l'origine dei tempi**.

2) **Dal punto di vista matematico** *il tempo viene dunque considerato come un insieme ordinato* Θ *di infiniti elementi t (istanti) omeomorfo allo spazio numerico* $\mathbb{R}$.

3) **Dal punto di vista fisico** *l'identificazione e la misura dei tempi presuppone la disponibilità di opportuni dispositivi come* **meridiane, clessidre ed orologi** *opportunamente costruiti ed in opportune condizioni di funzionamento.*

4) Si ritiene inoltre che *sia possibile regolare gli orologi, che nei vari sistemi cartesiani misurano il tempo, in modo tale che il valore di questo, riferito ad un determinato avvenimento, sia lo stesso in ogni sistema di coordinate*
$$t = t'.$$

Parte integrante del concetto di tempo assoluto è anche la seguente

5) **Proprietà d'invarianza degli intervalli temporali**:
È determinata in modo unico e completamente autonomo la **durata** *di un certo fenomeno, o* **l'intervallo temporale** *tra due avvenimenti, nel senso che la durata di un avvenimento è la medesima in ogni sistema di coordinate.*

Nella Meccanica ed in tutta la Fisica Classica, per quanto concerne lo spazio ed il tempo, ci si attiene sistematicamente agli Assiomi I e II.
La fisica classica attribuisce allo spazio Σ e al tempo Θ un'esistenza a priori, indipendente l'uno dall'altro e indipendente dai fenomeni che in essi si svolgono.
Una terna di assi metrizzati $\mathcal{T} = Oxyz$, e il tempo metrizzato t costituiscono un **riferimento fisico spazio-temporale completo** $\mathcal{R} \equiv (Oxyz, t) \equiv (\mathcal{T}, t)$ per lo studio di tutti i fenomeni fisici, nel senso che **ad ogni avvenimento** si associano tre coordinate spaziali x, y, z ed una coordinata temporale t, atte a individuare l'avvenimento stesso nella storia dell'universo, precisandone il **luogo** e la **data**. Alla quaterna numerica spazio-temporale (x, y, z, t), associata ad un dato avvenimento, si dà il nome di **evento puntiforme** o semplicemente **evento.**
Gli assiomi generali **I** e **II** permettono di costruire la teoria descrittiva dei fenomeni di moto, cioè **la Cinematica nel riferimento** $\mathcal{R} \equiv (Oxyz, t)$: *Cinematica del punto, Cinematica dei sistemi di punti*, in particolare *Cinematica dei corpi rigidi.*
Le nozioni cinematiche si applicano a tutti i fenomeni fisici che ad ogni istante siano ben localizzabili nello spazio; esse si adattano quindi sia al moto di un punto materiale, sia alla propagazione di un impulso elettromagnetico, in particolare luminoso.
La cinematica dei moti rigidi istituita in un determinato riferimento $\mathcal{R} \equiv (Oxyz, t)$ può ugualmente stabilirsi in qualunque altro riferimento $\mathcal{R}' \equiv (O'x'y'z', t')$, generalmente mobile rispetto ad $\mathcal{R} \equiv (Oxyz, t)$, in cui è da intendersi che si possa ritenere $t = t'$ essendo il tempo assoluto.
Gli assiomi generali **I** e **II** permettono anche di *confrontare* i moti simultanei di un medesimo punto o sistema rispetto a due riferimenti diversi $\mathcal{R}$ ed $\mathcal{R}'$, di stabilire cioè *la Cinematica dei moti relativi* con le note:

a) legge di composizione delle velocità (o teorema dei moti relativi)

$$\mathbf{v} = \mathbf{v}' + \mathbf{v}_\tau, \tag{1.1}$$

b) legge di composizione delle accelerazioni (o teorema di Coriolis)

$$\mathbf{a} = \mathbf{a}' + \mathbf{a}_\tau + \mathbf{a}_{cor}, \tag{1.2}$$

dove $\mathbf{v}$, è la velocità di un arbitrario punto P rispetto al riferimento $\mathcal{R}$ (la cosiddetta **velocità assoluta**), $\mathbf{v}'$, l'analoga velocità di P rispetto al riferimento $\mathcal{R}'$ (**velocità relativa**), mentre $\mathbf{v}_\tau$, è la **velocità di trascinamento** di P; $\mathbf{a}, \mathbf{a}', \mathbf{a}_\tau$ le analoghe accelerazioni, mentre $\mathbf{a}_{cor} = 2\boldsymbol{\omega}_\tau \wedge \mathbf{v}'$ è **l'accelerazione di Coriolis**, con $\boldsymbol{\omega}_\tau$ **velocità angolare di trascinamento** di $\mathcal{R}'$ rispetto ad $\mathcal{R}$.

Osserviamo che questi due teoremi cinematici sono una conseguenza diretta dell'assioma **II** del tempo assoluto: questa osservazione sarà essenziale, soprattutto nei riguardi della legge di composizione delle velocità, nel momento in cui dovremo costruire la Teoria della Relatività Ristretta.

1.3 II° Gruppo - Assiomi della Meccanica Newtoniana

La **Meccanica Newtoniana** può essere riassunta nelle seguenti due parti:

1. **Meccanica newtoniana in assenza di gravitazione;**
2. **Meccanica newtoniana in presenza di gravitazione.**

La prima parte si può compendiare in 5 leggi fondamentali:

1) **- Le tre leggi fondamentali della dinamica** (I°, II°, III°);

2) **- Il principio classico di propagazione delle interazioni** (IV°);

3) **- Il Principio Galileiano di Relatività** (V°).

La seconda parte consiste sostanzialmente, oltre che delle 5 leggi precedenti, delle seguenti altre 4 leggi inerenti al fenomeno della gravitazione:

a) **- La legge di Newton della gravitazione universale** (VI°);

b) **- La proprietà fondamentale del campo gravitazionale (di Galileo)** (VII°);

c) **- L'equivalenza fra massa inerziale e massa gravitazionale** (VIII°).

d) **- L'equazione di Poisson del campo gravitazionale** (IX°).

Di esse ci occuperemo con qualche dettaglio nei paragrafi seguenti. Cominciamo con la prima parte:

1.3.1 II° Gruppo (A) - Assiomi della Meccanica Newtoniana in assenza di Gravitazione

Ai concetti di spazio e di tempo e alle loro implicazioni la Meccanica aggiunge le nozioni primitive di **massa (inerziale),** $m = cost. > 0$ e quindi di **punto materiale o particella o elemento materiale** (P, m) e di **forza f**. Riferendo quiete e moto ad un **sistema di riferimento inerziale** $\mathcal{R} \equiv (Oxyz, t)$ - la cui definizione, che scaturisce dal Principio d'inerzia, sarà data poco più avanti - **munito del tempo assoluto** t, gli assiomi della Meccanica Classica del punto materiale si possono enunciare nella forma tradizionale data ad essi da Newton come sintesi e perfezionamento di tutti i risultati sperimentali:

P1) Lex I^a (**Principio d'inerzia o di Galilei**):
Un punto materiale isolato[10], *cioè non sottoposto all'azione di alcuna forza, si trova in quiete o si muove di moto rettilineo uniforme.*
Cioè in assenza di forze si ha $\mathbf{v} = \mathbf{cost.}$, quindi quiete se $\mathbf{v} = \mathbf{cost.} = \mathbf{0}$ e moto rettilineo uniforme se $\mathbf{v} = \mathbf{cost.} \neq \mathbf{0}$.
Osservazione di Einstein:
*Questa legge non dice soltanto qualcosa intorno al movimento dei corpi, ma indica altresì i corpi di riferimento o sistemi di coordinate (***sistemi inerziali***) accettabili nella meccanica, che possono venir usati nella descrizione del moto.*
Le stelle fisse sono corpi per i quali il Principio d'inerzia è certamente applicabile con un alto grado di approssimazione[11].
In virtù di questa osservazione di Einstein possiamo dare la seguente
1^a **Definizione di sistema inerziale**:
Dicesi **sistema di riferimento inerziale o galileiano**[12], *e lo indicheremo con* $\mathcal{R} \equiv (Oxyz, t)$,

[10] Osserviamo che **un corpo non soggetto ad alcuna forza** viene chiamato da Landau e Lifschitz, e dai fisici in generale, **corpo libero** anziché **corpo isolato**. Nell'ambito della Meccanica Razionale, però per **corpo libero** s'intende generalmente un corpo **non vincolato** quindi non soggetto a **forze reattive' o 'reazioni vincolari'**, ma sul quale possono agire benissimo delle **forze attive**; per es. sono **corpi liberi**, nel senso dato da noi a questa locuzione, i pianeti che gravitano attorno al sole, le stelle ed i corpi celesti in generale, sui quali agiscono, come forze attive, le interazioni gravitazionali. Perciò un corpo sottratto all'azione di qualsiasi forza si può chiamare, forse più opportunamente, **'corpo isolato** come abbiamo fatto noi più sopra.

[11] Albert Einstein, *Relatività, Esposizione divulgativa*, Boringhieri, Torino (1967), pag. 51.

[12] L'appellativo di **sistemi galileiani**, fu proposto da Einstein in onore di Galileo, ed esso appare non solo giustificato, ma per così dire doveroso, in quanto nelle opere di Galileo si trova espresso, in termini mirabilmente limpidi e precisi, il riconoscimento che i fenomeni meccanici si svolgono con le stesse leggi per due osservatori, animati l'uno rispetto all'altro di un moto traslatorio uniforme (**Principio di Relatività di Galileo**).

una terna di assi cartesiani $\mathcal{T} \equiv (Oxyz)$ in cui le relazioni spaziali, determinate da regoli rigidi in quiete rispetto al sistema, sono euclidee, ed in cui esiste un **tempo universale** *t in funzione del quale le* **particelle isolate** *rimangono in quiete o si muovono di moto rettilineo uniforme.*
In base a questa definizione il principio d'inerzia può essere enunciato anche nel seguente modo:

P1) Lex I^a (**Principio d'inerzia o di Galilei**):
Esistono ∞^9 sistemi di coordinate cartesiane ortonormali, ciascuno in moto traslatorio rettilineo uniforme (TRU), l'uno rispetto all'altro, detti brevemente **riferimenti inerziali o galileiani**, *ed un* **tempo**, *detto* **tempo assoluto**, *rispetto ai quali un punto materiale isolato o si trova in quiete o si muove di moto rettilineo uniforme.*[13]

P2) Lex II^a(2^a **legge di Newton o Legge fondamentale della dinamica**):
In un qualunque sistema di riferimento inerziale $\mathcal{R} \equiv (Oxyz, t)$, la derivata temporale della quantità di moto $\mathbf{p} = m\mathbf{v}$ di una particella è, istante per istante, uguale alla somma vettoriale $\mathbf{f} = \sum_{h=1}^{n} \mathbf{f}_h$ di tutte le forze, di qualsiasi natura, applicate ad essa, cioè:

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}(P, \mathbf{v}, t), \tag{1.3}$$

essendo m la massa della particella, P la sua posizione, e $\mathbf{v}$ la sua velocità.[14]

P3) Lex III^a(**Principio di azione e reazione o di mutua azione**):
*In un sistema inerziale $\mathcal{R} \equiv (Oxyz, t)$, sia in condizione di quiete sia in condizioni di moto, se un punto materiale P_1 esercita una forza $\mathbf{F_{21}}$ (**azione**) su di un altro punto materiale P_2, diretta secondo la retta P_1P_2, quest'ultimo esercita* **simultaneamente** *sul primo una forza $\mathbf{F_{12}}$ (**reazione**) avente la stessa retta d'azione, la stessa intensità e verso contrario all'azione.*

Come conseguenza del III principio possiamo enunciare quest'altro:

P4) **Principio classico di propagazione delle interazioni fra corpi**:
Nell'ambito della Meccanica Classica, le interazioni fra corpi si propagano istantaneamente, cioè con velocità infinita.
Ad esso è opportuno aggiungere anche i seguenti altri due principi:

P5) **Principio del parallelogrammo**:
Qualunque sia il numero delle forze agenti sopra un punto materiale P, esse sono sempre sostituibili, nei riguardi del moto del punto, con un'unica forza, rappresentata dalla loro somma vettoriale, che dicesi forza totale applicata al punto.

---

[13] **I solidi galileiani di riferimento**, che indicheremo con S_g, e che si muovono l'uno rispetto all'altro di moto TRU sono evidentemente ∞^3, quante sono le possibili velocità di trascinamento vettorialmente costanti, e cioè quanti sono i vettori dello spazio. In un medesimo S_g, le terne cartesiane associate alla stessa unità di lunghezza, ma distinte tra loro, sono ∞^6, come già si è detto in una nota precedente (la nota 9). Di conseguenza, comunque si assegni una terna inerziale $\mathcal{T}_g \equiv (Oxyz)$, esistono ∞^9 terne, associate alla stessa unità di lunghezza, che si muovono rispetto a quella assegnata, e quindi anche l'una rispetto all'altra, di moto TRU, il che significa che **in complesso si hanno ∞^9 terne inerziali distinte**.

[14] In tutto il libro, sia i vettori dello spazio euclideo 3-dimensionale (spazio fisico ordinario) chiamati talvolta 3-vettori, sia quelli dello spazio 4-dimensionale di Minkowski, saranno indicati, com'è d'uso nei testi a stampa, con lettere minuscole o maiuscole in carattere grassetto, per es. $\mathbf{p}$, $\mathbf{v}$, $\mathbf{f}$, $\mathbf{E}$, $\mathbf{H}$, ecc. Se si tratta di 3-vettori o di 4-vettori si capisce sempre dal contesto.
Mentre nei testi scritti a mano (manoscritti) preferiamo indicare i 3-vettori con una sottolineatura: $\underline{p}$, $\underline{v}$, $\underline{f}$, $\underline{E}$, $\underline{H}$, mentre i 4-vettori saranno indicati con lettere, minuscole o maiuscole, sormontate da una freccia, es. $\vec{a}$, $\vec{b}$,...., $\vec{F}$, $\vec{U}$, $\vec{V}$, ecc.

P6) **Principio d'indipendenza dell'azione simultanea di più forze**:
La simultaneità di più forze applicate ad un punto materiale non modifica le loro azioni individuali. In altre parole ciascuna di esse seguita a provocare sul moto del punto considerato la stessa accelerazione che produrrebbe agendo da sola.

In pratica poi torna molto utile il seguente concetto:
Definizione: Dicesi **atto di moto o stato cinetico** o semplicemente **stato** di un sistema materiale, l'insieme di tutte le posizioni e le velocità dei punti materiali dei quali è costituito, ad un dato istante.
Ai principi precedenti dobbiamo aggiungere anche quest'altro:

P7) **Principio di determinismo di Newton-Laplace**:
Lo stato cinetico iniziale (cioè ad un determinato istante) di un sistema meccanico, sottoposto ad un assegnato insieme di forze (sollecitazione), determina univocamente tutto il suo moto (passato, presente e futuro).
Infine dobbiamo concentrare la nostra attenzione su un altro principio, che senza dubbio **è il più importante di tutta la Meccanica Classica**, la cui estensione all'intera Fisica, operata da Einstein, porterà, assieme ad un altro principio riguardante la velocità della luce, alla formulazione della Teoria della Relatività Ristretta.

P8) **Principio di Relatività di Galileo** (1ª forma):
Nessuna esperienza meccanica svolta all'interno di un riferimento galileiano può individuare il riferimento medesimo nella classe di tutti i riferimenti galileiani.
È questo **un fatto sperimentale** ben noto e messo in luce, per la prima volta e in maniera mirabile, proprio da Galileo (tramite Salviati), in un celebre passo, *della Giornata Seconda*, nella sua opera immortale *Dialogo sui due massimi sistemi del mondo, tolemaico e copernicano* (1632)[15]. Galileo, immagina un veliero che si muove di moto traslatorio rettilineo uniforme rispetto alla terraferma. Tutte le esperienze di meccanica svolte all'interno di esso (per esempio, la caduta dei gravi, il moto di un pendolo, ecc.) presenteranno, afferma Galileo, le stesse caratteristiche che avrebbero, *a parità di condizioni iniziali*, se eseguite sulla Terra. Con la conclusione che, basandosi unicamente su di esse, lo sperimentatore non avrebbe modo di accorgersi del moto del veliero.[16]

Torneremo con maggiori dettagli su questo Principio, nei paragrafi, 1.3.3, 1.3.7, 1.3.8.

1.3.2 Altre proprietà dello spazio e del tempo: omogeneità ed isotropia

Le proprietà dello spazio e del tempo, come sono concepiti nella Fisica Classica, non si esauriscono certamente nel contenuto dei due assiomi generali, visti in precedenza, in maniera separata ed indipendentemente l'uno dall'altro. Nonostante Newton presenti lo spazio ed il tempo, appunto come concetti separati ed indipendenti *ciò non è del tutto vero nemmeno nell'ambito della Meccanica Classica*. Per esempio l'unità di misura del tempo, **il secondo**, com'è noto, è attualmente definito, in base ad una convenzione internazionale del 1967, **come l'intervallo di tempo durante il quale avvengono 9.192.631.770 vibrazioni (di specificata lunghezza d'onda) della radiazione emessa da un atomo di cesio**. Utilizzando poi la proprietà della *costanza della velocità della luce nel vuoto*, nel 1983 la 17ª Conferenza Generale dei Pesi e Misure stabilì che: **il**

[15]Si consiglia vivamente il lettore di leggere tale passo meraviglioso perché si tratta di un vero capolavoro di prosa scientifica. Lo studioso può consultare: Galileo Galilei, *Dialogo dei massimi Sistemi*, Oscar Classici Mondadori, 2014, pagg. 196-197. Oppure, l'edizione Nazionale delle Opere di Galileo, Firenze, Barbera, Vol. VII, pagg. 212-213.

[16]Galileo suppone naturalmente che la Terra costituisca un riferimento inerziale. Più precisamente, egli non pensa di eseguire sul veliero esperienze di Meccanica, così raffinate, da risentire del fatto che una terna solidale alla Terra non è esattamente inerziale.

metro è la distanza percorsa dalla luce nel vuoto in un intervallo di tempo di $1/299.792.458$**-esimo di secondo**. Così facendo, la velocità della luce nel vuoto viene ad assumere il valore esatto $c = 299.792.458\,\mathrm{m/s}$.
Se si fa un'analisi critica rigorosa di queste definizioni ci si rende conto che i concetti di **spazio, tempo, velocità, moto, e moto uniforme**, sono legati fra di loro, come afferma il Bridgman, in maniera pressoché inestricabile.
Un altro esempio in cui risulta difficile, se non impossibile, separare lo spazio dal tempo è la nozione di **sistema di riferimento inerziale**, già vista nel paragrafo precedente, nella cui definizione tali concetti fondamentali sono entrambi coinvolti.
Qui vogliamo ricordare, in maniera succinta, altre notevoli proprietà dello spazio e del tempo, in relazione ai sistemi di riferimento, non incluse in ciò che abbiamo già detto.
Rispetto ad un riferimento arbitrario, lo spazio Σ appare "non uniforme o eterogeneo" ed "anisotropo" ed il tempo non omogeneo.[17]
Ciò vuol dire che per un generico corpo le sue differenti posizioni nello spazio e le sue differenti orientazioni non sono equivalenti dal punto di vista meccanico.
Ed inoltre per quanto riguarda il tempo, si può dire che in generale, esso non è omogeneo (non è uniforme), cioè i diversi istanti non sono equivalenti.
La complicazione che queste proprietà dello spazio e del tempo porterebbero nella descrizione dei fenomeni meccanici è evidente, per esempio un **corpo isolato** (cioè non sollecitato da alcuna azione esterna) non potrebbe perseverare nello stato di quiete: pur essendo nulla la velocità del corpo in un dato istante, il corpo comincerebbe a muoversi nell'istante successivo in una determinata direzione"(Non sarebbe valido il principio d'inerzia!).
Ciò succede, per esempio, se il corpo è posto con velocità nulla in un punto arbitrario, non appartenente all'asse di rotazione, di una piattaforma ruotante priva di attrito, oppure su di un treno in fase di accelerazione o durante una frenata.

Tra gli infiniti sistemi di riferimento reperibili in Σ, particolare rilevanza assume, come già si è detto, la classe privilegiata dei **riferimenti inerziali o galileiani** $\mathcal{R} \equiv (Oxyz, t)$, che ora possiamo definire, seguendo Landau e Lifschitz, in maniera diversa ma equivalente a quella già data.
2^a **Definizione di sistema inerziale**:
Dicesi **sistema inerziale o galileiano** *una terna di assi cartesiani* $\mathcal{T} \equiv (Oxyz)$ *munita di un regolo rigido per la misura delle lunghezze e di un orologio per la misura del tempo, tale che rispetto ad essa* **lo spazio sia omogeneo ed isotropo ed il tempo omogeneo**. Possiamo perciò aggiungere le seguenti proprietà dello spazio e del tempo in riferimento alla classe dei sistemi inerziali.

2) Σ è omogeneo (o uniforme) rispetto ad un qualunque sistema di riferimento inerziale $\mathcal{R} = (Oxyz, t)$:
   a) *dal punto di vista fisico* ciò significa che nell'ambito di un riferimento inerziale $\mathcal{R} = (Oxyz, t)$ tutti i punti di Σ sono equivalenti, ovvero che il risultato di qualunque esperienza svolta in un laboratorio $\mathcal{L}$ in quiete in $\mathcal{R}$ non dipende - a parità di modalità sperimentali - dall'ubicazione di $\mathcal{L}$ in $\mathcal{R}$ ovvero di $\mathcal{R}$ in Σ;
   b) *dal punto di vista matematico* ciò significa che le leggi della fisica classica pensate formulate in un determinato sistema inerziale, sono formalmente invarianti per qualunque traslazione di assi (trasformazioni lineari di coordinate non involgenti il tempo).
   Conseguenza fondamentale dell'omogeneità dello spazio, rispetto ai sistemi di riferimento inerziali, è la *legge di conservazione della quantità di moto per sistemi isolati* (cfr. Landau-Lifschitz, vol. 1, pag. 44).

[17]Per le questioni che riguardano *l'omogeneità e l'isotropia dello spazio e l'omogeneità del tempo* ci atteniamo alla trattazione che Landau e Lifschitz presentano nel loro ottimo libro *Fisica Teorica 1, Meccanica*, Editori Riuniti, Roma, 1976, pagg. 31-33 .

3) Σ è isotropo rispetto ad un qualunque sistema di riferimento galileiano:
il che significa che non esistono, rispetto a tali riferimenti, direzioni privilegiate. Più precisamente si vuol dire che:
a) *dal punto di vista fisico* il risultato di qualunque esperienza fisica svolta in un laboratorio $\mathcal{L}$, solidale con un sistema di riferimento inerziale, non dipende - a parità di modalità sperimentali - dall'orientamento di $\mathcal{L}$ in Σ;
b) *dal punto di vista matematico* si vuol dire che le leggi della fisica classica, pensate formulate in un qualunque sistema inerziale sono formalmente invarianti per rotazioni degli assi (trasformazioni lineari omogeneee di coordinate non involgenti il tempo).
Conseguenza fondamentale dell'isotropia dello spazio, rispetto ad un riferimento inerziale, è la *legge di conservazione del momento angolare per sistemi isolati* (Landau-Lifschitz, vol. 1, pag. 49).

4) Il tempo è omogeneo o uniforme rispetto ad un qualsiasi riferimento inerziale:
a) *dal punto di vista fisico* ciò significa che il risultato di qualsiasi esperienza eseguita entro un riferimento inerziale $\mathcal{R} = (Oxyz, t)$ è indipendente dall'istante in cui l'esperienza ha inizio;
b) *dal punto di vista matematico* ciò significa che tutte le leggi della fisica sono formalmente invarianti per una traslazione dei tempi, cioè per un cambiamento dell'origine temporale $\tau = t - t_0$ e dell'unità di misura.
Conseguenza fondamentale dell'uniformità del tempo, nei sistemi inerziali, è la *legge di conservazione dell'energia per sistemi isolati* (Landau-Lifschitz, vol. 1, pag. 43).

1.3.3 Ulteriori considerazioni su omogeneità ed isotropia e loro legame con i riferimenti inerziali ed il Principio di Relatività

Osserviamo ancora che **"omogeneità"** ed **"isotropia"** non sono proprietà intrinseche dello spazio ma sono piuttosto **proprietà spaziali dei sistemi di riferimento galileiani (o inerziali)**; ed analogamente **"l'uniformità (o omogeneità) del tempo"** non è una proprietà intrinseca del tempo ma **una proprietà temporale dei sistemi galileiani!**
A tal proposito scrive Rindler:[18]
"*Ogni riferimento inerziale ideale è perfettamente simmetrico. Con ciò intendiamo dire che ogni riferimento inerziale è spazialmente omogeneo ed isotropo non soltanto nei riguardi della sua geometria euclidea, ma per tutta la fisica, ed inoltre è anche temporalmente omogeneo. In altri termini, un dato esperimento fisico può effettuarsi dovunque in un sistema di riferimento inerziale (omogeneità), in ogni direzione (isotropia), ed essere ripetuto in ogni istante (omogeneità temporale), ottenendo gli stessi risultati. Tutto ciò è una diretta conseguenza del Principio di Relatività di Galileo (PRG). Per vederlo, facciamo una distinzione logica fra riferimenti inerziali (RI) e sistemi di coordinate inerziali (SCI) (distinzione che di seguito, in genere ignoreremo). I primi, cioè i riferimenti inerziali (RI), sono mere estensioni (reali o immaginate) di corpi rigidi non rotanti e mobili l'uno rispetto all'altro di moto traslatorio rettilineo uniforme (moto TRU) in un mondo senza gravità, come in Relatività Ristretta. Un sistema di coordinate inerziale è un riferimento inerziale più, in esso, una scelta di coordinate standard x, y, z, t in unità standard. Per un dato riferimento questi sistemi differiscono fra di loro tutt'al più per rotazioni, traslazioni spaziali e traslazioni temporali. Ora, strettamente parlando, PRG concerne i sistemi inerziali di coordinate.*
Principio di Relatività di Galileo (2^a forma): Le leggi della meccanica sono invarianti per un cambiamento di sistema di coordinate inerziali.
Cosicché si vede che questa equivalenza di sistemi di coordinate inerziali, quando viene applicata ad un

[18] Wolfgang Rindler, *Relativity, Special, General, and Cosmological*, second edition, Oxford University Press, 2006, pagg. 39-41.

dato riferimento inerziale RI, ne garantisce l'omogeneità e l'isotropia. È forse meno noto che, inversamente, l'omogeneità e l'isotropia di tutti i riferimenti inerziali implica il principio di relatività, come è stato specialmente sottolineato da Dixon. Sicché è veramente una questione di gusto quale delle due ipotesi si prenda per assioma e quale come conseguenza".[19]
Concludiamo questo paragrafo osservando che dai postulati ammessi, e cioè dalle proprietà di isotropia ed omogeneità spaziale ed omogeneità temporale dei riferimenti galileiani e dal principio di relatività si può dedurre rigorosamente una proprietà di reciprocità tra $\mathcal{R}$ ed $\mathcal{R}'$ che per il suo carattere intuitivo viene generalmente accettata senza dimostrazione, oppure viene assunta come postulato aggiuntivo.[20] Si tratta del seguente:
Principio di reciprocità del moto:
Dati due riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$ in moto TRU l'uno rispetto all'altro, è equivalente, dal punto di vista cinematico,[21] *supporre il primo in quiete ed $\mathcal{R}'$ in moto con velocità* $\mathbf{v}$*, oppure il secondo in quiete ed il primo in moto con velocità* $-\mathbf{v}$;
che utilizzeremo nel paragrafo successivo.

1.3.4 Relazioni tra coordinate di un medesimo evento in riferimenti galileiani diversi

L'elemento oggettivo più semplice di un fenomeno o avvenimento è il fatto che esso *accade*. Di qui il concetto astratto di **evento**, avente un'esistenza oggettiva indipendente dall'osservatore. Caso limite del concetto sopraddetto è *l'evento puntiforme*, anch'esso assunto come elemento primitivo intrinseco, e caratterizzato dal fatto di avere, in ogni riferimento $\mathcal{R} \equiv (Oxyz, t)$, una ben determinata *posizione* (x, y, z) e una ben determinata *data* t.
Il carattere oggettivo di un evento $\mathcal{E}$ assicura che, note le sue coordinate spazio-temporali (x, y, z, t) in un riferimento galileiano $\mathcal{R}$, devono risultare determinate anche le sue coordinate (x', y', z', t') in un qualunque altro riferimento galileiano $\mathcal{R}'$. Ciò equivale a dire che x', y', z', t' sono funzioni *univoche ed invertibili* di x, y, z, t:

$$x'_i = f_i(x_1, x_2, x_3, x_4) \quad (i = 1, 2, 3, 4) \tag{1.4}$$

dove, per brevità, si è posto $x_1 = x, x_2 = y, x_3 = z, x_4 = t$.
Non è possibile ottenere la forma delle funzioni $f_i(x_1, x_2, x_3, x_4)$, $(i = 1, 2, 3, 4)$ senza utilizzare alcune delle proprietà dello spazio e del tempo relative ai sistemi galileiani.
Nelle deduzioni che seguono consideriamo due sistemi di riferimento galileiani $\mathcal{R}$, $\mathcal{R}'$ in moto traslatorio rettilineo uniforme (TRU) l'uno rispetto all'altro lungo l'asse x (in tal caso si dice che i due riferimenti sono in **configurazione tipica o x-standard**, come mostrato in Fig.3.4), ed utilizzeremo, in un primo tempo, soltanto proprietà inerenti ai sistemi galileiani; tale modo di procedere ci consente di ottenere trasformazioni di coordinate che sono valide sia in meccanica classica che in Relatività Ristretta. Sarà un postulato aggiuntivo, diverso per le due teorie, che ci farà imboccare la strada che porta alle trasformazioni di Galilei ovvero alle trasformazioni di Lorentz-Einstein.

Le proprietà dei sistemi galileiani a cui si allude sono quelle già discusse in precedenza nei paragrafi 1.3.2, 1.3.3 e cioè:

[19] La distinzione che fa Rindler tra **riferimenti inerziali (RI)** e **sistemi di coordinate inerziali (SCI)** altro non è che la distinzione, ben nota, che si fa nei corsi di Meccanica Razionale tra **solidi galileiani di riferimento (S_g)** e **terne cartesiane galileiane di riferimento ($\mathcal{T}_g$)** come già abbiamo osservato in precedenza nelle note 9 e 13 a piè di pagina.

[20] Per una dimostrazione di questa proprietà si veda: Carlo Cattaneo, *Appunti di Meccanica Relativistica*, La Goliardica Editrice, Roma, 1972, pagg. 42-44.

[21] Per *punto di vista cinematico* si vuol qui intendere che l'ipotesi assunta riguarda unicamente la struttura delle formule di trasformazione per il passaggio dalle coordinate (x, y, z, t) di $\mathcal{R}$ a quelle (x', y', z', t') di $\mathcal{R}'$, mentre non si fa assolutamente nessuna asserzione sulla relazione che lega le leggi di un fenomeno osservato in $\mathcal{R}$ e quelle dello stesso fenomeno osservate in $\mathcal{R}'$.

Figura 1.2: Sistemi di riferimento galileiani in configurazione x-standard.

I) *In qualunque sistema galileiano lo spazio è omogeneo (o uniforme) ed isotropo ed il tempo omogeneo (o uniforme) rispetto a tutti i fenomeni fisici;*

II) *Principio di Relatività;*

III) *Principio di reciprocità del moto.*

In virtù dell'assioma I) cioè per *l'omogeneità dello spazio e del tempo* nei sistemi inerziali, le trasformazioni debbono essere lineari ed invertibili[22].
Infatti, se le trasformazioni non fossero lineari verrebbe violata l'ipotesi di omogeneità dello spazio. Supponiamo, per esempio, che x' dipenda dal quadrato di x, cioè sia $x' = a_{11}x^2$, allora la distanza fra due punti posti sugli assi $x \equiv x'$ nel sistema $\mathcal{R}'$ sarebbe legata alla posizione degli stessi punti nel sistema $\mathcal{R}$ dalla relazione $x'_2 - x'_1 = a_{11}(x_2^2 - x_1^2)$. Supponiamo ora che un'asta rigida di lunghezza unitaria in $\mathcal{R}$ abbia i suoi estremi in $x_2 = 3$ e $x_1 = 2$, in $\mathcal{R}'$ si avrebbe $x'_2 - x'_1 = 5a_{11}$. Se, invece, la stessa asta fosse posta con gli estremi in $x_2 = 5$ e $x_1 = 4$ avremmo ottenuto $x'_2 - x'_1 = 9a_{11}$. Cioè, la lunghezza dell'asta dipenderebbe dalla sua posizione nello spazio e perciò verrebbe meno, come si è detto, la sua omogeneità. Similmente, possiamo escludere ogni dipendenza da t che non sia lineare, in quanto la durata temporale di un ben determinato avvenimento non deve dipendere dagli istanti iniziale e finale ma solo dalla loro differenza.
Le trasformazioni cercate saranno dunque del tipo

$$\begin{cases} x' = a_{11}x + a_{12}y + a_{13}z + a_{14}t + a_{10}, \\ y' = a_{21}x + a_{22}y + a_{23}z + a_{24}t + a_{20}, \\ z' = a_{31}x + a_{32}y + a_{33}z + a_{34}t + a_{30}, \\ t' = a_{41}x + a_{42}y + a_{43}z + a_{44}t + a_{40}. \end{cases} \tag{1.5}$$

Si tratta ora di determinare i 20 coefficienti a_{ik} della (1.5). Una volta fissati i riferimenti $\mathcal{R}$ e $\mathcal{R}'$, questi coefficienti dipenderanno tutt'al più dalla componente algebrica v della velocità di $\mathcal{R}'$ rispetto ad $\mathcal{R}$, ma non dipenderanno né da $\mathcal{R}$ né da $\mathcal{R}'$, stante il principio di relatività. Senza ledere la generalità fisica possiamo supporre che all'istante $t = t' = 0$ le due terne $\mathcal{R}$ ed $\mathcal{R}'$ risultino sovrapposte, il che significa che per $x = 0, y = 0, z = 0, t = 0$ deve essere anche $x' = 0$, $y' = 0$, $z' = 0$, $t' = 0$, e ciò comporta che sia

$$a_{10} = a_{20} = a_{30} = a_{40} = 0,$$

cioè che la trasformazione sia lineare omogenea. Come già abbiamo detto all'inizio del paragrafo, supponiamo che $\mathcal{R}'$ sia in moto traslatorio rettilineo uniforme con velocità $\mathbf{v}$ rispetto ad $\mathcal{R}$ lungo l'asse x. Consideriamo ora la terza delle (1.5), per il modo in cui sono stati scelti gli

[22]La linearità delle trasformazioni si può dedurre anche nel seguente modo: valendo nei due sistemi $\mathcal{R}$ ed $\mathcal{R}'$ il principio d'inerzia, un punto materiale isolato o si trova in quiete o si muove di moto rettilineo uniforme sia in $\mathcal{R}$ sia in $\mathcal{R}'$. Ciò è possibile soltanto se (x', y', z', t') sono funzioni lineari invertibili di (x, y, z, t).

assi, ogni evento che in $\mathcal{R}$ abbia luogo sul piano xy (piano $z = 0$), con x, y, t arbitrari, in $\mathcal{R}'$ deve aver luogo sul piano $z' = 0$. Ciò implica che sia

$$a_{31} = a_{32} = a_{34} = 0.$$

Per analogo motivo, considerando la seconda delle (1.5), risulta che ogni evento che in $\mathcal{R}$ abbia luogo sul piano xz (piano $y = 0$), con x, z, t arbitrari, in $\mathcal{R}'$ deve aver luogo sul piano $y' = 0$. Ciò implica che sia

$$a_{21} = a_{23} = a_{24} = 0.$$

Sempre in base alle convenzioni fatte, risulta che ogni evento che in $\mathcal{R}$ abbia luogo all'istante $t = 0$ sul piano yz piano ($x = 0$), con y, z arbitrari, in $\mathcal{R}'$ ha luogo sul piano $x' = 0$ all'istante $t' = 0$. Ciò implica che nella prima delle (1.5) si abbia

$$a_{12} = a_{13} = 0,$$

mentre nell'ultima si avrà

$$a_{42} = a_{43} = 0.$$

Perciò la trasformazione (1.5) si riduce a

$$\begin{cases} x' = a_{11}(v)x + a_{14}(v)t \\ y' = a_{22}(v)y \\ z' = a_{33}(v)z \\ t' = a_{41}(v)x + a_{44}(v)t. \end{cases} \tag{1.6}$$

Ora, essendo gli assi y e z trattati allo stesso modo (data *l'isotropia dello spazio*), avremo:

$$a_{22}(v) = a_{33}(v);$$

indicando il valore comune di questi coefficienti con $\lambda(v)$, la 2^a e la 3^a delle (1.6) si riduranno a:

$$y' = \lambda(v)y, \quad z' = \lambda(v)z. \tag{1.7}$$

Ma, per l'ipotesi della reciprocità del moto, gli stessi risultati si dovranno avere considerando $\mathcal{R}$ in moto con velocità $-\mathbf{v}$ rispetto ad $\mathcal{R}'$ (vedi Fig.1.3).

Figura 1.3: Reciprocità del moto.

Dovranno quindi valere le relazioni:

$$y = \lambda(-v)y', \quad z = \lambda(-v)z',$$

che confrontate con le precedenti, danno:

$$\lambda(v)\lambda(-v) = 1. \tag{1.8}$$

D'altra parte un'inversione del verso positivo di x e x' porta a sostituire $-\mathbf{v}$ a $\mathbf{v}$, senza però che nulla muti nei riguardi di y e z. Dovendo pertanto le equazioni che in questo modo si ottengono, e cioè:

$$y' = \lambda(-v)y, \qquad z' = \lambda(-v)z,$$

coincidere con le (1.7), si avrà:

$$\lambda(v) = \lambda(-v)$$

onde dalla (1.8):

$$[\lambda(v)]^2 = 1,$$

ossia:

$$\lambda(v) = \pm 1.$$

Avendo però supposto gli assi y' e z' rispettivamente equiversi con y e z, dovremo prendere senz'altro il segno $+$, per cui in definitiva avremo:

$$y' = y, \qquad z' = z.$$

Teniamo ora conto del fatto che ogni punto solidale ad $\mathcal{R}'$, in $\mathcal{R}$ deve apparire muoversi parallelamente all'asse x con velocità (algebrica) v. In particolare, ogni punto di $\mathcal{R}'$ appartenente al piano $x' = 0$ deve muoversi, rispetto ad $\mathcal{R}$, con legge $x = vt$, $y = \text{cost}$, $z = \text{cost}$, t arbitrario. In altri termini, ogni evento avente in $\mathcal{R}$ coordinate del tipo $x = vt$, $y = \text{cost}$, $z = \text{cost}$, t arbitrario, deve avere in $\mathcal{R}'$ la prima coordinata nulla, ovvero $x' = 0$. Da ciò si ricava che, per la 1ª delle (1.6), deve risultare

$$a_{14} = -a_{11}v.$$

La nostra trasformazione (1.6) si riduce perciò alla

$$\begin{cases} x' = a_{11}(v)(x - vt) \\ y' = y \\ z' = z \\ t' = a_{41}(v)x + a_{44}(v)t. \end{cases} \tag{1.9}$$

In questa trasformazione sono rimasti incogniti ancora tre coefficienti $a_{11}(v), a_{41}(v), a_{44}(v)$, è però possibile dimostrare che deve essere $a_{11}(v) = a_{44}(v)$ e che $a_{41}(v)$ si può esprimere in funzione di $a_{11}(v)$ e di $a_{11}(-v)$.
Infatti, con un'ulteriore applicazione del principio di reciprocità del moto, la prima delle (1.9) è ancora valida scambiando fra di loro x con x', t con t' e v con $-v$, avremo

$$x = a_{11}(-v)(x' + vt'),$$

da cui risolvendo rispetto a t' si ottiene

$$t' = \frac{1}{v}\left[\frac{x}{a_{11}(-v)} - x'\right].$$

Sostituendo in quest'ultima equazione al posto di x' la sua espressione data dalla prima delle (1.9) si ottiene

$$t' = \frac{1}{v}\left[\frac{1}{a_{11}(-v)} - a_{11}(v)\right]x + a_{11}(v)t \tag{1.10}$$

che deve identificarsi con la quarta delle (1.9), cioè si dovrà avere

$$a_{41}(v) = \frac{1}{v}\left[\frac{1}{a_{11}(-v)} - a_{11}(v)\right], \qquad a_{44}(v) = a_{11}(v). \tag{1.11}$$

In definitiva la nostra trasformazione avrà la forma

$$\begin{cases} x' = a_{11}(v)(x - vt) \\ y' = y \\ z' = z \\ t' = a_{41}(v)x + a_{11}(v)t, \end{cases} \tag{1.12}$$

con $a_{41}(v)$ dato dalla prima delle (1.11). Possiamo infine notare che il coefficiente $a_{11}(v)$ è una funzione pari di v, mentre $a_{41}(v)$ è una funzione dispari. Infatti, utilizzando ancora una volta l'isotropia spaziale dei riferimenti galileiani possiamo contemporaneamente invertire le direzioni dell'asse x e del moto, cioè sostituire nell'ultima delle (1.12) x con $-x$ e v con $-v$ per renderci conto che $a_{11}(-v) = a_{11}(v)$ e $a_{41}(-v) = -a_{41}(v)$.
Ci siamo perciò ridotti ad una trasformazione in cui è rimasto incognito sostanzialmente il solo coefficiente $a_{11}(v)$. Non è possibile semplificare ulteriormente questa trasformazione, né determinare i coefficienti $a_{41}(v), a_{11}(v)$ sulla sola base del principio di relatività, ovvero, il che è lo stesso, utilizzando ulteriormente le proprietà di isotropia ed omogeneità dei sistemi galileiani come abbiamo fatto finora. É necessario introdurre qualche altro assioma: la Meccanica Classica introduce, senza alcuna verifica sperimentale, come 2° assioma l'esistenza di un TEMPO ASSOLUTO $t = t'$ e così facendo determina le trasformazioni di Galileo, mentre la Meccanica Relativistica non postula a priori nessuna relazione fra il tempo t del riferimento $\mathcal{R}$ ed il tempo t' di $\mathcal{R}'$, ma introduce, più realisticamente, un assioma verificabile sperimentalmente (si veda l'esperimento di Michelson-Morley nel Cap.2) e cioè il POSTULATO DELL'INVARIANZA DELLA VELOCITÁ DELLA LUCE NEL VUOTO rispetto ai riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$, cioè $c = c'$, e determina le trasformazioni di Lorentz-Einstein.
Nel prossimo paragrafo ricaviamo le trasformazioni di Galileo ed in seguito, nel Cap.3, quelle di Lorentz-Einstein.

1.3.5 Trasformazioni di Galileo

Se ammettiamo l'ipotesi di un **tempo assoluto** $t' = t$, dalla quarta delle (1.12) si ottiene immediatamente che deve essere $a_{41}(v) = 0$ e $a_{11}(v) = 1$ e perciò le (1.12) si riducono alla cosìdetta *trasformazione speciale di Galileo*:[23]

$$(G) \qquad \begin{cases} x' = x - vt, \\ y' = y, \\ z' = z, \\ t' = t, \end{cases} \tag{1.13}$$

la cui inversa è:

$$(G^{-1}) \qquad \begin{cases} x = x' + vt, \\ y = y', \\ z = z', \\ t = t'. \end{cases} \tag{1.14}$$

Se il riferimento galileiano $\mathcal{R}'$ si muove di moto traslatorio uniforme rispetto ad $\mathcal{R}$ con velocità $\mathbf{v}$ lungo una direzione arbitraria, è facile dimostrare, sempre sotto l'ipotesi che all'istante iniziale

[23] La qualifica di "speciale" si riferisce alla particolare scelta degli assi nei due riferimenti galileiani considerati in mutua relazione x-standard.

$t = t' = 0$ i due riferimenti siano coincidenti, $\mathcal{R} \equiv \mathcal{R}'$, che la trasformazione di Galileo sarà

$$\begin{cases} \mathbf{x}' = \mathbf{x} - \mathbf{v}t \\ t' = t, \end{cases} \tag{1.15}$$

nella quale $\mathbf{x}$ e $\mathbf{x}'$ sono i vettori di posizione di un punto arbitrario in $\mathcal{R}$ ed $\mathcal{R}'$ rispettivamente. Se invece all'istante iniziale $t = t' = 0$, comune ai due riferimenti, essi non coincidono ma sono equiorientati ed inoltre l'origine O' di $\mathcal{R}'$ sia individuata dal vettore $\mathbf{x}_0$ rispetto ad $\mathcal{R}$, si avrà la trasformazione più generale:

$$\begin{cases} \mathbf{x}' = \mathbf{x} - \mathbf{v}t - \mathbf{x}_0 \\ t' = t. \end{cases} \tag{1.16}$$

Infine se abbandoniamo le ipotesi restrittive necessarie per ricavare le (1.15) e (1.16) le trasformazioni di Galileo che otteniamo non sono più omogenee ed incorporano anche una rotazione degli assi di uno dei sistemi inerziali rispetto all'altro. Interpretando i vettori $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{x}_0$ come matrici 3×1 delle coordinate che esse individuano[24], avremo:

$$\begin{cases} \mathbf{x} = R\,\mathbf{x}' + \mathbf{v}t + \mathbf{x}_0 \\ t = t' - t'_0, \end{cases} \tag{1.17}$$

la cui trasformazione inversa è:

$$\begin{cases} \mathbf{x}' = R^{-1}\left(\mathbf{x} - \mathbf{v}t - \mathbf{x}_0\right) \\ t' = t + t'_0. \end{cases} \tag{1.18}$$

Il significato dei simboli che figurano nelle (1.17), (1.18) è il seguente:

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

è il vettore di posizione di una particella mobile nel riferimento $\mathcal{R}$,

$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

l'analogo vettore di posizione nel riferimento $\mathcal{R}'$,

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$$

la velocità di traslazione di $\mathcal{R}'$ rispetto ad $\mathcal{R}$, ed

$$\mathbf{x_0} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

il vettore di posizione dell'origine O' del sistema $\mathcal{R}'$ rispetto ad $\mathcal{R}$, all'istante iniziale $t = 0$ segnato dall'orologio di O; t'_0 è invece l'istante segnato dall'orologio di O' quando quello di O

[24] La moltiplicazione fra matrici sarà sempre intesa righe per colonne.

segna $t = 0$. Se si tiene presente che il tempo è assoluto, possiamo sempre supporre che gli orologi dei due riferimenti siano sincronizzati in modo tale che risulti $t'_0 = 0$. Infine, abbiamo indicato con R la *matrice ortogonale* 3×3 (cioè tale che $R^{-1} \equiv R^T$) di rotazione che porta il riferimento inerziale $\mathcal{R}$ a coincidere con l'analogo $\mathcal{R}'$ qualora essi all'istante iniziale avessero in comune le origini $O \equiv O'$ ma orientamento diverso.
Le (1.17) e (1.18) sono **le più generali trasformazioni di Galileo non omogenee fra due sistemi inerziali**. Il loro insieme forma **un gruppo a 10 parametri (il gruppo di Galileo)**: le 3 componenti di R che individuano asse ed angolo di rotazione e possono essere espresse, per esempio, per mezzo degli *angoli di Eulero* (ψ, θ, φ); le 3 componenti della velocità $\mathbf{v}$ di traslazione; le 3 componenti di $\mathbf{x}_0$, ed infine t'_0. Rispetto alle (1.15) esse includono una *rotazione degli assi*, *una traslazione spaziale*, ed una *traslazione temporale*; ovviamente la traslazione temporale può essere eliminata sincronizzando gli orologi di O e di O', riducendo così $t'_0 = 0$, nel qual caso il gruppo di Galileo si riduce ad avere 9 parametri conformemente al fatto, già visto in precedenza (nota 13), che complessivamente si hanno ∞^9 terne inerziali distinte.
La rotazione che porta la terna di riferimento $\mathcal{R}$ a coincidere con la terna $\mathcal{R}'$ può essere spiegata dal teorema di Eulero sulle rotazioni finite (nell'ambito della Cinematica dei moti rigidi), che per comodità del lettore riportiamo qui sotto:
Teorema di Eulero:
Il più generale spostamento di un corpo rigido (riferimento $\mathcal{R}$) con un punto fisso $O \equiv O'$ è una rotazione attorno ad un asse, uscente da O.
Precisamente, se indichiamo con α $(0 \leq \alpha < \pi)$ **l'angolo di rotazione** e con $\mathbf{u}$ il versore dell'asse di rotazione, si definisce **vettore di rotazione** il vettore $\mathbf{w}$ dato da

$$\mathbf{w} = \tan\left(\frac{\alpha}{2}\right)\mathbf{u}.$$

Inoltre l'asse di rotazione ha la direzione dell'autovettore associato all'unico autovalore reale della matrice di rotazione R, e l'angolo α di rotazione soddisfa la relazione

$$tr\, R = 1 + 2\cos\alpha,$$

avendo indicato con $tr\, R$ la **traccia** della matrice R.

Osservazione 1:
Se la rotazione del sistema $\mathcal{R}'$ rispetto ad $\mathcal{R}$ si descrive utilizzando gli angoli di Eulero (ψ, θ, ϕ), la matrice di rotazione, dipendendo dall'angolo α, possiamo ora indicarla con $R(\alpha)$ e si può decomporre nel prodotto delle tre rotazioni:

$$R_\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$R_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix},$$

$$R_\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

precisamente si ha $R(\alpha) = R_\phi R_\theta R_\psi$, ed in forma esplicita:

$$R(\alpha) = \begin{pmatrix} \cos\psi\,\cos\phi - \sin\psi\,\cos\theta\,\sin\phi & \sin\psi\,\cos\phi + \cos\psi\,\cos\theta\,\sin\phi & \sin\theta\,\sin\phi \\ -\cos\psi\,\sin\phi - \sin\psi\,\cos\theta\,\cos\phi & -\sin\psi\,\sin\phi + \cos\psi\,\cos\theta\,\cos\phi & \sin\theta\,\cos\phi \\ \sin\psi\,\sin\theta & -\cos\psi\,\sin\theta & \cos\theta \end{pmatrix}.$$

Inoltre, la relazione che lega l'angolo di rotazione α con gli angoli di Eulero e che ci fornisce il modulo del vettore di rotazione è[25]

$$|\mathbf{w}|^2 = \tan^2\left(\frac{\alpha}{2}\right) = \tan^2\left(\frac{\theta}{2}\right)\frac{1}{\cos^2\left(\frac{\psi+\phi}{2}\right)} + \tan^2\left(\frac{\psi+\phi}{2}\right).$$

Osservazione 2:
Esprimiamo ora la più generale trasformazione di Galileo in termini di coordinate cartesiane. A tal fine, consideriamo una *terna cartesiana di riferimento ortonormale* $\mathcal{R} \equiv (O, \mathbf{e}_i)$, di origine O e versori $\mathbf{e}_i$, rispetto alla quale la posizione di un generico punto P dello spazio è individuata dal vettore $OP = x^i\mathbf{e}_i$ (somma sottintesa rispetto all'indice ripetuto i). Fissato il riferimento ortonormale, brevemente la terna $\mathcal{R}$, le coordinate x^i sono determinate univocamente.
Passando ad un altro riferimento ortonormale arbitrario $\mathcal{R}' \equiv (O', \mathbf{e}'_i)$, lo stesso punto P sarà rappresentato da altre coordinate x'^i: $O'P = x'^i\mathbf{e}'_i = x^{i'}\mathbf{e}_{i'}$.[26] Qual'è il legame fra le x^i e le $x^{i'}$? Per far ciò è necessario tener conto che i versori della base $\mathbf{e}_{i'}$ si esprimono come combinazione lineare dei versori della base $\mathbf{e}_i$ e viceversa:

$$\mathbf{e}_{i'} = \alpha^i_{i'}\mathbf{e}_i, \qquad \mathbf{e}_i = \alpha^{i'}_i\mathbf{e}_{i'}, \tag{1.19}$$

dove le matrici $(\alpha^i_{i'})$ ed $(\alpha^{i'}_i)$ (indice in alto di riga, in basso di colonna) sono l'una l'inversa dell'altra ed i loro elementi sono i coseni direttori dei versori di una base rispetto all'altra.
Essendo i versori di base ortonormali sono soggetti alle condizioni di ortonormalità

$$\mathbf{e}_i \cdot \mathbf{e}_k = \delta_{ik}, \qquad \mathbf{e}_{i'} \cdot \mathbf{e}_{k'} = \delta_{i'k'}, \tag{1.20}$$

dove $\delta_{ik} \equiv \delta_{i'k'}$ è il simbolo di Kronecker così definito:

$$\delta_{ik} = 1, \quad \text{per } i = k, \qquad \delta_{ik} = 0, \quad \text{per } i \neq k.$$

Le precedenti relazioni comportano perciò le condizioni di ortogonalità:

$$\alpha^i_{i'}\alpha^k_{k'}\delta_{ik} = \delta_{i'k'}, \qquad \alpha^{i'}_i\alpha^{k'}_k\delta_{i'k'} = \delta_{ik}.$$

Figura 1.4: Trasformazione generale di Galileo.

In definitiva, dalla relazione triangolare $OP = OO' + O'P$ (vedi Fig.1.4), posto $OO' = T^i\mathbf{e}_i$ ed invertendo i termini della somma, avremo

$$x^i\mathbf{e}_i = x^{i'}\mathbf{e}_{i'} + T^i\mathbf{e}_i = x^{i'}\alpha^i_{i'}\mathbf{e}_i + T^i\mathbf{e}_i,$$

[25]Per la dimostrazione di questa formula si può consultare il seguente volume: F. Bampi, M. Benati, A. Morro, *Problemi di Meccanica Razionale*, Casa Editrice: ECIG (Edizioni Culturali Internazionali Genova), 1988, pagg. 48-60.
[26]Nel Calcolo Tensoriale è usuale mettere l'apice sull'indice anziché sulla lettera.

da cui, essendo i vettori della base $\mathbf{e}_i$ linearmente indipendenti, ne segue il legame cercato (ovviamente invertibile):

$$x^i = \alpha^i_{i'} x^{i'} + T^i, \tag{1.21}$$

dove $\alpha^i_{i'}$, ripetiamo, sono i coefficienti di una matrice ortogonale (*matrice di rotazione*).
Per ottenere *la più generale trasformazione di Galileo* basta che i coseni direttori $\alpha^i_{i'}$, che figurano nella (1.21), siano costanti e invece le T^i funzioni lineari di t (esse sono le coordinate dell'origine O' della terna $\mathcal{R}'$, punto in moto rettilineo uniforme rispetto ad $\mathcal{R}$), avremo:

$$x^i = \alpha^i_{i'} x^{i'} + v^i t + s^i \qquad (i = 1, 2, 3). \tag{1.22}$$

Il PRG può essere perciò rienunciato così:
Le leggi differenziali della Meccanica sono formalmente invarianti per trasformazioni del tipo (1.22).
Esercizio: Scrivere in forma cartesiana esplicita la più generale trasformazione di Galileo.

1.3.6 Conseguenze delle trasformazioni di Galileo

In virtù della trasformazione speciale (1.13) di Galileo, derivano notevoli conseguenze; di seguito ne enunciamo tre abbastanza importanti, ma la più notevole sarà esaminata nel paragrafo successivo.

I. **Invarianza delle lunghezze o delle distanze**.
Supponiamo di avere una sbarra lineare rigida AB in quiete in una determinata posizione nel riferimento galileiano $\mathcal{R}$, i suoi estremi avranno certe coordinate ben determinate $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$. In $\mathcal{R}$ la sua lunghezza si ottiene facilmente, come sappiamo dalla Geometria Analitica, applicando il teorema di Pitagora:

$$\ell = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}.$$

In $\mathcal{R}'$ avremo in maniera analoga

$$\ell' = \sqrt{(x'_B - x'_A)^2 + (y'_B - y'_A)^2 + (z'_B - z'_A)^2}.$$

Per la trasformazione speciale di Galileo (1.13) si avrà

$$x'_A = x_A - vt, \;\; y'_A = y_A, \;\; z'_A = z_A, \;\; x'_B = x_B - vt, \;\; y'_B = y_B, \;\; z'_B = z_B,$$

che sostituite nelle precedenti relazioni ci danno

$$\ell' = \ell. \tag{1.23}$$

In altre parole, **la lunghezza della sbarra è invariante per trasformazioni speciali di Galileo**, ma lo stesso risultato si ottiene per una trasformazione galileiana generale, in quanto le rotazioni puramente spaziali e le traslazioni non alterano le lunghezze.
Il principio d'invarianza delle lunghezze, essendo una conseguenza diretta della trasformazione di Galileo, dipende sostanzialmente dal fatto che *il tempo si ritiene assoluto*!

II. **Teorema dei moti relativi di Galileo (o legge di composizione delle velocità**.
Derivando le (1.13) rispetto al tempo si ottiene:

$$u'_x = u_x - v, \qquad u'_y = u_y, \qquad u'_z = u_z,$$

o anche, in notazione vettoriale

$$\mathbf{u}' = \mathbf{u} - \mathbf{v}, \tag{1.24}$$

ovvero

$$\mathbf{u} = \mathbf{u}' + \mathbf{v}, \tag{1.25}$$

dove si è posto, per definizione (ricordando che $t' = t$):

$$\begin{array}{lll} u_x \equiv \dot{x}, & u_y \equiv \dot{y}, & u_z \equiv \dot{z}, \\ u'_x \equiv \dot{x}', & u'_y \equiv \dot{y}', & u'_z \equiv \dot{z}'. \end{array}$$

È bene avvertire che tradizionalmente, a scopo di distinzione linguistica, e quindi in senso puramente convenzionale, la $\mathbf{u}$ viene chiamata **velocità assoluta** della particella, la $\mathbf{u}'$ la sua **velocità relativa** e $\mathbf{v}$ la **velocità di trascinamento** di $\mathcal{R}'$ rispetto ad $\mathcal{R}$. Sicché la (1.25) può essere così enunciata:
Teorema dei moti relativi: *Per una particella in moto,* **la sua velocità assoluta** *(cioè rispetto ad $\mathcal{R}$) è uguale alla somma vettoriale della* **sua velocità relativa** *(cioè rispetto ad $\mathcal{R}'$) più la* **velocità di trascinamento** *di $\mathcal{R}'$ rispetto ad $\mathcal{R}$.*
È immediato trovare la legge di composizione delle velocità anche, derivando rispetto a t, la trasformazione di Galileo (1.15) in forma vettoriale; lo stesso risultato si ottiene, in maniera meno semplice, derivando, sempre rispetto a t, la prima delle (1.17) (in questo caso è necessario ricorrere alla teoria dei moti relativi, utilizzando le formule di Poisson, come lo studente avrà certamente visto nel corso di Meccanica Razionale).
Derivando la (1.25), rispetto al tempo t si ottiene un risultato notevolissimo
III.**Teorema d'invarianza dell'accelerazione**:
L'accelerazione di una particella è invariante nel passaggio da un sistema di riferimento inerziale $\mathcal{R}$ ad un qualsiasi altro $\mathcal{R}'$ anch'esso inerziale, cioè

$$\mathbf{a} = \mathbf{a}'. \tag{1.26}$$

1.3.7 Principio di Relatività e invarianza della 2^a legge della dinamica per trasformazioni di Galileo

Stabilito che, nell'ambito della Meccanica Classica, le formule che legano le coordinate spazio-temporali di un dato **evento** in un riferimento inerziale $\mathcal{R}$ alle analoghe in un altro riferimento inerziale $\mathcal{R}'$, sono date da una trasformazione di Galileo (1.15), (1.16) o (1.17), siamo in grado di riformulare il PRG in una terza forma matematicamente più soddisfacente delle prime due già enunciate, che, per comodità del lettore riportiamo qui di seguito:

- **Principio di Relatività di Galileo** (1ª forma):
  Nessuna esperienza meccanica svolta all'interno di un riferimento galileiano può individuare il riferimento medesimo nella classe di tutti i riferimenti galileiani.

- **Principio di Relatività di Galileo** (2ª forma):
  Le leggi della meccanica sono identiche in tutti i riferimenti inerziali.

- **Principio di Relatività di Galileo** (3ª forma):
  Le leggi della Meccanica sono formalmente invarianti di fronte ad una qualunque trasformazione di Galileo.

Sappiamo che, per una singola particella di massa $m = cost.$, la legge fondamentale della dinamica in un riferimento galileiano $\mathcal{R}$ è espressa dalla ben nota 2^a legge di Newton:

$$m\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t), \tag{1.27}$$

dove $\ddot{\mathbf{x}} = \mathbf{a}$ è l'accelerazione della particella e la $\mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t) = \sum_h \mathbf{f}_h$ è la somma di tutte le forze, di qualsiasi tipo, che agiscono su di essa. Inoltre non ci si deve dimenticare che, in natura, non esistono *forze isolate* applicate alle singole particelle ma esse sono sempre dovute all'azione di altre particelle, cioè fanno sempre parte di una **interazione**. Osserviamo che, nell'ambito della Meccanica Classica, le interazioni fra punti materiali generalmente dipendono dalla loro distanza $|\mathbf{x}_2 - \mathbf{x}_1|$ e dalla **differenza delle loro velocità (la cosiddetta velocità relativa)** $\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1$, ed eventualmente anche dal tempo t.
Alla fine del paragrafo precedente abbiamo visto che l'accelerazione $\ddot{\mathbf{x}} = \mathbf{a}$ è invariante nel passaggio da un sistema inerziale $\mathcal{R}$ ad un altro $\mathcal{R}'$, cioè $\ddot{\mathbf{x}} = \ddot{\mathbf{x}}'$. *Ma come si trasformano le forze in un tale cambiamento di riferimento?*
Consideriamo il seguente semplice esempio.
Siano date due particelle mobili $\mathcal{P}_1$ e $\mathcal{P}_2$ che, in un riferimento $\mathcal{R}$, interagiscono con una forza che dipenda dalla loro distanza, dalla loro **velocità relativa** e dal tempo, sia dunque del tipo:

$$\mathbf{f} = \mathbf{f}(|\mathbf{x}_2 - \mathbf{x}_1|, \dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1, t). \tag{1.28}$$

Se applichiamo la trasformazione di Galileo (1.15) alla forza (1.28) si ottiene

$$\mathbf{f} = \mathbf{f}(|\mathbf{x}_2 - \mathbf{x}_1|, \dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1, t) = \mathbf{f}(|\mathbf{x}'_2 - \mathbf{x}'_1|, \dot{\mathbf{x}}'_2 - \dot{\mathbf{x}}'_1, t') = \mathbf{f}', \tag{1.29}$$

con $t = t'$, dato che il tempo è assoluto. Di conseguenza l'equazione di Newton (1.27) applicata alla prima o alla seconda particella ci dà:

$$m_i\ddot{\mathbf{x}}_i = \mathbf{f}_i(|\mathbf{r}|, \dot{\mathbf{r}}, t), \quad (i = 1, 2), \quad \mathbf{f}_2 = -\mathbf{f}_1, \quad \mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1, \quad \dot{\mathbf{r}} = \dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1,$$

e passando da $\mathcal{R}$ ad $\mathcal{R}'$ essa si trasforma per invarianza conservando la stessa forma

$$m_i\ddot{\mathbf{x}}'_i = \mathbf{f}_i(|\mathbf{r}'|, \dot{\mathbf{r}}', t), \qquad (i = 1, 2), \quad \mathbf{r}' = \mathbf{x}'_2 - \mathbf{x}'_1, \quad \dot{\mathbf{r}}' = \dot{\mathbf{x}}'_2 - \dot{\mathbf{x}}'_1,$$

cioè rispetta il principio di relatività galileiano.
È importante rendersi conto che l'invarianza della forza espressa dalla (1.28) nel passaggio dal riferimento inerziale $\mathcal{R}$ all'altro $\mathcal{R}'$ dipende dalla sua struttura matematica, cioè dal fatto che essa è funzione esclusivamente della distanza $|\mathbf{x}_2 - \mathbf{x}_1|$ fra le particelle e della loro velocità relativa $\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1$. Ma non sempre una legge di forza si presenta nella forma (1.28), ci sono leggi di forza che pur essendo funzioni della posizione delle particelle in gioco e delle loro velocità non si trasformano per invarianza nel passare da un sistema inerziale ad un altro, utilizzando una trasformazione di Galileo: una forza di tal tipo è la forza di Lorentz che si manifesta su particelle cariche in un campo elettromagnetico; di essa ci occuperemo nell'ambito della Teoria della Relatività Ristretta. Nel prossimo paragrafo daremo un esempio, anche se un po' artificioso, di forza che non si trasforma per invarianza galileiana.

1.3.8 Il Principio di Relatività come Principio di Simmetria

A questo punto diventa importante chiarire il rapporto logico esistente tra le leggi della dinamica ed il principio di relatività galileiano.[27] Sottolineiamo subito che **il PRG non deriva affatto dalle**

[27] Per le considerazioni che seguono, riguardanti la gerarchia della conoscenza del mondo fisico, abbiamo tratto ispirazione dalla presentazione che ne fa Vincenzo Barone nel suo bel libro *Relatività*, Bollati Boringhieri, pagg. 9-10, Torino (2004).

leggi della dinamica come si potrebbe credere. Non è infatti possibile dedurre, per via logica, il PRG dalla seconda legge di Newton

$$m\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t), \tag{1.30}$$

dove m è la massa, supposta costante, di una data particella, $\mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t)$ è **l'espressione matematica della forza**, o della somma di tutte le forze che agiscono su di essa, $\mathbf{x}$ il suo vettore di posizione, $\dot{\mathbf{x}}$ la sua velocità e $\ddot{\mathbf{x}}$ la sua accelerazione.
Al contrario, è proprio il Principio di Relatività, e cioè la richiesta che le leggi della meccanica siano le stesse in ogni sistema inerziale, che permette di limitare la forma dell'espressione matematica della legge della forza e quindi delle equazioni del moto: tale principio, si badi bene, non determina in modo univoco le equazioni del moto, ma consente di eliminare perlomeno quelle forme che non sono ammissibili proprio in virtù della validità del principio stesso. Nella gerarchia della conoscenza del mondo fisico, il PRG, si situa dunque al disopra delle leggi dinamiche, caratterizzandosi come una **super-legge** che regola le leggi della dinamica: esso costituisce, come si dice, **un principio di simmetria o d'invarianza**, precisamente **il PRG è il principio di simmetria della meccanica newtoniana** o, se si vuole, è **il principio d'invarianza per le trasformazioni di Galileo**.
Per rendersi conto in che modo il PRG agisca come principio di simmetria, supponiamo che due particelle $\mathcal{P}_1, \mathcal{P}_2$, di masse m_1, m_2, interagiscano con una forza posizionale, avente la direzione della retta $\mathcal{P}_1\mathcal{P}_2$ espressa, in un dato sistema galileiano $\mathcal{R}$, dalla seguente legge[28]

$$\mathbf{f}_{12} = k\frac{\mathbf{x}_2 - \mathbf{x}_1}{|\mathbf{x}_2 + \mathbf{x}_1|}, \qquad (\mathbf{f}_{21} = -\mathbf{f}_{12}) \tag{1.31}$$

dove $\mathbf{f}_{12}$ è la forza che la particella $\mathcal{P}_2$ esercita sulla particella $\mathcal{P}_1$, $\mathbf{x}_1$ ed $\mathbf{x}_2$ i loro vettori di posizione e $k = cost. \neq 0$. Nel dato sistema inerziale $\mathcal{R}$, l'equazione newtoniana di moto, per esempio, della particella $\mathcal{P}_1$ è

$$m_1\ddot{\mathbf{x}}_1 = \mathbf{f}_{12} = k\frac{\mathbf{x}_2 - \mathbf{x}_1}{|\mathbf{x}_2 + \mathbf{x}_1|}. \tag{1.32}$$

Si può constatare che in un altro sistema inerziale $\mathcal{R}'$, legato ad $\mathcal{R}$ dalla trasformazione di Galileo $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$ la legge del moto della stessa particella non ha più la medesima forma, si ha infatti

$$m_1\ddot{\mathbf{x}}'_1 = \mathbf{f}'_{12} = k\frac{\mathbf{x}'_2 - \mathbf{x}'_1}{|\mathbf{x}'_2 + \mathbf{x}'_1 + 2\mathbf{v}t|}, \tag{1.33}$$

da cui si vede che la forza $\mathbf{f}'_{12}$ dipende dalla velocità relativa dei due riferimenti inerziali! Perciò la (1.31) non è una legge di forza accettabile e conseguentemente la (1.32) non è una legge del moto ammissibile perché non rispetta il principio di relatività galileiano.
Basta una lieve modifica alla (1.31) e cioè fare in modo che essa dipenda dalla sola distanza $|\mathbf{x}_2 - \mathbf{x}_1|$ fra le due particelle:

$$\mathbf{f}_{12} = k\frac{\mathbf{x}_2 - \mathbf{x}_1}{|\mathbf{x}_2 - \mathbf{x}_1|}, \qquad (\mathbf{f}_{21} = -\mathbf{f}_{12}) \tag{1.34}$$

per renderla una legge che rispetta il principio di relatività galileiano!
Si può verificare facilmente che l'interazione gravitazionale newtoniana fra due particelle è del tipo (1.34) e perciò dà luogo, per es. per la particella $\mathcal{P}_1$ (ma ovviamente anche per la $\mathcal{P}_2$), ad un'equazione di moto

$$m_1\ddot{\mathbf{x}}_1 = Gm_1m_2\frac{\mathbf{x}_2 - \mathbf{x}_1}{|\mathbf{x}_2 - \mathbf{x}_1|^3}, \tag{1.35}$$

[28] L'esempio sopra riportato di una legge di forza del tipo (1.31) è puramente accademico (artificioso) in quanto non si tratta di una reale legge di forza. Le forze utili nelle applicazioni hanno tutte carattere **intrinseco** cioè sono per loro stessa natura invarianti rispetto ad un qualsiasi cambiamento di riferimento non solo inerziale (si veda la successiva **Osservazione 1**).

che rispetta il principio di relatività galileiano, cioè è tale che, applicando la trasformazione di Galileo, essa si trasforma per invarianza

$$m_1\ddot{\mathbf{x}}'_1 = Gm_1m_2\frac{\mathbf{x}'_2 - \mathbf{x}'_1}{|\mathbf{x}'_2 - \mathbf{x}'_1|^3}. \tag{1.36}$$

In generale, leggi di forza dipendenti dalla sola distanza $|\mathbf{x}_2 - \mathbf{x}_1|$ fra due punti, sono invarianti rispetto alle trasformazioni di Galileo, e rispettano quindi il PRG.

Osservazione 1:

A proposito della discussione fatta, in questo paragrafo e nel precedente, sulla **legge di una forza** è opportuno tenere presente quanto segue. Nella meccanica newtoniana, una volta definita la massa e l'accelerazione di una particella, in un dato riferimento $\mathcal{R}$, la nozione di forza viene introdotta per mezzo della definizione

$$\mathbf{f} \stackrel{\text{def}}{=} m\mathbf{a}. \tag{1.37}$$

La (1.37) diviene successivamente **una legge di moto**, non appena sia specificata **la legge di forza**, la quale, nei problemi riconducibili allo schema **punto materiale**, *l'esperienza insegna*, che essa è del tipo:

$$\mathbf{f} = \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t). \tag{1.38}$$

Si tenga presente che una forza come la (1.38) non si presenta mai isolata, ma fa sempre parte di un'interazione fra due o più corpi.

La precedente espressione della legge di forza è in effetti il risultato di un lungo, approfondito e delicato esame delle *leggi sperimentali delle forze*[29] che porta a ritenere che generalmente le forze che si considerano, nell'ambito strettamente meccanico, debbano rispettare il seguente:

Principio d'invarianza della legge di forza:

La legge di forza $\mathbf{f} = \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t)$ *deve avere carattere invariantivo rispetto alla scelta del sistema di coordinate cartesiane, qualunque esso sia, cioè:*

$$f^i(x|\dot{x}|t)\mathbf{e}_i = f^{i'}(x'|\dot{x}'|t)\mathbf{e}_{i'}, \qquad (i, i' = 1, 2, 3) \tag{1.39}$$

dove f^i, $f^{i'}$ *sono le componenti cartesiane di* $\mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t)$ *in due generici riferimenti* $\mathcal{R}$, $\mathcal{R}'$ *dove* $\mathbf{e}_i$, $\mathbf{e}_{i'}$ *sono i rispettivi versori degli assi e* x^j, $x^{j'}$ *le rispettive coordinate.*

Esempi di forze, importanti nelle applicazioni, conformi al precedente principio sono per esempio:

i) **Le forze costanti** $\mathbf{f} = \mathbf{cost.}$;
   come
   1. la forza peso $\mathbf{p} = m\mathbf{g}$;

ii) **Le forze posizionali** $\mathbf{f} = \mathbf{f}(\mathbf{r})$,
   avendo indicato, come spesso si fa, il vettore di posizione con $\mathbf{r}$ anziché con $\mathbf{x}$;
   come
   2. le forze centrali $\mathbf{f} = \varphi(r)\dfrac{\mathbf{r}}{r}$;
   in particolare:
   3. le forze elastiche $\mathbf{f} = -k\mathbf{r}, \quad (k > 0)$;
   4. le forze gravitazionali $\mathbf{f} = G\dfrac{mM}{r^2}\dfrac{\mathbf{r}}{r}, \quad r = |\mathbf{r}|$;
   5. le forze Coulombiane $\mathbf{f} = -k_c\dfrac{q_1q_2}{r^2}\dfrac{\mathbf{r}}{r}, \quad r = |\mathbf{r}|$;
   6. le forze nucleari $\mathbf{f} = k_n\dfrac{e^{-\lambda r}}{r^2}\dfrac{\mathbf{r}}{r}, \quad \lambda > 0, \quad r = |\mathbf{r}|$;

[29] Lo studioso desideroso di approfondire le questioni inerenti alla formulazione, nell'ambito della Meccanica Classica, della *legge matematica delle forze*, può consultare i seguenti volumi: Antonio Signorini, *Elementi di Meccanica Razionale con elementi di Statica Grafica*, Perrella, Roma, 1954, pagg. 20-27; Giorgio Ferrarese e Luigi Stazi, Lezioni di Meccanica Razionale, vol.1, Pitagora Editrice, Bologna, 1989, pagg, 205-215.

iii) **le forze dipendenti dalla sola velocità $\mathbf{f} = \mathbf{f}(\mathbf{v})$;**
come:
7. la forza di Coriolis $\mathbf{f}_{Cor.} = -2m\boldsymbol{\omega} \wedge \mathbf{v}_r$;
8. la resistenza del mezzo $\mathbf{R} = -\alpha A\mu \,\text{vers}\mathbf{v} \quad (\alpha, \mu > 0)$;
in particolare:
9. la resistenza viscosa $\mathbf{R}_{visc.} = -\alpha A\mu\lambda v \,\text{vers}\mathbf{v}, \quad (\alpha, \mu, \lambda > 0; v = |\mathbf{v}|)$;
10. la resistenza idraulica $\mathbf{R}_{idr.} = -\alpha A\mu\lambda v^2 \,\text{vers}\mathbf{v} \quad (\alpha, \mu, \lambda > 0)$;
11. la resistenza balistica $\mathbf{f} = -\varphi(v)\mathbf{v}$;
12. le forze periodiche $\mathbf{f} = \mathbf{A}\sin(\lambda t), \quad \mathbf{A} = \mathbf{cost}$.

iv) **le forze dipendenti esplicitamente soltanto dal tempo $\mathbf{f} = \mathbf{f}(t)$;**
come:
13. forza di Coriolis $\mathbf{f}_{Cor.} = -2m\boldsymbol{\omega}(t) \wedge \mathbf{v}_r$ nel caso che $\mathbf{v}_r = \mathbf{cost}$. e $\boldsymbol{\omega} = \boldsymbol{\omega}(t)$;

14. forza esercitata, su di una particella $\mathcal{P}$, da una persona che, ad ogni istante, modifichi a suo piacimento tale forza secondo una legge prestabilita;

v) **la più generale legge di forza è $\mathbf{f} = \mathbf{f}(OP, \dot{OP}, t) \equiv \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t)$;**
ad esempio:
15. il campo delle forze apparenti nel suo complesso:
$\mathbf{f}_{App.} = -m\mathbf{a}_\tau - m\mathbf{a}_{Cor.} = -m\Big(\mathbf{a}_O + \dot{\boldsymbol{\omega}}(t) \wedge OP + \boldsymbol{\omega}(t) \wedge (\boldsymbol{\omega}(t) \wedge OP)\Big) + 2m\boldsymbol{\omega}(t) \wedge \mathbf{v}_r$.

Ora il Principio di Relatività Galileiano è subordinato al *carattere assoluto dell'accelerazione* nonché alla definizione dinamica della massa e della forza. Pertanto, una volta ammesso che l'espressione matematica della legge di forza rispetti il suo carattere assoluto, cioè rispetti il precedente **Principio d'ivarianza della legge di forza** se questo viene applicato ai sistemi inerziali, cioè a qualsiasi trasformazione galileiana, il PRG appare non già come un postulato, ma come un teorema.
Diverso sarà il suo ruolo logico nella teoria della Relatività Ristretta, in cui sarà ammesso a priori, non già limitato alla sola Meccanica, ma in forma estesa a tutta la Fisica.
Osservazione 2:
Sottolineiamo ancora una volta, che le forze che generalmente si considerano in meccanica, dipendono sia **dalla distanza** fra le particelle in gioco (per esempio: le forze gravitazionali, le forze di interazione elettrica, le forze elastiche, ecc.), sia **dalle velocità relative** (resistenza del mezzo, viscosa, idraulica, ecc.), grandezze che per loro natura hanno carattere assoluto sono cioè invarianti nel passaggio da un riferimento generico ad un altro. Inoltre, se si considera un insieme di particelle interagenti, possiamo includere la meccanica dei corpi rigidi, dei corpi elastici ed anche l'idrodinamica, e così via. Quindi le forze che considera la meccanica rimangono inalterate rispetto alle trasformazioni di Galileo, salvo lo scambio delle coordinate non accentate con le analoghe accentate.
Dobbiamo invece escludere le forze che intervengono nell'elettrodinamica (es. la forza di Lorentz) in quanto l'interazione fra cariche elettriche in movimento (cioè fra cariche e campi magnetici) mette in gioco forze che dipendono non solo dalle posizioni delle cariche ma anche dalle loro velocità, e le cui direzioni non sono lungo la retta congiungente le cariche. Sulla forza di Lorentz torneremo con più dettagli nell'ambito della Relatività Ristretta.
Osserviamo infine che l'equazione fondamentale della dinamica (1.27), scritta in forma vettoriale, è equivalente ad un sistema di tre equazioni differenziali del 2° ordine nelle tre funzioni incognite $x(t), y(t), z(t)$, che, assegnate le condizioni iniziali $[x(t_0), y(t_0), z(t_0); \dot{x}(t_0), \dot{y}(t_0), \dot{z}(t_0)]$, ci fornisce come soluzione uno ed un solo moto. Affinché le soluzioni della (1.27) ci diano lo stesso moto nei due sistemi galileiani $\mathcal{R}$ e $\mathcal{R}'$ è necessario che coincidano le condizioni iniziali.

L'invarianza delle equazioni fondamentali della dinamica assicura che i fenomeni meccanici si svolgano nello stesso identico modo in tutti i sistemi in moto traslatorio rettilineo uniforme gli uni rispetto agli altri, purché, in tali riferimenti, coincidano le condizioni iniziali: se le condizioni iniziali sono diverse si otterranno moti differenti!

1.3.9 II° Gruppo (B) - Assiomi della Meccanica Newtoniana in presenza di Gravitazione

L'occasione venne dalla caduta di una mela,
mentre stava seduto in uno stato d'animo contemplativo....
(William Stukeley, *Memorie della vita di Isaac Newton*)

P9) **Legge di Newton della gravitazione universale**[30]:
Nell'Universo due corpi materiali quali si vogliono P_1, P_2 (supposti puntiformi) si attraggono vicendevolmente con una forza diretta secondo la loro congiungente, di intensità direttamente proporzionale alle loro masse, ed inversamente proporzionale al quadrato della loro distanza r:

$$F = G\frac{m_1 m_2}{r^2}. \tag{1.40}$$

La costante G di attrazione gravitazionale (o costante di Gauss), *misura della mutua attrazione di due masse unitarie all'unità di distanza*, fu per la prima volta determinata con misure di laboratorio dal Cavendish (1797); in seguito ne furono compiute parecchie altre determinazioni con mezzi sempre più precisi, ed esse concordano tutte nell'assegnare a G, in unità C.G.S., il valore numerico

$$G = (6,6726 \pm 0,0008) \times 10^{-8}\,\mathrm{dyn} \cdot \mathrm{cm}^2 \cdot \mathrm{g}^{-2},$$

ed in unità SI il valore

$$G = (6,6726 \pm 0,0008) \times 10^{-11}\,\mathrm{N} \cdot \mathrm{m}^2 \cdot \mathrm{kg}^{-2},$$

tali valori, nell'uno o nell'altro sistema di misura, sono i valori di G adottati nella lista più recente, a nostra conoscenza (20 Maggio 2019), dei migliori valori delle costanti fisiche pubblicate dal NIST (National Institute of Standards and Technology). [31]

Osservazione 1:
Osserviamo che nella legge di gravitazione universale, Newton concepisce le forze che si scambiano i due corpi P_1 e P_2 come **azioni simultanee immediate a distanza**; successivamente però, quando Michael Faraday per primo e James Clerk Maxwell poi, introdussero nella fisica il concetto di **campo elettrico e magnetico**, tale concetto venne esteso anche alle forze newtoniane e si cominciò a parlare di **campo gravitazionale** generato da una o più masse ed agente sempre **istantaneamente e simultaneamente** sulle altre.

[30] Osserviamo che la **Legge della gravitazione universale** è stata formulata da Newton in base all'imponente massa di osservazioni astronomiche effettuate dal famoso Astronomo danese Ticho Brahe (1546-1601), alla corte imperiale di Praga, che portarono il suo allievo Johann Kepler (1571-1630), alla formulazione delle sue tre famosissime leggi:

1. Le orbite dei pianeti sono ellissi ed il Sole ne occupa uno dei fuochi;
2. Le aree descritte dal raggio vettore che va dal Sole ad un pianeta, sono proporzionali ai tempi impiegati a percorrerle;
3. I quadrati dei tempi impiegati dai vari pianeti a percorrere le loro orbite (periodi delle rivoluzioni) sono proporzionali ai cubi dei semiassi maggiori.

[31] Una fonte aggiornata e con tutti i riferimenti bibliografici per ogni costante fisica è il NIST: https://physics.nist.gov/cuu/Units/index.html. Inoltre nel 2018 le unità fondamentali del SI sono cambiate. Si veda http://www.bipm.org/en/CGPM/db/26/1/.

Osservazione 2:
Un'altra osservazione importante è che le forze gravitazionali sono estremamente deboli. Questo fatto è ben noto dallo studio delle orbite planetarie. Invero, le forze gravitazionali sono molto deboli in confronto delle altre tre interazioni fondamentali che agiscono fra particelle: interazioni elettromagnetiche, interazioni forti ed interazioni deboli. Per esempio, se calcoliamo il rapporto fra l'interazione gravitazionale e l'interazione elettrica fra due elettroni di carica e e massa m_e posti ad una distanza arbitraria r l'uno dall'altro, avremo

$$\frac{F_{gravit.}}{F_{elettr.}} = \frac{Gm_e^2}{r^2} : \frac{ke^2}{r^2} = \frac{G}{k}\frac{m_e^2}{e^2}$$

essendo, nel sistema SI[32]

$$\begin{aligned} G &= 6,6726 \times 10^{-11}\,\mathrm{N}\cdot\mathrm{m}^2\cdot\mathrm{kg}^{-2}, \\ k &= \frac{1}{4\pi\epsilon_0} = 8,99 \times 10^9 \mathrm{N}\cdot\mathrm{m}^2\cdot\mathrm{C}^{-2}, \\ m_e &= 9,11 \times 10^{-31}\mathrm{kg}, \\ e &= 1,60 \times 10^{-19}\mathrm{C}, \end{aligned}$$

dove k è la costante dielettrica e C = Coulomb l'unità di misura della carica elettrica.
Si ottiene allora:

$$\frac{F_{gravit.}}{F_{elettr.}} = \frac{6,6726 \times 10^{-11}\,\mathrm{N}\cdot\mathrm{m}^2\cdot\mathrm{kg}^{-2}}{8,99 \times 10^9\mathrm{N}\cdot\mathrm{m}^2\cdot\mathrm{C}^{-2}}\frac{(9,11 \times 10^{-31}\mathrm{kg})^2}{(1,60 \times 10^{-19}\mathrm{C})^2} = 0,24 \times 10^{-42}.$$

Tutti gli altri campi che conosciamo sono molto più intensi del campo gravitazionale!
Osserviamo che, nei trattati classici di Meccanica Celeste, la costante di gravitazione è generalmente indicata con f, inoltre alcuni autori la indicano anche con k o k^2.

Com'è noto, gli esperimenti fatti da Galileo, dalla torre pendente di Pisa, sulla caduta dei gravi hanno messo in luce la seguente:

P10 **Proprietà fondamentale del campo gravitazionale:**
Oggetti aventi differenti masse inerziali, a parità di condizioni iniziali, cadono tutti con la medesima accelerazione g derivante dall'attrazione gravitazionale della Terra, accelerazione che non dipende minimamente né dalla costituzione materiale dei corpi, né dal loro stato fisico.

Questa notevole proprietà dei corpi in caduta libera in un campo gravitazionale implica l'equivalenza fra *massa inerziale* e *massa gravitazionale* di cui ci occuperemo in dettaglio nel prossimo paragrafo.

1.3.10 La massa nella teoria newtoniana

È opportuno, prima di proseguire nelle nostre discussioni, di riflettere sul concetto fondamentale di *"massa"* per quel che riguarda le sue *proprietà inerziali e gravitazionali* così come vengono messe in luce dalla teoria newtoniana. Infatti nella teoria newtoniana il concetto di massa di un corpo viene usato con tre significati differenti:

1) *la massa inerziale* (m^I): misura la resistenza che hanno i corpi al cambiamento del loro stato di quiete o di moto (la loro inerzia appunto), essa è la massa che compare nella 2^a legge della dinamica $\mathbf{F} = m^I\mathbf{a}$;

[32]Si veda Halliday-Resnick-Krane, *Fisica 2*, pag. 639 ed Appendice B, Quarta edizione 1994.

2) *la massa gravitazionale passiva* (m^{GP}): misura la reazione di un corpo all'azione di un campo gravitazionale esterno;

3) *la massa gravitazionale attiva* (m^{GA}): misura la capacità di un corpo di produrre un campo gravitazionale che agisce attivamente su altri corpi.

Le masse gravitazionali attiva e passiva intervengono ambedue nella legge di gravitazione universale sia che ci si riferisca ad un corpo o all'altro: se si pensa all'attrazione che subisce il corpo P_1 da parte del corpo P_2 la massa gravitazionale del primo corpo si comporta in modo passivo, mentre la massa gravitazionale del secondo agisce in modo attivo e viceversa scambiando il ruolo dei due corpi.
Dimostriamo dapprima che massa gravitazionale passiva e massa gravitazionale attiva sono la stessa cosa cioè che esiste una sola massa gravitazionale e poi faremo vedere che essa si identifica anche con la massa inerziale, sicché *il concetto di massa è unico*.
Partiamo dalla constatazione sperimentale che *nulla può essere sottratto all'azione di un campo gravitazionale ("attrazione fatale")* o equivalentemente che vale la **Legge newtoniana della gravitazione universale** che abbiamo enunciato nel paragrafo precedente. Più precisamente la forza $\mathbf{F}_1$ che il corpo P_2 esercita sul corpo P_1 è:

$$\mathbf{F}_1 = G\frac{m_1^{GP} m_2^{GA}}{r^2}\mathbf{u},$$

dove $\mathbf{u} = \text{vers}\,\overrightarrow{P_1P_2}$ mentre la forza $\mathbf{F}_2$ che il corpo P_1 esercita sul corpo P_2 è:

$$\mathbf{F}_2 = -G\frac{m_1^{GA} m_2^{GP}}{r^2}\mathbf{u}.$$

Per la 3ª legge di Newton *(Principio di azione e reazione)* si ha $\mathbf{F}_1 = -\mathbf{F}_2$, avremo perciò uguagliando i moduli delle due forze:

$$G\frac{m_1^{GP} m_2^{GA}}{r^2} = G\frac{m_1^{GA} m_2^{GP}}{r^2}$$

da cui si deduce che:

$$\frac{m_1^{GP}}{m_1^{GA}} = \frac{m_2^{GP}}{m_2^{GA}} = costante\,(uguale\,per\,tutti\,i\,corpi).$$

Cioè il rapporto fra la massa gravitazionale passiva e la massa gravitazionale attiva è lo stesso per ogni corpo, cioè è invariante nel passare da un corpo ad un altro! Scegliendo opportunamente le unità di misura si può fare in modo che la costante di proporzionalità sia uguale ad 1, cioè:

$$m^{GP} = m^{GA}.$$

Per ogni corpo, esiste perciò una sola massa gravitazionale che indichiamo con m^G.
Facciamo ora vedere che anche la massa inerziale m^I di un qualsiasi corpo è uguale alla sua massa gravitazionale m^G. Notiamo che, a priori, *la massa inerziale m^I non ha nulla a che fare direttamente con la gravitazione*. Per esempio, se proviamo a spingere un blocco di ghiaccio, qui sulla Terra, inizialmente fermo sulla superficie di un piano orizzontale, privo di attrito, ci accorgiamo che ciò richiede sforzo. Il blocco sembra essere inerte e tende a rimanere nella posizione di riposo, o, se sta muovendosi, tende a mantenere il movimento. La gravità qui non entra per nulla. Si richiederebbe lo stesso sforzo per accelerare il blocco di ghiaccio in uno spazio privo di gravità. È la massa del blocco (concepita come quantità di materia) che rende

necessario esercitare una forza per cambiare il suo stato di quiete o di moto: è la massa che compare nella 2^a legge di Newton

$$\mathbf{F} = m^I \mathbf{a}$$

e che chiamiamo *massa inerziale*.
C'è una situazione diversa in cui interviene nuovamente la massa del blocco. Se non lo sosteniamo il blocco cadrà a terra con moto accelerato. La forza necessaria per tenere il blocco è uguale in grandezza alla forza di attrazione gravitazionale tra esso e la Terra. Qui l'inerzia non gioca alcun ruolo; è la proprietà dei corpi materiali di essere attratti verso altri oggetti come la Terra, che importa. Il modulo della forza è dato da

$$F = G\frac{m^G m_T^G}{r^2}$$

dove m^G è la *massa gravitazionale del blocco*, e m_T^G la *massa gravitazionale della Terra*.
La *massa gravitazionale* m^G e la *massa inerziale* m^I del blocco sono, oppure no, la stessa cosa?

Per rispondere a tale domanda, consideriamo due corpi puntiformi P_1, P_2 di masse inerziali m_1^I, m_2^I e masse gravitazionali m_1^G, m_2^G e facciamoli cadere dalla stessa altezza nel campo gravitazionale della Terra, si avrà per la seconda legge di Newton

$$F_1 = m_1^I a_1, \qquad F_2 = m_2^I a_2,$$

ma in virtù della legge fondamentale del campo gravitazionale $a_1 = a_2 = g$. Per la legge della gravitazione universale si ha invece

$$F_1 = G\frac{m_1^G m_T^G}{r^2}, \qquad F_2 = G\frac{m_2^G m_T^G}{r^2}.$$

Dalle relazioni precedenti, dividendole membro a membro, avremo rispettivamente:

$$\frac{F_1}{F_2} = \frac{m_1^I}{m_2^I}, \qquad \frac{F_1}{F_2} = \frac{m_1^G}{m_2^G}$$

da cui otteniamo

$$\frac{m_1^I}{m_1^G} = \frac{m_2^I}{m_2^G} = costante.$$

Perciò le masse inerziali e le masse gravitazionali dei corpi sono proporzionali, di fatto esse risultano uguali. Infatti si possono scegliere le unità di misura in modo tale che il fattore di proporzionalità sia uguale ad 1.
Tuttavia è possibile dimostrarne l'uguaglianza attraverso un esperimento progettato da Newton e consistente nella misura del periodo di oscillazione di un pendolo. Si dimostra infatti teoricamente, utilizzando le equazioni intrinseche del moto di un pendolo su traiettoria prestabilita, che il periodo di oscillazione del pendolo è dato da

$$T = 2\pi\sqrt{\frac{m^I}{m^G}\frac{l}{g}}$$

D'altra parte, sperimentalmente si può constatare che è:

$$T = 2\pi\sqrt{\frac{l}{g}},$$

e ciò è possibile soltanto se

$$m^I = m^G.$$

abbiamo allora la seguente legge fondamentale:

P11 **Equivalenza fra massa inerziale e massa gravitazionale:**
Per qualsiasi corpo si ha:

$$\textbf{massa inerziale} = \textbf{massa gravitazionale}.$$

Dunque esiste una sola massa che, nel seguito, indicheremo semplicemente con m.
A proposito di questa conclusione Einstein dice[33]: *".....Questa importante legge è stata registrata dalla meccanica classica ma non interpretata. Un'interpretazione soddisfacente si può avere soltanto se riconosciamo il fatto seguente*: **la stessa qualità di un corpo si manifesta a seconda delle circostanze come inerzia o come pesantezza**."
La vera interpretazione di questa legge è stata data, per la prima volta, proprio da Einstein, nell'ambito della sua teoria della Relatività Generale, col fondamentale **Principio di equivalenza**, come vedremo in seguito.
Ci rimane ora da illustrare la fondamentale **equazione di Poisson** che costituisce l'equazione (alle derivate parziali) del campo gravitazionale descritto dal cosiddetto **potenziale newtoniano.** Per far ciò è necessario definire quest'ultimo concetto di cui ci occuperemo, nel modo più breve possibile, nel seguente paragrafo.

1.3.11 Digressione sulla Teoria del Potenziale Newtoniano

La Teoria del Potenziale Newtoniano[34] è un capitolo importantissimo ma abbastanza difficile della Fisica Matematica classica e ad essa sono dedicati interi volumi. Qui, noi ci limiteremo ad un cenno fugace, il più semplice possibile, ma avvertiamo il lettore che, come diceva il grande matematico italiano Francesco Severi, *"il riassumere è sempre pieno d'insidie, onde tenerne il debito conto!"*
La Teoria del Potenziale Newtoniano è fondata sulla **legge della gravitazione universale di Newton** la quale esprime, come abbiamo visto, che due punti materiali di masse M, m, posti alla distanza r, esercitano l'uno sull'altro un'attrazione di intensità

$$F = G\frac{Mm}{r^2}. \tag{1.41}$$

Se indichiamo con Q, P i due punti di masse M, m, la forza di attrazione che il punto Q (*punto potenziante*) esercita sul punto P (*punto potenziato*), espressa vettorialmente, sarà

$$\mathbf{F} = -G\frac{Mm}{r^2}\text{vers}\,\mathbf{r} \tag{1.42}$$

essendo $\text{vers}\,\mathbf{r} = \dfrac{\mathbf{r}}{|\mathbf{r}|} = \dfrac{\mathbf{r}}{r}$, il versore del vettore $\mathbf{r} = \overrightarrow{QP} = P - Q$ di posizione di P rispetto a Q. Tale forza è diretta da P a Q, cioè orientata come il vettore $\overrightarrow{PQ} = -\overrightarrow{QP}$, ciò che giustifica la presenza del segno $-$ al secondo membro della (1.42). Naturalmente, le veci del punto potenziante e di quello potenziato possono essere ovviamente scambiate fra di loro (Fig.1.5).

Se nel punto P poniamo una **massa unitaria** $(m = 1)$, l'attrazione che esercita il punto Q di massa M sul punto P di massa unitaria, risulta

$$\mathcal{G}(P) = -G\frac{M}{r^2}\text{vers}\,\mathbf{r}, \tag{1.43}$$

[33] A. Einstein, *Relatività, Esposizione divulgativa*, Universale Scientifica Boringhieri (1967), pagg. 95-98.
[34] Per un'esposizione più estesa ed approfondita della teoria del potenziale newtoniano si veda il seguente volume: Cataldo Agostinelli, *Istituzioni di Fisica Matematica*, vol. I, Parte II, Cap.II, pagg. 149-364, Zanichelli Editore Bologna, 1962. Un'ottima esposizione, molto più succinta della precedente, si può trovare anche in Tullio Levi-Civita e Ugo Amaldi, *Lezioni di Meccanica Razionale*, vol.1, Cap.XI, pagg. 489-532, Zanichelli Editore, Bologna, 1974.

Figura 1.5: Campo gravitazionale generato da una massa puntiforme Q (punto potenziante).

e se teniamo fisso *il punto potenziante* Q e facciamo variare *il punto potenziato* P, il vettore $\boldsymbol{\mathcal{G}}(P)$, funzione di P, definisce in tutto lo spazio un campo vettoriale che chiameremo *il campo delle forze di attrazione newtoniane o (***campo gravitazionale***) generato dal punto Q (corpo centrale) di massa M*. Il modulo $|\boldsymbol{\mathcal{G}}(P)|$ del vettore $\boldsymbol{\mathcal{G}}(P)$ è **l'intensità del campo gravitazionale** generato dalla massa M.
Val la pena di osservare che, poiché la forza di attrazione fra due masse puntiformi è proporzionale alle masse, se nel punto P variabile poniamo, anziché la massa unitaria, una massa arbitraria m avremo che la forza di attrazione in P sarà data da

$$\mathbf{F} = m\,\boldsymbol{\mathcal{G}}, \tag{1.44}$$

e questa mostra che *il campo gravitazionale non è propriamente una forza, ma il rapporto tra una forza ed una massa*, avente perciò le dimensioni di una accelerazione $[LT^{-2}]$.
Possiamo ora far vedere facilmente che il campo gravitazionale $\boldsymbol{\mathcal{G}}$ espresso dalla (1.43) **è conservativo**, deriva cioè da un **potenziale**. Infatti la (1.43), essendo G ed M delle costanti, tenendo conto di note proprietà del **gradiente di una funzione**, possiamo trasformarla nel modo seguente[35]

$$\boldsymbol{\mathcal{G}}(P) = GM\Big(-\frac{1}{r^2}\text{vers}\,\mathbf{r}\Big) = GM\Big(-\frac{1}{r^2}\text{grad}_P r\Big) = GM\text{grad}_P\frac{1}{r} = \text{grad}_P\frac{GM}{r}, \tag{1.45}$$

nella quale si è tenuto conto che, tenendo fisso Q e facendo variare P si ha

$$\text{vers}\,\mathbf{r} = \text{grad}_P\, r,$$

ed inoltre è facile verificare che vale l'identità

$$\text{grad}_P\,\frac{1}{r} = -\frac{1}{r^2}\text{grad}_P\, r.$$

Infine, la (1.45) può essere riscritta nella forma

$$\boldsymbol{\mathcal{G}}(P) = -\text{grad}_P\Big(-\frac{GM}{r}\Big), \tag{1.46}$$

e questa mostra che:
Prop.1: *Il campo gravitazionale generato da una massa puntiforme M posta nel punto fisso Q deriva dal potenziale*[36]

$$\Phi(P) = -\frac{GM}{r}, \tag{1.47}$$

[35]Ricordiamo che il **gradiente di una funzione** di tre variabili $f(x,y,z)$ è un vettore le cui componenti cartesiane sono le derivate parziali di f rispetto alle variabili x,y,z, cioè $\text{grad}\, f \equiv \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}\right)$; inoltre con $\text{grad}_P r$ indichiamo il gradiente calcolato nel punto $P \equiv (x,y,z)$ della funzione $r = r(P,Q) = \sqrt{(x-x_Q)^2 + (y-y_Q)^2 + (z-z_Q)^2}$ nella quale si tiene fisso il punto Q, che potrebbe essere assunto come origine del sistema di coordinate, e variabile P.

[36]A proposito del **potenziale gravitazionale**, dobbiamo osservare che la definizione da noi data più sopra è quella universalmente adottata dai fisici, che considera il potenziale negativo in analogia con la definizione di **potenziale del campo elettrico** che si dà in elettrostatica. Ma è opportuno osservare che i matematici, gli astronomi e gli studiosi di Meccanica Celeste, generalmente definiscono il potenziale gravitazionale col segno positivo: $U = \frac{GM}{r}$, nel qual caso **il campo gravitazionale** sarà dato dall'espressione $\boldsymbol{\mathcal{G}} = \text{grad}\, U$. Le due definizioni sono ovviamente entrambe legittime e scegliere l'una piuttosto che l'altra è soltanto una questione di convenzione.

che chiameremo **potenziale gravitazionale.**
La (1.46) si scriverà perciò nella forma

$$\mathcal{G}(P) = -\text{grad}_P\,\Phi. \tag{1.48}$$

Ciò che abbiamo detto nel caso di un solo punto potenziante si generalizza al caso di più punti ed al caso, ancora più generale, di una distribuzione continua di masse con densità $\mu(Q) = \mu(x, y, z)$ in un volume S, in quest'ultimo caso il **campo gravitazionale**, in un punto P **esterno** ad S, è definito dalla:

$$\mathcal{G}(P) \stackrel{\text{def}}{=} -G \int_S \frac{\mu(Q)}{r^2} \frac{P-Q}{r} dS = G \int_S \text{grad}_P \frac{\mu(Q)}{r(P,Q)} dS, \qquad (P \notin S) \tag{1.49}$$

e dato che l'operazione d'integrazione è invertibile con l'operatore gradiente possiamo riscrivere così:

$$\mathcal{G}(P) = \text{grad}_P \left[G \int_S \frac{\mu(Q)}{r(P,Q)} dS \right] = -\text{grad}_P \left[-G \int_S \frac{\mu(Q)}{r(P,Q)} dS \right]. \tag{1.50}$$

Perciò possiamo definire **il potenziale del campo gravitazionale** generato da una distribuzione continua di masse in un volume S, in un punto **esterno** ad S, col seguente integrale triplo:

$$\Phi(P) \stackrel{\text{def}}{=} -G \int_S \frac{\mu(Q)}{r(P,Q)} dS, \qquad (P \notin S) \tag{1.51}$$

in cui l'integrazione è fatta rispetto al punto Q variabile in S e dove P funge da parametro (Fig.1.6).

Figura 1.6: Campo gravitazionale generato da una distribuzione continua di materia.

Quando il punto P è **interno** al volume S la definizione del campo gravitazionale e del corrispondente potenziale è estremamente delicata ed è necessaria un'analisi minuziosa e rigorosa della questione che qui, in questi pochi cenni, non possiamo assolutamente trattare; quel che possiamo dire è che vale il seguente teorema:
Teorema 1: *Se la densità di massa $\mu(Q)$ è* **continua nel dominio** S **sino alla superficie** σ **che lo delimita**, *allora anche nei* **punti interni al corpo** *si può definire sia il* **il campo gravitazionale** *che* **il potenziale** *ed essi hanno la medesima espressione che si ha all'esterno, cioè:*

$$\mathcal{G}(P) = G \int_S \text{grad}_P \frac{\mu(Q)}{r} dS = -\text{grad}_P\,\Phi,$$

$$\Phi(P) = -G \int_S \frac{\mu(Q)}{r(Q,P)} dS, \tag{1.52}$$

Avremo allora, come nel caso discreto, che **il campo gravitazionale** in P è dato da

$$\mathcal{G}(P) = -\text{grad}_P\,\Phi. \tag{1.53}$$

In conclusione, tenuto conto delle precedenti considerazioni, si può dimostrare che dalla legge di gravitazione universale è possibile dedurre **l'equazione di campo della teoria gravitazionale di Newton** compendiata nel seguente fondamentale:

P12) **Teorema di Poisson:**
Il potenziale gravitazionale $\Phi(P)$ *di una distribuzione continua di masse in un volume S, la cui densità* $\mu(x, y, z)$ *sia continua con derivate parziali del 1° ordine continue all'interno di S, ammette in S derivate parziali prime e seconde continue. Inoltre, all'esterno della massa attirante, il potenziale* $\Phi(P)$ *soddisfa all'equazione di LAPLACE:*

$$\Delta_2\Phi(P) = 0, \quad (\forall P \notin S), \tag{1.54}$$

e all'interno all'equazione di POISSON:

$$\Delta_2\Phi(P) = 4\pi G\mu(P), \quad (\forall P \in S). \tag{1.55}$$

dove $\Delta_2 \equiv \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}$ *è* **l'operatore di Laplace**, G **la costante newtoniana di gravitazione** *e* $\mu(P) = \mu(x, y, z)$ **la densità di massa** *in* S.

Se si osserva che all'esterno della massa attirante la densità è nulla, cioè $\mu(P) = \mu(x, y, z) = 0$, la prima delle precedenti equazioni è contenuta nella seconda (1.55) che costituisce **l'equazione di campo della teoria newtoniana della gravitazione**.
Le due equazioni (1.54) e (1.55) sono equazioni alle derivate parziali lineari del 2° ordine nella funzione incognita $\Phi(x, y, z)$. **Le soluzioni** dell'equazione di Laplace (1.54) sono, com'è noto, le cosiddette **funzioni armoniche**.
Si dimostra che una **soluzione particolare** dell'equazione di Poisson è data proprio dall'integrale triplo che definisce il potenziale:

$$\Phi(P) = -G\int_S \frac{\mu(Q)}{r(P, Q)}dS.$$

Per ottenere la **soluzione generale** dell'equazione (1.55) occorre aggiungere ad essa una qualunque funzione armonica.
Osservazione 1ª:
Il segno + nel secondo membro dell'equazione di Poisson (1.55) dipende da due convenzioni:

1) il potenziale Φ è legato **all'intensità** $\mathcal{G}(P)$ **del campo gravitazionale** dalla $\mathcal{G}(P) = -\text{grad}_P\Phi$ (come fa la generalità dei fisici in analogia con la legge di Coulomb dell'elettrostatica);

2) la normale n alla superficie σ che racchiude il volume S è orientata verso l'esterno di S, cioè $n \equiv n_e$.

Se l'intensità del campo gravitazionale viene invece definita col segno opposto di quello precedente $\mathcal{G}(P) = \text{grad}_P\Phi$ (come fanno la maggior parte dei matematici e degli astronomi) e si mantiene la convenzione 2) sull'orientamento della normale a σ, allora il 2° membro dell'equazione di Poisson cambia segno, si avrebbe dunque $\Delta_2\Phi = -4\pi G\mu(P)$. Talvolta, alcuni autori, anziché considerare la normale esterna n_e a σ, introducono invece la normale interna n_i che ha orientamento opposto alla normale esterna, e ciò comporta ancora un cambiamento del segno al 2° membro dell'equazione di Poisson. Bisogna dunque prestare un po' di attenzione a queste convenzioni se si vuole scrivere correttamente il segno che deve competere al 2° membro dell'equazione di Poisson.
Osservazione 2ª:
Osserviamo che le due leggi VI° e VII° (cioè rispettivamente la legge della gravitazione di Newton e la proprietà fondamentale del campo gravitazionale di Galileo) non intervengono nella Teoria della Relatività Ristretta (RR) in quanto si tratta di una teoria costruita volutamente in assenza di gravitazione.

L'equazione di Poisson (1.55) ha invece un ruolo importantissimo, nell'ambito della Relatività Generale (RG) perché è servita da guida ad Einstein per giungere ad una formulazione corretta delle equazioni gravitazionali (equazioni di campo della RG), in modo tale da essere contenuta, in un'approssimazione opportuna, detta limite newtoniano, nelle più generali equazioni di campo einsteiniane.
Esempio importante:
Potenziale e campo gravitazionale generato da una sfera piena omogenea o formata da strati sferici omogenei concentrici
Uno degli esempi più importanti per le applicazioni, specialmente in Astronomia, Astrofisica e Meccanica Celeste ed anche per la trattazione del fondamentale problema di Schwarzschild, che studieremo a suo tempo, è la determinazione sia del **potenziale che del campo gravitazionale** di una massa sferica omogenea, o formata da strati sferici omogenei concentrici, come è il caso del nostro Sole o di una stella qualsiasi.
Se R è il raggio della sfera, $\mu = cost.$ la sua densità, $M = \mu \frac{4}{3}\pi R^3$ la sua massa, se inoltre, indichiamo con r la distanza di un generico punto potenziato P della sfera dal suo centro O, si dimostra che [37]

$$\Phi = \begin{cases} -\dfrac{GM}{r}, & nei\ punti\ esterni\ (r > R), \\ -\dfrac{GM}{R}\left[\dfrac{3}{2} - \dfrac{1}{2}\dfrac{r^2}{R^2}\right] & nei\ punti\ interni\ (0 \leq r < R), \\ \\ valendo\ indifferentemente\ entrambe\ le\ espressioni\ per\ r = R. \end{cases}$$

Abbiamo perciò l'importante risultato utile nelle applicazioni:
Il potenziale gravitazionale di una sfera omogenea (o formata da strati sferici omogenei concentrici), in un punto P esterno ad essa, ed alla distanza r dal suo centro, è

$$\Phi(P) = -\frac{GM}{r}, \tag{1.56}$$

cioè è lo stesso che si avrebbe se tutta la sua massa fosse concentrata nel centro O della sfera.
L'attrazione gravitazionale è in ogni caso diretta verso il centro e la sua intensità $|\mathcal{G}| = |-d\Phi/dr|$ è data da GM/r^2 nei punti esterni, come se tutta la massa fosse concentrata nel centro; da $\frac{4}{3}\pi G\mu r = \frac{GM}{R^3}r$ nei punti interni, risultando proporzionale alla distanza dal centro. Sulla superficie della sfera $(r = R)$ si ha il limite comune GM/R^2..

1.4 III° Gruppo - Assiomi dell'elettromagnetismo

1.4.1 Carica elettrica, campo elettrico, campo magnetico, densità di carica e densità di corrente

Anche l'elettromagnetismo, come la meccanica, introduce nuovi concetti primitivi.[38] Questi sono essenzialmente la **carica elettrica** q, il **campo elettrico** $\mathbf{E}(\mathbf{x}, t)$ e il **campo magnetico** $\mathbf{H}(\mathbf{x}, t)$. Concetti derivati sono invece la **densità di carica** $\rho(\mathbf{x}, t)$ e la **densità di corrente elettrica** $\mathbf{j}(\mathbf{x}, t)$,

[37] Si veda: T. Levi-Civita e U. Amaldi, *Lezioni di Meccanica Razionale*, vol.1, pagg. 512-513, Zanichelli, Bologna (1974).

[38] Per i concetti fondamentali sull'elettromagnetismo si può consultare un qualunque testo di *Fisica 2*, per esempio l'ottimo libro di: Edoardo Amaldi, Romano Bizzarri, Guido Pizzella, *Fisica Generale, Elettromagnetismo, Relatività, Ottica*, Zanichelli Editore, Bologna, 1986; oppure il molto diffuso: D. Halliday, R. Resnick, K. S. Krane, *Fisica 2*, Casa Editrice Ambrosiana, Milano (1994).

dove $\mathbf{x}$ è il vettore di posizione rispetto al riferimento adottato, e t è la variabile temporale, il che significa che tali grandezze si considerano generalmente variabili da punto a punto ed eventualmente variabili anche col tempo.
Ricordiamo che le grandezze $\mathbf{E}$ ed $\mathbf{H}$ sono definite come la forza agente sulla carica elettrica unitaria e sul polo magnetico unitario, rispettivamente, in un taglio di spessore piccolissimo, praticato nel mezzo materiale in cui i campi agiscono, in direzione parallela alle linee di forza del campo. Supporremo che i mezzi in cui $\mathbf{E}$ ed $\mathbf{H}$ sono definiti *siano isotropi, non ferromagnetici né ferroelettrici*; così se indichiamo con ϵ e con μ la **costante dielettrica** e la **permeabilità magnetica** del mezzo[39], **l'induzione elettrica**, $\mathbf{D}$, e **l'induzione magnetica**, $\mathbf{B}$, sono legate ad $\mathbf{E}$ ed $\mathbf{H}$ dalle relazioni

$$\mathbf{D} = \epsilon \mathbf{E}, \qquad \mathbf{B} = \mu \mathbf{H}, \tag{1.57}$$

ϵ e μ essendo grandezze non dipendenti da $\mathbf{E}$ e da $\mathbf{H}$.
A queste equazioni dobbiamo aggiungere la relazione tra $\mathbf{E}$ e $\mathbf{j}$. Se σ è la **conduttività** del mezzo, cioè l'inverso della **resistività**, tale relazione è

$$\mathbf{j} = \sigma \mathbf{E}. \qquad (\textit{Legge di Ohm}) \tag{1.58}$$

Ricordiamo che per le sostanze conduttrici la conduttività σ è relativamente grande, mentre per sostanze isolanti essa è relativamente piccola.
In generale le caratteristiche del mezzo materiale cambiano da punto a punto e cambiano nel tempo in conseguenza della sua evoluzione; quindi ϵ, μ e σ sono funzioni di $\mathbf{x}$ e di t, cioè: $\epsilon(\mathbf{x}, t)$, $\mu(\mathbf{x}, t)$ e $\sigma(\mathbf{x}, t)$ e perciò non ci si deve ingannare quando per ϵ si parla di costante dielettrica.

1.4.2 Equazioni di Maxwell nel sistema di Gauss

Conveniamo di svolgere le nostre considerazioni con riferimento ad una terna cartesiana $\mathcal{T}^* \equiv (O^*x^*y^*z^*)$ **solidale allo spazio delle stelle fisse** Σ^*, ed al tempo assoluto t (cioè nel *riferimento inerziale privilegiato* $\mathcal{R}^* \equiv (O^*x^*y^*z^*, t)$, noto anche come *sistema di riferimento dell'etere*. Se allora, in tale riferimento, supponiamo di misurare $\mathbf{E}$, $\mathbf{D}$ in unità del sistema CGS elettrostatico, $\mathbf{j}$, $\mathbf{B}$, $\mathbf{H}$ in unità del sistema CGS elettromagnetico ed indichiamo con

$$c = \frac{u(CGS)_{es}}{u(CGS)_{em}} \tag{1.59}$$

il rapporto fra le corrispondenti unità di misura nei due sistemi, le **equazioni di Maxwell**, nelle quali sostanzialmente si compendia tutto l'elettromagnetismo, sono, in forma vettoriale:

Equazioni di Maxwell (nel sistema di unità di misura di Gauss):

$$\begin{cases} \text{I.} \quad \operatorname{div} \mathbf{D} \equiv \nabla \cdot \mathbf{D} = 4\pi\rho, & (\textit{Teorema di Gauss}) \\ \text{II.} \quad \operatorname{div} \mathbf{B} \equiv \nabla \cdot \mathbf{B} = 0, & (\textit{Inesistenza di poli magnetici isolati}) \\ \text{III.} \quad \operatorname{rot} \mathbf{E} \equiv \nabla \wedge \mathbf{E} = -\dfrac{1}{c}\dfrac{\partial \mathbf{B}}{\partial t}, & (\textit{Legge di Faraday} - \textit{Neumann} - \textit{Lenz}) \\ \text{IV.} \quad \operatorname{rot} \mathbf{H} \equiv \nabla \wedge \mathbf{H} = \dfrac{4\pi}{c}\mathbf{j} + \dfrac{1}{c}\dfrac{\partial \mathbf{D}}{\partial t} & (\textit{Legge di Ampère} - \textit{Maxwell}). \end{cases} \tag{1.60}$$

[39] Si ricordi, a scanso di equivoci, che anche la densità di massa è stata indicata con la stessa lettera μ della permeabilità magnetica, ma dal contesto in cui tale lettera viene usata non c'è mai pericolo di confusione.

Le equazioni (1.60) sono chiamate **equazioni di Maxwell propriamente dette** e le equazioni (1.57) e (1.58) **relazioni costitutive**. Le equazioni di Maxwell propriamente dette hanno tutte un importante significato fisico che è quello delle leggi che ciascuna di esse esprime, e che abbiamo scritto accanto a ciascuna di esse.
Le equazioni di Maxwell, dal punto di vista matematico, costituiscono un sistema di equazioni alle derivate parziali del 1° ordine che per essere risolto completamente necessita di opportune condizioni al contorno.
Nel caso limite di un **mezzo omogeneo indefinitamente esteso** le quantità $\epsilon(\mathbf{x}, t)$, $\mu(\mathbf{x}, t)$ e $\sigma(\mathbf{x}, t)$ si possono tuttavia supporre costanti, nel qual caso, le equazioni di Maxwell (1.60) assumono la forma particolarmente semplice

Equazioni di Maxwell semplificate:

$$\begin{cases} \text{I.} \quad \operatorname{div} \mathbf{E} = \dfrac{4\pi}{\epsilon}\rho, & (\textit{Teorema di Gauss}) \\ \text{II.} \quad \operatorname{div} \mathbf{H} = 0, & (\textit{Inesistenza di poli magnetici isolati}) \\ \text{III.} \quad \operatorname{rot} \mathbf{E} + \dfrac{\mu}{c}\dfrac{\partial \mathbf{H}}{\partial t} = 0, & (\textit{Legge di Faraday} - \textit{Neumann} - \textit{Lenz}) \\ \text{IV.} \quad \operatorname{rot} \mathbf{H} - \dfrac{\epsilon}{c}\dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{4\pi}{c}\mathbf{j} & (\textit{Legge di Ampère} - \textit{Maxwell}). \end{cases} \tag{1.61}$$

Equazioni del tipo (1.61) continuano a valere per un mezzo non omogeneo se si ammette che ϵ, μ e σ varino con $\mathbf{x}$ e t così lentamente che le loro derivate siano trascurabili rispetto a quelle dei campi.

1.4.3 Forza di Lorentz, conservazione della carica, energia elettromagnetica e vettore di Poynting

Per completare il quadro delle equazioni di Maxwell, è necessario aggiungere un'altra importantissima relazione che costituisce, per così dire, **l'anello di congiunzione tra elettromagnetismo e meccanica**, la cosiddetta:

V. **Forza di Lorentz**:
Se, rispetto al sistema di riferimento inerziale $\mathcal{R}^*$, una particella di carica q si muove in un campo elettromagnetico con velocità $\mathbf{u}$, essa è sottoposta all'azione di una forza detta forza di Lorentz

$$\mathbf{F} = q\left(\mathbf{E} + \frac{1}{c}\mathbf{u} \wedge \mathbf{B}\right). \tag{1.62}$$

Le equazioni di Maxwell implicano inoltre **l'equazione di continuità** che stabilisce un legame fra $\rho(\mathbf{x}, t)$ e $\mathbf{j}(\mathbf{x}, t)$ ed esprime matematicamente **la legge di conservazione della carica**.
Infatti, derivando membro a membro la prima delle (1.60) rispetto al tempo si ottiene

$$\frac{\partial}{\partial t}\operatorname{div} \mathbf{D} = 4\pi\frac{\partial \rho}{\partial t}; \tag{1.63}$$

d'altra parte, se si applica l'operatore divergenza alla quarta delle (1.60) e si usa la proprietà che la divergenza di un rotore è nulla ($\operatorname{div}\operatorname{rot}\mathbf{H} = 0$) e che i due operatori div e $\dfrac{\partial}{\partial t}$ sono permutabili, si trova

$$-\frac{1}{c}\frac{\partial}{\partial t}\operatorname{div} \mathbf{D} = \frac{4\pi}{c}\operatorname{div} \mathbf{j}. \tag{1.64}$$

Eliminando il termine che contiene la divergenza di **D** fra le due precedenti relazioni si ottiene la voluta

VI. **Equazione di continuità o legge di conservazione della carica**:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} = 0. \tag{1.65}$$

Vi sono due altri importanti risultati dell'elettromagnetismo classico che val la pena di ricordare. Il primo, è che il campo elettromagnetico ha in sé una certa energia, e la densità di tale energia è

VII. **L'energia elettromagnetica per unità di volume**:

$$\mathcal{W} = \frac{1}{8\pi}(\epsilon \mathbf{E}^2 + \mu \mathbf{H}^2). \tag{1.66}$$

Il secondo, consiste nella possibilità di introdurre un vettore, che rende conto della propagazione dell'energia elettromagnetica

VIII. **Vettore di Poynting**:
Esiste un vettore **S**, *detto* **vettore di Poynting**, *che col suo orientamento individua la direzione ed il verso di propagazione dell'energia e con il suo modulo dà* **l'intensità energetica** *dell'onda, cioè l'energia che fluisce nell'unità di tempo attraverso una superficie di area unitaria, perpendicolare alla direzione di propagazione*

$$\mathbf{S} = \frac{c}{4\pi} \mathbf{E} \wedge \mathbf{H}. \tag{1.67}$$

1.4.4 Equazioni di Maxwell in mezzi omogenei non conduttori e legge di propagazione delle onde elettromagnetiche. Equazioni di Maxwell nel vuoto

A conclusione di questo paragrafo vogliamo esaminare quali semplificazioni subiscono le equazioni di Maxwell nel caso di mezzi particolari, per esempio **non conduttori**, com'è il vetro, per il quale si può porre nella (1.58) $\sigma = 0$, il che implica $\mathbf{j} = 0$; supponiamo inoltre che i mezzi siano anche **omogenei**, talché ϵ e μ sono costanti in senso stretto, come accade certamente **nel vuoto**. Se poi, come si verifica usualmente in tali mezzi, non vi sono cariche elettriche libere, possiamo porre anche $\rho = 0$. Con queste semplificazioni, le equazioni di Maxwell diventano:

Equazioni di Maxwell per mezzi omogenei non conduttori in assenza di cariche:

$$\begin{cases} \text{I.} \quad \operatorname{div} \mathbf{E} = 0, & (\textit{Teorema di Gauss}) \\ \text{II.} \quad \operatorname{div} \mathbf{H} = 0, & (\textit{Inesistenza di poli magnetici isolati}) \\ \text{III.} \quad \operatorname{rot} \mathbf{E} = -\dfrac{\mu}{c}\dfrac{\partial \mathbf{H}}{\partial t}, & (\textit{Legge di Faraday} - \textit{Neumann} - \textit{Lenz}) \\ \text{IV.} \quad \operatorname{rot} \mathbf{H} = \dfrac{\epsilon}{c}\dfrac{\partial \mathbf{E}}{\partial t} & (\textit{Legge di Ampère} - \textit{Maxwell}). \end{cases} \tag{1.68}$$

Se ora teniamo conto della seguente identità vettoriale

$$\operatorname{rot}\operatorname{rot} \mathbf{u} = \operatorname{grad}\operatorname{div} \mathbf{u} - \Delta_2 \mathbf{u},$$

dove $\Delta_2\mathbf{u}$ è il laplaciano di $\mathbf{u}$, e prendiamo il rotore del 1° e 2° membro della terza delle (1.68) otteniamo

$$\text{grad}\,\text{div}\mathbf{E} - \Delta_2\mathbf{E} = -\frac{\mu}{c}\,\text{rot}\,\frac{\partial\mathbf{H}}{\partial t} = -\frac{\mu}{c}\,\frac{\partial}{\partial t}\,\text{rot}\mathbf{H}.$$

Sostituendo in questa equazione i valori di div$\mathbf{E}$ e rot$\mathbf{H}$ dati dalle (1.68) avremo

$$\Delta_2\mathbf{E} = \frac{\epsilon\mu}{c^2}\,\frac{\partial^2\mathbf{E}}{\partial t^2}, \tag{1.69}$$

che, com'è noto, è l'equazione del moto ondoso relativa al campo elettrico, cioè il campo elettrico $\mathbf{E}$ si propaga come un'onda. Con analogo procedimento possiamo eliminare $\mathbf{E}$ anziché $\mathbf{H}$, ottenendo l'equazione del moto ondoso per $\mathbf{H}$, cioè

$$\Delta_2\mathbf{H} = \frac{\epsilon\mu}{c^2}\,\frac{\partial^2\mathbf{H}}{\partial t^2}. \tag{1.70}$$

Tenendo conto di ciò che si studia nell'ambito della teoria delle equazioni della Fisica Matematica possiamo dunque dire che le (1.69), (1.70) significano che onde costituite da contemporanee oscillazioni di $\mathbf{E}$ e di $\mathbf{H}$, quindi **onde elettromagnetiche**, possono propagarsi nel mezzo con velocità

$$v = \frac{c}{\sqrt{\epsilon\mu}}. \tag{1.71}$$

Nel vuoto, essendo $\epsilon = \mu = 1$, la velocità di propagazione è c e le equazioni di Maxwell saranno le più semplici possibili:

Equazioni di Maxwell nel vuoto in assenza di cariche:

$$\begin{cases} \text{I.} \quad \text{div}\,\mathbf{E} = 0, & (\textit{Teorema di Gauss}) \\ \text{II.} \quad \text{div}\,\mathbf{H} = 0, & (\textit{Inesistenza di poli magnetici isolati}) \\ \text{III.} \quad \text{rot}\,\mathbf{E} = -\dfrac{1}{c}\dfrac{\partial\mathbf{H}}{\partial t}, & (\textit{Legge di Faraday} - \textit{Neumann} - \textit{Lenz}) \\ \text{IV.} \quad \text{rot}\,\mathbf{H} = \dfrac{1}{c}\dfrac{\partial\mathbf{E}}{\partial t} & (\textit{Legge di Ampère} - \textit{Maxwell}). \end{cases} \tag{1.72}$$

La costante c che abbiamo definito con la (1.59) come rapporto tra unità corrispondenti dei sistemi $(CGS)_{es}$ e $(CGS)_{em}$, è suscettibile di misura diretta e risulta avere le dimensioni di una velocità ed un valore di circa 300.000 km/s. D'altra parte, anche **la velocità della luce nel vuoto** è suscettibile di misurazione diretta, e risulta avere lo steso valore di c. Perciò, da Maxwell in poi, i fisici sono stati indotti a concludere che **la luce è costituita da onde elettromagnetiche**: conclusione che è d'accordo con tutti i fatti sperimentali osservati. I raggi γ, i raggi X, i raggi ultravioletti, le radiazioni infrarosse e le radioonde sono, come la luce, onde elettromagnetiche e differiscono tra loro soltanto per il valore della lunghezza d'onda o della frequenza. Come si dimostra in elettromagnetismo si tratta in ogni caso di **onde trasversali**, cioè i campi elettrico $\mathbf{E}$ e magnetico $\mathbf{H}$ si propagano in direzione perpendicolare al loro piano di vibrazione.
Possiamo dunque concludere questi nostri pochi richiami di elettromagnetismo enunciando la seguente

IX. **Legge di propagazione delle onde elettromagnetiche**:
Nel riferimento $\mathcal{R}^*$, **cioè nel riferimento dell'etere**, *il fronte d'onda elettromagnetica (sede o confine di una perturbazione elettromagnetica) di equazione $\varphi(x, y, z, t) = 0$ si propaga nel vuoto*

con una velocità finita c costante e indipendente dalle modalità con cui la perturbazione viene provocata, secondo la legge

$$\Delta_2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = 0 \quad \Longleftrightarrow \quad \Box_c\varphi = 0, \tag{1.73}$$

dove $\Box_c$ è il cosiddetto **operatore di d'Alembert o d'Alembertiano** di φ.
Quest'equazione differenziale alle derivate parziali del 2° ordine - **equazione differenziale dei fronti d'onda** - esprime appunto che il fronte d'onda, qualunque sia la sua posizione e la sua forma iniziale, si propaga in ogni suo punto con velocità c. Essa individua univocamente il fronte d'onda ad ogni istante t non appena sia conosciuto il fronte medesimo all'istante $t = 0$.
Dal complesso delle IX leggi enunciate più sopra, discendono in particolare tutte le leggi dell'Ottica Fisica ed in particolare, come caso limite, quelle dell'Ottica Geometrica.

1.5 APPENDICE: Letture su spazio e tempo nella Fisica Classica

1.5.1 Spazio e tempo classici commentati da Einstein

Sentiamo ora **che cosa ne pensa Einstein dei concetti fondamentali di spazio, tempo e movimento** - pur rimanendo nell'ambito della Meccanica Classica - riportando un brano estremamente istruttivo ed illuminante tolto dalle prime pagine della sua *Relatività, Esposizione divulgativa.*[40]
"Accumulerei sulla mia coscienza dei gravi peccati contro il sacro spirito della chiarezza se affermassi, senza una seria riflessione e delle spiegazioni particolareggiate, che lo scopo della meccanica è di descrivere in qual modo i corpi mutano col tempo la loro posizone nello spazio. Non è chiaro che cosa si debba intendere qui per posizione e spazio. Io sto al finestrino di un vagone ferroviario che viaggia a velocità uniforme, e lascio cadere una pietra sulla banchina, senza imprimerle alcuna spinta. Allora, prescindendo dalla resistenza dell'aria, vedo discendere la pietra in linea retta. Un pedone che osserva il fattaccio dal sentiero lungo la ferrovia, vede che la pietra cade a terra descrivendo un arco di parabola. Domando ora: le posizioni percorse dalla pietra stanno in realtà su di una retta o su di una parabola? Inoltre, che cosa si intende qui per moto nello spazio?...
La risposta risulta di per sé evidente. In primo luogo evitiamo del tutto la vaga parola spazio[41]*, sotto la quale, dobbiamo onestamente riconoscerlo, non riusciamo ad immaginarci alcunchè, e sostituiamo l'espressione moto nello spazio con l'altra movimento rispetto ad un corpo di riferimento praticamente rigido...*
Se in luogo dell'espressione corpo di riferimento" noi introduciamo il concetto, utile per la descrizione matematica, di sistema di coordinate", siamo in grado di dire: la pietra percorre una linea retta relativamente ad un sistema di coordinate rigidamente collegato al vagone mentre descrive una parabola relativamente ad un sistema di coordinate rigidamente collegato al terreno (banchina). Da questo esempio si vede chiaramente che non esiste una traiettoria in sé, vale a dire una curva lungo cui si muove il corpo, ma soltanto una traiettoria rispetto ad un particolare sistema di riferimento.
Per ottenere una descrizione completa del movimento, dobbiamo però specificare in che modo il corpo altera la propria posizione col **tempo***; vale a dire, per ogni punto della traiettoria si deve specificare in quale istante il corpo si trova ivi situato. Questi dati devono venire completati con una definizione di tempo tale che, in virtù di essa, i valori del tempo possano venir considerati come grandezze per principio osservabili (risultati di misurazioni). Se ci poniamo sul terreno della Meccanica Classica possiamo, per il nostro esempio, soddisfare questo requisito nel modo seguente. Immaginiamo due orologi di costruzione*

[40] A. Einstein, *Relatività, Esposizione divulgativa*, Boringhieri-Torino (1967), pag. 50-51
[41] A tal proposito anche Max Born esprime lo stesso punto di vista dicendo: *.....poiché non è molto appropriato parlare di numerosi spazi che si muovono l'uno rispetto all'altro, preferiamo per quanto è possibile evitare l'uso del vocabolo spazio.* cfr. *La sintesi Einsteiniana*, Boringhieri, Torino (1969), pag. 95.

esattamente identica; la persona che si trova al finestrino del vagone ne tiene in mano uno, e la persona sulla banchina l'altro. Ciascuno degli osservatori determina in quale precisa posizione del rispettivo corpo di riferimento si trova la pietra ad ogni tic dell'orologio che ha in mano".

1.5.2 Spazio e tempo secondo l'operativismo di Bridgman

Torniamo ancora a Bridgman, limitandoci a riportare appena qualche stralcio, del suo pensiero personale a proposito dei concetti fondamentali di **spazio e tempo.**[42]
Spazio (secondo Bridgman): *"E' senza dubbio difficile dare una definizione logicamente soddisfacente di ciò che intendiamo col concetto di spazio, ma non andremo molto lontani dalla meta pensando ad esso come all'insieme di tutti i concetti relativi alla posizione. La posizione delle cose viene determinata con un qualche sistema di misura; il più semplice è forse quello implicito nel sistema cartesiano di coordinate con le sue tre misure di lunghezza.*
...Vediamo dunque che dal punto di vista operativo il sistema di riferimento della geometria cartesiana, spesso immaginato in senso matematico ideale, è di fatto un sistema fisico, e che per proprietà spaziali non intendiamo altro che le proprietà di questo sistema di riferimento. Dicendo che lo spazio è euclideo tridimensionale, intendiamo che lo spazio fisico dei regoli misuratori è euclideo (spazio tattile): non ha senso chiedersi se lo spazio vuoto è euclideo. La geometria, in quanto si vogliano applicare i suoi risultati al mondo fisico esterno e in quanto non la si consideri solo un sistema logico costruito sulla base di postulati, risulta pertanto una scienza sperimentale".
Osserviamo che anche Einstein,[43] dopo aver introdotto il concetto di corpo rigido e di regoli rigidi, dice: **La geometria a cui sia stata fatta quest'aggiunta deve quindi venir trattata come un ramo della fisica.**
Val la pena di osservare che con Bridgman ed Einstein, molti altri famosi scienziati, come Max Born, Arthur Eddington, Lev Davidovic Landau, Pascual Jordan, ecc., concepiscono in maniera analoga **la geometria come scienza sperimentale**.
Tempo (secondo Bridgman): *"Secondo il nostro punto di vista (operativo), il concetto di tempo è determinato dalle operazioni con cui misuriamo il tempo stesso. Dobbiamo distinguere due specie di tempo: il tempo degli eventi che hanno luogo uno vicino all'altro nello spazio, o tempo locale, e il tempo degli eventi che hanno luogo in punti dello spazio notevolmente lontani, o tempo esteso. Come ora sappiamo, il concetto di tempo esteso è legato inestricabilmente al concetto di spazio.*[44] *Questa non è affatto un'affermazione sulla Natura, e vi si può giungere con la semplice affermazione che le operazioni con cui si misura il tempo esteso implicano quelle mediante cui si misura lo spazio. Naturalmente in sede storica è alla Teoria della Relatività che si deve l'atteggiamento critico da cui si è giunti all'esame delle operazioni di misura del tempo,...., l'estendere nello spazio le coordinate temporali implica per definizione misure spaziali e la misura del tempo non risulta una faccenda autonoma".*

[42] P.W. Bridgman, *La logica della fisica moderna*, Ediz. Scientifiche Einaudi, Torino (1957), pag. 73-76. Allo studente desideroso di saperne di più, consigliamo la lettura sia di questo libro che di quest'altro già precedentemente citato: P.W. Bridgman, *La natura della teoria fisica*, La Nuova Italia Editrice, Firenze (1965).
[43] A. Einstein, *Relatività, Esposizione divulgativa*, Boringhieri-Torino (1967), pag.47.
[44] Noi aggiungiamo che: il concetto di tempo è legato inestricabilmente anche al concetto di moto uniforme, infatti le varie definizioni di un campione di durata per la misura del tempo sono tutte basate sulla regolarità (o uniformità) di certi movimenti: moto della sabbia nelle clessidre, piccole oscillazioni di un pendolo, moto diurno della Terra attorno al suo asse, moto delle lancette di un orologio, vibrazioni di un diapason, vibrazioni di un atomo di cesio, ecc.

Capitolo 2

CRISI DELLA FISICA CLASSICA

Al fine di raggiungere la verità, è necessario, una volta
nella propria vita, porre in dubbio ogni cosa, fin dove è possibile.
(René Descartes)[1]

2.1 La ricerca dell'etere cosmico

Nel 1° capitolo abbiamo riassunto in maniera concisa, ma pressoché completa, quali erano le leggi della Fisica Classica (sostanzialmente Meccanica ed Elettromagnetismo) così come si presentavano alla fine del 1800 inizio 1900. Nell'una o nell'altra di queste due classi si pensava allora potersi inquadrare ogni fenomeno fisico.
In questo capitolo ci occuperemo invece della crisi subita dalla fisica classica a causa di *certe difficoltà presentatesi nell'elettrodinamica di Maxwell* che, come dice lo stesso Einstein nel suo celebre lavoro del 1905:[2] *conduce, nella sua applicazione a corpi in movimento, ad* **asimmetrie** *che non sembrano conformi ai fenomeni.* Ciò che Einstein si propone di fare è stabilire una **simmetria** tra due descrizioni fisicamente equivalenti ma formalmente diverse - fornite dalla fisica classica - dello stesso fenomeno consistente nell'interazione fra un magnete ed un filo conduttore.[3]
Ci furono comunque altre, ed altrettanto gravi, difficoltà in cui la fisica classica si dibatteva, e delle quali ora ci occuperemo, che vennero affrontate da grandi fisici e matematici di fine secolo XIX°, specialmente dal celebre fisico olandese Hendrik Antoon Lorentz e dal grande matematico francese Jules Henri Poincaré.
Il primo segnale di allarme inerente alle difficoltà in cui venne a trovarsi la fisica classica arrivò proprio con la teoria elettromagnetica di Maxwell; l'allarme, sorprendentemente, risiede nel fatto che **le leggi dell'elettromagnetismo, cioè le equazioni di Maxwell, non risultano invarianti per trasformazioni di Galileo e perciò non verificano il Principio di Relatività Galileiano ben consolidato in meccanica da secoli di esperienze.** La teoria elettromagnetica di Maxwell, peraltro, uscirà indenne, in maniera quasi miracolosa, dalla critica einsteiniana nel senso che le equazioni di Maxwell risulteranno invarianti per trasformazioni di Lorentz - che sono le giuste trasformazioni da prendere in considerazione secondo la teoria einsteiniana - e perciò non ci sarà bisogno di modificarle. Saranno invece le venerande equazioni della Meccanica, che sembravano le più consolidate, a dover essere modificate alla luce **dei nuovi concetti di spazio**

[1] Questa citazione di Cartesio è riportata da Arthur Stanley Eddington nel suo libro *Spazio, Tempo e Gravitazione*, Universale Scientifica Boringhieri, 1971, pag. 31.

[2] Albert Einstein, *Zur Elektrodynamik bewegter korper*, Annalen der Physik, **17**, 1905, pp.891-921; traduzione italiana a cura di Paolo Straneo: *Sull'elettrodinamica dei corpi in moto*, inserito nel volume "Cinquant'anni di Relatività", pp.480, Vallecchi Editore, Firenze (1955).

[3] Per capire di che cosa si tratta si consiglia di leggere il lavoro originale di Einstein citato nella nota precedente ed anche, a complemento, le prime pagine del libro di Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), pagg.4-5, dove il pensiero di Einstein è spiegato chiaramente.

e tempo elaborati da Einstein per superare tutte le difficoltà.
Quindi, si presenta una situazione estremamente imbarazzante e poco credibile: mentre **in Meccanica, essendo valido il principio di relatività galileiano**, un riferimento inerziale $\mathcal{R}^*$, solidale con il cosiddetto **sistema delle stelle fisse** (identificato fin dall'epoca di Newton con lo **spazio assoluto** Σ^*), è affiancato da ∞^3 altri riferimenti inerziali $\mathcal{R}$, tutti ugualmente privilegiati e meccanicamente indistinguibili da $\mathcal{R}^*$, in moto TRU l'uno rispetto all'altro, **in elettromagnetismo $\mathcal{R}^*$ sembra rivestire l'ufficio di unico riferimento privilegiato**. D'altra parte, si sapeva che il fenomeno della propagazione per onde, nell'aria (come il suono), nei liquidi o in altre sostanze, come i metalli, ecc., avviene proprio grazie alla presenza di questi mezzi materiali: non si conosceva nessuna propagazione di onde nel vuoto. Perciò, fu abbastanza spontaneo ipotizzare l'esistenza di un mezzo estremamente tenue ed invisibile (dato che nessuno lo aveva mai osservato) che pervadesse tutto l'universo, *solidale con le stelle fisse*, sede della propagazione delle onde elettromagnetiche e quindi anche della luce, a cui venne dato un nome ad hoc: **etere cosmico.**[4] Il fatto che l'etere fosse concepito come solidale alle stelle fisse comporta che esso sia solidale all'unico sistema di riferimento privilegiato $\mathcal{R}^*$, rispetto al quale vengono studiati i fenomeni elettromagnetici.
Se tale riferimento è unico, allora deve anche essere fisicamente reperibile. L'individuazione dell'etere cosmico non potrà di certo essere perseguita mediante esperimenti puramente meccanici, stante il principio di relatività galileiano, bensì attraverso esperimenti ottici ed elettromagnetici. Dall'esame di alcuni celebri esperimenti ottici, con l'intento di cogliere da essi qualche informazione per l'individuazione dell'etere cosmico in natura, nacquero altre allarmanti incongruenze in seno alla teoria classica.
In questo capitolo ci limiteremo ad esaminare alcuni fatti fondamentali, fra altri che ce ne sono (per es. l'effetto Fresnel-Fizeau, ecc.), le cui spiegazioni messe a confronto, hanno evidenziato alcune contraddizioni che hanno determinato una crisi profonda della fisica classica. Questi fatti, che prenderemo in considerazione, possono riassumersi nei seguenti due punti:

1) l'apparente incompatibilità fra le leggi di Maxwell dell'elettromagnetismo, inclusa la legge di propagazione di un'onda elettromagnetica - come la luce - ed il principio di relatività;

2) la contraddizione fra le interpretazioni classiche di due esperimenti ottici: **l'aberrazione della luce stellare e l'effetto Doppler applicato alle onde luminose** da una parte ed il celeberrimo **esperimento di Michelson-Morley** dall'altra.

L'aberrazione della luce stellare, più che un esperimento è stata una notevole scoperta astronomica fatta già nella prima metà del Settecento dall'astronomo inglese James Bradley ed interpretata applicando la teoria dei moti relativi alla luce, supposta propagantesi nell'etere in quiete rispetto alle stelle fisse.
L'effetto Doppler, scoperto dal fisico austriaco Christian Doppler nel 1843 per le onde sonore, venne applicato, ingenuamente, alla propagazione della luce nel presunto etere supposto, anche in questo caso, in quiete rispetto alle stelle fisse.
L'esperienza di Michelson-Morley fu invece un vero esperimento ottico di laboratorio, cruciale

[4]Dal punto di vista storico, è doveroso precisare che il termine **etere** e le prime ipotesi sulla sua esistenza e la sua natura risalgono già al XVII° secolo, quando Christiaan Huygens formulò, per primo, nel suo *Traitè de la lumière (1690)* **la teoria ondulatoria della luce.** L'altra ipotesi antagonista sulla natura della luce, cioè **la teoria corpuscolare**, fu formulata più tardi da Newton nel suo trattato *Optiks (1704)* e per tutto il secolo XVIII° fu la più accreditata. Nel successivo secolo XIX°, con l'affermarsi della precedente **teoria ondulatoria della luce** per opera di Young e Fresnel, *l'esigenza di postulare un mezzo materiale (l'etere) per la sua propagazione si impose con sempre maggiore insistenza.* La natura di questo mezzo materiale fu sin dall'inizio fonte di numerosi problemi. Il fatto che le onde luminose fossero onde trasversali richiedeva un etere solido, invece che gassoso; l'elevatissima velocità di propagazione della luce richiedeva una rigidità corrispondentemente elevata per l'etere; infine il fenomeno astronomico dell'aberrazione della luce delle stelle indicava che l'etere dovesse restare immobile su distanze, appunto, astronomiche e perciò solidale al sistema delle stelle fisse.

per le conseguenze che ne derivarono, anche se non tutti gli storici della fisica sono concordi nell'attribuirle un ruolo determinante ai fini della formulazione della teoria della relatività da parte di Einstein. Lo stesso Einstein, un anno prima di morire, scrisse in una lettera indirizzata allo storico F. C. Davenport: *"L'esito dell'esperimento di Michelson non ebbe una grande influenza sull'evoluzione delle mie idee. Tanto che non riesco neppure a ricordare se ne avessi già sentito parlare, allorché scrissi la mia prima memoria sull'argomento (nel 1905). La spiegazione di ciò sta nel fatto che ero, per ragioni di carattere generale, fermamente convinto che non esista il moto assoluto, e il mio unico problema era come ciò potesse conciliarsi con quello che sapevamo dell'elettrodinamica".*[5]
Noi seguiremo la strada consistente nello studiare nei dettagli sia il fenomeno dell'aberrazione della luce stellare che l'effetto Doppler ottico, sia l'esperimento di Michelson-Morley, in quanto riteniamo che il capire questi fatti, specialmente l'esperimento di Michelson-Morley è comunque molto istruttivo in sé ed illuminante ai fini della comprensione della teoria della Relatività Ristretta. Inoltre, fra tutti gli approcci che si possono seguire, questo ci sembra particolarmente agevole dal punto di vista didattico e perciò molto adatto per principianti.

2.2 Non-invarianza delle equazioni di Maxwell per trasformazioni di Galileo

Nella sezione 1.4 del Cap.1, abbiamo riportato le leggi fondamentali dell'elettrodinamica classica, cioè le equazioni di Maxwell, la forza di Lorentz e la legge di propagazione delle onde elettromagnetiche dalla quale si deduce che queste si propagano nel vuoto con una velocità data dalla costante $c = \dfrac{u(CGS)_{es}}{u(CGS)_{em}}$ che rappresenta il rapporto fra unità del sistema CGS elettrostatico ed unità del sistema CGS elettromagnetico. Weber e Kohlrausch nel 1856, determinarono sperimentalmente il valore di c e scoprirono, come abbiamo già detto nel capitolo precedente, che tale rapporto era uguale a quello della velocità della luce nel vuoto misurato con altri esperimenti, ed è per questo motivo che tale costante è stata indicata con la stessa lettera con cui si indica la velocità della luce nel vuoto (si noti che la lettera c e l'iniziale della parola latina *celeritas* che significa *velocità*). **Ciò permise di affermare che i fenomeni ottici sono di natura elettromagnetica**.
Ci si chiese allora: rispetto a quale sistema di riferimento la velocità della luce vale $c \simeq 300.000\,\mathrm{km/s}$? La risposta che i fisici davano alla fine del 1800 era che: *il sistema di riferimento in cui la luce viaggia con velocità c è un sistema solidale all'etere, detto semplicemente* **il sistema dell'etere**. Questo mezzo non era mai stato osservato direttamente, perciò la sua esistenza era semplicemente postulata, e le sue presunte caratteristiche venivano dedotte in maniera indiretta dalla fenomenologia dell'elettromagnetismo. D'altra parte, secondo la meccanica newtoniana, un tale sistema di riferimento deve essere necessariamente inerziale, perché legato allo spazio delle stelle fisse, e perciò rappresentativo di quella classe di sistemi inerziali per i quali vale il fondamentale **principio di relatività di Galileo**. Questa idea tuttavia incontra una gravissima difficoltà. In accordo con **la legge classica di composizione delle velocità** (1.25), se in un dato sistema inerziale, diciamo il sistema dell'etere, un'onda elettromagnetica si propaga con velocità $\mathbf{c}$, in un altro sistema di riferimento inerziale in moto rispetto al precedente con velocità di trascinamento $\mathbf{v}$, essa si propaga con una velocità diversa $\mathbf{c}' = \mathbf{c} - \mathbf{v}$. In altri termini, **si può dimostrare che applicando una trasformazione di Galileo alle equazioni di Maxwell ed alla legge di propagazione di un'onda elettromagnetica, queste non si trasformano per invarianza, cioè non soddisfano al principio di relatività galileiano!**
Infatti consideriamo, per esempio, la terza delle equazioni di Maxwell (1.60), cioè la legge di

[5] Lettera a F.C. Davenport, 9 febbraio 1954, citato da Abraham Pais, nella sua ottima biografia di Einstein: *Sottile è il Signore.....*, Boringhieri, Torino (1986), pag. 188.

Faraday-Neumann-Lenz

$$\text{rot}\,\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{2.1}$$

e prendiamone la componente lungo l'asse x

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{1}{c}\frac{\partial B_x}{\partial t}, \tag{2.2}$$

che supponiamo valga in un dato sistema di riferimento inerziale $\mathcal{R}$. Per passare ad un altro sistema $\mathcal{R}'$ in configurazione x-standard rispetto ad $\mathcal{R}$, applichiamo la trasformazione speciale di Galileo (1.13) tenendo conto delle seguenti relazioni fra derivate parziali espresse nelle coordinate x, y, z, t di $\mathcal{R}$ e quelle x', y', z', t' (con $t = t'$) di $\mathcal{R}'$:

$$\frac{\partial}{\partial x} = \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} = \frac{\partial}{\partial x'} \tag{2.3}$$

$$\frac{\partial}{\partial y} = \frac{\partial y'}{\partial y}\frac{\partial}{\partial y'} = \frac{\partial}{\partial y'} \tag{2.4}$$

$$\frac{\partial}{\partial z} = \frac{\partial z'}{\partial z}\frac{\partial}{\partial z'} = \frac{\partial}{\partial z'} \tag{2.5}$$

$$\frac{\partial}{\partial t} = \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} = -v\frac{\partial}{\partial x'} + \frac{\partial}{\partial t'}. \tag{2.6}$$

La (2.2) diventa quindi

$$\frac{\partial E_z}{\partial y'} - \frac{\partial E_y}{\partial z'} = -\frac{1}{c}\left(\frac{\partial B_x}{\partial t'} - v\frac{\partial B_x}{\partial x'}\right). \tag{2.7}$$

Ora ci chiediamo: è possibile che i campi elettrico e magnetico si trasformino in modo tale che la (2.7) possa trasformarsi per invarianza, cioè possa scriversi nella stessa forma della (2.2), cioè

$$\frac{\partial E'_{z'}}{\partial y'} - \frac{\partial E'_{y'}}{\partial z'} = -\frac{1}{c}\frac{\partial B'_{x'}}{\partial t'}, \tag{2.8}$$

dove $\mathbf{E}' \equiv \mathbf{E}(x', y', z', t')$ e $\mathbf{B}' \equiv \mathbf{B}(x', y', z', t')$ sono i campi elettrico e magnetico in $\mathcal{R}'$? La risposta è no. Infatti qualunque sia la trasformazione di B_x, non c'è modo di eliminare il termine $v\frac{\partial B_x}{\partial x'}$ nella (2.7), in quanto le equazioni di Maxwell non contengono nessun'altra relazione tra $\frac{\partial B_x}{\partial t'}$ e $\frac{\partial B_x}{\partial x'}$.
Le equazioni di Maxwell, dunque, non rimangono invariate passando da un sistema di riferimento inerziale ad un altro legato al primo da una trasformazione di Galileo: **il principio di relatività galileiano è violato dall'elettromagnetismo!**

2.3 Non-invarianza della legge di propagazione delle onde elettromagnetiche

Dimostriamo ora che la legge (1.64) di propagazione delle onde elettromagnetiche che abbiamo enunciato nel paragrafo 1.4.4 del Cap.1, non è invariante rispetto alla trasformazione speciale di Galileo, al pari delle equazioni di Maxwell di cui è conseguenza.
Per fare questo, consideriamo due riferimenti inerziali in configurazione x-standard $\mathcal{R} \equiv (Oxyz, t)$ e $\mathcal{R}' \equiv (O'x'y'z', t')$ e supponiamo che in $\mathcal{R}$ si propaghi, nel vuoto, con velocità c, un fronte d'onda elettromagnetica (superficie mobile sede o confine di una perturbazione

elettromagnetica) di equazione $\varphi(x, y, z, t,) = 0$ soddisfacente all'equazione differenziale alle derivate parziali del 2° ordine:

$$\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = 0. \tag{2.9}$$

Se tale legge fosse invariante per trasformazioni di Galileo, allora esprimendo x, y, z, t nella (2.9) per mezzo delle (1.13) dovremmo ottenere nel riferimento $\mathcal{R}'$:

$$\frac{\partial^2\varphi}{\partial x'^2} + \frac{\partial^2\varphi}{\partial y'^2} + \frac{\partial^2\varphi}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t'^2} = 0. \tag{2.10}$$

Dimostriamo che ciò non avviene. Infatti, tenendo presenti le (1.13) del Cap.1, si ottengono le relazioni:

$$\begin{aligned}
&\frac{\partial\varphi}{\partial x} = \frac{\partial\varphi}{\partial x'}\frac{\partial x'}{\partial x} = \frac{\partial\varphi}{\partial x'},\\
&\frac{\partial\varphi}{\partial y} = \frac{\partial\varphi}{\partial y'},\\
&\frac{\partial\varphi}{\partial z} = \frac{\partial\varphi}{\partial z'},\\
&\frac{\partial\varphi}{\partial t} = \frac{\partial\varphi}{\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial\varphi}{\partial t'}\frac{\partial t'}{\partial t} = -v\frac{\partial\varphi}{\partial x'} + \frac{\partial\varphi}{\partial t'},
\end{aligned}$$

dalle quali discende che

$$\begin{aligned}
&\frac{\partial^2\varphi}{\partial x^2} = \frac{\partial^2\varphi}{\partial x'^2}\frac{\partial x'}{\partial x} = \frac{\partial^2\varphi}{\partial x'^2},\\
&\frac{\partial^2\varphi}{\partial y^2} = \frac{\partial^2\varphi}{\partial y'^2},\\
&\frac{\partial^2\varphi}{\partial z^2} = \frac{\partial^2\varphi}{\partial z'^2},\\
&\frac{\partial^2\varphi}{\partial t^2} = -v\left[\frac{\partial^2\varphi}{\partial x'^2}\frac{\partial x'}{\partial t} + \frac{\partial^2\varphi}{\partial x'\partial t'}\frac{\partial t'}{\partial t}\right] + \left[\frac{\partial^2\varphi}{\partial t'\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial^2\varphi}{\partial t'^2}\frac{\partial t'}{\partial t}\right] =\\
&= v^2\frac{\partial^2\varphi}{\partial x'^2} - 2v\frac{\partial^2\varphi}{\partial x'\partial t'} + \frac{\partial^2\varphi}{\partial t'^2}.
\end{aligned}$$

Sostituendo queste ultime nella (2.9), infine, si ha:

$$\frac{\partial^2\varphi}{\partial x'^2} + \frac{\partial^2\varphi}{\partial y'^2} + \frac{\partial^2\varphi}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t'^2} - \frac{v^2}{c^2}\frac{\partial^2\varphi}{\partial x'^2} + 2\frac{v}{c^2}\frac{\partial^2\varphi}{\partial x'\partial t'} = 0, \tag{2.11}$$

la quale non si identifica affatto con la (2.10). In conclusione **la legge di propagazione delle onde elettromagnetiche non è invariante rispetto alla trasformazione di Galileo**.
Perciò, tenuto conto delle considerazioni fatte nel paragrafo precedente e della dimostrazione fatta in questo, possiamo dire che: **in fisica classica il principio di relatività, valido per i fenomeni meccanici, non si estende ai fenomeni elettromagnetici.** Infatti se un tale principio valesse anche per le leggi dell'elettromagnetismo, un medesimo segnale luminoso - dovendo ubbidire alle stesse leggi in due riferimenti galileiani diversi $\mathcal{R}$ ed $\mathcal{R}'$ - dovrebbe possedere la stessa velocità c rispetto ad essi, violando in tal modo la legge classica di composizione delle velocità (teorema dei moti relativi).
Questo non significa che l'elettromagnetismo non soddisfi ad un principio di relatività, quel che possiamo dire è che non soddisfa al principio di relatività di Galileo; vedremo infatti, nel prossimo capitolo, che **l'elettromagnetismo soddisfa invece al principio di relatività di Einstein**, il cui nucleo matematico è la trasformazione di Lorentz e non quella di Galileo.

2.4 L'aberrazione annua delle stelle: spiegazione classica

A) Cenni storici e nozioni preliminari
Per capire bene l'argomento di cui ci occuperemo in questo paragrafo, sarebbe auspicabile che il lettore fosse a conoscenza di alcune semplici nozioni di Astronomia Sferica, che comunque, per sua comodità, riportiamo nell'APPENDICE 1 alla fine del capitolo.
La cosiddetta **aberrazione annua della luce stellare** costituisce una notevole scoperta astronomica che giocherà un ruolo importante per ciò che diremo in seguito.
È esperienza comune di chiunque, che se ci si muove camminando sotto la pioggia che cade verticalmente, per non bagnarsi è necessario inclinare l'ombrello nel senso del moto, in altri termini a causa del nostro moto sembra che la pioggia non cade più verticalmente ma cade inclinata e tanto più inclinata quanto è maggiore la nostra velocità. Un fenomeno analogo si osserva per un raggio di luce: se un osservatore in moto rettilineo misura l'angolo formato da un raggio di luce con la sua direzione di moto trova un angolo diverso da quello che misurerebbe se fosse in quiete rispetto al raggio. Il fenomeno si chiama **aberrazione** ed è noto da molto tempo come ora spiegheremo.
Tra il 1725 e il 1728 l'astronomo inglese James Bradley (Direttore del reale Osservatorio astronomico di Greenwich)[6] si accorse, osservando la stella γ Draconis, che *le stelle non restano assolutamente fisse sulla sfera celeste, ma ciascuna di esse descrive annualmente una piccola ellisse (Fig. 2.1a) il cui semiasse maggiore è sempre parallelo al piano dell'eclittica ed ha ampiezza angolare A costantemente uguale, per ogni stella, a circa* $20'',50$; *l'ampiezza angolare B del semiasse minore, invece, è uguale a quella del semiasse maggiore moltiplicata per il seno della latitudine celeste* φ *dell'astro* (Fig. 2.1b):

$$B = A \sin \varphi. \tag{2.12}$$

A tale fenomeno Bradley diede il nome di *aberrazione della luce stellare* (dal latino *aberratio* =

Figura 2.1: a) Ellisse di aberrazione; b) Coordinate eclitticali (λ, φ) di una stella S sulla sfera celeste.

scostamento, illusione), giacché si tratta appunto di una illusione ottica, dovuta al moto annuo della Terra attorno al Sole ed al fatto che la luce si propaga con velocità finita attraverso un ipotetico **etere** (come si credeva allora) *supposto solidale con le stelle fisse*. Essa costituisce di fatto *una prova tangibile* del moto annuo della Terra intorno al Sole.

Il fenomeno dell'aberrazione può essere spiegato con le leggi della meccanica classica. Tale spiegazione è dovuta sostanzialmente allo stesso Bradley, e di essa ci occuperemo in questo paragrafo. Inoltre, questo fenomeno è spiegato altrettanto bene dalla Teoria della Relatività Ristretta e di esso ci occuperemo in seguito nel Cap.4, paragrafo 4.2.1.
Una spiegazione accurata del fenomeno richiederebbe lo studio particolareggiato delle onde luminose nelle lenti e nei prismi degli strumenti astronomici di osservazione; ma, se si concepiscono i raggi luminosi come le traiettorie dei fotoni che trasportano l'energia delle onde

[6] J. Bradley, *Phil. Trans.* **35**, 637 (1728).

luminose, ci si può attenere alla spiegazione più semplice ed intuitiva fornita dalla teoria corpuscolare e che porta ai medesimi risultati.

Avvertiamo il lettore, che la maggior parte dei testi che discutono il fenomeno dell'aberrazione della luce stellare lo fanno, spesso, in maniera molto succinta ed eccessivamente schematica. Noi presentiamo due approcci, il primo dei quali è strettamente fedele alla pratica delle osservazioni astronomiche e ci condurrà di fatto all'ellisse di aberrazione scoperta da Bradley, il secondo, un pò più teorico, ha il vantaggio di seguire un procedimento analogo a quello che forniremo nell'ambito della teoria della Relatività Ristretta.[7]

B) Primo approccio: ellisse di aberrazione

Immaginiamo che *un raggio di luce*, dopo essere partito da una stella S, arrivi *attraverso l'etere, supposto in quiete,* all'occhio dell'osservatore O.

Poiché l'osservatore, trovandosi sulla Terra, si muove rispetto all'etere, con una certa velocità $\mathbf{v}_O$, il raggio giunge al suo occhio con una *velocità relativa* $\mathbf{c}'$, che, per la legge di composizione delle velocità, è data da:

$$\mathbf{c}' = \mathbf{c} - \mathbf{v_O}, \tag{2.13}$$

essendo $\mathbf{c}$ la velocità assoluta della luce rispetto all'etere. In conseguenza l'osservatore O, vedrà l'astro non già secondo la direzione OS, ma secondo la direzione della velocità relativa $\mathbf{c}'$, cioè secondo la direzione OS', essendo S' la posizione apparente della stella cioè la sua proiezione sulla sfera celeste.

D'altra parte, è opportuno notare che la velocità $\mathbf{v_O}$ dell'osservatore terrestre si compone di tre velocità, e cioè:

1) velocità dovuta al moto annuo di rivoluzione della Terra intorno al Sole che indicheremo con $\mathbf{v}_T$ (che produce la cosiddetta *aberrazione annua*);

2) velocità dovuta al moto di rotazione diurno della Terra attorno al proprio asse (che produce la cosiddetta *aberrazione diurna*);

3) velocità dovuta al moto del Sole nello spazio (che produce la cosiddetta *aberrazione secolare*).

Ma la seconda velocità è molto piccola (giungendo appena a $\frac{1}{2}$ km/s.) rispetto alla velocità della luce nel vuoto che è di circa 300.000 km/s.: possiamo quindi generalmente trascurare l'effetto a cui dà luogo che non supera $0'',3$ all'equatore terrestre ed è ancora più debole nelle altre regioni. Quanto alla terza velocità, essa si conserva per lungo tempo costante in grandezza e direzione onde dà luogo a spostamenti costanti; vale a dire non produce nessun movimento apparente nella posizione delle stelle. Ci resta dunque da tener conto soltanto della prima delle velocità soprannominate, cioè della velocità $\mathbf{v}_T$ della Terra intorno al Sole. A tale scopo, è necessario tenere presente il seguente teorema della teoria dei moti planetari (Fig.2.2):

la velocità di un pianeta $\mathbf{v}_P$ nel suo moto di rivoluzione attorno al Sole è uguale alla somma vettoriale di due componenti,[8] e cioè (vedi Fig.2.2):

[7]Per una buona trattazione in senso classico del fenomeno dell'aberrazione stellare si consiglia di consultare: Giuseppe Armellini, *I fondamenti scientifici dell'Astronomia*, Ulrico Hoepli Editore, (1952), pagg. 197-200; ovvero Francesco Zagar, *Astronomia Sferica e Teorica*, Zanichelli Editore, Bologna (1948), pagg. 171-175, ai quale ci siamo ispirati per la nostra trattazione.

[8]Cfr. ancora G. Armellini, *I fondamenti scientifici dell'Astronomia*, Ulrico Hoepli Editore, Milano (1952), pagg. 122-125, ovvero F. Zagar citato nella nota precedente.

Figura 2.2: Decomposizione della velocità di un pianeta.

a) Un componente **u** costante in grandezza e direzione, ortogonale all'asse maggiore dell'orbita, il cui modulo è $u = e\dfrac{C}{p}$, dove e, C, p indicano l'eccentricità, la costante delle aree (doppio della velocità areolare) e il parametro dell'orbita; [9]

b) Un secondo componente **v** di direzione variabile, ma di modulo costantemente uguale a $v = \dfrac{C}{p}$.

Osserviamo che, tenuto conto delle formule riportate nella nota 9 a piè di pagina, è facile dedurre che

$$u = \frac{2\pi ae}{T\sqrt{1-e^2}}, \qquad v = \frac{2\pi a}{T\sqrt{1-e^2}}. \tag{2.14}$$

Per la Terra, il precedente teorema è espresso perciò dalla seguente relazione vettoriale

$$\mathbf{v}_T = \mathbf{u} + \mathbf{v}.$$

Ma, per quanto si è detto, il primo componente **u**, essendo costantemente ortogonale all'asse maggiore dell'orbita, non dà luogo a movimenti apparenti delle stelle ed inoltre ha modulo costante ed estremamente piccolo (se l'orbita fosse circolare, quindi con eccentricità nulla, $(e = 0)$[10] sarebbe addirittura $\mathbf{u} = 0$), quindi può essere trascurato. In ultima analisi, dovremo perciò tener conto soltanto del secondo componente **v** che **coincide praticamente, con buonissima approssimazione, con la velocità orbitale della Terra**, dunque possiamo porre

$$\mathbf{v}_T = \mathbf{v}.$$

In conseguenza, la legge dei moti relativi (2.13), che vogliamo applicare, risulta:

$$\mathbf{c}' = \mathbf{c} - \mathbf{v},$$

nella quale, ripetiamolo, **c** ($|\mathbf{c}| = c \simeq 300.000\,\mathrm{km/s}$), è la velocità della luce nell'etere supposto in quiete rispetto alle stelle fisse, $\mathbf{c}'$ è l'analoga velocità rispetto all'osservatore terrestre O, in

[9] Ricordiamo al lettore che, essendo l'orbita di un pianeta ellittica, se usiamo coordinate polari (r, θ) col polo coincidente con uno dei fuochi, per esempio col Sole, e l'asse polare diretto come l'asse maggiore dell'ellisse verso il vertice più vicino, e si denota con *a il semiasse maggiore*, con *b il semiasse minore*, con $e = \sqrt{1 - b^2/a^2}$ *l'eccentricità*, con $p = b^2/a$ *il parametro*, l'equazione polare dell'ellisse è: $r = p/(1 + e\,cos\theta)$ in cui θ è la cosiddetta *anomalia* del pianeta, cioè l'angolo formato dall'asse polare con il raggio vettore che va dal Sole al pianeta. Infine, se T è *il periodo di rivoluzione* del pianeta allora $C = 2\pi ab/T$ è *la costante delle aree*, doppio della velocità areolare.

[10] L'orbita della Terra è quasi circolare in quanto la sua eccentricità è $e = 0,0167 \approx 0,017$.

moto rispetto all'etere, e $\mathbf{v}$ ($|\mathbf{v}| = v \simeq 30\,\mathrm{km/s}$) è la velocità orbitale della Terra[11]. Notiamo che il punto della sfera celeste a cui punta $\mathbf{v}$, è chiamato dagli astronomi *apice del moto*.
Indichiamo ora con θ e θ' gli angoli, convessi compresi fra 0 e π, formati dal vettore $\mathbf{v}$ con i vettori $\mathbf{c}$ e $\mathbf{c}'$ rispettivamente.

Figura 2.3: Posizione reale ed apparente di una stella: angolo di aberrazione $\Delta\theta$.

In riferimento alla Fig.2.3, sia S la posizione effettiva della stella, sulla sfera celeste, come la vedrebbe un osservatore in quiete nell'etere, e sia invece S' la sua posizione apparente rispetto ad un osservatore terrestre O in moto nell'etere con velocità $\mathbf{v}$, nell'istante dell'osservazione.
Tenendo conto della legge di composizione delle velocità $\mathbf{c} = \mathbf{c}' + \mathbf{v}$, possiamo costruire, con origine in O, il parallelogramma $OBAC$ avente i lati $\overrightarrow{OB} = \mathbf{v}$, $\overrightarrow{OC} = \mathbf{c}'$ e diagonale $\overrightarrow{OA} = \mathbf{c}$. Il prolungamento di OA dalla parte di O passa ovviamente per la posizione reale della stella S, mentre l'analogo prolungamento di OC passa per la sua posizione apparente S'.
Ciò posto, mandiamo da S *una parallela alla velocità orbitale* $\mathbf{v}$, cioè ad OB, e prolunghiamola fino ad intersecare il prolungamento di OC nel punto S'' ottenendo così il triangolo OSS'' simile ad OAC; notiamo, per inciso, che la visuale OS' coincide con la visuale OS''.
Osserviamo ora che il triangolo OSS'' ha per costruzione il lato SS'' parallelo ed equiorientato come la velocità orbitale della Terra e quindi *parallelo al piano dell'eclittica.* Se indichiamo con k la lunghezza del segmento SS'' ed r la distanza dell'astro dall'osservatore O, e cioè la lunghezza del segmento OS, per avere la dimensione angolare di SS'' basterà dividere k per r.[12]
I due triangoli OAC, OSS'' essendo, come già detto, simili ci permettono di scrivere:

$$\frac{k}{r} = \frac{v}{c}.$$

Ma, come visto al precedente punto b), $v = \dfrac{C}{p}$, ed inoltre C, p, c si conservano costanti nel tempo, perciò avremo:

$$\frac{k}{r} = \frac{v}{c} = \frac{C}{pc} = cost. \tag{2.15}$$

Cioè l'ampiezza angolare del segmento SS'' rimarrà costante e sempre molto piccola essendo $v \ll c$.

[11]La velocità orbitale della Terra si calcola facilmente con la 2ª delle formule (2.14) ove vi si ponga $a = 149.600.000\,\mathrm{km}$, $T = 365,25\,giorni$, $e = 0,017$; si ottiene $v \simeq 30\,\mathrm{km/s}$.

[12]Si tenga presente che la distanza lineare $OS \equiv r$ da noi assunta come raggio della sfera celeste è del tutto arbitraria ed in conseguenza è arbitraria anche la distanza lineare $SS'' \equiv k$. Queste distanze, nello studio dei moti apparenti dei corpi celesti e quindi anche nello studio dell'aberrazione, non hanno nessuna importanza: quello che conta è la distanza angolare fra la posizione reale S di una stella osservata e la sua posizione apparente S' (misurata dall'angolo $\widehat{SOS'} \equiv \widehat{SOS''}$) che è sempre piccolissima ed è valutata, a meno di infinitesimi di ordine superiore, dal rapporto $\dfrac{k}{r}$.

In conseguenza, il punto S'' descriverà annualmente un piccolo cerchio posto in un piano parallelo a quello dell'eclittica, avente per raggio l'ampiezza angolare $A = \frac{k}{r} = \frac{v}{c}$ e per centro la posizione vera dell'astro S.
L'osservatore terrestre O vedrà la proiezione di tale cerchio sul piano tangente alla sfera celeste nel punto S occupato dall'astro. Perciò è necessario determinare l'angolo formato da questo piano tangente col piano dell'eclittica. Ma, com'è noto dalla Geometria, l'angolo fra due piani è definito come l'angolo formato dalle rispettive normali. Queste normali sono: *l'asse dell'eclittica* OE che unisce O col suo polo boreale E, e la semiretta OS (raggio della sfera celeste passante per S). Quest'ultima semiretta forma col piano dell'eclittica l'angolo φ che altro non è che *la latitudine celeste della stella*, dunque tali normali formano un angolo che è il complementare della latitudine celeste cioè $\frac{\pi}{2} - \varphi$.
Nella Geometria Analitica si dimostra il seguente teorema:
la proiezione ortogonale di un cerchio di raggio R su di un piano formante un angolo α col piano del cerchio, è un'ellisse il cui semiasse maggiore a è uguale al raggio del cerchio, cioè $a = R$, e il semiasse minore è uguale al semiasse maggiore moltiplicato per il coseno dell'angolo formato dai due piani, cioè: $b = R \cos\alpha$.[13]
Nel nostro caso, l'ellisse proiezione ha il semiasse maggiore uguale all'ampiezza angolare A del raggio del cerchio di centro S, descritto dal punto S'', parallelo al piano dell'eclittica ed il semiasse minore uguale a $A \cos\left(\frac{\pi}{2} - \varphi\right) = A \sin\varphi$.
Potremo quindi concludere dicendo che:
ogni stella molto lontana S sembra descrivere annualmente in cielo una piccola ellisse avente per centro la sua posizione vera S, il cui semiasse maggiore, parallelo all'eclittica, ha un'ampiezza angolare $A = \frac{v}{c}$ dove $\frac{v}{c} = \beta = 10^{-4}$ radianti $\approx 20'', 50$ è la cosiddetta **costante di aberrazione**, *ed il semiasse minore ha un' ampiezza angolare $B = A \sin\varphi$.*
Nelle formule precedenti si è indicata con $\beta = \frac{v}{c}$ la cosiddetta *velocità römeriana*.[14] Abbiamo così ritrovato **l'ellisse di aberrazione**, come è stata determinata per la prima volta da Bradley e come ancora oggi ogni astronomo può constatarne l'esistenza.
Concludendo, se ora applichiamo il teorema dei seni al triangolo OAC; si ha:

$$\frac{\sin\Delta\theta}{v} = \frac{\sin(\pi - \theta')}{c},$$

cioè:

$$\sin\Delta\theta = \beta \sin\theta', \tag{2.16}$$

che costituisce quella che comunemente viene indicata come **formula classica esatta di aberrazione della luce stellare**.
Da essa si può calcolare immediatamente l'angolo di aberrazione $\Delta\theta$

$$\Delta\theta = \arcsin(\beta \sin\theta'), \tag{2.17}$$

non appena si sia misurato θ'.
La formula esatta (2.16) può essere riscritta in un'altra forma, basta tener conto che $\Delta\theta = \theta' - \theta$, per avere

$$sin\theta' \cos\theta - \sin\theta \cos\theta' = \beta \sin\theta,'$$

[13]Questo teorema si può trovare nel libro di Oscar Chisini, *Lezioni di geometria Analitica e Proiettiva*, Zanichelli Editore, Bologna, 1960, pagg. 433-434.

[14]Ovviamente β non è una velocità, ma il rapporto di una velocità - in questo caso la velocità orbitale della Terra - con la velocità della luce. Questo rapporto è stato chiamato impropriamente *velocità römeriana* in onore dell'astronomo danese Ole Christensen Römer che è stato il primo ad effettuare, nel 1676, una misura della velocità della luce nel vuoto.

dividendo ora membro a membro per $\cos\theta'$ e portando $\sin\theta$ al 2° membro avremo

$$\tan\theta' = \frac{\sin\theta}{\cos\theta - \beta'}, \tag{2.18}$$

che costituisce una **seconda formula classica esatta di aberrazione.**
Siccome θ' e θ differiscono di pochissimo, e perciò $\Delta\theta$ è sempre molto piccolo, se esprimiamo questi angoli in radianti, la formula esatta (2.16) può essere scritta in forma approssimata, al primo ordine in β, così

$$\Delta\theta \approx \beta \sin\theta', \tag{2.19}$$

che costituisce la cosiddetta **formula classica approssimata di aberrazione**.
Essendo $\sin\theta' \leq 1$ si ha per l'angolo di aberrazione $\Delta\theta \approx \frac{v}{c} = 10^{-4}$ rad $\approx 20'',50$ in ottimo accordo con i risultati sperimentali ottenuti da Bradley e confermati da altri sperimentatori.
Tale accordo sembrerebbe confermare la plausibilità dell'ipotesi adottata che **l'etere si trovi in quiete rispetto alle stelle fisse**. Va tuttavia rilevato che la precisione delle osservazioni non è sufficiente a confermare o negare la validità né della formula esatta (2.16) né della formula approssimata (2.19).

C) Secondo approccio: uso di due riferimenti inerziali in configurazione x-standard
Consideriamo un sistema di riferimento galileiano $\mathcal{R} \equiv (Oxyz, t)$ nella cui origine O sia situato l'occhio dell'osservatore terrestre nella posizione occupata sulla propria orbita attorno al Sole ad un determinato istante che consideriamo come iniziale $t = 0$ per gli orologi di $\mathcal{R}$.

Figura 2.4: Studio dell'aberrazione utilizzando due riferimenti inerziali in configurazione x-standard.

Prendiamo l'asse x, uscente da O, parallelo ed equiorientato come la velocità orbitale $\mathbf{v}$ della Terra, l'asse z ortogonale ad x ed al piano dell'orbita terrestre ed orientato verso l'emisfero celeste in cui si trova la stella da osservare, l'asse y giacente su tale piano ed orientato in modo che $Oxyz$ sia una terna destra (Fig.2.4).
Supponiamo però che il riferimento $\mathcal{R}$ non venga trascinato dalla Terra ma sia immobile nello

spazio ovvero sia **solidale con l'etere cioè con lo spazio delle stelle fisse.**[15] Consideriamo inoltre un secondo sistema $\mathcal{R}' \equiv (O'x'y'z', t)$ che all'istante $t = 0$, precedentemente fissato, sia coincidente con $\mathcal{R} \equiv (Oxyz, t)$ ma trascinato dalla Terra con orientamento invariabile, e che perciò possiamo ritenere, almeno per un breve periodo di tempo, in moto traslatorio uniforme con velocità costante $\mathbf{v}$ rispetto ad $\mathcal{R}$ e quindi anch'esso galileiano. $\mathcal{R}$ ed $\mathcal{R}'$ sono dunque due sistemi di riferimento galileiani in configurazione x-standard. Osserviamo che il riferimento $\mathcal{R}'$ non lo possiamo considerare solidale alla Terra perché questa, oltre che essere animata dal moto annuo attorno al Sole, è animata dal moto rotatorio diurno attorno al proprio asse e perciò se $\mathcal{R}'$ fosse ad essa solidale non potrebbe essere inerziale anche per un breve intervallo di tempo. Supponiamo inoltre, per semplicità, di effettuare l'osservazione di una stella nell'istante $t = 0$ in cui i due sistemi di riferimento sono coincidenti e che tale stella giaccia sul piano xz e quindi anche sul piano $x'z'$ ad esso sovrapposto (vedi Fig.2.4). L'osservatore terrestre O' vedrà la stella in una certa posizione apparente S' sulla sfera celeste; sia S la posizione vera della stessa stella come si vedrebbe dal sistema $\mathcal{R}$ solidale con l'etere. Quindi l'osservatore O' vede la luce arrivare dall'astro con una certa velocità $\mathbf{c}'$ che, per la legge classica di composizione delle velocità, è data in forma vettoriale, da:

$$\mathbf{c}' = \mathbf{c} - \mathbf{v}. \tag{2.20}$$

Proiettando la (2.20) sugli assi del riferimento $\mathcal{R}$, otteniamo:

$$c'_x = c_x - v, \;\; c'_y = c_y = 0, \;\; c'_z = c_z, \tag{2.21}$$

dove c_x, c_y, c_z e c'_x, c'_y, c'_z sono le componenti cartesiane, in $\mathcal{R}$, dei vettori $\mathbf{c}$ e $\mathbf{c}'$; le componenti c_y, c'_y sono nulle perché i vettori $\mathbf{c}$ e $\mathbf{c}'$ giacciono, all'istante $t = 0$ dell'osservazione, nel piano $xz \equiv x'z'$ ortogonale all'asse y. Indichiamo, come abbiamo fatto nel primo approccio, al punto **B)**, con θ e θ' gli angoli, convessi compresi fra 0 e π, formati dal vettore $\mathbf{v}$ con i vettori $\mathbf{c}$ e $\mathbf{c}'$ rispettivamente, e scriviamo le componenti cartesiane di questi vettori secondo gli assi x, z, avremo:

$$\begin{cases} c_x = \mathbf{c}\cdot\mathbf{i} = c\,\cos\theta, \;\; c_z = \mathbf{c}\cdot\mathbf{k} = c\,\cos\left[\dfrac{\pi}{2} + (\pi - \theta)\right] = c\,\cos\left(\dfrac{3}{2}\pi - \theta\right) = -c\,\sin\theta \\ c'_x = \mathbf{c}'\cdot\mathbf{i} = c'\,\cos\theta', \;\; c'_z = \mathbf{c}'\cdot\mathbf{k} = c'\,\cos\left[\dfrac{\pi}{2} + (\pi - \theta')\right] = c'\cos\left(\dfrac{3}{2}\pi - \theta'\right) = -c'\,\sin\theta', \end{cases} \tag{2.22}$$

dove $\mathbf{i}$ e $\mathbf{k}$ sono i versori degli assi x e z e $c = |\mathbf{c}|$, $c' = |\mathbf{c}'|$ i moduli della velocità della luce nel vuoto rispettivamente in $\mathcal{R}$ ed $\mathcal{R}'$.
Sostituiamo ora le (2.22) nella prima e terza equazione di (2.21) e dividiamo membro a membro quest'ultima per la prima, si ottiene

$$\tan\theta' = \frac{\sin\theta}{\cos\theta - \beta}, \qquad \left(\beta = \frac{v}{c}\right). \tag{2.23}$$

Abbiamo così ritrovato, con procedimento completamente diverso e con un'esperienza ipotetica, puramente pensata, *la formula classica esatta di aberrazione* (2.17) già ricavata seguendo il primo approccio. La (2.23), con facili calcoli, può essere riscritta nella forma (2.16) più conveniente, basta sostituire $\tan\theta'$ con $\dfrac{\sin\theta'}{\cos\theta'}$ e ridurre l'espressione a forma intera, si avrà:

$$\sin\Delta\theta = \beta\sin\theta', \tag{2.24}$$

[15] Facciamo notare al lettore che la considerazione di un tale riferimento non è realistica nel senso che noi abitanti della Terra ci muoviamo con essa e perciò non ci è consentito nemmeno un istante di essere in quiete nell'etere. Tuttavia la sua considerazione è utile perché ci consente, dal punto di vista teorico, di effettuare un'esperienza puramente ipotetica che a posteriori si può constatare che fornisce gli stessi risultati di un'esperienza effettiva.

la quale fornisce immediatamente l'angolo di aberrazione $\Delta\theta$ non appena si sia misurato θ'. Da essa segue anche ovviamente la **formula classica approssimata di aberrazione**

$$\Delta\theta \approx \beta \sin\theta'.$$

2.5 Effetto Doppler ottico: spiegazione classica

Introduzione

La velocità con cui le onde si propagano in un mezzo è generalmente indipendente dalla velocità della sorgente delle onde medesime; così se un sasso è lanciato in uno stagno con una certa velocità orizzontale, onde, in forma di cerchi concentrici, si propagano nell'acqua dal punto di caduta del sasso con una velocità che è indipendente da quella del sasso che le ha generate.

Se una sorgente va irradiando onde **mentre si muove**, *la velocità delle onde è in genere quella stessa che si avrebbe se la sorgente fosse ferma (essa è leggermente variabile se c'è dispersione), ma* **la lunghezza d'onda e la frequenza, quali appaiono ad un osservatore fermo, possono risultare diverse**.

Così, se per esempio, consideriamo il fischio di un treno, la sirena di un'auto della polizia o di un'autoambulanza, il rombo di un aeroplano, ecc., cambiano altezza a seconda che la sorgente sonora si stia avvicinando o allontanando rispetto a chi ascolta. Questo fenomeno acustico si chiama **effetto Doppler**, dal nome del fisico austriaco Christian Doppler (1803-1853) che per primo lo ha spiegato nel 1843. Si verifica ogni qualvolta una sorgente di onde sonore *ma anche elettromagnetiche, in particolare una sorgente di luce,* e l'osservatore sono in moto relativo l'uno rispetto all'altro[16]. In questo paragrafo ci occuperemo, dal punto di vista della fisica classica, dell'effetto Doppler applicato alla propagazione della luce nell'etere. Per far ciò con profitto è necessario avere ben presenti alcuni concetti fondamentali sulla propagazione ondosa che abbiamo raccolto nell'APPENDICE 2 alla fine del capitolo (paragrafo 2.10), per venire incontro agli studenti che li avessero dimenticati.

Consideriamo due sistemi di riferimento inerziali $\mathcal{R}$ ed $\mathcal{R}'$ in configurazione x-standard, che si muovono l'uno rispetto all'altro con velocità di modulo v.

Siano $\mathbf{x} \equiv (x, y, z)$ ed $\mathbf{x}' \equiv (x', y', z')$ i vettori di posizione di un generico punto P ed $\mathbf{n} \equiv (\alpha, \beta, \gamma)$, $\mathbf{n}' \equiv (\alpha, \beta, \gamma)$ due vettori unitari che individuano la direzione di propagazione di un'onda rispettivamente nei riferimenti inerziali $\mathcal{R}$ ed $\mathcal{R}'$. È anche comodo introdurre i due **vettori d'onda** $\mathbf{k} = k\mathbf{n}$ e $\mathbf{k}' = k'\mathbf{n}'$ aventi la direzione di propagazione dell'onda in $\mathcal{R}$ ed $\mathcal{R}'$ ed il cui modulo coincide col rispettivo **numero d'onda**.

Supponiamo ora che un'onda elettromagnetica piana monocromatica[17] sia emessa da una **sorgente luminosa** $\mathcal{S}$ posta nell'origine O di $\mathcal{R}$ (vedi Fig.2.5).

Supponiamo ancora che l'onda abbia frequenza ν, pulsazione $\omega = 2\pi\nu$, lunghezza d'onda $\lambda = \dfrac{2\pi}{k}$ e si muova con velocità $\mathbf{c}$ ($\lambda\nu = |\mathbf{c}| = c$) in $\mathcal{R}$, e rispettivamente frequenza ν',

[16]Sulla propagazione delle onde e sull'effetto Doppler acustico si consiglia il lettore di consultare uno dei tanti testi universitari di Fisica generale, per esempio il molto diffuso: D. Halliday, R. Resnick, K.S. Krane, *Fisica 1*, 4^a edizione, Casa Editrice Ambrosiana, Milano, 1993, pp. 486-488 per quel che riguarda l'effetto Doppler per le onde sonore. Per la spiegazione dell'effetto Doppler per la luce, sia in ambito classico che relativistico, basta la nostra trattazione esposta nelle pagine che seguono.

[17]**Un'onda piana** è quella nella quale la perturbazione Φ, qualunque essa sia, ha lo stesso valore in tutti i punti di ogni piano perpendicolare alla direzione di propagazione individuata dal versore $\mathbf{n}$. Ognuno di tali piani è detto **fronte d'onda**, e si sposta parallelamente a sé stesso con la velocità $\mathbf{v}$ di propagazione dell'onda. L'equazione dei fronti d'onda è ad ogni istante, ovviamente

$$\mathbf{x} \cdot \mathbf{n} = \alpha x + \beta y + \gamma z = cost,$$

con $|\mathbf{n}| = \alpha^2 + \beta^2 + \gamma^2 = 1$.

Figura 2.5: Effetto Doppler ottico: spiegazione classica.

pulsazione $\omega' = 2\pi\nu'$, lunghezza d'onda $\lambda' = \dfrac{2\pi}{k'}$ e si muova con velocità $\mathbf{c}'$ ($\lambda'\nu' = |\mathbf{c}'| = c'$) in $\mathcal{R}'$.
Ciò posto, in $\mathcal{R}$ l'equazione dell'onda sarà del tipo:

$$\Phi = a\ \cos\left(\omega t - \mathbf{k}\cdot\mathbf{x}\right) = a\ \cos\ \omega\Big(t - \frac{\mathbf{n}\cdot\mathbf{x}}{c}\Big), \tag{2.25}$$

la cui fase è definita dalla relazione

$$\phi = \omega\big(t - \frac{\mathbf{x}\cdot\mathbf{n}}{c}\big), \tag{2.26}$$

nella quale si è tenuto conto che $\mathbf{k} = k\mathbf{n}$ e $k = \dfrac{\omega}{c}$.
Nel riferimento $\mathcal{R}'$ i fronti d'onda saranno ancora piani, poiché la trasformazione di coordinate che fa passare dal riferimento inerziale $\mathcal{R}$ al riferimento $\mathcal{R}'$ è una trasformazione di Galileo che essendo lineare, trasforma piani in piani.
Perciò la propagazione dell'onda in $\mathcal{R}'$ è rappresentata da un'equazione dello stesso tipo della (2.25), cioè

$$\Phi' = a'\ \cos\left(\omega' t' - \mathbf{k}'\cdot\mathbf{x}'\right) = a'\ \cos\ \omega'\Big(t' - \frac{\mathbf{n}'\cdot\mathbf{x}'}{c'}\Big), \tag{2.27}$$

la cui fase sarà espressa, analogamente alla (2.79), dalla

$$\phi' = \omega'\big(t' - \frac{\mathbf{x}'\cdot\mathbf{n}'}{c'}\big). \tag{2.28}$$

A questo punto dobbiamo ricordarci un fatto fondamentale:
La fase di un'onda piana è una quantità invariante *che rimane la stessa in tutti i sistemi di riferimento*, cioè:

$$\phi = \phi'. \tag{2.29}$$

Questo è vero perché la misura della fase di un'onda può essere ricondotta al semplice conteggio delle creste d'onda in un treno d'onde che sono passate nella posizione occupata dall'osservatore: trattandosi di una pura operazione di conteggio, il risultato deve essere indipendente dal sistema di coordinate.
Se dunque consideriamo un'onda elettromagnetica nel vuoto, la sua fase, misurata nei due riferimenti inerziali $\mathcal{R}$ e $\mathcal{R}'$ deve soddisfare alla seguente uguaglianza

$$\omega\Big(t - \frac{\mathbf{x}\cdot\mathbf{n}}{c}\Big) = \omega'\Big(t' - \frac{\mathbf{x}'\cdot\mathbf{n}'}{c'}\Big). \tag{2.30}$$

Ricordando che $\omega = 2\pi\nu$, $\omega' = 2\pi\nu'$ potremo scriverla nella forma:

$$\nu\Big(t - \frac{\mathbf{x}\cdot\mathbf{n}}{c}\Big) = \nu'\Big(t' - \frac{\mathbf{x}'\cdot\mathbf{n}'}{c'}\Big), \tag{2.31}$$

dove ω, ν, t, $\mathbf{x}$, $\mathbf{n}$ e c sono rispettivamente: la pulsazione, la frequenza dell'onda, il tempo, il vettore di posizione di un generico punto sul fronte d'onda, il versore di avanzamento del fronte d'onda ed il modulo della velocità della luce nel riferimento inerziale $\mathcal{R}$; mentre ω', ν', t', $\mathbf{x}'$, $\mathbf{n}'$ e c' sono le corrispondenti grandezze nel riferimento $\mathcal{R}'$, con $t = t'$ essendo il tempo assoluto. Se ora esprimiamo $\mathbf{x}$ e t, nel 1° membro della (2.31), mediante la trasformazione di Galileo

$$\begin{cases} \mathbf{x} = \mathbf{x}' + \mathbf{v}t \\ t = t' \end{cases} \tag{2.32}$$

otteniamo:

$$\nu\left[t'\left(1 - \frac{\mathbf{v}\cdot\mathbf{n}}{c}\right) - \frac{\mathbf{x}'\cdot\mathbf{n}}{c}\right] = \nu'\left(t' - \frac{\mathbf{x}'\cdot\mathbf{n}'}{c'}\right). \tag{2.33}$$

Siccome questa uguaglianza deve essere verificata per tutti i valori di t' e di $\mathbf{x}'$, se introduciamo le coordinate del vettore di posizione, $\mathbf{x}' \equiv (x', y', z')$, ed i coseni direttori di $\mathbf{n} \equiv (\alpha, \beta, \gamma)$ ed $\mathbf{n}' \equiv (\alpha', \beta', \gamma')$, è necessario che i coefficienti di t', x', y', z' a 1° e 2° membro della (2.33) siano a due a due uguali. Troviamo intanto la seguente uguaglianza

$$\nu\left(1 - \frac{\mathbf{v}\cdot\mathbf{n}}{c}\right)t' - \frac{\nu}{c}[\alpha x' + \beta y' + \gamma z'] = \nu' t' - \frac{\nu'}{c'}[\alpha' x' + \beta' y' + \gamma' z'], \tag{2.34}$$

da cui si deduce che

$$\nu' = \nu\left(1 - \frac{\mathbf{v}\cdot\mathbf{n}}{c}\right), \qquad \frac{\alpha}{\alpha'} = \frac{\beta}{\beta'} = \frac{\gamma}{\gamma'} = \frac{\nu' c}{\nu c'}. \tag{2.35}$$

Tenuto ora conto che essendo $\mathbf{n}$, $\mathbf{n}'$ dei versori, deve essere $\alpha^2 + \beta^2 + \gamma^2 = 1$ e $\alpha'^2 + \beta'^2 + \gamma'^2 = 1$, e dovendo inoltre verificarsi la seconda delle (2.35) si avrà che non solo è $\mathbf{n} = \mathbf{n}'$, ma anche $\frac{\nu'}{\nu} = \frac{c'}{c}$. Combinando ancora quest'ultima con la prima delle (2.35) ed indicando con θ, θ' gli angoli formati da da $\mathbf{v}$ con $\mathbf{n}$ e con $\mathbf{n}'$ rispettivamente, avremo il risultato definitivo con ($\theta = \theta'$):

$$\begin{cases} \nu' = \nu\left(1 - \frac{\mathbf{v}\cdot\mathbf{n}}{c}\right) = \nu(1 - \beta\cos\theta) = \nu(1 - \beta\cos\theta') \\ \mathbf{n} = \mathbf{n}' \\ c' = c - \mathbf{v}\cdot\mathbf{n} \end{cases} \tag{2.36}$$

Queste sono le **formule dell'effetto Doppler ottico della Relatività Galileiana**: come si vede si tratta di un'effetto del 1° ordine in $\beta = \frac{v}{c}$.
Osserviamo che la terza di tali formule è una conseguenza della legge galileiana di composizione delle velocità $\mathbf{c} = \mathbf{c}' + \mathbf{v}$.
La prima delle (2.36) mostra, che in ambito classico, *non esiste un effetto Doppler trasversale* perché in tal caso dovrebbe essere $\mathbf{v}\cdot\mathbf{n} = 0$, ovvero $\theta = 90°$, ciò che implica $\nu' = \nu$ e perciò assenza di effetto Doppler. *L'effetto è soltanto di tipo longitudinale, e risulta del 1° ordine nel rapporto $\beta = \frac{v}{c}$, come l'aberrazione della luce stellare.*
Se introduciamo la variazione di frequenza $\Delta\nu = \nu' - \nu$, la formula (2.36) dell'effetto Doppler si può esprimere anche così:

$$\frac{\Delta\nu}{\nu} = -\frac{\mathbf{v}\cdot\mathbf{n}}{c}. \tag{2.37}$$

Si osservi che per la validità delle (2.36), (2.37), insieme al tempo assoluto ($t = t'$), *si è considerata la lunghezza d'onda λ (distanza fra due creste d'onda successive) come un* **invariante** nel passaggio dal sistema dell'etere al riferimento $\mathcal{R}'$, in conformità con l'invarianza classica delle distanze spaziali.

Le considerazioni precedenti e le formule dell'effetto Doppler (2.36), (2.37) a cui siamo pervenuti, sono perfettamente accettabili dal punto di vista teorico, ma non vanno bene dal punto di vista del controllo sperimentale. Infatti per verificarle si dovrebbe essere in grado, non solo di misurare la frequenza ν' nel sistema mobile costituito dal nostro laboratorio solidale alla Terra[18], cosa perfettamente fattibile, ma saper anche effettuare misure di frequenza ν nel sistema dell'etere, ciò che è praticamente impossibile.
Supponiamo, come già detto, che l'etere coincida con lo spazio delle stelle fisse, che abbiamo identificato con il sistema inerziale $\mathcal{R}$, ed assumiamo come sorgente una stella molto lontana posta sul **piano dell'eclittica**. Sia poi $\mathcal{R}'$ solidale alla Terra, ed immaginiamo di eseguire una prima misura di frequenza quando la stella è in avvicinamento verso la Terra. Essendo in tal caso $\theta = 180°$, la prima delle (2.36) ci dà la frequenza

$$\nu' = \nu\left(1 + \frac{v}{c}\right). \tag{2.38}$$

Sei mesi dopo quando la stella è in allontanamento dalla Terra[19], essendo allora $\theta = 0°$, la sua luce avrà una frequenza minore ν'' data dalla formula analoga:

$$\nu'' = \nu\left(1 - \frac{v}{c}\right), \tag{2.39}$$

essendo v il modulo della velocità orbitale terrestre, che com'è noto, è $v \approx 30\,\mathrm{km/s}$. Ponendo ora $\Delta\nu = \nu' - \nu'' > 0$, ne consegue che sommando e sottraendo membro a membro le (2.38) e (2.39) si ottiene facilmente:

$$\nu = \frac{1}{2}(\nu' + \nu''), \qquad \Delta\nu = \frac{1}{2}(\nu' - \nu''), \tag{2.40}$$

e pertanto, dividendo membro a membro la seconda per la prima, avremo il legame voluto:

$$\frac{\Delta\nu}{\nu} = \frac{\nu' - \nu''}{\nu' + \nu''} = \frac{v}{c} = \beta \approx 10^{-4}. \tag{2.41}$$

La (2.41) con i valori ν' e ν'' **osservabili** risolve il nostro problema che è poi quello che hanno sempre fatto in pratica gli Astronomi effettuando misure Doppler ottiche nell'ambito della fisica classica, confermando sempre con buonissima approssimazione la (2.41).
Dunque, **anche l'effetto Doppler, come il fenomeno dell'aberrazione della luce stellare, sembra convalidare, o almeno non smentire, la presunzione di un etere in quiete rispetto alle stelle fisse.**

2.6 Un esercizio istruttivo di Cinematica Classica: regata su un fiume

Risolviamo, in questo paragrafo, un esercizio di Cinematica classica sui moti relativi[20] che presenta una notevole analogia con il fondamentale esperimento ottico di Michelson-Morley e perciò ci aiuterà a capire meglio quest'ultimo.
Ecco l'enunciato del problema:

[18]Osserviamo che la Terra, pur essendo animata dal *moto rotatorio diurno attorno al proprio asse* e mobile attorno al Sole (*moto annuo*), può essere considerata, per brevi intervalli di tempo, come un sistema **quasi inerziale**.

[19]Si ammette che l'orbita terrestre sia quasi circolare (come è in effetti essendo la sua eccentricità di appena $17/1000$, cioè $e = 0,017$). È ovvio che se ad un dato istante la Terra, nel suo moto annuo, si allontana dalla stella (red shift), dopo sei mesi essa sarà in avvicinamento (blue shift).

[20]Questo esercizio sui moti relativi si può trovare in Bruno Finzi e Paolo Udeschini, *Esercizi di Meccanica Razionale*, Tamburini Editore-Milano, Ristampa della terza edizione, 1964, Es. 8, pagg.123-124. Esso si trova anche nel famoso libro divulgativo sulla relatività di Clement V. Durell, *La Relatività con le quattro operazioni*, Boringhieri, Torino,1966, pagg. 30-33.

La corrente d'acqua di un fiume a sponde rettilinee e parallele, distanti ℓ fra di loro, si muove con velocità costante **v** *(Fig.2.6).*[21] *Due barche (a remi o a motore), partono entrambe, allo stesso istante, da un punto A situato su una delle due rive del fiume: una, che indicheremo con L, naviga prima contro corrente fino ad un punto B, situato a monte di A, e ad una distanza da esso pari alla larghezza ℓ del fiume, quindi ritorna immediatamente in A navigando ora nello stesso verso della corrente. L'altra, che chiameremo T, naviga tagliando trasversalmente la corrente fino al punto C della riva opposta e anch'essa ritorna immediatamente in A. Entrambe le barche si muovono* **rispetto alla corrente** *(cioè rispetto ad un sistema di riferimento solidale con la corrente), con velocità aventi lo stesso modulo $u = cost$.*
Calcolare i tempi di andata e ritorno delle due barche ed il loro rapporto. Chi vince la regata?
Soluzione:
Osserviamo subito che non è facile né intuitivo prevedere chi vincerà la regata:

1) perché la barca L, nel percorso di andata, da A a B, muovendosi contro corrente va molto più lenta che nel percorso di ritorno da B ad A, in quanto ora si muove nello stesso senso della corrente. Precisamente nel percorso di andata, la velocità della barca rispetto alla corrente si sottrae con quella della corrente stessa, mentre al ritorno le due velocità si sommano;

2) perché la barca T tagliando trasversalmente la corrente viene da questa trascinata verso valle, perdendo quindi velocità rispetto alla riva, e perciò, chi sta al timone, deve continuamente inclinare la barra in modo da puntare la prua della barca dalla parte in cui proviene la corrente, se vuole fare un percorso rettilineo perpendicolare alle sponde del fiume in modo da raggiungere il punto C.

Figura 2.6: Una regata su un fiume.

Come abbiamo detto, nella nota 21 a piè di pagina in questo capitolo, supponiamo, per semplicità, che il fiume scorra su di un piano orizzontale e introduciamo due sistemi di riferimento inerziali: uno di essi, diciamolo $\mathcal{R} \equiv Oxyz$ sia mobile solidalmente con la corrente (perciò in quiete rispetto ad essa): possiamo dunque pensarlo, per fissare le idee, come solidale ad una *zattera rigida* trascinata dalla corrente. Prendiamo l'origine O in un punto arbitrario della zattera (se questa è rettangolare possiamo, per semplicità, prendere O coincidente con un suo vertice)

[21] Per fissare le idee, supponiamo che la corrente si muova da destra verso sinistra rispetto ad un osservatore che si trovi in piedi sulla riva da cui partono le barche e che guardi verso di esse. Inoltre, per semplicità, supponiamo che il fiume scorra su di un piano orizzontale.

e supponiamo che l'asse x rimanga costantemente parallelo e di verso opposto alla corrente e parallelo alle rive rettilinee del fiume; l'asse y sia ad esso perpendicolare, giacente sul piano orizzontale ed orientato come il segmento AC, l'asse z verticale ed orientato verso l'alto, in modo che $Oxyz$ formi una terna destra.
L'altro sistema $\mathcal{R}' \equiv O'x'y'z'$ sia invece solidale con la Terra, equiorientato con $\mathcal{R} \equiv Oxyz$ ed avente l'origine $O' \equiv A$.
Indichiamo con $\mathbf{u}(L)$, $\mathbf{u}(T)$ e con $\mathbf{u}'(L)$, $\mathbf{u}'(T)$, le velocità delle due barche rispetto ai due sistemi di riferimento $\mathcal{R}$ ed $\mathcal{R}'$ rispettivamente; $\mathbf{v}$, come è stabilito nell'enunciato del problema, è la velocità della corrente rispetto alla Terra, cioè rispetto al riferimento $\mathcal{R}'$, sarà perciò $\mathbf{v}_\tau = -\mathbf{v}$ la velocità di trascinamento del sistema $\mathcal{R}'$ rispetto al sistema $\mathcal{R}$.
Possiamo allora applicare, ad entrambe le barche, il **teorema dei moti relativi (o legge di composizione delle velocità)**.
Per la barca L, che fa il tragitto di andata e ritorno da A a B, prima contro corrente e poi nel verso della corrente, si ha

$$\mathbf{u}(L) = \mathbf{u}'(L) + \mathbf{v}_\tau = \mathbf{u}'(L) - \mathbf{v},$$

ovvero

$$\mathbf{u}'(L) = \mathbf{u}(L) + \mathbf{v}. \tag{2.42}$$

Ricordando che le velocità delle due barche rispetto alla corrente, cioè rispetto al sistema $\mathcal{R}$, hanno, per ipotesi, lo stesso, modulo, cioè $|\mathbf{u}(L)| = |\mathbf{u}(T)| = cost. = u$, se proiettiamo la (2.42) sull'asse x' otteniamo

$$|\mathbf{u}'(L)| = |\mathbf{u}(L)| \mp |\mathbf{v}| = u \mp v, \tag{2.43}$$

dove vale il segno - per l'andata da A a B, cioè contro corrente, ed il segno + per il ritorno da B ad A, cioè nel verso della corrente, com'è del tutto intuitivo.
Avremo allora, valutando i tempi in $\mathcal{R}'$,

$$t_{AB} = \frac{|AB|}{|\mathbf{u}'(L)|_{And.}} = \frac{\ell}{u-v}, \qquad t_{BA} = \frac{|BA|}{|\mathbf{u}'(L)|_{Rit.}} = \frac{\ell}{u+v},$$

ed il tempo totale di andata e ritorno per la barca L sarà[22]

$$t(L) = \frac{\ell}{u-v} + \frac{\ell}{u+v} = \frac{2\ell u}{u^2 - v^2} = \frac{2\ell}{u}\frac{1}{1-\dfrac{v^2}{u^2}} = \frac{2\ell}{u}\gamma^2, \tag{2.44}$$

dove abbiamo indicato con

$$\gamma = \frac{1}{\sqrt{1-\dfrac{v^2}{u^2}}} \tag{2.45}$$

un'espressione, che si rivelerà in seguito molto importante per la Teoria della Relatività Ristretta, e che, fin d'ora, chiameremo **fattore di Lorentz**.
Per la barca T avremo analogamente:

$$\mathbf{u}(T) = \mathbf{u}'(T) + \mathbf{v}_\tau = \mathbf{u}'(T) - \mathbf{v}, \tag{2.46}$$

perciò è possibile costruire un diagramma delle velocità (vedi Fig.2.6) in cui $\mathbf{u}'(T) \equiv AD$ è diagonale del parallelogramma avente per lati i vettori $\mathbf{u}(T) \equiv AE$ e $\mathbf{v} \equiv AF$ supposti entrambi spiccati dal punto $A \equiv O'$. Tenendo conto che $\mathbf{u}'(T) \perp \mathbf{v}$ il triangolo ADF è rettangolo in A con cateti $|\mathbf{u}'(T)|$, $|\mathbf{v}|$ ed ipotenusa $|\mathbf{u}(T)|$.

[22] Si tenga presente che in meccanica classica **il tempo è assoluto** $t = t'$, perciò i tempi di andata e ritorno per le due barche possono essere valutati indifferentemente in $\mathcal{R}$ o in $\mathcal{R}'$.

Ricordando che la velocità di T, rispetto alla corrente, cioè rispetto al sistema $\mathcal{R}$ è $|\mathbf{u}(T)| = u$, applicando il il teorema di Pitagora al triangolo ADF si avrà:

$$|\mathbf{u}'(T)| = \sqrt{|\mathbf{u}(T)|^2 - |\mathbf{v}|^2} = \sqrt{u^2 - v^2}.$$

Perciò il tempo che impiega la barca T per andare da A a C (valutato in $\mathcal{R}'$) sarà allora:

$$t_{AC} = \frac{|AC|}{|\mathbf{u}'(T)|} = \frac{\ell}{\sqrt{u^2 - v^2}} = \frac{\ell}{u}\frac{1}{\sqrt{1 - \frac{v^2}{u^2}}} = \frac{\ell}{u}\gamma. \tag{2.47}$$

Il tempo t_{CA}, di ritorno da C ad A (essendo il ritorno del tutto simmetrico rispetto all'andata) sarà ovviamente uguale a t_{AC}, perciò il tempo totale di andata e ritorno per la barca T è

$$t(T) = \frac{2\ell}{\sqrt{u^2 - v^2}} = \frac{2\ell}{u}\frac{1}{\sqrt{1 - \frac{v^2}{u^2}}} = \frac{2\ell}{u}\gamma. \tag{2.48}$$

Come si può constatare dalle (2.44), (2.48) i tempi $t(L)$ ed $t(T)$ di andata e ritorno delle due barche sono diversi, ed anzi $t(L) > t(T)$, perché è immediato riconoscere che il fattore di Lorentz $\gamma > 1$, dunque **la regata sarà vinta dalla barca T che taglia trasversalmente la corrente!**
In altri termini: *Se le due barche hanno uguali velocità rispetto alla corrente, a parità di cammino percorso per viaggi di andata e ritorno, impiega meno tempo chi viaggia trasversalmente alla corrente di chi viaggia invece parallelamente ad essa.*
Questo risultato teorico è perfettamente coerente con le misure sperimentali dei tempi $t(L)$ e $t(T)$.
Il rapporto dei tempi, chiesto dal problema, è allora

$$\frac{t(L)}{t(T)} = \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{u^2}}}. \tag{2.49}$$

Affinché il fattore di Lorentz γ sia reale, deve essere $1 - \frac{v^2}{u^2} > 0$, cioè $v < u$, il che significa che: **affinché il problema ammetta soluzione reale è necessario che la velocità della corrente sia sempre inferiore alla velocità comune delle due barche rispetto alla corrente stessa.**
Osservazione 1:
È ovvio che se si conosce la velocità u delle barche rispetto alla corrente ed il rapporto $\gamma = \frac{t(L)}{t(T)}$ dei tempi di andata e ritorno in A, si può determinare la velocità v della corrente, si ha infatti con facili calcoli

$$v = u\sqrt{1 - \frac{1}{\gamma^2}}. \tag{2.50}$$

In modo analogo, conoscendo la velocità v della corrente ed il fattore di Lorentz γ, si può calcolare la velocità comune delle barche rispetto alla corrente

$$u = \frac{v}{\sqrt{1 - \frac{1}{\gamma^2}}}. \tag{2.51}$$

Osservazione 2:
Che il fattore di Lorentz sia $\gamma > 1$ si vede subito dalla formula (2.49), ove si tenga conto del

fatto che, affinché esso sia reale, dev'essere, come abbiamo detto più sopra, $v < u$, perciò $0 < 1 - \frac{v^2}{u^2} < 1$, la radice quadrata di un numero positivo minore di 1 è un numero positivo ancora minore di 1, ed infine il quoziente di 1 con un numero minore di 1 è un numero maggiore di 1.

2.7 L'esperimento di Michelson-Morley

Cominciamo con l'osservare che una volta scoperto, con la teoria di Maxwell, che la luce è un fenomeno di origine elettromagnetica e che quindi deve propagarsi per onde, e tenuto conto che tutti gli altri tipi di fenomeni ondosi (o vibratori) si propagano in un mezzo materiale, sembrò ovvio ai fisici postulare - come già detto nel preambolo introduttivo a questo capitolo - l'esistenza di un mezzo, detto **etere** in seno al quale si potesse propagare una qualsiasi onda elettromagnetica, compresa la luce. Si suppose anche che l'etere avesse insolite proprietà, come una densità nulla e una perfetta trasparenza, per giustificare il fatto che non si riuscisse a rivelarlo. Inoltre si postulava che esso riempisse tutto lo spazio, fosse in quiete rispetto alle cosiddette **stelle fisse** e fosse il mezzo in cui la luce ha velocità $\mathbf{c}$. Se si suppone, con Lorentz, che l'etere non venga trascinato dalla materia, la nostra Terra dovrebbe, almeno in qualche periodo dell'anno, trovarsi in moto rispetto all'etere, dato che essa è animata da moto di rivoluzione attorno al Sole e da moto di rotazione attorno al proprio asse. Un osservatore sulla Terra avvertirebbe un **vento d'etere** di velocità $\mathbf{v}$ avente lo stesso modulo della velocità orbitale terrestre $\mathbf{v}_T$ ma verso opposto, $\mathbf{v} = -\mathbf{v}_T$, con $|\mathbf{v}| = |\mathbf{v}_T| \simeq 30\,\mathrm{km/s} = 108.000\,\mathrm{km/h}$, che è una velocità piuttosto rilevante[23]. Inoltre, la sua velocità orbitale cambia in continuazione di direzione. In tali condizioni la velocità della luce rispetto alla Terra dovrebbe risultare diversa nelle diverse direzioni. Ne segue quindi che un osservatore in movimento attraverso l'etere con velocità $\mathbf{v}$, rispetto ad esso, per un raggio di luce dovrebbe misurare una velocità $\mathbf{c}'$ diversa da $\mathbf{c}$, dove $\mathbf{c}' = \mathbf{c} - \mathbf{v}$.
Per mettere in evidenza questa anisotropia e per misurare la velocità del presunto vento d'etere, il fisico americano, di origine prussiana, Albert Abraham Michelson ideò e costruì un *interferometro ottico* di grandissima precisione e fece, nel 1881, la prima esperienza per stabilire sperimentalmente questi fatti. È ovvio che Michelson, come qualsiasi altro fisico di fine ottocento, ragionasse in termini di fisica classica applicando, come abbiamo fatto noi nell'esercizio del paragrafo precedente, la teoria dei moti relativi e la legge di composizione delle velocità alla corrente d'etere, alla velocità della luce nell'etere ed alla velocità della luce rispetto alla Terra, ed effettuando la misura del rapporto dei tempi impiegati dalla luce su percorsi diversi entro il suo interferometro. Che cosa si aspettava Michelson da questa esperienza? Si aspettava che dall'osservazione dello spostamento delle frange di interferenza causato dalla variazione della velocità della luce rispetto alla Terra si potesse rivelare il vento d'etere e misurarne la velocità. Quest'ultima, ci si aspettava che fosse uguale alla velocità orbitale della Terra, almeno in qualche periodo dell'anno, basandosi naturalmente sulla conoscenza della velocità c della luce rispetto all'etere come era previsto dalla teoria elettromagnetica.
Con sua grandissima sorpresa il risultato fu assolutamente inatteso e sbalorditivo. Michelson **non riuscì a rivelare nessun vento d'etere!** L'esperimento fu perfezionato e da lui ripetuto nel 1887 in collaborazione con un altro fisico americano Edward Williams Morley ottenendo ancora lo stesso **risultato negativo**.
Per la sua invenzione dell'interferometro e per i molti altri suoi esperimenti di ottica, Michelson

[23]Se si vuole avere un confronto con altre velocità, si osservi che una palla di cannone o di un fucile non supera i 1500 m/s=1,5 km/s, cioè 1/20 della velocità orbitale terrestre; la velocità periferica della Terra all'equatore è piuttosto piccola e cioè circa 1/2 km/s $\simeq$ 17/1000 della velocità orbitale terrestre; la velocità di fuga di un missile dal campo gravitazionale della Terra è di circa 11,2 km/s.

ricevette il premio Nobel per la Fisica nel 1907, e fu il primo americano a ricevere questo premio. In questo paragrafo forniamo una spiegazione molto semplificata e schematica, ma concettualmente corretta, che si riallaccia all'esercizio di Cinematica visto nel paragrafo precedente.
Per chi volesse una spiegazione un pò più dettagliata e realistica di questo fondamentale esperimento ottico consigliamo la lettura del testo di Robert Resnick, *Introduzione alla Relatività Ristretta*, Casa Editrice Ambrosiana, Milano, 1979, pag.18-26, o di un qualsiasi trattato di Ottica.
Sostituiamo al fiume, dell'esercizio svolto nel paragrafo precedente, un *apparecchio interferometrico (interferometro di Michelson)* formato schematicamente da due aste rigide AB, AC di ugual lunghezza ℓ, perpendicolari fra di loro in A, montate su di una solida piattaforma orizzontale, e portanti alle loro estremità B e C due specchi con la superficie riflettente rivolta verso A (Fig.2.7).
Supponiamo che l'asta AB, orientata da A a B, venga disposta in modo tale da avere lo stesso verso della velocità orbitale della Terra $\mathbf{v}_T$, cioè nel verso opposto alla presunta corrente d'etere, mentre l'asta AC sia perpendicolare ad essa e giacente nel piano orizzontale.
In uno stesso istante t vengono fatti partire da A due raggi di luce, uno nella direzione AB e l'altro nella direzione AC; questi colpiscono gli specchi, ne vengono riflessi e tornano in A. Il loro moto corrisponde al moto delle due barche L ed T dell'esercizio precedente, mentre la corrente del fiume è qui sostituita dal **vento d'etere**, che almeno in qualche periodo dell'anno, dovrebbe avere, come si è detto sopra, velocità di modulo uguale alla velocità orbitale della Terra e verso opposto.

Figura 2.7: Descrizione schematica dell'esperimento di Michelson-Morley.

Introduciamo due sistemi di riferimento inerziali: il primo di essi, diciamolo $\mathcal{R} \equiv Oxyz$, sia in quiete nell'etere e perciò solidale con l'ipotetica corrente, avente l'asse x orientato come la velocità orbitale della Terra $\mathbf{v}_T$, e perciò equiorientato come il braccio AB dell'interferometro; l'asse y orientato come il braccio AC e l'asse z orientato verticalmente in modo tale che $Oxyz$ sia una terna destra.
Il secondo sistema inerziale, diciamolo $\mathcal{R}' \equiv O'x'y'z'$ sia solidale con l'interferometro, e quindi con la Terra, avente l'origine $O' \equiv A$, l'asse x' orientato come l'asse x cioè come il braccio AB, l'asse y', ortogonale alla corrente d'etere, ed orientato come il braccio AC ed infine l'asse z' sarà scelto in modo che le due terne $Oxyz$ e $O'x'y'z'$ siano equiorientate.
Indichiamo con $\mathbf{c}(L)$, $\mathbf{c}(T)$ e con $\mathbf{c}'(L)$, $\mathbf{c}'(T)$, le velocità dei raggi di luce rispetto ai due sistemi di riferimento $\mathcal{R}$ ed $\mathcal{R}'$ rispettivamente e con $\mathbf{v}$ la velocità della corrente d'etere rispetto alla Terra, cioè rispetto al riferimento $\mathcal{R}'$, sarà perciò $\mathbf{v}_\tau = -\mathbf{v}$ la velocità di trascinamento del sistema $\mathcal{R}'$ rispetto al sistema $\mathcal{R}$, coincidente con la velocità orbitale della Terra rispetto all'etere, $\mathbf{v}_\tau = \mathbf{v}_T$.
Possiamo allora applicare, ad entrambi i raggi di luce, il **teorema dei moti relativi (o legge di composizione delle velocità)**.

Per il raggio di luce L, che nel tragitto di andata da A a B viaggia contro l'ipotetica corrente d'etere ed al ritorno da B ad A nello stesso verso, avremo

$$\mathbf{c}(L) = \mathbf{c}'(L) - \mathbf{v},$$

ovvero

$$\mathbf{c}'(L) = \mathbf{c}(L) + \mathbf{v}. \tag{2.52}$$

Tenendo presente che le velocità dei due raggi di luce L e T rispetto alla corrente, cioè rispetto al sistema $\mathcal{R}$ solidale all'etere, hanno, per ipotesi, lo stesso, modulo, cioè $|\mathbf{c}(L)| = |\mathbf{c}(T)| = cost. = c$, se proiettiamo la (2.52) sull'asse x' otteniamo

$$|\mathbf{c}'(L)| = |\mathbf{c}(L)| \mp |\mathbf{v}| = c \mp v, \tag{2.53}$$

dove vale il segno - per l'andata da A a B, cioè quando la luce viaggia controcorrente, ed il segno + per il ritorno da B ad A quando essa viaggia nello stesso senso della corrente.
Avremo allora (valutando i tempi in $\mathcal{R}'$)[24]

$$t_{AB} = \frac{|AB|}{|\mathbf{c}'(L)|_{And.}} = \frac{\ell}{c-v}, \qquad t_{BA} = \frac{|BA|}{|\mathbf{c}'(L)|_{Rit.}} = \frac{\ell}{c+v},$$

ed il tempo totale di andata e ritorno per il raggio L sarà

$$t(L) = \frac{\ell}{c-v} + \frac{\ell}{c+v} = \frac{2\ell c}{c^2 - v^2} = \frac{2\ell}{c}\gamma^2, \tag{2.54}$$

dove abbiamo indicato con

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

il **fattore di Lorentz** già introdotto nel paragrafo precedente.
Per il raggio T avremo analogamente:

$$\mathbf{c}(T) = \mathbf{c}'(T) + \mathbf{v}_\tau = \mathbf{c}'(T) - \mathbf{v}, \tag{2.55}$$

perciò è possibile costruire un diagramma delle velocità (analogo a quello fatto nel paragrafo precedente nel caso del fiume: si riveda la Fig.2.6 oltre che la Fig.2.7) in cui $\mathbf{c}'(T) \equiv AD$ è diagonale del parallelogramma avente per lati i vettori $\mathbf{c}(T) \equiv AE$ e $\mathbf{v} \equiv AF$ supposti entrambi spiccati dal punto $A \equiv O'$. Tenendo conto che $\mathbf{c}'(T) \perp \mathbf{v}$ il triangolo ADF è rettangolo in A con cateti $|\mathbf{c}'(T)|$, $|\mathbf{v}|$ ed ipotenusa $|\mathbf{c}(T)|$.
Ricordando che la velocità di T, rispetto alla corrente, cioè rispetto al sistema $\mathcal{R}$ è $|\mathbf{c}(T)| = cost. = c$, applicando il il teorema di Pitagora al triangolo ADF si avrà:

$$|\mathbf{c}'(T)| = \sqrt{|\mathbf{c}(T)|^2 - |\mathbf{v}|^2} = \sqrt{c^2 - v^2}.$$

Perciò il tempo che impiega il raggio di luce T per andare da A a C (valutato in $\mathcal{R}'$) sarà allora:

$$t_{AC} = \frac{|AC|}{|\mathbf{c}'(T)|} = \frac{\ell}{\sqrt{c^2 - v^2}} = \frac{\ell}{c}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\ell}{c}\gamma. \tag{2.56}$$

[24]Come si è già detto, in meccanica classica **il tempo è assoluto**, perciò i tempi di andata e ritorno per i due raggi L e T possono essere valutati indifferentemente in $\mathcal{R}$ o in $\mathcal{R}'$.

Il tempo t_{CA}, di ritorno del raggio da C ad A, sarà ovviamente uguale a t_{AC}, quindi il tempo totale di andata e ritorno del raggio T è

$$t(T) = \frac{2\ell}{\sqrt{c^2 - v^2}} = \frac{2\ell}{c}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{2\ell}{c}\gamma. \tag{2.57}$$

Come si può constatare dalle (2.54), (2.57) i tempi $t(L)$ ed $t(T)$ di andata e ritorno dei due raggi sono diversi, ed anzi $t(L) > t(T)$, perché, come abbiamo già dimostrato, si ha $\gamma > 1$.
In altri termini: *Secondo la meccanica classica, a parità di cammini percorsi in andata e ritorno, impiegherebbe meno tempo il raggio che si muove perpendicolarmente alla presunta corrente d'etere rispetto a quello che invece si muove parallelamente ad essa.*
Il rapporto dei tempi, come nel caso delle barche, è anche in questo caso uguale al fattore di Lorentz:

$$\frac{t(L)}{t(T)} = \gamma(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{2.58}$$

Affinché tale fattore $\gamma(v)$ sia reale, è necessario che $1 - \frac{v^2}{c^2} > 0$, cioè $v < c$, il che significa che: **affinché il problema ammetta soluzione reale è necessario che la velocità della corrente d'etere sia sempre inferiore alla velocità c con cui la luce si propaga nell'etere stesso: ciò è senz'altro vero per la Terra.**
Osservazione:
È ovvio che, nota la velocità c della luce nell'etere, se si riuscisse a misurare con l'esperienza di Michelson-Morley, che abbiamo appena descritto molto sommariamente, il rapporto $\gamma = \frac{t(L)}{t(T)}$ dei tempi di andata e ritorno in A, dei due raggi di luce, si potrebbe determinare la velocità v della corrente d'etere. Si ha infatti per la (2.58)

$$v = c\sqrt{1 - \frac{1}{\gamma^2}}. \tag{2.59}$$

2.8 Conseguenze dell'esperimento di Michelson-Morley

Come abbiamo già osservato nel paragrafo precedente, Michelson, con grandissimo stupore, constatò per primo nel 1881 e con Morley, in un esperimento perfezionato nel 1887,[25] che il risultato di questa fondamentale esperienza ottica non era quello che ci si aspettava, cioè: *la luce nel vuoto non si comporta affatto come le due barche dell'esercizio svolto in precedenza, nel senso che essa, anziché impiegare tempi diversi, nei tragitti di andata e ritorno da A a B, e da A a C,* **impiega invece lo stesso tempo**, cioè

$$t(L) = t(T), \qquad \Leftrightarrow \qquad \frac{2\ell}{c}\gamma^2 = \frac{2\ell}{c}\gamma,$$

ciò che implica $\gamma = 1$. Essendo il fattore di Lorentz dato dalla (2.58), se esso è uguale ad 1, ne segue $\frac{v}{c} = 0$. Questo è possibile solo se $v = 0$ ovvero $c = \infty$. E siccome quest'ultima ipotesi non è verificata sperimentalmente, **se ne deduce che dovrebbe essere $v = 0$,** cioè l'etere dovrebbe essere in quiete rispetto alla Terra ovvero **la Terra dovrebbe essere in quiete nell'etere**!
Questa conclusione, a parte il suo carattere addirittura tolemaico, è incompatibile con le interpretazioni classiche sia del fenomeno dell'aberrazione astronomica, sia dell'effetto Doppler, le

[25] Lo stesso risultato si ebbe nei 50 anni successivi in cui l'esperimento fu ripetuto moltissime volte con tecniche sempre più moderne e sofisticate, da altri sperimentatori in tutto il mondo.

cui spiegazioni richiedono un etere in quiete rispetto alle stelle fisse. **E siccome la Terra è in moto rispetto alle stelle fisse, l'etere (se esistesse) non potrebbe essere in quiete rispetto al sistema delle stelle fisse e nello stesso tempo in quiete rispetto alla Terra!**
Osserviamo infine che, sia nel caso dell'esperienza di Michelson-Morley, sia nei fenomeni dell'aberrazione astronomica e dell'effetto Doppler, **abbiamo usato la legge classica di composizione delle velocità - la quale, è bene osservarlo, è conseguenza del concetto di tempo assoluto - la cui validità non era mai stata messa in dubbio**. *È stato con la Teoria della Relatività Ristretta di Einstein del 1905 - che ha comportato una profonda revisione dei concetti di* **spazio e tempo** *- che le precedenti contraddizioni scompaiono perché la legge di composizione delle velocità viene radicalmente cambiata in quanto viene abolito il concetto di tempo assoluto*.
Dall'esperienza di Michelson-Morley si possono dedurre alcune conseguenze estremamente importanti che ora esamineremo.
Nella realtà sembra che le cose vanno come se fossero rispettati i seguenti fatti:

1) **la velocità della luce ha lo stesso valore assoluto (lo stesso modulo) in qualunque sistema di riferimento inerziale**
$$|\mathbf{c}| = |\mathbf{c}'| = c, \tag{2.60}$$
dove $\mathbf{c}$ è la velocità della luce nel riferimento inerziale $\mathcal{R} = Oxyz$ solidale con l'etere, e $\mathbf{c}'$ la velocità della luce nell'altro riferimento $\mathcal{R}' = O'x'y'z'$ (quasi inerziale) solidale con la Terra. *Essa non si somma vettorialmente in maniera classica con nessun'altra velocità;*

2) **Non esiste nessun etere ove si propaga la luce**: essa si propaga nel vuoto senza il bisogno di alcun mezzo materiale!

3) Se la legge classica di composizione delle velocità non è applicabile alla luce, siccome essa è stata ricavata considerando il tempo assoluto, **abbiamo un motivo fondato per ritenere che l'ipotesi di un tempo assoluto da sempre accettata in fisica, sia ora, alla luce dei risultati sperimentali discussi, da mettere in dubbio!** Questa è una conclusione, a dir poco, sconvolgente, che richiederà un'analisi molto approfondita di tutto l'apparato sperimentale e teorico in nostro possesso.

A conclusione di questo capitolo, ed a completamento di ciò che abbiamo detto più sopra, ci piace riportare - a proposito del risultato negativo dell'esperimento di Michelson-Morley - alcune belle ed interessanti considerazioni del grande matematico italiano Tullio Levi-Civita che fu tra i più autorevoli interlocutori di Einstein, a livello mondiale, specialmente per quanto riguarda lo sviluppo della sua Teoria della Relatività Generale.[26]
"Durante parecchi lustri molto si arrovellarono i fisici per render conto del risultato negativo testé riferito. Una almeno delle premesse, su cui si fondava la previsione, doveva essere rigettata onde stabilire il buon accordo coi fatti osservati. Ma quale delle premesse, tutte così ovvie e spontanee? Non sembrano attaccabili le nozioni geometriche, né, per se stesse, quelle cinematiche, si dubitò in primo luogo che fosse troppo semplicistico il trattare la propagazione della luce come un ordinario movimento di velocità costante c; e, in particolare, che nei corpi in movimento intervenga qualche modificazione delle leggi ottiche, per cui non sia più il caso di applicare la regola di composizione delle velocità (Stokes e Ritz). Queste ipotesi portano però a notevoli complicazioni, ed anzi peggiorano, per altri fenomeni, l'accordo coi risultati sperimentali, che si voleva invece assicurare più generalmente, sacrificando se indispensabile, la semplicità del modello. Una spiegazione fu trovata, come è ben noto da Lorentz, in base alla teoria elettromagnetica della luce. Ma non tutti se ne appagarono. Le formule denunciavano **certi effetti (contrazione di tempo e di lunghezze)** *che, per quanto inapprezzabili nell'esperienza macroscopica quotidiana, apparivano in netto*

[26]Tullio Levi-Civita, *Nuova impostazione elementare della Relatività*, Atti del 1° Congresso dell'Unione Matematica Italiana, Firenze, 1937, pp. 69-87; cfr. anche *Opere Matematiche, Memorie e Note*, vol. V, Memoria XL, pp, 638-655, Zanichelli Editore, Bologna, 1970.

contrasto con **le abituali concezioni cinematiche**. *L'accogliere insieme queste e quelli costituiva un compromesso poco soddisfacente: il famoso* **coup de pouce**[27] *del Poincaré, accettabile soltanto perché riesce conforme all'esperienza, così nei limiti di approssimazione raggiungibili in meccanica, come nei più delicati fenomeni elettro-magnetico-ottici.*
Una interpretazione dell'ibrido connubio, senz'altro **luminosa sotto l'aspetto speculativo, spontanea per il pensiero galileiano che la informa, e grandiosa nelle sue conseguenze concrete**, *è quella fornita nel 1905 dall'Einstein,* **col suo principio di relatività.** *Egli ebbe il coraggio di romperla coi ripieghi, lasciandosi guidare dal convincimento che il principio ordinario (galileiano) di relatività, a norma del quale le leggi dei fenomeni meccanici sono le stesse rispetto a quanti si vogliano riferimenti, animati l'uno rispetto all'altro da traslazione uniforme, doveva, in qualche maniera,* **potersi estendere a tutti i fenomeni**, *in particolare al più elementare dei fenomeni ottici: la propagazione rettilinea della luce.*
Nello schema classico non è così, a norma del principio $\mathbf{v} = \mathbf{v}' + \mathbf{v}_\tau$ *del moto relativo. Einstein si propose di* **modificare lo schema classico in modo da poter estendere la relatività galileiana anche ai fenomeni luminosi**: *si intende che la modificazione doveva essere abbastanza insignificante, nell'ambito delle velocità piccole di fronte a c, che intervengono nell'ordinaria meccanica (terrestre e celeste), da riportare sensibilmente alla solita legge* $\mathbf{v} = \mathbf{v}' + \mathbf{v}_\tau$ *di composizione delle velocità.*
Un tale programma riconobbe l'Einstein essere realizzabile, anzi, in certo senso, addirittura realizzato dalla trasformazione di Lorentz. Bastò, come già per Copernico di fronte al sistema tolemaico, che egli ne desse una nuova e subito avvincente interpretazione".

2.9 APPENDICE 1: Alcune nozioni utili di Astronomia Sferica

Per capire bene l'argomento relativo al fenomeno **dell'aberrazione annua delle stelle**, sviluppato nel paragrafo 2.4, esponiamo qui, molto succintamente, alcune semplici (anche se non assolutamente necessarie) nozioni di Astronomia Sferica che ogni studente di matematica, fisica o ingegneria dovrebbe comunque conoscere. [28]
Dicesi **sfera celeste o sfera rappresentativa** Σ, una sfera ideale, avente il centro O nell'occhio dell'Osservatore (che può essere comunque un punto arbitrario dello spazio) e di raggio arbitrario (per comodità il raggio può pensarsi arbitrariamente grande) sulla quale vengono proiettati, nei punti rappresentativi $A, B, C, ...$, tutti i corpi celesti $A^*, B^*, C^*, ...$, che si trovano a distanze disparate dall'osservatore O. Su Σ si svolgono tutti quei fenomeni del cielo **che dipendono dalle sole direzioni** in cui sono visti gli astri (vedi Fig.2.8).

Ad ogni semiretta a uscente dal centro O della sfera si fa corrispondere il punto A in cui essa incontra la superficie, e reciprocamente ad ogni punto A della superficie si fa corrispondere la direzione $a = OA$. Ad ogni piano passante per il centro O della sfera si fa poi corrispondere il cerchio massimo secondo il quale ne taglia la superficie e reciprocamente ad ogni cerchio massimo sulla superficie della sfera si fa corrispondere un piano diametrale: quello contenente il cerchio massimo e passante per il centro O della sfera.
Allora all'angolo fra due semirette uscenti dal centro corrisponde l'arco di circolo massimo compreso fra i due punti che rappresentano le due semirette; ed all'angolo fra due piani diametrali corrisponde l'angolo sferico fra i due circoli massimi che li rappresentano, uguale all'angolo fra le due tangenti ai cerchi ora detti nei loro punti d'incontro, od uguale anche all'angolo tra le due rette d'intersezione dei due piani predetti con ciascuno dei piani tangenti alla sfera nei punti (diametralmente opposti) d'incontro dei due cerchi rappresentativi.

[27] La locuzione coup de pouce, nella lingua francese, significa letteralmente colpo di pollice ma anche, in senso figurato, "spintarella": si tratta della "spintarella" apportata da Poincaré alla teoria per renderla accettabile.

[28] Per queste nozioni, si può vedere qualunque trattato di Astronomia generale, per es. Cesare Barbieri, *Lezioni di Astronomia*, Zanichelli Editore, Bologna, 1999, Cap. 1° e Cap. 2°; oppure: Francesco Zagar, *Astronomia sferica e teorica*, Zanichelli Editore, Bologna, 1984, Cap. 1° e Cap. 2°; o ancora: P. Bakulin, E. Kononovic, V. Moroz, *Astronomia generale*, Editori Riuniti - Edizioni Mir (Mosca), 1984, Cap.1°.

Figura 2.8: Sfera Celeste come sfera rappresentativa.

Sulla sfera rappresentativa, come abbiamo detto, vengono proiettati tutti i corpi celesti per mezzo delle loro **direzioni o visuali** condotte dal centro, **senza tenere conto delle loro distanze**, e queste posizioni vengono chiamate **posizioni sferiche** o anche **posizioni apparenti**. Anche l'insieme dei moti celesti verrà allora proiettato sulla sfera ed il moto di ciascun astro così ottenuto sarà il suo **moto apparente**, il quale è evidentemente un moto angolare e **deve misurarsi per angoli e non in unità lineari**.

Nella rappresentazione mediante un globo o mediante un disegno la sfera celeste si immagina generalmente vista dall'esterno Fig.2.8, ma ci sono problemi in cui è più conveniente rappresentare la porzione di sfera che interessa, come vista dall'interno.
In relazione alla sfera celeste si definiscono tre direzioni fondamentali:

1) La prima di queste *dicesi* **verticale** *del luogo di osservazione.*
   Essa è definita come la direzione del *filo a piombo.* La verticale incontra la sfera celeste Σ in due punti diametralmente opposti, dei quali quello sopra l'osservatore è chiamato **Zenit** (Z), e quello sotto **Nadir** (Z'), voci queste derivanti dall'arabo.

Figura 2.9: Sfera Celeste, punti e cerchi fondamentali.

   Il piano normale alla verticale, passante per il luogo di osservazione e quindi per il centro O della sfera rappresentativa, è detto **piano orizzontale**. Esso è facilmente identificabile quale piano tangente alla superficie libera di un liquido in quiete nel luogo di osservazione. Il circolo massimo in cui esso incontra la sfera celeste si chiama **orizzonte astronomico** ed ha evidentemente per poli lo Zenit ed il Nadir. Esso può essere materialmente identificato col circolo che, per un occhio situato al livello del mare, separa cielo e mare; nella Fig.2.9 è indicato con $NESW(Nord - Est - Sud - Ovest)$.

2) La seconda direzione fondamentale è quella definita **dall'asse di rotazione della Terra.**
   Supponiamo di prolungare questo asse fino ad intersecare la sfera celeste, allora questa

sembra ruotare attorno alla retta che individua l'asse di rotazione e che viene chiamata **asse del mondo**. I suoi punti d'incontro con la sfera sono i *poli celesti*, chiamati **polo Nord celeste o boreale o artico** P (quello visibile dalle nostre regioni) o **polo Sud celeste o australe o antartico** P' (quello invisibile dalle nostre regioni). Essendo l'asse del mondo generalmente inclinato sul piano orizzontale, il polo boreale sta sopra questo piano ed è visibile, mentre il polo australe sta sotto ed è invisibile.
Il piano normale all'asse del mondo e passante per il luogo di osservazione si chiama **piano dell'equatore**; esso determina sulla sfera un circolo massimo, detto **equatore celeste**, il quale ha per poli il polo Nord P ed il polo Sud P' celesti. Insomma l'equatore celeste non è altro che la proiezione, sulla sfera celeste, dell'equatore terrestre.
- L'angolo $\widehat{PON}$ che l'asse OP della sfera celeste forma col piano orizzontale, ovvero l'angolo che la verticale del luogo forma col piano dell'equatore celeste, si chiama *latitudine del luogo di osservazione*; la indichiamo con ℓ, e la sua misura soddisfa alla limitazione $0^\circ \leq \ell \leq 90^\circ$.
- I cerchi della sfera celeste paralleli all'equatore celeste si chiamano semplicemente *paralleli celesti*; i cerchi paralleli all'orizzonte si chiamano, con parola araba, *almucantarat*.
- Il cerchio massimo della sfera celeste passante per i due poli celesti e per lo Zenit ed il Nadir, si chiama *meridiano celeste*.
- Il cerchio massimo passante per i due poli celesti e per la posizione apparente A di un astro qualsiasi, si chiama *cerchio orario dell'astro*.

3) La terza direzione fondamentale è **l'asse dell'eclittica**.
Si chiama **piano dell'eclittica** il piano passante per l'osservatore O e contenente l'orbita della Terra attorno al Sole. Il termine *eclittica* deriva dal fatto che le eclissi di Sole avvengono su questo piano. L'intersezione di questo piano con la sfera celeste determina un circolo massimo detto **eclittica.**
Osservando giorno per giorno il Sole nel suo cammino nel cielo, si trova che in un anno esso si sposta apparentemente fra le costellazioni, descrivendo un circolo massimo che è proprio l'eclittica (non si confonda questo moto apparente annuo del Sole con l'altro, più facile da osservare, quello diurno, cioè quello seguito dal Sole durante una giornata). Possiamo anche dire che l'eclittica è il luogo geometrico descritto dal punto proiezione del Sole sulla sfera celeste fatto dalla Terra, istante per istante, durante un intero anno.
La retta passante per il centro della sfera celeste e normale al piano dell'eclittica si chiama **asse dell'eclittica** e determina sulla sfera celeste **il polo Nord o boreale dell'eclittica** E (visibile dalle nostre regioni) ed il **polo Sud o australe dell'eclittica** E' (invisibile).

Essendo l'asse dell'eclittica inclinato sull'equatore celeste, quest'ultimo taglia obliquamente il circolo massimo rappresentato dall'eclittica e l'angolo formato da questi due piani fondamentali dicesi **l'obliquità dell'eclittica** ed il suo valore è $\varepsilon = 23^\circ\, 27'$ circa. L'obliquità dell'eclittica è ovviamente uguale anche all'angolo formato dall'asse del mondo con l'asse dell'eclittica.
I punti d'intersezione dell'equatore celeste con l'eclittica, cioè i cosiddetti **nodi dell'eclittica sull'equatore** sono chiamati **equinozi**, e precisamente, tenendo presente il moto apparente del Sole sull'eclittica, *il nodo ascendente* **equinozio di primavera o punto Vernale** γ (da Ver che in latino significa primavera) ed *il nodo discendente* **equinozio di autunno** γ'. Gli equinozi, sulla sfera celeste sono diametralmente opposti.[29]
Chiamasi **circolo di latitudine** qualunque circolo massimo passante per i due poli E (Nord) ed E' (Sud) dell'eclittica, in particolare per **circolo di latitudine di un astro** A della sfera celeste

[29] L'equinozio di primavera (Punto Vernale) si dice *nodo ascendente* perché il Sole, nel suo moto annuo sull'eclittica, in quel punto, passa da sotto a sopra l'equatore celeste e quindi "sale" ovvero "ascende". L'equinozio di autunno si dice invece *nodo discendente* perché il Sole passa da sopra a sotto l'equatore celeste e quindi "discende".

s'intende il circolo massimo EAE' che dal polo boreale E dell'eclittica va al polo australe E' passando per l'astro A.

Figura 2.10: Sistema eclitticale.

Chiamasi inoltre **parallelo di latitudine** qualunque circolo minore parallelo al cerchio massimo dell'eclittica. Nella Fig.2.10 il cerchio massimo $Q\gamma Q'$ rappresenta l'equatore celeste. Il cerchio massimo $S_i\gamma S_e$ rappresenta l'eclittica ed i suoi due punti S_i, S_e diametralmente opposti rappresentano rispettivamente **il solstizio d'inverno** ed **il solstizio d'estate**. *Equinozi e solstizi* $(S_i\gamma S_e\gamma')$ *formano i quattro punti cardinali dell'eclittica.*
In Astronomia si usano quattro sistemi fondamentali di coordinate, definiti sopra la sfera celeste: per i nostri scopi ne serve uno soltanto chiamato **il sistema eclitticale** che è quello necessario per lo studio dell'aberrazione della luce stellare e del quale ci serviremo.
Per **sistema eclitticale** s'intende il sistema di coordinate sferiche avente come primo cerchio massimo fondamentale, per la misura delle ascisse sferiche, **l'eclittica** e come secondo circolo massimo fondamentale, per la misura delle ordinate sferiche, il cosiddetto **cerchio di latitudine** $E\gamma E'$ relativo all'equinozio di primavera: questi due cerchi s'intersecano in γ che viene perciò scelto come origine del sistema di coordinate.
Le coordinate ortogonali sferiche di questo sistema sono: la **longitudine celeste** che è l'arco dell'eclittica tra l'equinozio di primavera γ ed il punto H d'intersezione dell'eclittica col semicerchio di latitudine EAE', cioè $\lambda \equiv \gamma H$, ; **la latitudine celeste** $\varphi \equiv HA$, arco del cerchio di latitudine passante per l'astro A compreso tra l'eclittica ed A. La longitudine celeste si conta da $0°$ a $360°$ sull'eclittica positivamente da γ nel verso in cui il Sole si muove su di essa, passando da sotto a sopra l'equatore celeste; la latitudine si conta positivamente, da $0°$ a $+90°$, dal punto H, intersezione dell'eclittica col cerchio di latitudine passante per A verso il polo boreale E della stessa; negativamente, da $0°$ a $-90°$ verso il polo australe. **La longitudine e la latitudine celesti**, (λ, φ), considerate in questo ordine, sono dette **le coordinate eclitticali** dell'astro A. Nelle considerazioni necessarie per capire il fenomeno dell'aberrazione della luce stellare è la latitudine celeste φ che interviene.

2.10 APPENDICE 2: Richiami sulla propagazione delle onde

Lo scopo di questa APPENDICE è di richiamare alcuni concetti basilari sulla propagazione di onde di qualsiasi natura necessari per poter capire bene il fondamentale **effetto Doppler**, sia nell'ambito della fisica classica, sia nell'ambito della Teoria della Relatività Ristretta.
Propagazione delle onde[30]
L'idea di **onda** è familiare a tutti, in quanto suggerita dall'esperienza quotidiana. Come si legge

[30]Sulla propagazione delle onde si consiglia il lettore di consultare uno dei tanti testi universitari di Fisica generale, per esempio il molto diffuso: D. Halliday, R. Resnick, K.S. Krane, *Fisica 1*, 4^a edizione, Casa Editrice Ambrosiana, Milano, 1993, pagg. 443-489.

in tutti i libri elementari di fisica e come abbiamo constatato più volte personalmente, quando un sasso viene fatto cadere in uno stagno, la superficie dell'acqua s'increspa formando onde circolari che si irradiano dal punto di caduta del sasso propagandosi sulla superficie calma dell'acqua; quando si pizzica una corda di chitarra o di violino, questa si mette in vibrazione ed emette onde sonore che si propagano nell'aria circostante; quando si batte sui tasti di un pianoforte, le corde di questo vibrano, e onde sonore si propagano nella stanza; quando una stazione radiotrasmittente o televisiva è in funzione, onde elettromagnetiche s'irradiano tutt'intorno all'antenna; si potrebbe così continuare a fare moltissimi esempi di propagazione di onde.
Tutti questi sono esempi di **moti ondosi**, di **onde**, di **propagazione per onde** ed hanno in comune due caratteristiche fondamentali:

1) c'è una propagazione di energia dalla sorgente delle onde ai vari punti dello spazio circostante;

2) il mezzo in cui le onde si propagano, qualunque esso sia, non partecipa alla propagazione.

Così, nell'esempio del sasso gettato nello stagno, onde che trasportano energia si propagano sulla superficie dell'acqua, ma, come si può constatare osservando il movimento di un piccolo galleggiante, per esempio di legno o di sughero, nello stagno, esso ha un movimento oscillatorio verticale in corrispondenza del passaggio delle onde, ma l'acqua non si sposta insieme con le onde.
Si può verificare che, qualunque sia la natura del mezzo in cui le onde si propagano, sia esso l'aria, una corda tesa, un liquido, un cavo elettrico, ecc., o addirittura il vuoto, le due caratteristiche fondamentali, sopra citate, consentono di mettere in luce, che tutte le onde, qualunque sia la loro natura, sono governate da una stessa equazione differenziale lineare alle derivate parziali del 2° ordine, detta appunto **equazione del moto ondoso o equazione delle onde**

$$\frac{\partial^2\Phi}{\partial x^2}+\frac{\partial^2\Phi}{\partial y^2}+\frac{\partial^2\Phi}{\partial z^2}-\frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}=0, \quad \Longleftrightarrow \quad \Delta_2\Phi-\frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}=0, \tag{2.61}$$

dove la **perturbazione ondosa** è rappresentata matematicamente da una funzione $\Phi(\mathbf{x},t)\equiv\Phi(x,y,z,t)$ delle coordinate spaziali e del tempo, mentre $v=|\mathbf{v}|$ è il modulo della velocità con cui si propaga l'onda. Per le generalità, di cui vogliamo qui occuparci brevemente, non è necessario precisare in che cosa consista la funzione $\Phi(\mathbf{x},t)$: potrebbe essere per es., l'altezza di un'onda in un liquido, o l'intensità di un campo elettrico o di un campo magnetico variabile, o l'altezza di un suono, ecc.
Il caso più semplice è quello unidimensionale in cui l'onda si propaga in una certa direzione che supponiamo coincidere con l'asse delle x di un dato sistema di coordinate. Dal momento che *si propaga*, Φ dipenderà in genere dal tempo t, e dall'ascissa x del particolare punto che si considera lungo la direzione di propagazione. Per $t=0$, Φ si riduce ad una funzione della sola x, $\Phi(x,0)\equiv\Phi_0=f(x)$, che chiameremo **profilo d'onda**. La denominazione è giustificata dal fatto che, se immaginiamo di *fotografare l'onda* nell'istante iniziale, l'istantanea ottenuta, rappresentativa quindi dell'onda nell'istante $t=0$, è appunto la linea di equazione cartesiana $\Phi_0\equiv y=f(x)$. Supponiamo che l'onda si propaghi senza deformarsi: una fotografia scattata in un istante successivo t, sarà identica a quella ottenuta all'istante $t=0$, salvo che l'onda risulterà spostata di un tratto vt nella direzione delle x crescenti. Se assumiamo come nuova origine il punto di ascissa $x=vt$ e indichiamo con X le nuove ascisse, avremo $X=x-vt$; l'equazione del profilo d'onda sarà $\Phi=f(X)$ se riferita al nuovo sistema di ascisse, con origine mobile lungo la direzione di propagazione, e

$$\Phi=f(x-vt), \tag{2.62}$$

se riferita al vecchio sistema, con origine fissa.
La (2.62) è la più generale equazione rappresentativa di un'onda che si propaghi, senza deformarsi, con velocità v nel verso positivo dell'asse x, cioè di **un'onda progressiva**: v è la **velocità di propagazione**.
Se l'onda procede nel verso negativo dell'asse x, l'equazione rappresentativa è ancora la (2.62) pur di sostituire v con $-v$, e cioè:

$$\Phi = f(x + vt). \tag{2.63}$$

In questo caso l'onda si dice **regressiva**.
Nel caso unidimensionale che stiamo trattando, l'equazione differenziale della propagazione ondosa (2.61) si riduce alla

$$\frac{\partial^2 \Phi}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 \Phi}{\partial t^2} = 0, \tag{2.64}$$

la cui soluzione generale, come si può dimostrare, è del tipo

$$\Phi = f(x - vt) + g(x + vt), \tag{2.65}$$

con f e g funzioni arbitrarie.
Dai corsi di Fisica del 1° biennio universitario sappiamo che la forma più semplice di un'onda unidimensionale che si propaghi con velocità v, senza deformarsi, nel verso positivo dell'asse x, e che conveniamo di rappresentare sul piano xy, è **l'onda armonica** o **onda sinusoidale**, caratterizzata dal fatto che, all'istante iniziale $t = 0$, **il profilo d'onda è una sinusoide o una cosinusoide**, avente perciò equazione

$$y = a \sin kx \quad \text{oppure} \quad y = a \cos kx,$$

dove sia a che k sono numeri reali di cui daremo un'interpretazione fisica più avanti. Così, se il profilo d'onda all'istante iniziale $t = 0$ ha equazione cartesiana

$$y = a \cos kx,$$

all'istante generico t, dato che essa si propaga con velocità v sarà:

$$y = a \cos k(x - vt). \tag{2.66}$$

Il valore massimo di y, cioè la grandezza $y_{max} = a$ che compare nelle equazioni ora scritte, si chiama **ampiezza** dell'onda; l'argomento del coseno

$$\phi = k(x - vt), \tag{2.67}$$

è la **fase** o **fase istantanea** o **fase all'istante** t; il valore $\phi_0 = kx$ della fase per $t = 0$ quello di **fase inziale**.
Com'è evidente, y assume lo stesso valore in punti dell'asse x che distano tra loro di $2\pi/k$ o di multipli interi di $2\pi/k$: vale a dire **il profilo d'onda si ripete a distanza regolare** di $2\pi/k$. Infatti, essendo il coseno una funzione periodica di periodo 2π, per un generico valore x_1 di x si avrà

$$y = a \cos k(x_1 - vt) = a \cos [k(x_1 - vt) + 2\pi] = a \cos k[(x_1 + \frac{2\pi}{k}) - vt] = a \cos k(x_2 - vt)$$

con $x_2 = x_1 + \frac{2\pi}{k}$, sicché la distanza fra x_1 e x_2 è $x_2 - x_1 = \frac{2\pi}{k}$.
Tale distanza si chiama **lunghezza d'onda** e si indica abitualmente con λ ($\lambda = \frac{2\pi}{k}$). La (2.66) può allora scriversi nella forma

$$y = a \cos \frac{2\pi}{\lambda}(x - vt), \tag{2.68}$$

che è la **forma tipica dell'equazione cartesiana di un'onda armonica progressiva** che si propaga lungo la direzione positiva dell'asse x con velocità v.
All'intervallo di tempo T impiegato dall'onda per percorrere la distanza λ si dà il nome di **periodo**:

$$T = \frac{\lambda}{v}. \tag{2.69}$$

Il periodo, com'è chiaro, rappresenta anche l'intervallo di tempo in cui y compie un ciclo completo di variazione o, se si vuole, in cui la fase varia di 2π. All'inverso, ν, del periodo,

$$\nu = \frac{1}{T}, \tag{2.70}$$

si dà il nome di **frequenza**: essa *rappresenta il numero di cili completi di variazione di y che si hanno nell'unità di tempo, cioè in un secondo.*
Dalle (2.69) e (2.70) si ottiene

$$v = \frac{\lambda}{T} = \lambda\nu. \tag{2.71}$$

Alla quantità

$$\omega = \frac{2\pi}{T} = 2\pi\nu = kv, \tag{2.72}$$

che rappresenta **la variazione di fase nell'unità di tempo**[31], si dà invece il nome di **pulsazione** o **frequenza angolare**. Tenuto conto delle precedenti definizioni, la (2.68) può essere scritta anche in una delle seguenti forme, equivalenti fra loro

$$y = a \cos 2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right), \tag{2.73}$$

$$y = a \cos 2\pi\left(\frac{x}{\lambda} - \nu t\right). \tag{2.74}$$

Talvolta è utile introdurre il cosiddetto **numero d'onda**: esso non è altro che il numero k, che compare già nell'equazione (2.66) e che esprime quante lunghezze d'onda entrano in 2π. Si ha ovviamente,

$$k = \frac{2\pi}{\lambda}, \tag{2.75}$$

e la (2.74), tenuto conto della (2.72) e che il coseno è una funzione pari, può essere scritta anche così

$$y = a \cos(kx - \omega t) = a \cos(\omega t - kx), \tag{2.76}$$

che è la forma che useremo nel seguito.
Osserviamo che le definizioni di lunghezza d'onda, periodo, frequenza, pulsazione e numero d'onda date più sopra, conservano la loro validità anche nel caso che l'onda non sia armonica, purché il profilo d'onda, quale che sia, si ripeta regolarmente lungo la direzione di propagazione.
È possibile generalizzare la (2.76), che rappresenta onde in una sola dimensione, al caso di onde piane oppure onde sferiche (o anche più generali) in tre dimensioni. Limitiamoci alla considerazione di **onde piane**.
Sia $\mathbf{x}(x, y, z)$ il vettore di posizione di un generico punto P in un riferimento inerziale $\mathcal{R}$ ed $\mathbf{n}(\alpha, \beta, \gamma)$ un vettore unitario che individua la direzione di propagazione dell'onda. È anche comodo introdurre il **vettore d'onda** $\mathbf{k} = k\mathbf{n}$ avente anch'esso la direzione di propagazione

[31] Infatti, **la variazione di fase tra l'istante t e l'istante $t+1$** è data dalla differenza $\Delta\phi \equiv \phi(t+1) - \phi(t) = k[x - v(t+1)] - k(x - vt) = kv = 2\pi\nu = \omega$, perciò **la variazione di fase nell'unità di tempo, essendo proporzionale (con fattore 2π) alla frequenza ν rappresenta il numero di cicli completi dell'onda, cioè fisicamente il numero di creste d'onda.**

dell'onda ed il cui modulo coincide col **numero d'onda** ($|\mathbf{k}| = k$).
Un'onda piana è quella nella quale la perturbazione $\Phi(\mathbf{x}, t)$, qualunque essa sia, ha lo stesso valore in tutti i punti di ogni piano perpendicolare alla direzione di propagazione individuata dal versore $\mathbf{n}$. Ognuno di tali piani è detto **fronte d'onda**, e si sposta parallelamente a sé stesso con la velocità v di propagazione dell'onda. L'equazione dei fronti d'onda è ad ogni istante, ovviamente

$$\mathbf{x} \cdot \mathbf{n} = \alpha x + \beta y + \gamma z = cost, \tag{2.77}$$

con $|\mathbf{n}| = \alpha^2 + \beta^2 + \gamma^2 = 1$.
Un'onda armonica in 3 dimensioni sarà rappresentata, in analogia con la (2.76), con l'equazione

$$\Phi = a \cos(\omega t - \mathbf{k} \cdot \mathbf{x}), \tag{2.78}$$

essendo $\mathbf{k}$ il vettore d'onda ed $\mathbf{x}$ il vettore di posizione di un generico punto del fronte d'onda. Se si tratta di un'onda elettromagnetica che si propaga nel vuoto, la sua velocità sarà uguale alla velocità della luce c e la relazione fra la pulsazione ω ed il numero d'onda k sarà, per la (2.72), $\omega = kc$. Inoltre la fase che compete all'equazione d'onda 3-dimensionale (2.78), in analogia alla (2.67) e tenuto conto che $\mathbf{k} = k\mathbf{n}$, è data da:

$$\phi = \omega\big(t - \frac{\mathbf{x} \cdot \mathbf{n}}{c}\big). \tag{2.79}$$

Inoltre, come abbiamo osservato nella nota 31 della pagina precedente, **la variazione di fase nell'unità di tempo, essendo proporzionale alla frequenza ν, rappresenta il numero di cicli completi dell'onda, cioè fisicamente il numero di creste d'onda.**

Capitolo 3

FONDAMENTI DELLA RELATIVITÀ RISTRETTA

Raffiniert ist der Herrgott, aber boshaft ist er nicht.[1]
Sottile è il Signore, ma non malizioso.
(Albert Einstein 1879-1955)

Non vi è alcuna differenza tra una qualsiasi delle tre dimensioni spaziali e il tempo, se si esclude il fatto che la nostra consapevolezza si muove lungo esso.... L'uomo è riuscito a salire contro la forza di gravità negli aerostati. Perché non dovrebbe sperare di riuscire infine a fermare o accelerare la sua corsa lungo la dimensione temporale, o anche di riuscire a girare e muoversi in senso opposto?
(G. H. Wells, *La macchina del tempo*)[2]

3.1 Concezione unitaria della Fisica: I nuovi Assiomi

Il fallimento di ogni tentativo di spiegare il risultato negativo dell'esperimento di Michelson-Morley nell'ambito della fisica classica ha condotto a dubitare di questa, coinvolgendo nel dubbio l'intero blocco degli assiomi fondamentali. Si è visto come, nella fisica classica, il principio di relatività di Galileo, valido per i fenomeni meccanici, non sussista per i fenomeni elettromagnetici. Questo diverso comportamento attribuito ai due ordini di fenomeni non è apparso ad Einstein conciliabile con una visione unitaria di tutta la fisica. In verità sarebbe ben arduo individuare in natura fenomeni puramente meccanici e fenomeni puramente elettromagnetici. In virtù di questa concezione unitaria dei fenomeni fisici, Einstein fu condotto ad ammettere che un principio di relatività non poteva sussistere se non esteso a tutti i fenomeni fisici.
Ma sentiamo cosa dice lo stesso Einstein a proposito **dell'apparente incompatibilità fra la legge di propagazione della luce e il principio di relatività.**[3]
"Di fronte a tale dilemma [appunto questa apparente incompatibilità] *pare che non vi sia modo di*

[1]Questa frase di Einstein diventata celebre è oggi scolpita nella pietra, sopra il camino della sala comune dell'Istituto di Matematica dell'Università di Princeton; si veda P.A.M. Dirac, *La bellezza come metodo*, Indiana Editore srl, 2013, pag.177.

[2]Questa citazione è riportata da Arthur Stanley Eddington nel suo libro *Spazio, Tempo e Gravitazione*, Universale Scientifica Boringhieri, 1971, pag. 95.

[3]A. Einstein, *Relatività, Esposizione divulgativa*, Universale scientifica Boringhieri, Torino (1967), pagg. 57-58. Cfr. anche Albert Einstein, *Opere scelte*, a cura di Enrico Bellone, Bollati Boringhieri Editore, Torino (1988), pagg. 401-403. Si tenga presente che questa esposizione divulgativa della Relatività è stata pubblicata da Einstein, nel 1917, col titolo: *Über die Spezielle und die Allgemeine Relativitätstheorie (Gemeinverständlich)*,Vieweg, Braunschweig, 1917.

uscirne se non abbandonando o il principio di relatività o la semplice legge di propagazione della luce nel vuoto. Quanti fra i lettori hanno attentamente seguito la precedente discussione si attenderanno certamente che debba venir conservato il principio di relatività, il quale, per la sua semplicità e naturalezza, si raccomanda alla mente come pressoché irrefutabile, e che invece la legge di propagazione della luce nel vuoto debba venir sostituita da una legge più complicata che si conformi al principio di relatività. Lo sviluppo della fisica teorica ha dimostrato però che non possiamo seguire questa strada. Le indagini teoriche profondamente innovatrici di H. A. Lorentz, sui fenomeni elettrodinamici ed ottici nei corpi in movimento, hanno dimostrato che le esperienze in questi campi conducono in modo inequivocabile ad una teoria dei fenomeni elettromagnetici, che ha per conseguenza necessaria la legge della costanza della velocità della luce nel vuoto. Fisici teorici eminenti furono perciò maggiormente inclini a lasciar cadere il principio di relatività, sebbene non si fosse trovato alcun dato empirico in contraddizione con questo principio.
A questo punto entrò in campo la Teoria della Relatività. Da una analisi dei concetti fisici di tempo e di spazio, risultò evidente che nella realtà non esiste la minima incompatibilità fra il principio di relatività e la legge di propagazione della luce, e che attenendosi strettamente e sistematicamente a entrambe queste leggi si poteva pervenire ad una teoria logicamente ineccepibile. Ad essa si dà il nome di **Teoria della Relatività Ristretta (o Particolare)**, *per distinguerla dalla Teoria della Relatività Generale".*[4]
In questo brano si ravvisa chiaramente *il colpo di genio* di Einstein nell'intuire che per avere una teoria *logicamente ineccepibile* era necessario accettare sia il principio di relatività, sia il principio della costanza della velocità della luce e contemporaneamente fare una profonda revisione critica dei concetti di *spazio e tempo*, ciò che egli fece, con mirabile e cristallina chiarezza nel lavoro del 1905, più volte citato: *Sull'elettrodinamica dei corpi in movimento.*

3.2 I Postulati della Teoria della Relatività Ristretta

Siamo ora pronti per enunciare i due postulati fondamentali sui quali Einstein ha fondato la sua Teoria della Relatività.
L'estensione del principio galileiano di relatività fatta da Einstein si può riassumere nel modo seguente:

> **1° Postulato (Principio di Relatività di Einstein (PRE)):**
> *Esistono in Natura infiniti riferimenti privilegiati ciascuno dotato di un proprio tempo - che continueremo a denominare inerziali o galileiani - in moto traslatorio rettilineo uniforme l'uno rispetto all'altro. Essi sono tra loro equivalenti rispetto alla formulazione di tutte le leggi fisiche. Nessuna esperienza fisica, svolta all'interno di un laboratorio galileiano, può permettere di riconoscere in quale di essi ci si trovi.*[5]

[4] È doveroso osservare che riguardo la nascita della Teoria della Relatività è assolutamente necessario fare i nomi di due grandissimi scienziati che vi contribuirono in maniera determinante: l'eminente fisico e premio Nobel olandese, Hendrik Antoon Lorentz ed uno fra i più grandi matematici della sua epoca, il francese Jules Henri Poincaré. In particolare Lorentz, già nel 1904 - imponendo la condizione matematica d'invarianza delle equazioni di Maxwell nel passaggio da un riferimento inerziale ad un altro - trovò le trasformazioni, che ora portano il suo nome, datogli successivamente da Poincaré, e riconobbe, assieme a Fitzgerald, anche il fenomeno della contrazione delle lunghezze. Poincaré invece, pubblicò nel 1905, qualche settimana prima di Einstein, l'enunciato del Principio di relatività riconoscendone l'importanza. Tuttavia né Lorentz, né Poincaré riuscirono ad abbandonare la concezione classica dello spazio e del tempo, rimanendo legati entrambi psicologicamente alle vedute ottocentesche e perciò non riuscirono a fare il grande salto: ciò è stato fatto invece con grande coraggio, ed in maniera così chiara e limpida, dal giovane Einstein con il suo lavoro del 1905, fondando una nuova teoria dello spazio e del tempo logicamente coerente e conforme ai dati sperimentali.

[5] Per l'esattezza storica ed a scanso di equivoci, dobbiamo avvertire il lettore che gli enunciati sia del 1° che del 2° postulato, da noi più sopra formulati, non sono esattamente quelli presentati da Einstein nella memoria del 1905, ma una nostra elaborazione fatta unicamente a scopo didattico: il contenuto concettuale è ovviamente esattamente lo stesso. Il lettore è pregato di consultare il lavoro originale.

Per evitare equivoci facciamo notare che il principio ora enunciato non significa che se gli osservatori inerziali $\mathcal{R}$ e $\mathcal{R}'$ fanno delle misure di grandezze fisiche (per esempio di dimensioni, temperatura, carica elettrica, ecc.) sopra lo stesso oggetto, essi debbono trovare lo stesso risultato: ciò, anzi, è generalmente falso. Il Principio di Relatività significa invece che se gli osservatori $\mathcal{R}$ e $\mathcal{R}'$ eseguono eguali esperimenti su *oggetti uguali e nelle stesse condizioni di moto o di quiete rispetto all'osservatore* (per es. entrambi misurano la carica dell'elettrone, ma ciascuno servendosi di un elettrone in quiete rispetto al proprio riferimento) essi debbono trovare lo stesso risultato. Cosa ancora più importante è osservare che, per la prima volta nelle storia della fisica, Einstein usa il *Il Principio di Relatività come Principio di simmetria o d'invarianza* nel senso che abbiamo spiegato nel Capitolo 1°, cioè: *le leggi della fisica devono essere invarianti nella trasformazione di coordinate (trasformazione di Lorentz) che fa passare dal riferimento inerziale $\mathcal{R}$ a quello $\mathcal{R}'$.*
Il Principio di Relatività di Einstein potrebbe apparire, a prima vista, sostanzialmente equivalente a quello galileiano. Tuttavia Einstein, conformemente alla sua visione unitaria dell'intera fisica, non si limita a considerarlo valido soltanto nell'ambito della meccanica ma ne estende la validità a tutti i fenomeni fisici; inoltre, egli *non pone, a priori, alcuna restrizione sui tempi relativi a riferimenti galileiani diversi*. Einstein assume quindi che il tempo possa essere differente in riferimenti diversi. Perciò il principio di relatività einsteiniano è, contemporaneamente, l'estensione di un assioma classico (*il Principio di Relatività di Galileo*) e, quasi a compenso, l'attenuazione di un altro (*quello del Tempo Assoluto*).
Il secondo postulato enunciato da Einstein, nella ferma convinzione della validità della teoria elettromagnetica di Maxwell-Lorentz, è il seguente:

> **2° Postulato (Principio della costanza della velocità della luce nel vuoto):**
> *In un sistema di riferimento galileiano i segnali luminosi nel vuoto viaggiano sempre rettilineamente alla velocità costante c, in tutte le direzioni (sistema "otticamente isotropo"), indipendentemente dalle modalità di emissione del segnale ed indipendentemente dallo stato di quiete o di moto della sorgente rispetto al riferimento considerato.*

Osservazione:
Col senno di poi, oggi alcuni autori, affermano che il 2° postulato poggia su almeno due fatti sperimentali incontrovertibili:
1) primo fra tutti, sulla corretta interpretazione dell'esperienza di Michelson-Morley (1881-1887) che **esclude**, tra l'altro, come un fatto sperimentalmente accertato, l'ipotesi che la velocità della luce possa dipendere dalla direzione "nello spazio". A tal proposito è opportuno osservare che non è certo, come già dicemmo, che Einstein, nel 1905, fosse a conoscenza di tale esperimento cruciale;
2) sul risultato di alcune osservazioni ottiche eseguite dopo la formulazione della Relatività Ristretta e delle quali Einstein non poteva perciò tenerne conto. Ci riferiamo alle osservazioni fatte, nel 1913 sulle stelle doppie, dall'astronomo olandese Willem de Sitter, il quale riuscì a dimostrare che **la velocità di propagazione della luce non può dipendere dalla velocità della sorgente in moto che la emette ed inoltre è la stessa per tutte le frequenze** (*A proof of the constancy of the velocity of light*, by Willem de Sitter, Procceding of the Royal Netherlands Academy of Art and Sciences, 1913, **15** II: 1287-1298, online).
Di quest'ultima prova tuttavia oggi si dubita. Ci sono alcuni autori ipercritici, come J. G. Fox (*Am. J. Phys.*, **30**, 297, (1962), **33**, 1, (1965), *J. Opt. Soc.*, **57**, 967, (1967)) ma anche T. Alvager, A. Nilsson e J. Kjellman, *Ark. Fys.*, **26**, 209, (1963)) che sostengono che: *sembra abbastanza chiaro che la maggior parte delle prime prove a favore del 2° postulato non sono valide a causa dell'interazione della radiazione con la materia che essa attraversa prima di essere rivelata.*
Fortunatamente esistono alcuni esperimenti, relativamente recenti, che si sottraggono alla critica di Fox ed altri. Il più conclusivo è un bellissimo esperimento eseguito presso il CERN di Ginevra nel 1964. Entro gli errori sperimentali, si trovò che: alcune particelle elementari in moto estremamente rapido come i pioni neutri, possono emettere simultaneamente due fotoni

in direzioni opposte, uno formante un angolo acuto, e l'altro un angolo ottuso con la direzione del movimento della particella. Una misura delle velocità dei due fotoni, resa possibile dalle tecniche attuali di laboratorio, conferma il loro identico valore c, malgrado che essi siano emessi da una sorgente veloce in direzioni opposte. Se le velocità misurate vengono scritte nella forma $c' = c + kv$, dove v è la velocità della sorgente, dall'esperimento si deduce che $k = (0 \pm 1,3) \times 10^{-4}$. **L'esperimento del CERN ha stabilito in modo definitivo e su scala di laboratorio la validità del 2° postulato della Relatività Ristretta.** Altri esperimenti su particelle cariche e neutrini verificano in modo indipendente la validità del 2° postulato (G. R. Kalbfleisch, N.Baggett, E. C. Fowler e J. Alspector, *Phys. Rev. Lett.*, **43**, 1361, (1979).
Combinando il principio di relatività con il 2° postulato della costanza della velocità della luce, è facile dimostrare che il secondo può essere riformulato in maniera più semplice:

> **Principio d'invarianza della velocità della luce nel vuoto**:
> *I segnali luminosi nel vuoto si propagano rettilineamente, con la stessa velocità costante c ad ogni istante, in tutte le direzioni, in tutti i sistemi inerziali.*

La dimostrazione di questo principio è abbastanza semplice. Siano infatti $\mathcal{R}$ e $\mathcal{R}'$ due sistemi di riferimento galileiani in moto reciproco TRU ed in configurazione x-standard. Siano S ed S' due segnali luminosi, emessi da due sorgenti differenti in direzioni arbitrarie, l'una in quiete rispetto a $\mathcal{R}$ e l'altra in quiete rispetto a $\mathcal{R}'$, rispettivamente. Per il principio di costanza della velocità della luce, i due segnali si propagano, rispetto a $\mathcal{R}$, con la stessa velocità (in modulo):

$$v_{\mathcal{R}}(S) = v_{\mathcal{R}}(S') = c. \tag{3.1}$$

D'altra parte, per il principio di relatività, la velocità del segnale S' rispetto ad $\mathcal{R}'$ deve essere uguale alla velocità di S rispetto ad $\mathcal{R}$:

$$v_{\mathcal{R}'}(S') = v_{\mathcal{R}}(S) = c. \tag{3.2}$$

Dal confronto delle (3.1) e (3.2) si deduce che deve risultare

$$v_{\mathcal{R}}(S') = v_{\mathcal{R}'}(S') = c. \tag{3.3}$$

Si vede così che uno stesso segnale, quello emesso dalla sorgente S' in quiete in $\mathcal{R}'$, ha la stessa velocità c sia nel riferimento $\mathcal{R}'$ in cui la sorgente è in quiete, sia nel riferimento $\mathcal{R}$ in cui si trova in moto. Lo stesso discorso si può ripetere per il segnale emesso dalla sorgente S in quiete in $\mathcal{R}$ ma in moto in $\mathcal{R}'$, così che resta dimostrato l'asserto.
La teoria di Einstein sostituisce dunque ad una grandezza come **il tempo assoluto** della fisica classica, un altro assoluto, **la velocità della luce nel vuoto**:
la velocità della luce nel vuoto è finita ed ha lo stesso valore in ogni sistema di riferimento inerziale:

$$c = c'. \tag{3.4}$$

3.2.1 Il tempo è relativo

La prima e notevolissima conseguenza dei due postulati di Einstein enunciati nel paragrafo precedente è che: **il tempo non è assoluto!**

$$t \neq t'.$$

Questo fatto è naturalmente e ovviamente sconvolgente per chiunque sia stato educato alla Meccanica di Galileo-Newton e perciò abituato da sempre a considerare **il tempo assoluto**, **la simultaneità assoluta**, e l'esistenza di **un'unica variabile temporale valida in qualsiasi sistema di riferimento.**

Vediamo come si può giungere alla conclusione che non può esistere un "tempo unico" valido per tutti gli osservatori, riportando un passo notevole, per semplicità e chiarezza, di L. D. Landau e E. M. Lifschitz, *Fisica Teorica 2, Teoria dei Campi*, Editori Riuniti, Roma (1976), pagg.15-16.
"È facile a questo punto convincersi quanto sia profonda la contraddizione tra il concetto di tempo assoluto e **il Principio Einsteiniano di Relatività.**[6] *E' sufficiente ricordare a questo proposito che in Meccanica Classica, fondata sul concetto di tempo assoluto, è valida la legge universalmente nota di composizione delle velocità secondo la quale la velocità di un moto composto è semplicemente uguale alla somma (vettoriale) delle velocità componenti. Essendo universale, questa legge dovrebbe essere applicabile anche alla propagazione delle interazioni. Risulterebbe allora che la velocità di propagazione in diversi sistemi di riferimento deve essere diversa. L'esperienza però conferma interamente sotto questo aspetto il principio di relatività. Le misure eseguite per la prima volta da Michelson (1881) rivelarono la totale indipendenza della velocità della luce dalla direzione della sua propagazione; secondo i postulati della Meccanica Classica, la velocità della luce nel verso della traslazione della Terra dovrebbe essere differente dalla velocità nel verso opposto. Così, il principio di relatività conduce a risultati secondo i quali il tempo non è assoluto. Il tempo scorre diversamente in diversi sistemi di riferimento. Di conseguenza, l'asserzione che due eventi dati sono separati da un intervallo di tempo determinato acquista un significato solo se è indicato il sistema di riferimento al quale si riferisce questa asserzione. In particolare gli "eventi simultanei" in un certo sistema di riferimento non lo saranno in un altro sistema".*
Le precedenti considerazioni di Landau e Lifschitz mettono in luce una seconda notevole conseguenza dei due postulati di Einstein: **l'esistenza di una "velocità limite" di propagazione delle interazioni, coincidente con la velocità della luce nel vuoto!**
Sul concetto di *velocità limite non superabile in alcun moto fisico*, che si può attribuire alla velocità della luce nel vuoto, ritorneremo con più dettagli in seguito.

3.2.2 Il concetto einsteiniano di tempo ed il concetto di simultaneità

Data l'importanza che il concetto di tempo riveste in tutta la Fisica, ci soffermeremo in questo paragrafo e nei due successivi a discuterne dettagliatamente la genesi e la misura.
Come abbiamo rilevato nei primi due paragrafi di questo capitolo, Einstein, nel primo lavoro sulla Teoria della Relatività Ristretta del 1905, estende *il Principio di Relatività di Galileo* dalla Meccanica all'intera Fisica non postulando però l'esistenza di un tempo assoluto come fece Newton (senza alcun riscontro sperimentale), postulando invece, su base sperimentale, *il Principio della costanza della velocità della luce nel vuoto.* Egli conserva immutato il concetto classico di sistema di riferimento inerziale, a cui dà il nome di **"sistema di coordinate galileiano"**, in onore a Galileo che fu il primo a riconoscere il principio d'inerzia ed il principio di relatività; assume però, più realisticamente, che riferimenti galileiani diversi $\mathcal{R}$, $\mathcal{R}'$, $\mathcal{R}''$, ..., debbano definire, al loro interno, ciascuno il proprio tempo t, t', t'', ... lasciando impregiudicato il legame che di fatto sussisterà fra questi tempi e che verrà scoperto applicando i due postulati e cioè con le

[6]Per capire chiaramente il discorso fatto da Landau e Lifschitz, che abbiamo riportato più sopra nel testo, è necessario osservare che essi introducono dapprima due Postulati (a pag. 14):

1) *il Principio di Relatività* esteso a tutti i fenomeni fisici;

2) *il Postulato dell'esistenza di una velocità limite di propagazione delle interazioni.*

Perciò essi, nel 2° principio, anziché porre l'attenzione sull'invarianza della velocità della luce nel vuoto, mettono in risalto *l'esistenza di una velocità limite di propagazione delle interazioni* ed impostano tutti i loro ragionamenti attorno a questo concetto. Successivamente, dicono testualmente: *"Il Principio di Relatività, insieme al Postulato dell'esistenza di una velocità limite di propagazione delle interazioni, è chiamato* **Principio di Relatività di Einstein** *(fu enunciato da Einstein nel 1905); ricordiamo qui che il Principio di Relatività di Galilei considera infinita la velocità di propagazione delle interazioni".* Landau e Lifschitz perciò unificano in un unico principio, appunto quello che loro chiamano **il Principio di Relatività di Einstein** i loro due postulati 1) e 2). Essi dunque si discostano un po' da ciò che la maggior parte degli autori indicano come i due Postulati della Teoria della Relatività Ristretta.

trasformazioni di Lorentz.
Ma sentiamo ora che cosa ne pensa Einstein del tempo, soprattutto perché le sue considerazioni, contenute nella memoria fondamentale del 1905, sono un capolavoro di prosa scientifica, inoltre sono di una semplicità e di una chiarezza sbalorditive, per cui vale veramente la pena di leggerle:[7]
"Noi dobbiamo considerare che tutti i nostri giudizi in cui interviene il tempo sono sempre giudizi su **eventi simultanei.** *Se io per esempio dico: «Quel treno giunge qui alle ore 7», ciò equivale a dire, in pratica: «Il posizionamento della lancetta del mio orologio sul 7 e l'arrivo del treno sono* **eventi simultanei**».[8]
Sembrerebbe che per superare tutte le difficoltà connesse alla definizione di tempo basti sostituire, in luogo della parola tempo, la posizione della lancetta del mio orologio. Una tale definizione è soddisfacente quando si tratti di definire un tempo soltanto per il luogo ove l'orologio si trova, ma non lo è più quando si tratti di correlare nel tempo eventi che si verificano in luoghi differenti, oppure - il che è lo stesso - di stabilire i tempi di eventi che si producono in luoghi lontani dall'orologio. Noi potremmo, certo, accontentarci di determinazioni di tempo effettuate nel modo seguente: un osservatore, munito di orologio e situato nell'origine delle coordinate, associa ad ogni segnale luminoso che gli giunge attraverso lo spazio vuoto, e che testimonia di un evento da valutare, la corrispondente posizione delle lancette dell'orologio. Ma un tale ordinamento ha lo svantaggio di non essere indipendente dalla posizione di questo osservatore, come noi sappiamo dall'esperienza. A una assai più pratica determinazione giungiamo per mezzo della seguente considerazione.
Supponiamo che nel punto A dello spazio sia collocato un orologio: in A, un osservatore può effettuare determinazioni di tempo, per eventi che si verifichino nelle immediate vicinanze del punto, controllando le posizioni delle lancette dell'orologio negli istanti stessi in cui quegli eventi si producono. Parimenti, un osservatore che si trovi nel punto B e che disponi di un orologio - e qui diciamo "strutturalmente identico al precedente" - potrà stabilire i valori di tempo di eventi nell'immediato intorno di B. Ma non è possibile, senza ulteriori convenzioni, paragonare temporalmente un evento in A con un evento in B; noi abbiamo finora definito un **tempo** *A e un* **tempo** *B, ma non un tempo comune ad A e B. Quest'ultimo tempo può solo venir definito quando si stabilisca,* **per definizione**, *che il "tempo" che la luce impiega nel percorso da A a B è uguale al "tempo" che essa impiega nel percorso da B ad A. Supponiamo cioè che un raggio di luce parta da A, diretto verso B, al tempo A t_A, venga, in B, riflesso verso A al tempo B t_B, e giunga nuovamente in A al tempo A t'_A.* **Per definizione, i due orologi sono sincronizzati (ovvero camminano sincroni) quando**

$$t_B - t_A = t'_A - t_B. \tag{3.5}$$

Noi ammettiamo che questa definizione del sincronismo sia libera da contraddizioni ed applicabile ad un numero qualsivoglia di punti, e che siano universalmente valide le seguenti relazioni:

1) *Se l'orologio in B è sincronizzato con l'orologio in A, l'orologio in A è sincronizzato con l'orologio in B [***proprietà simmetrica della sincronizzazione***].*

2) *Se l'orologio in A è sincronizzato sia con l'orologio in B che con l'orologio in C, sono sincronizzati tra loro anche gli orologi in B e in C [***proprietà transitiva della sincronizzazione***].*

[7] A. Einstein, *Zur Elektrodymamik bewegter korper*, Annalen der Physik, **17**, 1905, pagg. 891-921; traduzione italiana a cura di Paolo Straneo: *Sull'elettrodinamica dei corpi in moto*, pagg. 480-482, inserito nel volume "Cinquant'anni di Relatività", Vallecchi Editore, Firenze (1955). Cfr. anche A. Einstein, *Opere scelte*, a cura di E. Bellone, Bollati Boringhieri Editore, Torino (1988), pagg. 150-151. Si tratta del primo lavoro in assoluto sulla Relatività; si raccomanda vivamente la lettura di questo meraviglioso lavoro che costituisce un pietra miliare nella fisica. Non è esagerato affermare, come dicono giustamente Hans Ohanian e Remo Ruffini, che: "esso contiene quasi tutto della Teoria della Relatività Speciale e che ai seguaci di Einstein non rimaneva altro che aggiungere l'interpretazione geometrica dello spaziotempo e la teoria della rappresentazione del gruppo di Lorentz!"

[8] Non tratteremo qui l'inesattezza inerente al concetto di simultaneità di due eventi che si verificano (approssimativamente) nello stesso luogo, e che può essere solo rimossa per astrazione.

Avendo così stabilito, grazie a esperimenti fisici ideali, che cosa si debba intendere con sistema di orologi in quiete [stazionari], disposti in luoghi diversi e sincronizzati tra loro, abbiamo evidentemente dato una definizione di **simultaneità** *e di* **tempo**.

Definizione einsteiniana del tempo:
Il tempo di un avvenimento è l'indicazione [la lettura, ovvero la posizione delle lancette], simultanea al prodursi dell'evento stesso, di un orologio stazionario, situato nel luogo dell'evento e sincronizzato, per ogni determinazione temporale, con un ben preciso orologio stazionario [orologio campione].

Stabiliamo ancora, secondo l'esperienza, che la grandezza

$$\frac{2\overline{AB}}{t'_A - t_A} = c, \tag{3.6}$$

sia una costante universale: la velocità della luce nello spazio vuoto".
Si osservi che tale grandezza è definita utilizzando soltanto il concetto geometrico di distanza fra due punti e la nozione di tempo locale definito in A. Tale risultato sperimentale è assolutamente indipendente dall'ubicazione, nel sistema cartesiano inerziale $Oxyz$, dei punti A e B.
Osservazione:
La definizione (3.5) di sincronismo (e quindi di simultaneità), data da Einstein, equivale ad ammettere che la differenza fra il tempo t_B, segnato dall'orologio di B quando il segnale luminoso giunge in B, ed il tempo t_A segnato dall'orologio di A alla partenza del segnale da A, è uguale alla metà del tempo totale di andata e ritorno del segnale partito da A, cioè:

$$t_B - t_A = \frac{t'_A - t_A}{2}. \tag{3.7}$$

Facciamo osservare che t_A e t'_A sono tempi segnati dallo stesso orologio in A, mentre t_A e t_B sono tempi segnati da due orologi situati in punti diversi ma di costruzione perfettamente identica e sincronizzati.
Se ora combiniamo la definizione di sincronismo, espressa nella forma (3.7), con la (3.6) che definisce la velocità della luce nel vuoto - ed esprime nel contempo **l'isotropia ottica dello spazio** in $\mathcal{R}$ - è facile dedurre che

$$t_B = t_A + \frac{\ell}{c}, \qquad (\ell = \overline{AB}) \tag{3.8}$$

la qual relazione, peraltro del tutto intuitiva, suggerisce il modo operativo di sincronizzare l'orologio di un arbitrario osservatore posto in B con un determinato orologio campione collocato in A. Basta che l'osservatore posto in A invii, all'istante t_A segnato dal proprio orologio, verso B un segnale elettromagnetico e che l'osservatore in B, nell'istante in cui riceve il segnale aggiusti il proprio orologio - di costruzione perfettamente identica a quella di A - sul tempo $t_B = t_A + \frac{\ell}{c}$ (è ovviamente sottinteso che i due osservatori conoscano la loro distanza $\overline{AB} = \ell$ e siano preventivamente d'accordo sul tempo t_A d'invio del segnale da parte di A).

3.2.3 Ulteriori riflessioni sulla sincronizzazione degli orologi

Tutto ciò che Einstein ha detto sul concetto di tempo, sulla simultaneità e sulla sincronizzazione degli orologi e che abbiamo riportato nel paragrafo precedente è sufficientemente chiaro e non ci sarebbe bisogno di ulteriori commenti.
Desideriamo però soffermarci ancora sull'argomento per puntualizzarne, dal punto di vista didattico ed a beneficio dello studente, i passi necessari per giungere ad una definizione coerente

di "tempo esteso" definito in luoghi distanti fra di loro.
Poniamoci entro un **sistema di coordinate galileiano** $\mathcal{R} \equiv Oxyz$, così com'è concepito nella meccanica di Galileo-Newton, e definiamo nell'ordine:

a) un tempo locale (o individuale) segnato da un dato **orologio campione**, in quiete nel sistema di riferimento $\mathcal{R}$, per esempio nella sua origine O, e anche una **unità di misura del tempo;**

b) la velocità di un segnale luminoso (*velocità della luce nel vuoto*) sulla sola base dei concetti di lunghezza e di tempo locale già introdotti in $\mathcal{R}$ e la nozione di **spazio otticamente isotropo in** $\mathcal{R}$;

c) un procedimento di sincronizzazione degli orologi in quiete in $\mathcal{R}$, e la definizione di un **tempo esteso o universale o pantopico**, che diremo semplicemente **il tempo di** $\mathcal{R}$.

Discutiamo ora dettagliatamente punto per punto le tre fasi precedenti.

1) **Costruzione di una variabile temporale (tempo locale)**
Per maggiore chiarezza conviene fissare dapprima l'attenzione su di **un ben determinato osservatore** $\mathcal{O}$, supposto in quiete nell'origine di un dato sistema inerziale di coordinate $Oxyz$. Alla successione delle sue sensazioni un tale osservatore coordinerà una successione di istanti e quindi, passando dal discreto al continuo, una variabile indipendente t (**tempo individuale o locale**). E se in origine tale coordinamento sarà avvenuto presumibilmente attraverso **fenomeni ritmici** eminentemente individuali e grossolani: (*battito del polso, avvicendarsi di giorni o di stagioni, ecc.,*), successivamente l'osservatore sentirà la necessità di criteri impersonali, quali possono essere forniti da *dispositivi del tipo delle clessidre o da pendoli o, comunque, da ben determinate vibrazioni spontanee od opportunamente disciplinate*. Infine, il nostro osservatore, per ottenere una coordinazione più raffinata, si costruirà un buon orologio, secondo certe modalità standard, in maniera tale da poter misurare **il suo "tempo locale"** e fisserà anche una **unità di misura** del tempo;
il "tempo locale" così definito è, per l'osservatore, la posizione delle lancette del suo orologio.

2) **Propagazione rettilinea della luce nel vuoto con velocità costante c: "isotropia ottica dello spazio"**
Il nostro problema è ora quello di unificare i tempi, in modo da arrivare ad una variabile unica t, *tempo universale o pantopico*, comune a tutti gli osservatori in quiete nel dato sistema di riferimento $\mathcal{R} \equiv Oxyz$.
Per far ciò è necessario utilizzare la (3.6) supponendo che A coincida con l'origine delle coordinate $A \equiv O$ e B sia un punto arbitrario in quiete nel sistema galileiano $\mathcal{R}$, sia inoltre nota la distanza $\overline{OB} = \ell$. Supponiamo che l'osservatore che si trova in O faccia partire verso B un raggio di luce all'istante t_O, indicato dal proprio orologio, questo giunga in B dove viene istantaneamente riflesso e torni in O all'istante t'_O, avremo allora:

$$c = \frac{2\overline{OB}}{t'_O - t_O} = \frac{2\ell}{t'_O - t_O}. \tag{3.9}$$

Tale risultato sperimentale è assolutamente indipendente dall'ubicazione, nel sistema cartesiano $Oxyz$, del punto B. Tutto ciò può essere espresso dicendo che si ritiene sperimentalmente provato che rispetto ad un dato sistema galileiano $Oxyz$, **"lo spazio è otticamente isotropo", cioè la luce - o un'onda elettromagnetica qualsiasi - si propaga in linea retta alla stessa velocità costante c in tutte le direzioni, e perciò impiega tempi uguali per percorrere distanze uguali.**
Il valore sperimentale esatto della velocità della luce nel vuoto è stato stabilito, con una convenzione internazionale del 1983, in $c = 299.792.458\,\mathrm{m/s}$ cioè circa $(300.000\,\mathrm{km/s})$.

3) **Sincronizzazione degli orologi in quiete in $\mathcal{R}$**, e definizione di un **"tempo esteso o universale o pantopico"**, che diremo semplicemente **il tempo di $\mathcal{R}$**.
Tutti sappiamo benissimo, anche se in maniera empirica ed approssimata, come in pratica si proceda alla sincronizzazione degli orologi; infatti si tratta di un'operazione che facciamo tutti i giorni quando accordiamo il nostro orologio – che va un po' indietro o un po' in avanti – col segnale orario trasmesso dalla radio o dalla televisione ed irradiato da una stazione centrale i cui orologi campione (precisissimi orologi atomici al cesio) riteniamo attendibili (in Italia tale stazione è l'Istituto Nazionale di Ricerca Metrologica (l'INRIM), ex Istituto Galileo Ferraris di Torino che ha appunto il compito di diffondere, su tutto il territorio nazionale, l'ora esatta). L'operazione che noi compiamo è semplice in quanto si tratta di spostare in avanti o indietro le lancette del nostro orologio di qualche minuto o di qualche secondo in maniera tale da accordare (sincronizzare) il nostro orologio su quello campione. In tale operazione trascuriamo però il tempo che impiega il segnale elettromagnetico fra la stazione che trasmette il segnale ed il luogo in cui noi lo riceviamo; tale tempo viene da noi trascurato consapevolmente perché sappiamo che esso è veramente molto piccolo data l'altissima velocità con cui viaggiano le onde elettromagnetiche e tenuto conto che le distanze fra le varie località sulla Terra che ricevono il segnale orario sono abbastanza limitate: per esempio la massima distanza fra l'INRIM di Torino, che trasmette il segnale, e la più remota località d'Italia, rispetto a Torino, che è l'isola di Lampedusa, non supera i 1200 km, sicché il tempo trascurato è di soli 1200/300.000 s = 0,004 s, cioè di appena quattro millesimi di secondo!

Traduciamo ora queste considerazioni intuitive ed approssimate in un procedimento rigoroso e realistico per definire un **tempo universale unico** per il sistema di riferimento $\mathcal{R}$.[9]
Supponiamo che due osservatori simili $\mathcal{O}$, $\mathcal{O}'$, in quiete relativa l'uno rispetto all'altro, posseggano degli orologi perfettamente identici e che si scambino dei segnali elettromagnetici.[10]
Supponiamo di assumere, per convenzione, che l'osservatore $\mathcal{O}$ funga da **osservatore fondamentale** avente il compito di diffondere i segnali orari a tutti gli altri osservatori supposti in quiete rispetto ad esso, e supponiamo di collocare tale osservatore nell'origine del nostro sistema $\mathcal{R}$ di coordinate. Supponiamo che $\mathcal{O}$ invii un segnale orario all'istante t_0, segnato dal proprio orologio, e che l'osservatore $\mathcal{O}'$ distante ℓ da $\mathcal{O}$, riceva tale segnale all'istante t'_0 del proprio orologio. Siccome il segnale per raggiungere $\mathcal{O}'$ impiega il tempo $\frac{\ell}{c}$, quando il segnale raggiunge $\mathcal{O}'$ l'orologio di $\mathcal{O}$ segnerà il tempo $t_0 + \frac{\ell}{c}$, perciò se capita che $t'_0 = t_0 + \frac{\ell}{c}$, i due orologi sono già sincronizzati e se sono di costruzione perfettamente identica noi ipotizziamo che, da quel momento in poi, essi marceranno sempre con lo stesso identico ritmo, cioè continueranno a mantenere la sincronizzazione indefinitamente. Questo caso, com'è ovvio, si verifica però eccezionalmente, nella stragrande maggioranza dei casi capiterà che $t'_0 \neq t_0 + \frac{\ell}{c}$, nel qual caso sarà

$$t'_0 - \left(t_0 + \frac{\ell}{c}\right) = T(\mathcal{O}, \mathcal{O}') > 0, \ \textit{oppure} \ t'_0 - \left(t_0 + \frac{\ell}{c}\right) = T(\mathcal{O}, \mathcal{O}') < 0, \tag{3.10}$$

[9]Cfr.Tullio Levi-Civita, *Nuova impostazione elementare della Relatività*, Atti del 1° Congresso dell'Unione Matematica Italiana, Firenze, 1937, pagg. 69-87; cfr. anche *Opere Matematiche, Memorie e Note*, vol. V, Memoria XL, pagg. 638-655, Zanichelli Editore, Bologna, 1970.

[10]Osserviamo che ai fini della sincronizzazione di orologi in quiete in un dato sistema di riferimento e fra di loro lontani non è necessario trasportare uno di essi nel luogo occupato da un altro, ma si può ipotizzare, almeno concettualmente, che ciascun osservatore fabbrica il proprio orologio con le stesse identiche modalità seguendo le istruzioni contenute in un apposito **Manuale di costruzione degli orologi**, del quale ogni osservatore deve possederne una copia.
La sincronizzazione di orologi che si trovano nello stesso luogo è del tutto ovvia e viene comunque assunta come concetto primitivo.

essendo ovviamente $T(\mathcal{O}, \mathcal{O}')$ la differenza fra il tempo segnato dall'orologio di $\mathcal{O}'$ e quello segnato dall'orologio di $\mathcal{O}$.

Se $T(\mathcal{O}, \mathcal{O}') > 0$, ciò equivale a dire che $t'_0 > t_0 + \frac{\ell}{c}$ e perciò l'orologio di $\mathcal{O}'$ è avanti rispetto a quello di $\mathcal{O}$, dunque, per ottenere la sincronizzazione, le lancette dell'orologio di $\mathcal{O}'$ devono essere portate indietro di T secondi. Se invece risulta $T(\mathcal{O}, \mathcal{O}') < 0$, ciò equivale a dire che $t'_0 < t_0 + \frac{\ell}{c}$ e perciò l'orologio di $\mathcal{O}'$ è indietro rispetto a quello di $\mathcal{O}$, dunque, per ottenere la sincronizzazione, le lancette dell'orologio di $\mathcal{O}'$ devono essere portate avanti di $|T| = -T$ secondi.
In tal modo abbiamo sincronizzato tutti gli orologi che si trovano in quiete in $\mathcal{R}$ e, nello stesso tempo, costruito un'unica variabile temporale t ("tempo esteso" o "universale" o "pantopico") valida per tutti gli osservatori associati ai singoli orologi nel dato sistema di riferimento $\mathcal{R}$: lo chiameremo semplicemente "il tempo di $\mathcal{R}$".
Gli osservatori di ogni altro sistema galileiano procederanno in maniera analoga alla definizione di un tempo esteso all'interno del loro sistema: avremo dunque tante variabili temporali quanti sono i sistemi galileiani in considerazione.

3.2.4 Ancora alcune riflessioni sul concetto einsteiniano di tempo e di simultaneità

Nei paragrafi precedenti abbiamo esposto il **concetto di "tempo"** così come Einstein lo ha introdotto nella fisica col lavoro fondamentale del 1905, a cui abbiamo aggiunto, per completezza, il procedimento di sincronizzazione di orologi situati in punti lontani di un dato sistema di riferimento inerziale.
Ora desideriamo, a completamento di tutto ciò che abbiamo esposto fin qui, riportare integralmente, perché molto interessante ed illuminante, quello che Einstein dice del **concetto di tempo e di simultaneità** nella sua esposizione divulgativa della Relatività.[11]

Figura 3.1: Sul concetto di simultaneità.

"Il fulmine ha colpito le rotaie di una linea ferroviaria in due punti A e B, molto lontani l'uno dall'altro (Fig.3.1). Aggiungo l'affermazione che i due colpi di fulmine sono avvenuti simultaneamente. Ove ponessi al lettore la domanda se tale mia affermazione è sensata, la risposta sarebbe un reciso "si". Se ora però gli chiedessi di spiegarmi in maniera più precisa il senso di quell'affermazione, egli, dopo un po' di riflessione, troverebbe che la risposta a tale domanda non è così facile come sembra a prima vista. Dopo un certo tempo gli verrebbe forse in mente di rispondere così: "Il significato dell'affermazione è di per sé chiaro e non necessita alcuna spiegazione ulteriore; dovrei naturalmente fare una certa attenzione qualora ricevessi l'incarico di determinare, mediante l'osservazione, se nel caso concreto i due eventi abbiano avuto luogo simultaneamente oppure no". Non posso dichiararmi soddisfatto di tale risposta per la seguente ragione. Supponendo che un abile meteorologo fosse riuscito a scoprire, dopo ingegnose considerazioni, che il fulmine deve sempre colpire simultaneamente i due punti A e B, si presenterebbe

[11] A. Einstein, *Relatività, esposizione divulgativa*, Universale Scientifica Boringhieri, Torino (1967), pagg. 58-61.

allora il compito di controllare se questo risultato teorico sia davvero o no in accordo con la realtà. Ci imbattiamo nella stessa difficoltà in tutti gli enunciati fisici ove entra in gioco il concetto di **simultaneo**. *Questo concetto non esiste per il fisico fino a quando egli non ha la possibilità di scoprire nel caso concreto se tale concetto si verifichi oppure no. Abbiamo perciò bisogno di una definizione di simultaneità capace di fornirci il metodo per mezzo del quale decidere sperimentalmente, nel caso attuale, se entrambi i colpi di fulmine sono avvenuti simultaneamente oppure no. Finché non viene soddisfatto tale requisito, io, come fisico (e lo stesso vale naturalmente anche per il non fisico), mi abbandono a un inganno, quando immagino di poter attribuire un significato all'affermazione di simultaneità. (Vorrei chiedere al lettore di non procedere oltre finché non sia pienamente convinto su questo punto). Dopo una certa riflessione, il lettore farà la seguente proposta, per verificare la simultaneità. Con una misurazione effettuata lungo le rotaie, verrà calcolato l'intervallo che collega i punti A e B, e verrà messo un osservatore nel punto di mezzo M dell'intervallo AB. Quest'osservatore verrà fornito di un dispositivo (per esempio due specchi inclinati a 90°) che gli permetta di osservare visualmente i due punti A e B contemporaneamente. Se l'osservatore percepisce i due bagliori del fulmine nel medesimo istante, essi saranno allora simultanei (Fig.3.1). Sono molto soddisfatto di questo suggerimento, ma ciononostante non posso considerare la questione completamente risolta, giacché sono costretto a sollevare la seguente obiezione: "La definizione sarebbe certamente corretta, se io sapessi di già che la luce, per mezzo della quale l'osservatore situato in M percepisce i bagliori del fulmine, viaggia lungo l'intervallo $A \to M$ con la stessa velocità con cui viaggia lungo l'intervallo $B \to M$. Una verifica di quest'ipotesi sarebbe però possibile se avessimo già a nostra disposizione i mezzi per misurare il tempo. Si direbbe dunque che ci muoviamo in un circolo vizioso." Dopo aver di nuovo riflettuto, il mio ipotetico interlocutore mi getta (a ben ragione) un'occhiata un po' sprezzante e dichiara: "Continuo nondimeno a sostenere la mia definizione precedente. Perché in realtà essa non suppone assolutamente nulla riguardo alla luce. Esiste soltanto una cosa che si deve richiedere alla definizione di simultaneità, e precisamente questa: che in ogni caso concreto essa ci fornisca un metodo empirico per decidere se il concetto da definirsi si verifichi oppure no. È indiscutibile che la mia definizione soddisfa questo requisito. Che la luce richieda lo stesso tempo tanto per compiere il percorso $A \to M$ quanto per compiere il percorso $B \to M$, non è nella realtà né una supposizione né un'ipotesi circa la natura fisica della luce,* **bensì una convenzione** *che io posso fare a mio arbitrio al fine di giungere a una definizione di simultaneità." È chiaro che questa definizione può venir usata per dare un significato esatto non soltanto a due eventi, ma a quanti eventi vogliamo, e indipendentemente dai luoghi in cui gli eventi sono situati rispetto al corpo di riferimento (in questo caso la linea ferroviaria). Siamo in tal modo condotti ad una* **definizione di tempo** *nella fisica. Supponiamo a questo scopo che orologi di costruzione identica vengano collocati nei punti A, B e C della linea ferroviaria (sistema di coordinate), e che essi vengano regolati in modo che le loro lancette abbiano simultaneamente (nel senso suddetto) le medesime posizioni. In queste condizioni* **noi intendiamo per tempo di un evento la lettura (posizione delle lancette) di quello fra tali orologi che si trova nell'immediata vicinanza (spaziale) dell'evento in esame**. *In tal modo viene associato ad ogni evento un valore del tempo per principio suscettibile di osservazione. Questa convenzione contiene un'ulteriore ipotesi fisica, la cui validità non potrà venir posta in dubbio salvo che sulla base di una prova empirica del contrario. Si è supposto che tutti quegli orologi camminino con la stessa rapidità se sono di costruzione identica. Più esattamente diremo: quando due orologi, collocati fermi in luoghi differenti di uno stesso corpo di riferimento, vengono regolati in maniera che una posizione particolare delle lancette di uno degli orologi sia simultanea (nel senso suddetto) con la stessa posizione delle lancette dell'altro orologio, allora le posizioni identiche delle lancette saranno sempre simultanee (nel senso della definizione più sopra riferita)".*

3.2.5 Una terza definizione di simultaneità più aderente al senso comune

Il concetto di simultaneità è stato definito da Einstein sia nel lavoro del 1905 sia nell'esposizione divulgativa della Relatività riportata nel paragrafo precedente, con due procedimenti diversi ma fra loro equivalenti.
In questo paragrafo desideriamo soffermarci ancora su questo importantissimo concetto, in

modo che sia ben assimilato dal lettore. A tal fine enunciamo esplicitamente tre definizioni, la prima delle quali è quella di Einstein vista nel paragrafo precedente, la seconda è quella data da Landau e Lifschitz e la terza è dovuta al relativista francese Henri Arzeliés.

1^a **Definizione di simultaneità:**[12]
Siano A e B due punti in quiete rispetto ad $\mathcal{R}$, e sia M il punto medio del segmento AB; supponiamo che da A, B vengano emessi verso M due segnali luminosi, se essi giungono in M nello stesso istante, diremo che i due eventi, consistenti nell'emissione di tali segnali da A e da B, sono **eventi simultanei in** $\mathcal{R}$.

2^a **Definizione di simultaneità:**[13]
Siano A e B due punti in quiete rispetto ad $\mathcal{R}$, e sia M il punto medio del segmento AB; supponiamo che da M vengano emessi, nello stesso istante, due segnali luminosi in direzioni opposte, l'uno verso A e l'altro verso B. Diremo che le ricezioni dei due segnali in A e in B sono **eventi simultanei in** $\mathcal{R}$.

3^a **Definizione di simultaneità:**[14]
Siano A e B due punti in quiete rispetto ad $\mathcal{R}$, nei quali si verificano due eventi, rispettivamente negli istanti t_A e t_B letti sui rispettivi orologi, supposti preventivamente sincronizzati. Diremo che questi eventi **sono simultanei nel riferimento** $\mathcal{R}$ *se*

$$t_A = t_B, \tag{3.11}$$

ovvero se le letture dei rispettivi orologi sono uguali.
Quest'ultima definizione è quella che si avvicina di più al senso comune in quanto banalmente la gente è portata a dire che due eventi sono simultanei se accadono nello stesso istante segnato dai rispettivi orologi preventivamente sincronizzati!
Utilizzando la (3.8) si può constatare che la 3ª definizione di simultaneità è del tutto equivalente alle prime due.
Osservazione 1:

a) **Constatazione sperimentale della simultaneità di due eventi A e B.**
Bisogna preliminarmente aver sincronizzato gli orologi in A e B. Quando i due eventi si producono, due osservatori situati in A e in B annotano gli istanti t_A e t_B; questi osservatori possono essere evidentemente anche degli apparecchi fotografici che registrano i due eventi. Infine un fisico confronta le letture t_A e t_B. Si tratta del metodo generale, vecchio quanto la fisica stessa, ed utilizzato, senza cambiamenti, anche in fisica relativistica.

b) **Messa in guardia contro una falsa idea.**
Supponiamo che due eventi simultanei nel senso della 3ª definizione consistano nell'emissione di due segnali luminosi rispettivamente in A e in B. Un osservatore posto nel punto medio M del segmento AB riceverà simultaneamente tali segnali, (se è valido il postulato d'isotropia ottica dello spazio) soddisfacendo così alla definizione data da Einstein. Ma ciò presenta l'inconveniente di far intervenire un osservatore situato in un punto particolare dello spazio e di rimpiazzare la simultaneità degli eventi A, B (emissione dei segnali) con quella di due altri eventi (ricezione dei segnali in M). Da ciò si potrebbe dedurre erroneamente che, utilizzando la definizione di Einstein, si può affermare che gli eventi A e B sono simultanei soltanto relativamente all'osservatore M. Un osservatore situato

[12]Questa 1ª definizione di simultaneità è data da Einstein nel suo libro *Relatività, esposizione divulgativa*, Universale Scientifica Boringhieri, Torino (1967), pp. 59-60.
[13]Quest'altra definizione di simultaneità è riportata da Landau e Lifschitz nel loro libro citato più volte *Fisica Teorica 2: Teoria dei campi*, Editori Riuniti, Roma (1976), pag. 16.
[14]Questa 3ª definizione di simultaneità è data da Henri Arzeliès nel suo libro *La Cinématique Relativiste*, pag. 18, Gauthier-Villars, Paris (1955).

in un punto diverso M' attribuirebbe a tali eventi istanti diversi. Ciò è evidentemente un'interpretazione scorretta della definizione di Einstein. Se l'osservatore M' non può verificare la simultaneità di A e B ricevendo i segnali, la può verificare applicando la 3ª definizione, ovvero il metodo generale esposto nel punto a).

Osservazione 2:
A prima vista sembrerebbe che **le definizioni di simultaneità**, che abbiamo appena dato - specialmente la terza - si possano accettare anche nell'ambito della teoria classica newtoniana in cui **il tempo è assoluto**. Ma una riflessione attenta ci convince che ciò non è possibile, perché tali definizioni sono formulate tenendo conto sia del 1° che del 2° postulato della Teoria della Relatività Ristretta di Einstein, e perciò risultano in netto contrasto con l'ipotesi newtoniana di un tempo assoluto. *Nell'apprezzamento del tempo, in ambito classico, si ritiene che il moto reciproco di due sistemi di riferimento qualsiasi $\mathcal{R}$ ed $\mathcal{R}'$ (non necessariamente inerziali) sia inessenziale, e, quindi, che un'unica variabile t possa fungere da tempo per tutti i possibili osservatori, siano essi in quiete od anche in moto gli uni rispetto agli altri.* Con tale ipotesi, **priva di fondamento sperimentale**, il concetto di simultaneità sarebbe assoluto come quello di tempo.

3.2.6 La relatività della simultaneità

Utilizzando la prima e la seconda definizione di simultaneità viste nel paragrafo precedente dimostreremo ora che tale concetto non è assoluto ma dipende dal sistema di riferimento galileiano al quale viene riferito, in altri termini **due eventi simultanei in un dato riferimento galileiano non lo sono in un altro che si muove di moto TRU rispetto al primo.**

Figura 3.2: Relatività del concetto di simultaneità secondo Einstein.

Vediamo la spiegazione che dà Einstein della relatività del concetto di simultaneità utilizzando la sua definizione che abbiamo esposto come Def. 1^a nel paragrafo 3.2.5. Le sue considerazioni sulla genesi del concetto di tempo, che abbiamo riportato nel paragrafo precedente, sono state svolte rispetto ad un particolare corpo di riferimento a cui è stato dato il nome di "banchina ferroviaria" e che possiamo considerare come un sistema inerziale $\mathcal{R}$. Per quel che riguarda **la simultaneità**, Einstein continua così:[15]

1°) *"Supponiamo che un treno molto lungo viaggi sulle rotaie con la velocità costante* **v** *e nella direzione indicata dalla (Fig.3.2). Le persone che viaggiano su questo treno useranno vantaggiosamente il treno come corpo rigido di riferimento (sistema inerziale $\mathcal{R}'$); esse considerano tutti gli eventi in riferimento al treno. Ogni evento, poi, che ha luogo lungo la linea ferroviaria ha pure luogo in*

[15] A. Einstein, *Relatività, esposizione divulgativa*, Universale Scientifica Boringhieri, Torino (1967), pp.61-63.

un determinato punto del treno. Anche la definizione di simultaneità può venir data rispetto al treno nello stesso preciso modo in cui venne data rispetto alla banchina. Ora però si presenta, come conseguenza naturale, la seguente domanda: Due eventi (per esempio i due colpi di fulmine A e B) che sono simultanei rispetto alla banchina ferroviaria saranno tali anche rispetto al treno? Mostreremo subito che la risposta deve essere negativa.
Allorché diciamo che i colpi di fulmine A e B sono simultanei rispetto alla banchina intendiamo: i raggi di luce provenienti dai punti A e B dove cadono i due fulmini si incontrano l'uno con l'altro nel punto medio M dell'intervallo AB della banchina. Ma gli eventi A e B corrispondono anche alle posizioni A' e B' sul treno. Sia M' il punto medio dell'intervallo $A'B'$ sul treno in moto. Proprio quando si verificano i bagliori del fulmine (giudicati dalla banchina), questo punto M' coincide naturalmente con il punto M, ma esso si muove verso la destra del diagramma con la velocità **v** *del treno. Se un osservatore seduto in treno nella posizione M' non possedesse questa velocità, allora egli rimarrebbe permanentemente in M e i raggi di luce emessi dai bagliori del fulmine A e B lo raggiungerebbero simultaneamente, vale a dire s'incontrerebbero proprio dove lui è situato. Tuttavia nella realtà (considerata con riferimento alla banchina ferroviaria), egli si muove rapidamente verso il raggio di luce che proviene da B, mentre corre avanti al raggio di luce che proviene da A. Pertanto l'osservatore vedrà il raggio di luce emesso da B prima di vedere quello emesso da A. Gli osservatori che assumono il treno come loro corpo di riferimento debbono perciò giungere alla conclusione che il lampo di luce B abbia avuto luogo prima del lampo di luce A. Perveniamo così al seguente importante risultato: gli eventi che sono simultanei rispetto alla banchina non sono simultanei rispetto al treno e viceversa (relatività della simultaneità); ogni corpo di riferimento (sistema di coordinate) ha il suo proprio tempo particolare: un'attribuzione di tempo è fornita di significato solo quando ci venga detto a quale corpo di riferimento tale attribuzione si riferisce".*
Sentiamo ora quel che dicono Landau e Lifschitz del concetto di simultaneità così come l'abbiamo enunciato nella Def. 2ª del paragrafo 3.2.5.[16]

2°) *"Consideriamo due sistemi di riferimento inerziali $\mathcal{R} \equiv (Oxyz, t)$ e $\mathcal{R}' \equiv (O'x'y'z', t')$; il sistema $\mathcal{R}'$ si sposta rispetto al sistema $\mathcal{R}$ verso destra lungo gli assi x ed x' (Fig.3.3).*

Figura 3.3: Relatività del concetto di simultaneità secondo Landau-Lifshitz.

Supponiamo che da un punto M' sull'asse x' vengano emessi due segnali in direzioni opposte. Siccome la velocità di propagazione del segnale nel sistema $\mathcal{R}'$, come in qualsiasi altro sistema inerziale, è la stessa in ambedue le direzioni ed è uguale a c, i segnali giungeranno nei punti A' e B' equidistanti da M' in uno stesso istante (nel sistema $\mathcal{R}'$).
Tuttavia, è facile vedere che questi due eventi (arrivo del segnale in A' e B') non saranno affatto simultanei per un osservatore che si trovi nel sistema $\mathcal{R}$. In effetti, la velocità dei segnali relativamente al sistema $\mathcal{R}$, in accordo con il principio di relatività di Einstein, è sempre uguale a c, e poiché il punto A' si muove (relativamente al sistema $\mathcal{R}$) incontro al segnale emesso, mentre il punto B' si allontana dal segnale (emesso da M' verso B'), nel sistema $\mathcal{R}$ il segnale arriverà prima

[16] L.D.Landau e E.M.Lifschitz, *Fisica Teorica 2, Teoria dei Campi*, Editori Riuniti, Roma (1976), pp.15-16.

nel punto A' e poi nel punto B'.
Quindi **il principio di relatività di Einstein** *introduce cambiamenti fondamentali nei concetti principali della Fisica. I concetti di spazio e di tempo che abbiamo appreso dall'esperienza di tutti i giorni sono approssimativi poiché nella vita comune noi abbiamo a che fare soltanto con velocità molto piccole rispetto alla velocità della luce".*
Riportiamo infine una terza dimostrazione della relatività della simultaneità dovuta al russo Vladimir A. Ugarov che utilizza la Def. 1^a del paragrafo 3.2.5.[17]

3°) *"Anche qui si tratta di un semplice esperimento ideale che illustra qualitativamente la relatività del sincronismo, se il 2° postulato di Einstein è soddisfatto. L'esperimento è condotto nel treno di Einstein; questo termine viene dato ad ogni treno in moto rettilineo ed uniforme ma (è desiderabile!) con velocità relativistica. Nell'esperimento ideale si può assumere anche quest'ultima richiesta. È facile trovare il punto di mezzo del treno (questo si fa nel sistema di riferimento del treno e non presenta difficoltà). L'osservatore 1 è posto nel punto di mezzo del treno e l'osservatore 2 alla stazione. Dalle estremità del treno, che sono ugualmente distanti dall'osservatore 1, verso di lui sono inviati due segnali luminosi. L'esperimento ideale è fatto in modo che i segnali che viaggiano dalle estremità del treno arrivino all'osservatore 1 proprio nel momento in cui questo raggiunge l'osservatore 2 (negli esperimenti ideali non importa come questo si realizzi in pratica). Sono essenziali per noi le conclusioni che i due osservatori traggono dal simultaneo arrivo dei segnali nel punto centrale del treno.*

- **Osservatore 1**. *I segnali luminosi devono percorrere un'eguale distanza per raggiungermi [e siccome viaggiano alla stessa velocità c] essi sono quindi inviati simultaneamente.*
- **Osservatore 2**. *I segnali luminosi mi raggiungono quando il punto medio del treno mi passa di fronte, di conseguenza, essi sono inviati qualche tempo prima. Ma "prima" la testa del treno era più vicina a me della coda. Quindi, [tenuto ancora conto che i segnali viaggiano alla stessa velocità c] il segnale dalla coda è stato inviato leggermente in anticipo per poter giungere a me contemporaneamente al segnale proveniente dalla testa.*

È evidente da questo semplice ragionamento che due eventi simultanei in un sistema di riferimento (il riferimento del treno nel nostro esempio) non lo sono affatto in un altro (nel nostro caso il sistema Terra)".

In tutti e tre gli esempi sopra riportati abbiamo ottenuto la stessa conclusione: **se due eventi sono simultanei in un riferimento inerziale, non lo sono più in un altro sistema inerziale in moto rispetto al primo!**
Questo risultato è uno dei più sorprendenti della teoria einsteiniana, ed è noto come **relatività della simultaneità**; *esso afferma, per l'appunto, che la simultaneità è un concetto relativo al sistema di riferimento in cui si trova l'osservatore*. Perciò l'aver ammesso la validità del principio di relatività esteso a tutti i fenomeni fisici, e la validità del principio della costanza della velocità della luce nel vuoto - unitamente alla data definizione di simultaneità - necessariamente ne consegue, come abbiamo visto con queste semplici esperienze, **la relatività del concetto di simultaneità e di tempo.**
Dunque il concetto di "tempo assoluto" newtoniano è in contrasto con la semplice definizione di simultaneità introdotta da Einstein e con l'evidenza sperimentale che la velocità della luce nel vuoto è un invariante nel passaggio da un sistema galileiano ad un altro in moto TRU rispetto al primo. Esso è soltanto *"un pregiudizio"* in quanto la sua validità non è dimostrabile sperimentalmente: si tratta perciò di un concetto che non soddisfa al *Principio di verifica sperimentale (di Mach-Einstein)* - che abbiamo introdotto, nel paragrafo 1.1.1, come 2° assioma metodologico - e

[17] Vladimir Alexandrovic Ugarov, *Teoria della Relatività Ristretta*, Edizioni Mir, Mosca, 1982, pag. 41.

perciò dobbiamo rigettarlo.
Osservazione:
Facciamo notare al lettore, come abbiamo ribadito più volte, che già nella meccanica classica **le posizioni spaziali sono relative**:
un'attribuzione di luogo è fornita di significato solo quando ci venga detto a quale sistema di riferimento tale attribuzione si riferisce. **Un'analoga proprietà non vale invece per il tempo essendo questo assoluto**.
Nella Teoria della Relatività invece sussiste, anche per il tempo, una proprietà simile - che potremmo chiamare duale della precedente - infatti possiamo dire che:
un'attribuzione di tempo è fornita di significato solo quando ci venga detto a quale sistema di riferimento tale attribuzione si riferisce.
A questo proposito Einstein afferma:[18]
" Prima dell'avvento della teoria della relatività, nella fisica si era sempre tacitamente ammesso, che le attribuzioni di tempo avessero un significato assoluto, cioè fossero indipendenti dallo stato di moto del corpo di riferimento (sistema di coordinate). Abbiamo però visto or ora che tale ipotesi risulta incompatibile con la più naturale definizione di simultaneità; **se abbandoniamo quest'ipotesi, scompare il conflitto fra la legge di propagazione della luce nel vuoto e il principio di relatività"**.
È interessante osservare che nella teoria della relatività vi è *"perfetta simmetria"* concettuale fra *"attribuzioni spaziali" (luoghi)* e *"attribuzioni temporali" (istanti)*, ciò non si verifica nella meccanica classica newtoniana in cui vi è invece una *"asimmetria"* fra tali attribuzioni.

3.2.7 Formulazione matematica dei postulati della relatività einsteiniana

Cerchiamo ora di esprimere in forma analitica i due assiomi einsteiniani oppure, il che è equivalente, di dare una formulazione matematica al principio d'invarianza della velocità della luce.
Per fare ciò, consideriamo i soliti due sistemi di riferimento galileiani $\mathcal{R}$ ed $\mathcal{R}'$ in configurazione x-standard, con $\mathcal{R}'$ che si muove di moto TRU con velocità $\mathbf{v} = cost$. nel verso positivo dell'asse $x \equiv x'$, e supponiamo che all'istante iniziale $t = t' = 0$, per entrambi i riferimenti, essi coincidano $\mathcal{R} \equiv \mathcal{R}'$ (Fig.3.4).

Figura 3.4: Sistemi di riferimento galileiani in configurazione x-standard.

Consideriamo un evento $\mathcal{A}$ ed un evento $\mathcal{B}$ (infinitamente vicino ad $\mathcal{A}$ sia in $\mathcal{R}$ che in $\mathcal{R}'$), le cui coordinate differiscano da quelle di $\mathcal{A}$ per dx, dy, dz, dt in $\mathcal{R}$ e per dx', dy', dz', dt' in $\mathcal{R}'$. Supponiamo che l'evento $\mathcal{A}$ consista nell'emissione di un lampo di luce e che una particella nello spazio, illuminata da quel lampo, sia l'evento $\mathcal{B}$.
In conformità alla legge di propagazione della luce, l'osservatore in $\mathcal{R}$ troverà che il fronte d'onda luminoso si propaga in modo da soddisfare l'equazione

$$dl \equiv \sqrt{dx^2 + dy^2 + dz^2} = c\,dt$$

[18] A. Einstein, *Relatività, Esposizione divulgativa*, Universale Scientifica Boringhieri, Torino (1967), pag. 62.

essendo $dl \equiv |d\mathbf{x}|$ **l'intervallo spaziale** fra i due eventi, e dt (con $dt > 0$) il loro **intervallo temporale**. Elevando primo e secondo membro al quadrato, portando tutti i termini al secondo membro e scambiando il primo col secondo membro, avremo

$$c^2dt^2 - dl^2 \equiv c^2dt^2 - dx^2 - dy^2 - dz^2 = 0, \qquad dt > 0. \tag{3.12}$$

Analogamente l'osservatore in $\mathcal{R}'$, in base ai due postulati di Einstein, troverà che per lui il fronte d'onda luminoso si propaga secondo una legge analoga

$$c^2dt'^2 - dl'^2 \equiv c^2dt'^2 - dx'^2 - dy'^2 - dz'^2 = 0, \qquad dt' > 0. \tag{3.13}$$

Viceversa, ogni evento vicino ad $\mathcal{A}$, le cui coordinate soddisfino la (3.12) (o la (3.13)), è illuminato dal lampo proveniente da $\mathcal{A}$, e quindi le sue coordinate soddisferanno entrambe le relazioni (3.12) e (3.13). Ora, indipendentemente da quali possano essere le trasformazioni tra le coordinate, le trasformazioni tra i differenziali per ogni dato evento $\mathcal{A}$ sono lineari ed omogenee.[19] In tal modo il membro di sinistra della (3.13) è uguale ad una forma quadratica omogenea che, come abbiamo visto, si deve annullare per tutti i valori reali dei differenziali che soddisfano la forma quadratica (3.12).
Per un noto teorema sulle forme quadratiche reali, le due forme quadratiche differenziali (3.12) e (3.13) possono differire al più per un fattore moltiplicativo costante λ, cioè[20]:

$$c^2dt'^2 - dx'^2 - dy'^2 - dz'^2 = \lambda \left(c^2dt^2 - dx^2 - dy^2 - dz^2\right). \tag{3.14}$$

Tenendo ora conto che l'orientamento degli assi ortogonali di $\mathcal{R}$ ed $\mathcal{R}'$ può essere arbitrario e che lo spazio è isotropo rispetto a sistemi di riferimento galileiani, possiamo invertire il verso degli assi x e y di $\mathcal{R}$ e x' e y' di $\mathcal{R}'$ e considerare il moto dal punto di vista di $\mathcal{R}'$ anziché da quello di $\mathcal{R}$; dal punto di vista di $\mathcal{R}'$ questa configurazione si presenta come mostra la Fig.3.5. Chiaramente, dal punto di vista di $\mathcal{R}'$, il ruolo dei sistemi $\mathcal{R}$ e $\mathcal{R}'$ è manifestamente scambiato: ora è $\mathcal{R}$ che si allontana (verso destra) da $\mathcal{R}'$ con velocità algebrica v, per cui la relazione fra i due riferimenti è perfettamente simmetrica.

Per il principio di relatività varrà dunque una relazione analoga alla (3.14), con le coordinate scambiate (ma con lo stesso valore di λ altrimenti si potrebbe distinguere un riferimento inerziale da un altro), dunque del tipo:

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = \lambda \left(c^2dt'^2 - dx'^2 - dy'^2 - dz'^2\right). \tag{3.15}$$

Sostituendo quest'ultima nella (3.14) si ottiene

$$\lambda^2 = 1$$

[19] Infatti, qualunque sia la trasformazione $x^{i'} = f^{i'}(x^0, x^1, x^2, x^3)$, $i' = 0, 1, 2, 3$, i differenziali delle coordinate sono espressi dalla $dx^{i'} = \left(\frac{\partial f^{i'}}{\partial x^k}\right)_{\mathcal{A}} dx^k$ che costituisce ovviamente una trasformazione lineare omogenea fra i differenziali $dx^{i'}$ e dx^k, in quanto la matrice delle derivate, essendo calcolata nell'evento $\mathcal{A}$ è una matrice i cui elementi sono numeri reali.

[20] Il teorema cui si allude è il seguente:
Teorema: *Si considerino le seguenti due forme quadratiche in 4 variabili reali (forme quaternarie)* $p(x_0, x_1, x_2, x_3) = x_0^2 - x_1^2 - x_2^2 - x_3^2$ *e* $q(x_0, x_1, x_2, x_3) = \sum_{i,j=0}^{2} a_{ij}x_ix_j$ *con* $a_{ij} \in \mathbb{R}$ *e* $a_{ij} = a_{ji}$. *Se ogni quaterna di numeri reali che annulla la forma quadratica* $p(x_0, x_1, x_2, x_3)$ *è tale da annullare anche la forma quadratica* $q(x_0, x_1, x_2, x_3)$, *allora le due forme differiscono al più per un fattore costante, cioè risulta:* $q(x_0, x_1, x_2, x_3) = \lambda p(x_0, x_1, x_2, x_3)$, *con* $\lambda \in \mathbb{R}$.
La dimostrazione di questo teorema è inserita nell'APPENDICE a questo capitolo. Osserviamo che molti autori ipotizzano che il fattore λ che compare nella (3.14), possa dipendere da v, quindi considerano, nei loro ragionamenti, una $\lambda(v)$, per poi concludere, con considerazioni un po' involute, invocanti l'omogeneità e l'isotropia dello spazio, ma assolutamente non necessarie, che λ è in effetti indipendente da v. *Che λ sia indipendente da v è invece un fatto puramente matematico come abbiamo chiaramente messo in luce col teorema sulle forme quadratiche sopra citato!*

Figura 3.5: Isotropia spaziale dei sistemi galileiani e reciprocità del moto.

e pertanto

$$\lambda = \pm 1.$$

Si può immediatamente scartare $\lambda = -1$, poiché la (3.14), essendo λ una costante indipendente da v, è valida per qualunque valore di v e perciò deve mantenersi valida, per ragioni di continuità, anche quando $v \to 0$, nel qual caso essa deve ridursi ad una identità, in quanto si avrà $\mathcal{R} \equiv \mathcal{R}'$. Di conseguenza, per ogni possibile coppia $\mathcal{A}$, $\mathcal{B}$ di eventi infinitamente vicini, avremo **l'importantissima relazione**:

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = c^2dt'^2 - dx'^2 - dy'^2 - dz'^2. \tag{3.16}$$

La forma quadratica differenziale che figura al primo membro della (3.16) sarà da noi così indicata

$$\mathbf{ds}^2 \equiv c^2dt^2 - dx^2 - dy^2 - dz^2 \equiv c^2dt^2 - d\mathbf{x}^2 = \textbf{Invariante}, \tag{3.17}$$

essa costituisce **l'invariante fondamentale della Relatività Ristretta** e verrà chiamata - scostandosi leggermente dalla denominazione in uso - **intervallo o divario spazio-temporale** fra i due eventi considerati. Essa si chiama anche **metrica dello spazio-tempo** come vedremo, con più precisione, nel seguito.[21] Abbiamo ottenuto così un risultato di capitale importanza:

> *L'intervallo spazio-temporale* $\mathbf{ds}^2$ *fra due eventi, in un dato riferimento galileiano* $\mathcal{R}$*, è una forma quadratica omogenea nei differenziali* dx, dy, dz, dt *a coefficienti costanti* $(+1, -1, -1, -1)$*, che assume la stessa espressione in tutti i sistemi di riferimento galileiani, cioè, è un invariante rispetto alla trasformazione lineare di coordinate che fa passare da un riferimento galileiano ad un altro*[22]:
>
> $$\mathbf{ds}^2 = \mathbf{ds}'^2. \tag{3.18}$$

Questa invarianza è l'espressione matematica dei due assiomi einsteiniani.

3.3 Trasformazioni speciali di Lorentz

Gli assiomi fondamentali che stanno alla base della deduzione delle trasformazioni di Lorentz sono quelli introdotti nella fisica, per la prima volta, da Einstein nel lavoro del 1905, e cioè

1) il Principio di Relatività di Einstein (1° Postulato),

[21] In Relatività, quasi tutti gli autori, indicano l'invariante fondamentale, o metrica dello spazio-tempo, con ds^2 e lo chiamano **il quadrato dell'intervallo spazio-temporale**, intendendo con **intervallo spazio-temporale** il ds. Noi non seguiremo questa terminologia, anche se estremamente diffusa, ed useremo invece la terminologia utilizzata da Cattaneo e la notazione vettoriale di Rindler $\mathbf{ds}^2$, per ragioni che appariranno nel seguito.

[22] Per essere precisi, osserviamo che questo risultato è stato ottenuto considerando due riferimenti galileiani in configurazione x-standard, ma può essere dimostrato anche nel caso generale di due riferimenti galileiani in configurazione qualunque (vedi Cap.5, paragrafo 5.1.1).

2) il Principio della costanza della velocità della luce nel vuoto (2° Postulato),

che abbiamo enunciato nel paragrafo 3.2.
Le conseguenze del principio di relatività di Einstein sono sostanzialmente le stesse che abbiamo utilizzato nel Cap.1, paragrafo 1.3.4, per determinare le relazioni che intercorrono tra le coordinate di un medesimo evento osservato in riferimenti galileiani diversi, al fine di ottenere le trasformazioni di Galileo. Queste ipotesi riguardano la scelta della configurazione reciproca dei due riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$, che non menoma peraltro la generalità fisica, e delle loro proprietà di omogeneità ed isotropia.
A tal fine è sempre possibile (vedi Fig.3.4):

1) Scegliere la traiettoria rettilinea dell'origine spaziale O' di $\mathcal{R}'$ coincidente con l'asse spaziale x di $\mathcal{R}$;

2) Scegliere le origini temporali in $\mathcal{R}$ e $\mathcal{R}'$ in modo che all'istante zero per i due riferimenti $(t = t' = 0)$ le loro origini spaziali coincidano $O \equiv O'$;
Il riferimento $\mathcal{R}'$ si muova con velocità costante $\mathbf{v}$ rispetto ad $\mathcal{R}$ e parallela all'asse x, ma non necessariamente concorde con esso, e diremo v la componente algebrica (positiva o negativa) di tal velocità secondo l'asse x. Gli assi y e z in $\mathcal{R}$ possono essere scelti arbitrariamente, purché $Oxyz$ sia una terna cartesiana ortonormale destra come abbiamo sempre convenuto.

Assumiamo ancora:

3) che i piani coordinati $x'y'(z' = 0)$ e $x'z'(y' = 0)$ di $\mathcal{R}'$ possono essere presi coincidenti con i rispettivi piani coordinati $xy(z = 0)$ e $xz(y = 0)$ di $\mathcal{R}$, e gli assi y', z' equiorientati con y, z;

4) il piano coordinato $y'z'(x' = 0)$ di $\mathcal{R}'$ si identifichi, per t arbitrario, col piano $x = vt$, parallelo al piano coordinato $yz(x = 0)$ di $\mathcal{R}$;

5) assumiamo infine che valga il già citato **principio di reciprocità del moto**:
Se v è la velocità algebrica di $\mathcal{R}'$ rispetto ad $\mathcal{R}$, la velocità algebrica di $\mathcal{R}$ rispetto ad $\mathcal{R}'$ vale $-v$.
In tal caso, come abbiamo già detto, i due riferimenti sono in configurazione tipica o *x-standard*.

Sotto tali ipotesi, la trasformazione di coordinate che si ottiene è la (1.12), cioè

$$\begin{cases} x' = a_{11}(v)(x - vt) \\ y' = y \\ z' = z \\ t' = a_{41}(v)x + a_{11}(v)t, \end{cases} \tag{3.19}$$

che riscriviamo sotto la forma

$$\begin{cases} x' = \gamma(v)(x - vt) \\ y' = y \\ z' = z \\ t' = \alpha(v)x + \gamma(v)t, \end{cases} \tag{3.20}$$

nella quale, per semplificare un pò le notazioni, abbiamo posto $a_{11}(v) = \gamma(v)$ e $a_{41}(v) = \alpha(v)$, quest'ultimo coefficiente, per la (1.11), è uguale a

$$\alpha(v) = \frac{1}{v}\left[\frac{1}{\gamma(-v)} - \gamma(v)\right], \tag{3.21}$$

ed inoltre $\gamma(-v) = \gamma(v)$, $\alpha(-v) = -\alpha(v)$.
La determinazione di $\gamma(v)$ e quindi di $\alpha(v)$ possiamo ottenerla utilizzando il 2° assioma di Einstein e cioè l'invarianza della velocità della luce nel passare da un sistema inerziale $\mathcal{R}$ ad un altro $\mathcal{R}'$ in configurazione *x-standard*.[23]
Supponiamo, a tal fine, che all'istante in cui le origini O ed O' dei due sistemi galileiani sopraddetti coincidono, parta da esse un segnale luminoso. Il suo fronte d'onda all'istante t sarà per l'osservatore $\mathcal{R}$ una sfera di centro O e raggio ct e quindi di equazione

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{3.22}$$

Per l'osservatore $\mathcal{R}'$, in virtù del 1° e del 2° assioma, lo stesso fronte d'onda deve essere la sfera

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{3.23}$$

Affinché le due equazioni (3.22) e (3.23) possano coesistere si dovrà avere che, sostituendo nella (3.23) ad x', y', z', t' le loro espressioni (3.19), la nuova equazione ottenuta

$$\gamma^2(x - vt)^2 + y^2 + z^2 - c^2(\alpha x + \gamma t)^2 = 0, \tag{3.24}$$

e la (3.22) siano equivalenti. Per questo, detto μ un coefficiente di proporzionalità, dovrà essere identicamente, cioè per valori arbitrari delle variabili x, y, z, t:

$$\gamma^2(x - vt)^2 + y^2 + z^2 - c^2(\alpha x + \gamma t)^2 = \mu(x^2 + y^2 + z^2 - c^2t^2). \tag{3.25}$$

Nelle due equazioni (3.24) e (3.25), per semplicità di scrittura, si è omessa la dipendenza di γ e di α da v. Ora, dall'uguaglianza dei coefficienti di y^2 e di z^2 nei due membri si ottiene

$$\mu = 1.$$

Sviluppando i calcoli nella (3.25), raccogliendo i termini simili e semplificando, avremo

$$(\gamma^2 - c^2\alpha^2)x^2 - 2\gamma(v\gamma + c^2\alpha)xt - \gamma^2(c^2 - v^2)t^2 = x^2 - c^2t^2, \tag{3.26}$$

ed identificando i coefficienti dei termini simili nei due membri otteniamo infine le relazioni

$$\gamma^2 - c^2\alpha^2 = 1, \quad v\gamma + c^2\alpha = 0, \quad \gamma^2(c^2 - v^2) = c^2. \tag{3.27}$$

Dalla terza delle (3.27) si deduce

$$\gamma^2 = \frac{c^2}{c^2 - v^2} = \frac{1}{1 - \dfrac{v^2}{c^2}} = \frac{1}{1 - \beta^2}, \tag{3.28}$$

avendo indicato con $0 < \beta = \dfrac{v}{c} < 1$ la cosìddetta velocità römeriana. La (3.28) fornisce

$$\gamma(v) = \pm\frac{1}{\sqrt{1 - \beta^2}}, \tag{3.29}$$

Osserviamo infine, che, essendosi scelti gli assi x ed x' equiversi, per $v = 0$ la prima delle (3.19) deve ridursi all'identità $x = x'$ e perciò deve essere $\gamma(0) = 1$, dunque, affinché non venga mai

[23] Allo stesso risultato si giunge usando l'invarianza $\mathbf{ds}^2 = \mathbf{ds}'^2$ dell'intervallo spazio-temporale fra due eventi qualsiasi, scritta in termini finiti anziché infinitesimi.

meno l'invertibilità della trasformazione la $\gamma(v)$, che si ritiene funzione continua di v, deve mantenersi sempre positiva al variare di v, perciò si avrà

$$\gamma(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}}, \tag{3.30}$$

con $\gamma(-v) = \gamma(v)$ funzione pari. La funzione $\gamma(v)$, *d'importanza fondamentale nella relatività speciale*, prende il nome di *fattore di Lorentz.*[24]
Col valore di $\gamma(v)$ dato dalla (3.30) possiamo determinare dalla 1^a o dalla 2^a delle (3.27) (ovvero utilizzando anche la (3.21)) la funzione dispari

$$\alpha(v) = -\frac{v}{c^2}\gamma = -\frac{\beta}{c}\gamma, \qquad \alpha(-v) = -\alpha(v). \tag{3.31}$$

Sostituendo i risultati ottenuti nella (3.19) si ottiene **il risultato più importante della Teoria della Relatività Ristretta** - equivalente all'invarianza dell'intervallo spaziotemporale - la celebre **trasformazione speciale di Lorentz-Einstein**[25]:

$$(L) \qquad \begin{cases} x' = \gamma(v)(x - vt) = \gamma(v)(x - \beta ct), \\ y' = y, \\ z' = z, \\ t' = \gamma(v)\left(t - \frac{v}{c^2}x\right) = \gamma(v)\left(t - \frac{\beta}{c}x\right). \end{cases} \tag{3.32a}$$

Si può constatare che la trasformazione 3.32a è invertibile essendo non nullo il determinante dei suoi coefficienti, si ha infatti

$$det \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{\beta\gamma}{c} & 0 & 0 & \gamma \end{pmatrix} = 1,$$

inoltre, in virtù della definizione (3.30) del fattore di Lorentz, essa è reale soltanto se il modulo v della velocità di traslazione di $\mathcal{R}$ rispetto ad $\mathcal{R}'$ è inferiore alla velocità c della luce nel vuoto, cioè deve essere $v < c$: **ciò costituisce una prima indicazione del valore limite che la velocità della luce nel vuoto assume in ambito relativistico**; ritorneremo su questa questione nel paragrafo 4.1.12.
Per scrivere la trasformazione inversa ovviamente si può risolvere il precedente sistema lineare rispetto alle variabili x, y, z, t, ma si può, più semplicemente, scambiare fra di loro le variabili

[24]Si osservi che abbiamo già incontrato il fattore di Lorentz nel Cap.2 nella spiegazione classica dell'esperimento di Michelson-Morley.

[25]Ricordiamo che la qualifica di "speciale" si riferisce alla particolare scelta degli assi nei due riferimenti galileiani considerati (in mutua relazione x-standard). Inoltre avvertiamo il lettore che la (3.32a) è chiamata generalmente **trasformazione di Lorentz**, ma ci sembra doveroso chiamarla, più correttamente, **trasformazione di Lorentz-Einstein**. Infatti è pur vero che fu Lorentz a ricavarla, per la prima volta, ponendosi il problema, puramente matematico, di trovare quali trasformazioni lasciassero invarianti le equazioni di Maxwell nel passaggio da un sistema inerziale ad un altro senza però riconoscerne il vero significato fisico: per Lorentz si trattava di un puro esercizio matematico. Riconoscere la giusta interpretazione fisica di quelle trasformazioni, che implicano soprattutto **la relatività del tempo**, fu il grande merito di Einstein.

accentate con quelle non accentate e cambiare contemporaneamente v in $-v$, si avrà:

$$(L^{-1}) \qquad \begin{cases} x = \gamma(v)(x' + vt') = \gamma(v)(x' + \beta ct'), \\ y = y', \\ z = z', \\ t = \gamma(v)\left(t' + \dfrac{v}{c^2}x'\right) = \gamma(v)\left(t' + \dfrac{\beta}{c}x'\right). \end{cases} \tag{3.32b}$$

Se ora introduciamo due nuove variabili "temporali" $T = ct$ e $T' = ct'$, aventi le dimensioni di due lunghezze, le precedenti trasformazioni di Lorentz risulteranno simmetriche, oltre che nelle variabili y, z e y', z' anche nelle variabili x, T e x', T', infatti si ha:

$$(L) \qquad \begin{cases} x' = \gamma(x - \beta T), \\ y' = y, \\ z' = z, \\ T' = \gamma(T - \beta x), \end{cases} \tag{3.33a}$$

la cui inversa è:

$$(L^{-1}) \qquad \begin{cases} x = \gamma(x' + \beta T'), \\ y = y', \\ z = z', \\ T = \gamma(T' + \beta x'). \end{cases} \tag{3.33b}$$

Applicando le trasformazioni di Lorentz a due eventi qualsiasi $\mathcal{A}$ e $\mathcal{B}$ di coordinate x_A, y_A, z_A, t_A e x_B, y_B, z_B, t_B nel riferimento galileiano $\mathcal{R}$ e di coordinate x'_A, y'_A, z'_A, t'_A e x'_B, y'_B, z'_B, t'_B nel riferimento galileiano $\mathcal{R}'$, sottraendo poi membro a membro le equazioni così ottenute ed indicando come segue le differenze fra coordinate omonime

$$\Delta x = x_B - x_A, \;\; \Delta y = y_B - y_A, \;\; \Delta z = z_B - z_A, \;\; \Delta t = t_B - t_A, \;\; ecc., \tag{3.34}$$

tenuto conto della linearità, otterremo le trasformazioni di Lorentz sotto la forma di differenze finite:

$$(L) \qquad \begin{cases} \Delta x' = \gamma(\Delta x - v\Delta t) = \gamma(\Delta x - \beta c\Delta t), \\ \Delta y' = \Delta y, \\ \Delta z' = \Delta z, \\ \Delta t' = \gamma\left(\Delta t - \dfrac{v}{c^2}\Delta x\right) = \gamma\left(\Delta t - \dfrac{\beta}{c}\Delta x\right), \end{cases} \tag{3.35a}$$

$$(L^{-1}) \qquad \begin{cases} \Delta x = \gamma(\Delta x' + v\Delta t') = \gamma(\Delta x' + \beta c\Delta t'), \\ \Delta y = \Delta y', \\ \Delta z = \Delta z', \\ \Delta t = \gamma\left(\Delta t' + \dfrac{v}{c^2}\Delta x'\right) = \gamma\left(\Delta t' + \dfrac{\beta}{c}\Delta x'\right). \end{cases} \tag{3.35b}$$

Se anziché considerare differenze finite si considerano differenze infinitesime, cioè eventi infinitamente vicini, allora si avranno ancora equazioni dello stesso tipo delle precedenti ma scritte in forma differenziale, cioè:

$$(L) \qquad \begin{cases} dx' = \gamma(dx - vdt) = \gamma(dx - \beta cdt), \\ dy' = dy, \\ dz' = dz, \\ dt' = \gamma\left(dt - \dfrac{v}{c^2}dx\right) = \gamma\left(dt - \dfrac{\beta}{c}dx\right), \end{cases} \tag{3.36a}$$

$$(L^{-1}) \qquad \begin{cases} dx &= \gamma(dx' + vdt') = \gamma(dx' + \beta c dt'), \\ dy &= dy', \\ dz &= dz', \\ dt &= \gamma\left(dt' + \frac{v}{c^2}dx'\right) = \gamma\left(dt' + \frac{\beta}{c}dx'\right). \end{cases} \tag{3.36b}$$

Concludendo, possiamo affermare che *le trasformazioni di Lorentz- Einstein costituiscono il fondamento matematico della TRR e per il Principio di Relatività di Einstein (PRE) tutte le leggi della Fisica devono essere invarianti per tali trasformazioni.*

3.3.1 Trasformazioni speciali di Lorentz come rotazioni nello spazio di Minkowski

Vogliamo ora osservare che le trasformazioni speciali di Lorentz possono essere interpretate come **rotazioni nello spazio di Minkowski**.
A tal fine, riprendiamo la trasformazione (3.19) che riscriviamo con le nuove variabili $T = ct$ e $T' = ct'$ introdotte nel paragrafo precedente

$$\begin{cases} x' = a_{11}(v)x - a_{11}(v)\beta T \\ y' = y \\ z' = z \\ T' = ca_{41}(v)x + a_{11}(v)T. \end{cases} \tag{3.37}$$

Semplificando ulteriormente le notazioni, ponendo:

$$a_{11}(v) = \gamma(v), \qquad ca_{41}(v) = \lambda(v),$$

la precedente trasformazione si scriverà così:

$$\begin{cases} x' = \gamma(v)x - \gamma(v)\beta T \\ y' = y \\ z' = z \\ T' = \lambda(v)x + \gamma(v)T. \end{cases} \tag{3.38}$$

Tenendo ora conto dell'invarianza dell'intervallo spaziotemporale, cioè che:

$$c^2t'^2 - x'^2 - y'^2 - z'^2 = c^2t^2 - x^2 - y^2 - z^2,$$

sostituendo al 1° membro della precedente uguaglianza le (3.38) e semplificando possiamo scrivere:

$$\gamma^2(1-\beta^2)T^2 - (\gamma^2 - \lambda^2)x^2 + 2\gamma(\beta\gamma + \lambda)xT = T^2 - x^2. \tag{3.39}$$

Applicando il principio d'identità dei polinomi al 1° ed al 2° membro, otteniamo le tre uguaglianze:

$$\gamma^2(1-\beta^2) = 1, \quad \gamma^2 - \lambda^2 = 1, \quad \beta\gamma + \lambda = 0. \tag{3.40}$$

La prima delle (3.40) ci dà il valore del fattore di Lorentz in funzione della velocità römeriana β

$$\gamma(v) = \pm\frac{1}{\sqrt{1-\beta^2}}. \tag{3.41}$$

La seconda delle (3.40) può essere soddisfatta con l'uso di **funzioni iperboliche**[26]ponendo

$$\gamma = \cosh\theta, \quad \lambda = \sinh\theta. \tag{3.42}$$

[26]Ricordiamo che **le funzioni iperboliche** sono così definite: $\cosh\theta = \frac{e^\theta + e^{-\theta}}{2}$, $\sinh\theta = \frac{e^\theta - e^{-\theta}}{2}$, $\tanh\theta = \frac{\sinh\theta}{\cosh\theta} = \frac{e^\theta - e^{-\theta}}{e^\theta + e^{-\theta}}$; inoltre esse soddisfano l'identità fondamentale $\cosh^2\theta - \sinh^2\theta = 1$.

Infine, la terza delle (3.40) diventa allora

$$\sinh\theta = -\beta\cosh\theta, \quad \text{cioè} \quad \tanh\theta = -\beta. \tag{3.43}$$

Dalle precedenti considerazioni segue che la trasformazione di Lorentz (3.38), a parte le $y' = y$ e $z' = z$, si può scrivere nella forma

$$\begin{cases} x' = x\cosh\theta + T\sinh\theta \\ T' = x\sinh\theta + T\cosh\theta, \end{cases} \tag{3.44}$$

che rappresenta una **rotazione nel piano iperbolico** (x, T) di Minkowski.[27]
Ci rimane da precisare il segno del fattore di Lorentz $\gamma(v)$ che, per la (3.41), a priori potrebbe essere positivo o negativo, ma tenuto conto della prima delle (3.42) essendo il coseno iperbolico sempre positivo $(\cosh\theta > 0)$ si avrà, come già sappiamo:

$$\gamma(v) = \frac{1}{\sqrt{1-\beta^2}}. \tag{3.45}$$

Osserviamo infine che il parametro θ, essendo legato alla velocità relativa v di $\mathcal{R}'$ rispetto ad $\mathcal{R}$, tramite la seconda delle (3.43), è spesso chiamato **parametro di velocità**: esso gioca un ruolo utile in alcune formule sulla composizione delle velocità ed anche in certi problemi di dinamica di una particella.

3.4 Trasformazioni generali di Lorentz

Come si è già detto, la particolare scelta dei due riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$ in mutua relazione $x - \textit{standard}$ non pregiudica la generalità ed il contenuto fisico delle trasformazioni speciali di Lorentz (3.32a), (3.32b). Tuttavia è utile conoscere la struttura delle trasformazioni stesse quando, come sistemi di riferimento inerziali, si adottino due terne cartesiane ortogonali completamente arbitrarie. Inizialmente, per tali terne, manterremo l'unica condizione che le rispettive origini O, O', abbiano, nella loro storia, un evento in comune, e che in corrispondenza di tale evento vengano scelte le origini dei tempi sia in $\mathcal{R}$ sia in $\mathcal{R}'$, cioè $t = t' = 0$.
Ci occupiamo ora, molto brevemente, delle trasformazioni generali di Lorentz, nella loro forma vettoriale, lasciando cadere le ipotesi semplificatrici che ci hanno condotto alle trasformazioni speciali (3.32a) e (3.32b), e cioè:

1) l'ipotesi che i due sistemi inerziali $\mathcal{R}$ ed $\mathcal{R}'$ siano in configurazione $x - \textit{standard}$, cioè si muovono di moto relativo traslatorio rettilineo uniforme lungo gli assi coincidenti $x \equiv x'$;

2) i sistemi $\mathcal{R}$ ed $\mathcal{R}'$ sono equiorientati ed all'istante iniziale per entrambi $t = t' = 0$ essi coincidano.

Supponiamo invece che i riferimenti $\mathcal{R}$ ed $\mathcal{R}'$ siano ancora equiorientati e che per $t = t' = 0$ si abbia, come prima, $\mathcal{R} \equiv \mathcal{R}'$, ma essi siano in moto traslatorio uniforme l'uno rispetto all'altro lungo una direzione qualsiasi con velocità relativa di modulo v. Scomponiamo il vettore di

[27] Si noti l'analogia delle (3.44) con le formule che definiscono le rotazioni in un piano euclideo in coordinate (x, y) nelle quali compaiono le funzioni trigonometriche ordinarie $\sin\theta$ e $\cos\theta$, cioè

$$\begin{cases} x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta. \end{cases}$$

posizione $\mathbf{x} = (x, y, z)$ di un generico evento, nel riferimento $\mathcal{R}$, in un vettore $\mathbf{x}_\parallel$ parallelo a $\mathbf{v}$ ed un vettore $\mathbf{x}_\perp$ perpendicolare a $\mathbf{v}$,

$$\mathbf{x} = \mathbf{x}_\parallel + \mathbf{x}_\perp. \tag{3.46}$$

Il componente parallelo a $\mathbf{v}$ si definisce facilmente così:

$$\mathbf{x}_\parallel = (\mathbf{x} \cdot \text{vers}\mathbf{v})\text{vers}\mathbf{v} = \left(\mathbf{x} \cdot \frac{\mathbf{v}}{v}\right)\frac{\mathbf{v}}{v}, \tag{3.47}$$

avendo indicato con vers$\mathbf{v}$ il versore della velocità. Invece il componente perpendicolare a $\mathbf{v}$, stante la (3.46), possiamo esprimerlo semplicemente con

$$\mathbf{x}_\perp = \mathbf{x} - \mathbf{x}_\parallel. \tag{3.48}$$

È pressoché ovvio che la trasformazione di Lorentz per $\mathbf{x}_\parallel$, che è il componente del vettore di posizione lungo la direzione del moto relativo di $\mathcal{R}$ ed $\mathcal{R}'$, sarà analoga alla prima formula che compare nella trasformazione speciale di Lorentz (3.32a), cioè:

$$\mathbf{x}_\parallel = \gamma(v)(\mathbf{x}_\parallel - \mathbf{v}t). \tag{3.49}$$

La trasformazione di Lorentz per $\mathbf{x}_\perp$ è analoga alle formule seconda e terza della (3.32a)

$$\mathbf{x}'_\perp = \mathbf{x}_\perp. \tag{3.50}$$

Sommando membro a membro la (3.49) con la (3.50), ed utilizzando le (3.47) e (3.48) otteniamo

$$\begin{aligned}\mathbf{x}' = \mathbf{x}'_\parallel + \mathbf{x}'_\perp = \gamma\mathbf{x}_\parallel - \gamma\mathbf{v}t + \mathbf{x}_\perp = \gamma\mathbf{x}_\parallel - \gamma\mathbf{v}t + \mathbf{x} - \mathbf{x}_\parallel = \\ = (\gamma - 1)\mathbf{x}_\parallel + \mathbf{x} - \gamma\mathbf{v}t = (\gamma - 1)\left(\mathbf{x} \cdot \frac{\mathbf{v}}{v}\right)\frac{\mathbf{v}}{v} + \mathbf{x} - \gamma\mathbf{v}t,\end{aligned}$$

cioè

$$\mathbf{x}' = \mathbf{x} + \left[(\gamma - 1)\frac{\mathbf{x} \cdot \mathbf{v}}{v^2} - \gamma t\right]\mathbf{v}, \qquad (\mathbf{v} \neq 0) \tag{3.51}$$

Infine, per t vale l'ultima delle (3.32a), con $\mathbf{x} \cdot \mathbf{v}$ al posto di vx:

$$t' = \gamma(v)\left(t - \frac{\mathbf{x} \cdot \mathbf{v}}{c^2}\right). \tag{3.52}$$

Perciò **la più generale trasformazione di Lorentz fra sistemi di riferimento inerziali equiorientati** sarà allora

$$\begin{cases}\mathbf{x}' = \mathbf{x} + \left[(\gamma - 1)\dfrac{\mathbf{x} \cdot \mathbf{v}}{v^2} - \gamma t\right]\mathbf{v}, \qquad (\mathbf{v} \neq 0), \\ t' = \gamma(v)\left(t - \dfrac{\mathbf{x} \cdot \mathbf{v}}{c^2}\right).\end{cases} \tag{3.53}$$

Si può verificare che le (3.53), si riducono alla trasformazione speciale di Lorentz (3.32a) nel caso in cui la velocità $\mathbf{v}$ sia diretta lungo l'asse x.
Osserviamo che la (3.51) ha senso se $v \neq 0$, se invece $v = 0$ - nel qual caso tale formula dovrebbe ridursi alla trasformazione identità $\mathbf{x}' = \mathbf{x}$ - essa perde senso in quanto la frazione che è entro parentesi quadre $\dfrac{\gamma - 1}{v^2}$ diventa indeterminata perché se $v \to 0$, il fattore di Lorentz $\gamma \to 1$ e la frazione tende a $\dfrac{0}{0}$. Possiamo però dimostrare subito che tale frazione per $v \to 0$ tende ad un valore finito e perciò la formula (3.51) può essere riscritta in modo tale da essere valida per qualunque valore della velocità $\mathbf{v}$ di traslazione, compreso il valore nullo.

Se ora introduciamo la cosiddetta *velocità römeriana* $\beta = \dfrac{v}{c}$ e scriviamo il fattore di Lorentz sotto forma di potenza $\gamma = (1-\beta^2)^{-1/2}$, possiamo svilupparlo utilizzando la serie binomiale, avremo:[28]

$$\gamma = 1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \frac{5}{16}\beta^6 + ... \tag{3.54}$$

Perciò possiamo scrivere:

$$\frac{\gamma-1}{v^2} = \frac{1}{c^2}\frac{c^2}{v^2}(\gamma-1) = \frac{1}{c^2}\beta^{-2}\left[\frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \frac{5}{16}\beta^6 + ...\right] = \frac{1}{c^2}\left[\frac{1}{2} + \frac{3}{8}\beta^2 + \frac{5}{16}\beta^4 + ...\right],$$

da cui si deduce che

$$\lim_{v\to 0}\frac{\gamma-1}{v^2} = \frac{1}{2c^2}. \tag{3.55}$$

È allora possibile trasformare il rapporto $\dfrac{\gamma-1}{v^2}$ nel seguente modo

$$\frac{\gamma-1}{v^2} = \frac{1}{v^2}\frac{\gamma^2-1}{\gamma+1},$$

e tenendo conto che $\gamma^2 - 1 = \dfrac{v^2}{c^2-v^2}$, con facili calcoli otteniamo

$$\frac{\gamma-1}{v^2} = \frac{1}{c^2}\frac{\gamma^2}{\gamma+1}. \tag{3.56}$$

È immediato constatare che il secondo membro della (3.56) per $v \to 0$ tende allo stesso valore a cui tende il primo membro cioè $\dfrac{1}{2c^2}$.
La trasformazione di Lorentz (3.53) si può scrivere perciò anche nel modo seguente

$$\begin{cases} \mathbf{x}' = \mathbf{x} + \gamma\left[\dfrac{\gamma}{\gamma+1}\dfrac{\mathbf{x}\cdot\mathbf{v}}{c^2} - t\right]\mathbf{v}, \qquad (\textit{valida qualunque sia } \mathbf{v}), \\ \\ t' = \gamma(v)\left(t - \dfrac{\mathbf{x}\cdot\mathbf{v}}{c^2}\right), \end{cases} \tag{3.57}$$

nel limite newtoniano, cioè per $c \to \infty$, sia la (3.53) sia la (3.57) si riducono, come ci si aspetta, alla trasformazione di Galileo

$$\begin{cases} \mathbf{x}' = \mathbf{x} - \mathbf{v}t, \\ t' = t. \end{cases} \tag{3.58}$$

Se ora si tralascia anche l'ipotesi del parallelismo degli assi di $\mathcal{R}$ ed $\mathcal{R}'$, **le più generali trasformazioni di Lorentz omogenee**, che combinano una trasformazione di tipo traslatorio come la (3.57) con una rotazione rappresentata da una matrice ortogonale $\mathbf{R}(\boldsymbol{\theta})$ dove il modulo di $\boldsymbol{\theta}$ è l'angolo di rotazione mentre il versore associato a $\boldsymbol{\theta}$ individua l'asse di rotazione, sono della forma

$$\begin{cases} \mathbf{x}' = \mathbf{R}(\boldsymbol{\theta})\mathbf{x} + \gamma\left[\dfrac{\gamma}{\gamma+1}\dfrac{\mathbf{v}\cdot\mathbf{R}(\boldsymbol{\theta})\mathbf{x}}{c^2} - t\right]\mathbf{v}, \\ \\ t' = \gamma(v)\left(t - \dfrac{\mathbf{v}\cdot\mathbf{R}(\boldsymbol{\theta})\mathbf{x}}{c^2}\right). \end{cases} \tag{3.59}$$

[28] Per sviluppare il fattore di Lorentz $\gamma(v)$ conviene considerarlo come funzione di β scrivendolo nella forma $\gamma(\beta) = (1-\beta^2)^{-1/2}$ ed utilizzare la serie binomiale $(1+x)^\alpha = \sum_{k=0}^{+\infty}\binom{\alpha}{k}x^k$, con $x, \alpha \in \mathbb{R}$, e $\binom{\alpha}{k} = \dfrac{\alpha(\alpha-1)(\alpha-2).....(\alpha-k+1)}{k!}$. Nel nostro caso specifico è $\alpha = -1/2$, $x = -\beta^2$, ed essendo $|x| < 1$ la serie è senz'altro assolutamente convergente; si ottiene allora $\gamma(\beta) = 1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \frac{5}{16}\beta^6 + ...$

Se infine non imponiamo che i due sistemi di riferimento $\mathcal{R}$ ed $\mathcal{R}'$ abbiano le origini coincidenti nell'istante $t = t' = 0$, otteniamo **le più generali trasformazioni di Lorentz non omogenee dette anche trasformazioni di Poincaré**, che compendiano, oltre che una rotazione degli assi, anche una traslazione spaziale ed una traslazione temporale, avremo perciò

$$\begin{cases} \mathbf{x}' = \mathbf{R}(\boldsymbol{\theta})\mathbf{x} + \gamma \left[\dfrac{\gamma}{\gamma+1} \dfrac{\mathbf{v} \cdot \mathbf{R}(\boldsymbol{\theta})\mathbf{x}}{c^2} - t \right] \mathbf{v} + \mathbf{x_0}, \\ \\ t' = \gamma(v) \left(t - \dfrac{\mathbf{v} \cdot \mathbf{R}(\boldsymbol{\theta})\mathbf{x}}{c^2} \right) + t_0, \end{cases} \tag{3.60}$$

dove $\mathbf{x_0}$ e t_0 sono costanti, la prima vettoriale e la seconda scalare.
Concludendo, avvertiamo il lettore che il modo migliore e più trasparente di scrivere le trasformazioni generali di Lorentz si ottiene ponendosi nello spaziotempo 4-dimensionale di Minkowski ed utilizzando il linguaggio del calcolo tensoriale come vedremo nel Cap.4.

3.5 APPENDICE: Un teorema sulle forme quadratiche reali in 4 variabili

Teorema:
Si considerino le seguenti due forme quadratiche in 4 variabili reali (forme quaternarie)

$$p(x_0, x_1, x_2, x_3) = x_0^2 - x_1^2 - x_2^2 - x_3^2$$

e

$$q(x_0, x_1, x_2, x_3) = \sum_{i,j=0}^{2} a_{ij} x_i x_j$$

con $a_{ij} \in \mathbb{R}$ e $a_{ij} = a_{ji}$.
Se ogni quaterna di numeri reali che annulla la forma quadratica $p(x_0, x_1, x_2, x_3)$ è tale da annullare anche la forma quadratica $q(x_0, x_1, x_2, x_3)$, allora le due forme differiscono al più per un fattore costante, cioè risulta:

$$q(x_0, x_1, x_2, x_3) = \lambda p(x_0, x_1, x_2, x_3),$$

con $\lambda \in \mathbb{R}$.
Dim. È immediato verificare che le seguenti quaterne annullano la forma quadratica $p(x_0, x_1, x_2, x_3)$:

$$\begin{array}{lll} p_0 = (1,1,0,0), & p_1 = (-1,1,0,0), & p_2 = (1,0,1,0) \\ p_3 = (-1,0,1,0), & p_4 = (1,0,0,1), & p_5 = (-1,0,0,1) \\ p_6 = \left(1, \dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right), & p_7 = \left(1, \dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right), & p_8 = \left(1, 0, \dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right). \end{array}$$

Determiniamo i coefficienti a_{ij} in modo che tali quaterne annullino anche la forma quadratica $q(x_0, x_1, x_2, x_3)$.
Se le quaterne p_0 e p_1 annullano anche $q(x_0, x_1, x_2, x_3)$ dovrebbe essere soddisfatto il seguente sistema lineare:

$$\begin{cases} a_{00} + 2a_{10} + a_{11} = 0 \\ a_{00} - 2a_{01} + a_{11} = 0 \end{cases}$$

che ammette come soluzione $a_{00} = -a_{11}, a_{01} = a_{10} = 0$.
Se le quaterne p_2 e p_3 annullano anche $q(x_0, x_1, x_2, x_3)$ dovrebbe essere soddisfatto il seguente sistema di equazioni:

$$\begin{cases} a_{00} + 2a_{20} + a_{22} = 0 \\ a_{00} - 2a_{02} + a_{22} = 0 \end{cases}$$

che ammette come soluzione $a_{00} = -a_{22}, a_{02} = a_{20} = 0$.
Se le quaterne p_4 e p_5 annullano anche $q(x_0, x_1, x_2, x_3)$ dovrebbe essere soddisfatto il seguente sistema di equazioni:

$$\begin{cases} a_{00} + 2a_{30} + a_{33} = 0 \\ a_{00} - 2a_{03} + a_{33} = 0 \end{cases}$$

che ammette come soluzione $a_{00} = -a_{33}, a_{03} = a_{30} = 0$.
Affinchè la quaterna p_6 annulli il polinomio $q(x_0, x_1, x_2, x_3)$ (si tenga conto dei valori fin qui determinati per semplificare i calcoli) deve essere $\frac{1}{2}a_{12} + \frac{1}{2}a_{21} = 0$ e poichè, per ipotesi $a_{21} = a_{12}$, si ottiene $a_{21} = a_{12} = 0$. In modo analogo, imponendo che le quaterne p_7 e p_8 annullino il polinomio $q(x_0, x_1, x_2, x_3)$ si ottiene $a_{31} = a_{13} = a_{23} = a_{32} = 0$.
Mettendo insieme tutti i risultati fin qui ottenuti si trova che, se si vuole che il polinomio $q(x_0, x_1, x_2, x_3)$ si annulli in corrispondenze delle quaterne $p_i, i = 0, 1, \ldots 8$ in cui si annulla il polinomio $q(x_0, x_1, x_2, x_3)$ deve essere:

$$\begin{aligned} q(x_0, x_1, x_2, x_3) &= a_{00}x_0^2 - a_{00}x_1^2 - a_{00}x_2^2 - a_{00}x_3^2 = a_{00}\left(x_0^2 - x_1^2 - x_2^2 - x_3^2\right) \\ &= a_{00}p(x_0, x_1, x_2, x_3) \end{aligned}$$

ovvero $q(x_0, x_1, x_2, x_3) = \lambda p(x_0, x_1, x_2, x_3)$ con $\lambda = a_{00}$. Risulta ora chiaro che qualunque sia la quaterna-eventualmente diversa da una delle p_i, $i = 0.1, \ldots 8$ prima considerate- che annulla $p(x_0, x_1, x_2, x_3)$ essa deve annullare anche il polinomio $q(x_0, x_1, x_2, x_3)$. c.d.d.

NOTA: La dimostrazione svolta consente di affermare che la costante di proporzionalità fra i due polinomi $p(x_0, x_1, x_2, x_3)$ e $q(x_0, x_1, x_2, x_3)$ é effettivamente una costante numerica e non una funzione dipendente da tutte le variabili in gioco. Questo consente di semplificare la parte finale della dimostrazione sull'invarianza del $\mathbf{ds}^2$, data nel paragrafo 3.2.7, perchè basta usare una sola volta la reciprocità del moto per ottenere $\lambda^2 = a_{00}^2 = 1$. Occorre poi escludere $a_{00} = -1$ come abbiamo visto nel paragrafo citato.

Capitolo 4

CINEMATICA RELATIVISTICA

"Ma non posso crederci", esclamò Alice.
Non puoi? - chiese la regina in tono compassionevole. - Provaci di nuovo: inspira profondamente e chiudi gli occhi.
Inutile provare: uno non può credere le cose impossibili.
Si direbbe che non hai fatto molto esercizio - disse la regina. - All'età tua facevo mezzora di esercizio al giorno. Figurati che qualche volta sono riuscita a credere sei cose impossibili prima di colazione!
(Lewis Carroll, *Alice attraverso lo specchio*) [1]

4.1 Conseguenze delle trasformazioni di Lorentz

4.1.1 Osservazioni fondamentali

Le trasformazione di Lorentz sono una conseguenza diretta dei nuovi assiomi introdotti da Einstein nella fisica. Affinché la nuova teoria sia una buona teoria deve verificare, come si è già detto a suo tempo (Cap.0, Premessa metodologica, paragrafo 0.1), almeno tre condizioni essenziali:

1) prevedere nuovi fenomeni;

2) contenere come caso limite la teoria precedente;

3) fornire risultati verificabili sperimentalmente (*cioè soddisfare al criterio dell'evidenza sperimentale*).

In più, tutte le leggi della Relatività Ristretta, devono soddisfare ad un quarto requisito di fondamentale importanza:
4) essere invarianti per trasformazioni di Lorentz (e perciò rispettare il Principio di Relatività di Einstein).
La teoria della Relatività di Einstein, come si vedrà, soddisfa, in modo eccellente, ai requisiti precedenti. In primo luogo perché prevede un'ampia gamma di nuovi fenomeni chiamati genericamente *effetti relativistici*, come la dilatazione dei tempi, la contrazione delle lunghezze, una nuova legge di composizione delle velocità, l'esistenza di una velocità limite, un nuovo modo di concepire la massa e l'energia, ecc., per non citarne che alcuni.
Inoltre soddisfa alla seconda condizione perché, nel cosiddetto **limite newtoniano (o galileiano)**, cioè per $c \to \infty$ le trasformazioni di Lorentz si riducono alle trasformazioni di Galileo e

[1] Questa citazione è riportata da Clement V. Durell nel suo libro *La Relatività con le quattro operazioni*, Universale Scientifica Boringhieri, 1976, pag. 15.

perciò le leggi relativistiche contengono come caso limite le leggi newtoniane.
Anche il terzo criterio è ampiamente soddisfatto perché tutte le previsioni della teoria sono state sottoposte ad innumerevoli verifiche sperimentali ottenendo sempre ottimi risultati: in effetti la teoria della Relatività Ristretta è fra le teorie fisiche una di quelle meglio verificate sperimentalmente. Infine il quarto criterio è verificato formulando la teoria nello spaziotempo di Minkowski (la cosiddetta **formulazione covariante**) in quanto le sue leggi venendo espresse in forma tensoriale - cioè in termini di tensori del primo ordine (4-vettori) o di ordine superiore - saranno automaticamente invarianti per trasformazioni di Lorentz.
Osserviamo ancora che il non aver postulato a priori alcuna relazione fra le variabili t e t' e l'aver assunto gli assiomi di Einstein alla base della fisica, implica come prima conseguenza fondamentale della teoria della relatività einsteiniana, che **il tempo non è assoluto** con tutte le conseguenze che ciò comporta e che studieremo dettagliatamente.
Infine, ricordando la definizione (3.30) del fattore γ, ed esaminando la trasformazione di Lorentz, si può notare che affinchè questa si mantenga reale, un riferimento galileiano non può muoversi, rispetto ad un altro, con velocità $v \geq c$, perciò deve essere sempre $v < c$ e ciò implica che si abbia sempre $0 < \beta < 1$ e $\gamma > 1$. Questa è una prima indicazione della **insuperabilità della velocità della luce nel vuoto**, la quale perciò assume il carattere di **velocità limite**: ciò verrà esaminato più dettagliatamente in seguito.
Esaminiamo ora **le principali proprietà e le conseguenze cinematiche più importanti delle trasformazioni di Lorentz.**

4.1.2 Simmetria nelle variabili y e z e nelle variabili x e $T = ct$

Le equazioni della trasformazione speciale di Lorentz godono di una notevole proprietà di simmetria, che manca alla trasformazione classica di Galileo, in quanto le (3.33a) e (3.33b) sono simmetriche non soltanto in y e z ma anche nelle variabili x e $T = ct$, con ciò si vuol dire che una trasformazione di Lorentz rimane invariata scambiando fra di loro le variabili y e z, x e $T = ct$ e le analoghe variabili accentate; infatti così facendo, la 1ª equazione diventa la 4^a, e la 2ª diventa la 3^a.

4.1.3 Limite Newtoniano

La trasformazione di Lorentz sostituisce la vecchia trasformazione di Galileo, alla quale tuttavia si approssima - come è desiderabile per una nuova teoria - quando $v \ll c$, ovvero $\beta \ll 1$. Questo spiega l'alta accuratezza della meccanica newtoniana - invariante per trasformazioni di Galileo - nel descrivere un largo dominio della Natura. Si ossevi però, per l'esattezza, che formalmente la trasformazione di Lorentz si riduce a quella di Galileo soltanto quando $c \to \infty$.
Vale la pena sottolineare che il vero limite newtoniano è proprio $c \to \infty$ e non semplicemente $v \ll c$. Se ci limitiamo infatti a richiedere che $v \ll c$, la trasformazione relativistica del tempo, nell'ultima formula di (3.32a), non si riduce necessariamente a $t' = t$, perché in tal caso avremo $\gamma(v) \approx 1$, ma potendo essere x arbitrariamente grande non è detto che si abbia $\frac{v}{c^2}x \approx 0$.

4.1.4 Proprietà del fattore di Lorentz

Per $v \neq 0$ il fattore di Lorentz $\gamma(v)$ è sempre maggiore di 1 ($\gamma > 1$) ma non di troppo finchè v è abbastanza piccola. Per esempio, se consideriamo la velocità orbitale della Terra $v \approx 30\,\mathrm{km/s}$ si ha $\beta = \frac{v}{c} = 10^{-4}$ perciò $\beta^2 = 10^{-8}$ nel qual caso $\gamma = 1,0000000050$, cio è $\gamma - 1 = 5 \times 10^{-9}$; se $v \geq \frac{1}{7}c$ (velocità alla quale il giro della Terra è fatto in un secondo) $\gamma < 1,01$; quando $\beta = 0,866$, $\gamma = 2$; e quando $\beta = 0,\underbrace{999...999}_{2n}5$, si ha $\gamma > 10^n$.

Infatti, partiamo da un esempio concreto che ci fa capire come vanno le cose:

$$\beta = 0,\underbrace{9999}_{2n=4}5 = 0,9999 + 0,00005 = 9999 \cdot 10^{-4} + 5 \cdot 10^{-5} = (10^4 - 1) \cdot 10^{-4} + \frac{1}{2} \cdot 10^{-4},$$

mettendo in evidenza, nell'ultimo passaggio, 10^{-4} possiamo scrivere

$$\beta = 10^{-4}(10^4 - 1 + \frac{1}{2}) = 1 - \frac{1}{2} \cdot 10^{-4}$$

perciò, con ragionamento induttivo, avremo

$$\beta = 0,\underbrace{999...999}_{2n}5 = 1 - \frac{1}{2}10^{-2n}. \tag{4.1}$$

Dalla (4.1) si ottiene

$$\beta^2 = (1 - \frac{1}{2}10^{-2n})^2 = 1 - 10^{-2n} + \frac{1}{4}10^{-4n}$$

trascurando l'ultimo termine che è molto piccolo avremo

$$1 - \beta^2 \approx 10^{-2n}. \tag{4.2}$$

Tenendo presente che $\gamma^2 = \dfrac{1}{1-\beta^2}$, avremo $\gamma^2 \approx 10^{2n}$, cioè $\gamma \approx 10^n$, in effetti è $\gamma > 10^n$ c.d.d.[2]

Figura 4.1: Andamento del fattore di Lorentz in funzione della velocità römeriana β.

Quando $v \to c$, $\gamma \to \infty$, ed inoltre $\gamma = 1$ per $v = 0$ oppure per $c = \infty$; se riportiamo su di un grafico, (vedi Fig.4.1), l'andamento di γ in funzione di β osserviamo subito che inizialmente, cioè per valori piccoli di v e quindi per valori piccoli di β fino a $\beta \approx 0,6$ nel qual caso $\gamma \approx 1,25$, la curva ha un andamento abbastanza piatto, essa è perciò quasi parallela all'asse delle ascisse. Per $\beta > 0,9$ essa si impenna molto rapidamente e tende asintoticamente a $+\infty$ per $\beta \to 1$. Infine, è facile verificare che il fattore di Lorentz, soddisfa alle seguenti utili identità:

$$\gamma^2 = \frac{1}{1-\beta^2} = \frac{c^2}{c^2 - v^2}, \quad \gamma v = c(\gamma^2 - 1)^{1/2}, \quad c^2 d\gamma = \gamma^3 v dv, \quad d(\gamma v) = \gamma^3 dv. \tag{4.3}$$

[2]Si può verificare con facili calcoli che il valore esatto del fattore di Lorentz, per l'esempio svolto più sopra, è $\gamma = 10^n(1 - \frac{1}{4} \cdot 10^{-2n})^{-1/2}$.

4.1.5 Esistenza di una sola velocità invariante

La trasformazione di Lorentz poteva essere ricavata anzichè con l'uso della velocità della luce utilizzando qualsiasi altro *"effetto fisico"* o qualsiasi *"interazione"* **la cui velocità nel vuoto fosse invariante nel passare da un riferimento inerziale ad un altro in relazione** $x - standard$ **col primo**, e ciò perchè nella deduzione si è fatto uso soltanto della proprietà d'invarianza di tale velocità e di nient'altro. *Poichè soltanto una trasformazione può essere valida, ne segue che tutti gli ipotetici "effetti" che potrebbero essere usati* (come per es. onde gravitazionali, onde elettromagnetiche, interazioni di tutti i tipi, in particolare interazioni deboli, interazioni forti, ecc,...) *devono propagarsi con la stessa velocità della luce.*

4.1.6 Relatività del concetto di simultaneità

In tutte le considerazioni che seguono i riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$ saranno considerati sempre in configurazione $x - standard$, inoltre, per semplificare le discussioni introduciamo le seguenti locuzioni: diremo che due eventi $\mathcal{A}$, $\mathcal{B}$ sono **equilocati**, in un determinato riferimento galileiano $\mathcal{R}$, **se essi accadono nello stesso luogo**, cioè se sono verificate le seguenti uguaglianze

$$x_B - x_A \equiv \Delta x = 0, \quad y_B - y_A \equiv \Delta y = 0, \quad z_B - z_A \equiv \Delta z = 0. \tag{4.4}$$

Se i due eventi verificano soltanto la $\Delta x = 0$, con $\Delta y \neq 0$, $\Delta z \neq 0$ diremo che essi sono $x -$ **equilocati**, cioè accadono su di un piano perpendicolare all'asse x o se si vuole su di un piano parallelo al piano coordinato yz; in maniera del tutto analoga si parlerà di eventi $y -$ **equilocati**, oppure $z -$ **equilocati** (Fig.4.2).

Figura 4.2: Relatività della simultaneità. Eventi x- equilocati.

Similmente, come già sappiamo, diremo che due eventi $\mathcal{A}$, $\mathcal{B}$ sono **simultanei**, in un determinato riferimento galileiano $\mathcal{R}$, **se essi accadono nello stesso istante**, cioè se è verificata l'uguaglianza

$$t_B - t_A \equiv \Delta t = 0. \tag{4.5}$$

Dalle trasformazioni speciali di Lorentz (3.35a) e (3.35b), scritte sotto forma di differenze finite, risulta immediatamente che se due eventi $\mathcal{A}$, $\mathcal{B}$ accadono nello stesso luogo, cioè (**sono equilocati**) e nello stesso istante, cioè (**sono simultanei**), in un riferimento galileiano $\mathcal{R}$, la stessa cosa si verifica in ogni altro riferimento galileiano $\mathcal{R}'$. Ciò è del resto ben naturale, ma anche banale, dato che in realtà i due eventi non sono che un unico evento.

Nel caso che i due riferimenti $\mathcal{R}$, $\mathcal{R}'$ siano in configurazione $x-$ *standard* come abbiamo ammesso, e perciò che le trasformazioni di Lorentz siano speciali, può però accadere che due eventi non equilocati in $\mathcal{R}$ ma ivi simultanei, siano simultanei anche nel riferimento $\mathcal{R}'$, si avrebbe cioè una simultaneità assoluta, almeno relativamente a quegli eventi. Ciò si verifica se i due eventi $\mathcal{A}$, $\mathcal{B}$ pur essendo distinti siano, in $\mathcal{R}$, $x-$ **equilocati** e simultanei. Infatti in tal caso avendosi $\Delta x = 0, \Delta y \neq 0, \Delta z \neq 0, \Delta t = 0$, le trasformazioni di Lorentz darebbero $\Delta x' = 0, \Delta y' \neq 0, \Delta z' \neq 0, \Delta t' = 0$, cioè gli eventi sarebbero non equilocati ma ancora simultanei anche in $\mathcal{R}'$.
Le cose vanno invece diversamente se in $\mathcal{R}$ i due eventi, pur essendo simultanei $(\Delta t = 0)$, non accadono su piani paralleli al piano yz, cioè non sono $x-$ **equilocati**, in tal caso essi avranno certamente diversa ascissa, cioè $\Delta x \neq 0$.
Dalla 1^a e dalla 4^a delle (3.35a) si vede immediatamente che anche nel riferimento $\mathcal{R}'$ i due eventi non sono né simultanei né $x'-$ **equilocati**. Infatti avremo

$$\Delta t' = -\gamma \frac{v}{c^2} \Delta x \neq 0, \qquad \Delta x' = \gamma \Delta x \neq 0.$$

Quindi la trasformazione di Lorentz contiene in sé la relatività del concetto di simultaneità. In definitiva, si può affermare che:

> *Se due eventi distinti accadono su di un piano perpendicolare all'asse x e sono simultanei in un certo riferimento galileiano, allora i due eventi sono simultanei anche in qualsiasi altro riferimento galileiano.*
> *Se invece i due eventi sono simultanei ma non $x-$ equilocati $(\Delta x \neq 0)$, in un dato sistema galileiano, in un altro arbitrario sistema galileiano essi non saranno né simultanei, né $x'-$ equilocati $(\Delta x' \neq 0)$.*

Analoghi risultati si hanno nel caso che i due eventi $\mathcal{A}$, $\mathcal{B}$ siano $y-$ **equilocati** ovvero $z-$ **equilocati** e simultanei in $\mathcal{R}$.

4.1.7 Relatività delle durate temporali (dilatazione dei tempi)

Consideriamo i due soliti riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$ in configurazione *x-standard*. Siano $\mathcal{A}$ e $\mathcal{B}$ due eventi che in $\mathcal{R}'$ abbiano coordinate (x'_A, y'_A, z'_A, t'_A), (x'_B, y'_B, z'_B, t'_B), inoltre supponiamo che $\mathcal{B}$ sia **un evento equilocato ma non simultaneo ad** $\mathcal{A}$, cioè un evento che avviene nello stesso luogo nel quale avviene $\mathcal{A}$ in $\mathcal{R}'$ $(x'_B = x'_A, y'_B = y'_A, z'_B = z'_A)$, ma non nello stesso istante $(t'_B \neq t'_A)$, avente perciò le stesse coordinate spaziali ed una differente coordinata temporale per cui sarà $\Delta x' = 0$, $\Delta y' = 0$, $\Delta z' = 0$, $\Delta t' \neq 0$.
Se si utilizzano le trasformazioni speciali di Lorentz (3.35a), (3.35b), scritte sotto forma di differenze finite, è facile determinare sia il Δx, sia il Δt misurato in $\mathcal{R}$; per la 1ª delle (3.35b) avremo

$$\Delta x = \gamma v \Delta t' \neq 0$$

e perciò gli eventi $\mathcal{A}$ e $\mathcal{B}$ in $\mathcal{R}$ non sono equilocati.
Il legame che intercorre fra l'intervallo temporale $\Delta t'$ misurato in $\mathcal{R}'$ e quello Δt misurato in $\mathcal{R}$ si ottiene dalla 4ª delle (3.35b), la quale, essendo $\Delta x' = 0$, ci dà immediatamente

$$\Delta t = \gamma \Delta t' = \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\Delta t'}{\sqrt{1 - \beta^2}}, \qquad (\Delta t > \Delta t'). \tag{4.6}$$

Il risultato (4.6), noto come **dilatazione dei tempi**, si può così enunciare:

Se in un riferimento galileiano $\mathcal{R}'$ **due eventi sono equilocati** *e sono separati da un intervallo temporale $\Delta t'$, essi in un altro riferimento galileiano $\mathcal{R}$, in moto rispetto ad $\mathcal{R}'$,* **non sono equilocati** *e l'intervallo temporale Δt dal quale sono separati risulta dilatato per il fattore di Lorentz γ rispetto a $\Delta t'$.*

Data l'importanza di questo effetto relativistico e tenuto conto delle complicazioni ed incomprensioni di natura psicologica a cui si presta ed ai paradossi a cui può dar luogo (per esempio il famosissimo **paradosso dei gemelli** che discuteremo in seguito) è opportuno soffermarsi ulteriormente su di esso per illustrarlo un pò più dettagliatamente.
Il risultato contenuto nella (4.6) ed enunciato subito dopo, possiamo spiegarlo, forse in maniera più espressiva, considerando che il tempo è misurato con degli orologi. A tal fine, supponiamo che un orologio sia in quiete nella posizione x' del riferimento $\mathcal{R}'$. Fissiamo l'attenzione sulla piccola lancetta che scandisce i secondi e che compie, di moto circolare uniforme, un giro intero in un minuto. Prendiamo come unità di tempo, il tempo che essa impiega a compiere un giro completo del quadrante, cioè 1 minuto. Consideriamo come eventi i due passaggi consecutivi della lancetta per una medesima posizione del quadrante alla quale facciamo corrispondere l'istante t', il passaggio successivo della lancetta per la stessa posizione avverrà all'istante $t'+1$. L'osservatore di $\mathcal{R}$, in base alla trasformazione di Lorentz, registra gli istanti in cui accadono tali eventi come

$$t_1 = \gamma\big(t' + \frac{\beta}{c}x'\big), \quad e \quad t_2 = \gamma\Big((t'+1) + \frac{\beta}{c}x'\Big).$$

Sottolineiamo che gli istanti t' e $t'+1$ relativi ai due eventi considerati **sono registrati da un unico orologio** che si trova in quiete, nel riferimento $\mathcal{R}'$, in una posizione fissa x', mentre i tempi t_1 e t_2, relativi agli stessi eventi, **son segnati da due orologi differenti del riferimento** $\mathcal{R}$, precisamente l'orologio stazionario in $\mathcal{R}$ che coincide con quello mobile, solidale ad $\mathcal{R}'$, all'inizio dell'intervallo, e l'orologio stazionario in $\mathcal{R}$ che coincide con quello mobile alla fine dell'intervallo. Questi orologi sono tuttavia sincronizzati, quindi l'intervallo di tempo che essi registrano fra i due eventi considerati, cioè fra i due passaggi consecutivi della lancetta dell'unico orologio di $\mathcal{R}'$ per la stessa posizione del suo quadrante, è in $\mathcal{R}$,

$$\Delta t = t_2 - t_1 = \gamma(v) = \frac{1}{\sqrt{1-\beta^2}} > 1. \tag{4.7}$$

Quindi il tempo unitario misurato dall'unico orologio di $\mathcal{R}'$, è registrato **come un tempo più lungo** dagli orologi di $\mathcal{R}$. Dal punto di vista dell'osservatore $\mathcal{R}$, l'orologio in movimento, solidale con $\mathcal{R}'$, risulta rallentato, cioè, batte ad un ritmo più lento del fattore $\sqrt{1-\beta^2}$ come mostra la (4.6) o meglio la (4.7). Questo risultato si applica a tutti gli orologi di $\mathcal{R}'$, osservati da $\mathcal{R}$, poiché nella nostra dimostrazione la posizione x' è arbitraria.
In relatività, comunemente si parla del riferimento in cui il corpo osservato è in riposo come del **riferimento proprio**. Perciò **l'intervallo di tempo proprio è l'intervallo di tempo registrato da un orologio solidale con il corpo osservato**. L'intervallo di tempo proprio può essere equivalentemente pensato come l'intervallo di tempo fra due eventi che avvengono nello stesso luogo in un dato riferimento (che può essere $\mathcal{R}'$ o anche $\mathcal{R}$) oppure come l'intervallo di tempo misurato da un unico orologio in un certo luogo di un dato riferimento. Un intervallo di tempo non proprio (o improprio) è un intervallo di tempo misurato da due differenti orologi in due luoghi differenti sempre di un determinato riferimento. Così, dalla precedente discussione vediamo che se $d\tau \equiv dt'$ rappresenta un intervallo infinitesimo di tempo proprio, misurato in $\mathcal{R}'$, allora le espressioni

$$dt = \gamma d\tau = \frac{d\tau}{\sqrt{1-\beta^2}}, \qquad d\tau = \sqrt{1-\beta^2}dt, \tag{4.8}$$

legano l'intervallo non proprio dt, misurato in $\mathcal{R}$, all'intervallo proprio $d\tau$, misurato in $\mathcal{R}'$, dalle quali si vede che $d\tau$ è una quantità sempre minore di dt.
In seguito, torneremo ancora sul concetto di tempo proprio e definiremo altre **grandezze proprie**, come *la lunghezza propria di un regolo rigido, la massa propria di una particella, la frequenza propria di una vibrazione elettromagnetica, ecc.*, e **troveremo che esse rappresentano grandezze invarianti nella teoria della relatività**.
Un'altra conseguenza notevole delle trasformazioni di Lorentz è questa: sebbene gli orologi di un riferimento mobile sembrano tutti rallentati quando sono osservati da un riferimento stazionario rispetto al quale essi sono in movimento, **gli orologi mobili sembrano differire l'uno dall'altro nelle loro letture di una fase costante che dipende dalle loro posizioni, cioè essi sembrano essere non sincronizzati**. Questo fatto discende subito in modo evidente dall'equazione di trasformazione

$$t = \gamma(v)(t' + \frac{\beta}{c}x').$$

Infatti, consideriamo un istante di tempo nel riferimento $\mathcal{R}$, cioè un dato valore di t. Allora, per soddisfare la precedente equazione, $t' + \frac{\beta}{c}x'$ deve avere un valore fisso ben definito. Ciò significa che più grande è x' (cioè, tanto più lontano sull'asse x' è posto un orologio di $\mathcal{R}'$) più piccolo dovrà essere t' (cioè, tanto più indietro apparirà questo orologio rispetto a quelli che lo precedono sull'asse x'). **Quindi gli orologi mobili (situati in $\mathcal{R}'$) sembrano sfasati, cioè non sincronizzati, l'uno rispetto all'altro per l'osservatore in $\mathcal{R}$.** Questa è un'altra manifestazione del fatto che due eventi che accadono simultaneamente nel riferimento $\mathcal{R}$ non sono, in generale, simultanei quando sono osservati dal riferimento $\mathcal{R}'$, e viceversa.
Tutti i risultati di questo paragrafo si possono rovesciare. Cioè, indipendentemente da quale riferimento prendiamo come proprio, l'osservatore di un qualsiasi altro riferimento - per gli eventi che avvengono nello stesso luogo nel riferimento proprio ed in luoghi diversi nel riferimento non proprio - misura un intervallo di tempo dilatato (allungato) rispetto a quello misurato nel riferimento proprio e trova che gli orologi in movimento, rispetto ad esso, non sono sincronizzati.

4.1.8 Relatività dell'ordine temporale

Ci possiamo chiedere ora se l'ordine temporale è un concetto assoluto oppure relativo al sistema di riferimento, in altri termini è possibile l'inversione dell'ordine temporale? Precisamente ci poniamo il seguente problema:
Siano $\mathcal{A}$ e $\mathcal{B}$ due eventi che per l'osservatore $\mathcal{R}$ non sono simultanei, per esempio $\mathcal{B}$ segua $\mathcal{A}$ ($t_B > t_A$), ed avvengono per $\mathcal{R}$ in due diversi punti, che per semplicità supponiamo entrambi sull'asse x. È possibile trovare un sistema di riferimento $\mathcal{R}'$ tale che riferendosi ad esso $\mathcal{B}$ preceda $\mathcal{A}$ ($t'_B < t'_A$)? A tale quesito risponde il seguente teorema:

Teorema 4.1 *C.N.E.S. affinché per due eventi $\mathcal{A}$ e $\mathcal{B}$ non simultanei ($\Delta t = t_B - t_A \neq 0$) nè equilocati nel riferimento galileiano $\mathcal{R}$ (giacenti, per semplicità, sull'asse x con $x_B > x_A$) sia possibile l'inversione dell'ordine temporale in un opportuno riferimento galileiano $\mathcal{R}'$, è che l'intervallo temporale $\Delta t = t_B - t_A$ che separa i due eventi in $\mathcal{R}$ sia minore del tempo τ_{AB} che impiega la luce per percorrere la distanza AB, cioè*

$$t_B - t_A < \tau_{AB}. \tag{4.9}$$

Dimostrazione. La condizione è necessaria
Supponiamo, per fissare le idee, che sia $t_B > t_A$ in $\mathcal{R}$ e che esista un $\mathcal{R}'$ tale che in esso si abbia l'inversione temporale e cioè:

$$t'_B < t'_A. \tag{4.10}$$

Per la trasformazione di Lorentz avremo

$$t'_B = \gamma\left(t_B - \frac{\beta}{c}x_B\right),$$
$$t'_A = \gamma\left(t_A - \frac{\beta}{c}x_A\right),$$

che sostituite nella (4.10), dividendo poi membro a membro per γ, si ottiene

$$t_B - \frac{\beta}{c}x_B < t_A - \frac{\beta}{c}x_A,$$

che possiamo riscrivere

$$t_B - t_A < \frac{\beta}{c}(x_B - x_A) = \beta\tau_{AB}, \tag{4.11}$$

τ_{AB} essendo il tempo che la luce impiega per andare da $\mathcal{A}$ a $\mathcal{B}$. Essendo $\beta < 1$, ne segue $t_B - t_A < \tau_{AB}$.

La condizione è sufficiente.

Supponiamo ora che in $\mathcal{R}$ sia

$$t_B - t_A < \tau_{AB},$$

con $t_B - t_A > 0$ e $\tau_{AB} > 0$, il che equivale a dire che

$$0 < \frac{t_B - t_A}{\tau_{AB}} < 1,$$

esiste allora, per ragioni di continuità, certamente un $0 < \beta < 1$ tale che

$$0 < \frac{t_B - t_A}{\tau_{AB}} < \beta < 1, \tag{4.12}$$

ovvero

$$t_B - t_A < \beta\tau_{AB},$$

cioè

$$t_B - t_A < \frac{\beta}{c}(x_B - x_A)$$

e ancora

$$t_B - \frac{\beta}{c}x_B < t_A - \frac{\beta}{c}x_A,$$

da cui moltiplicando membro a membro per $\gamma > 1$ si ottiene

$$\gamma(t_B - \frac{\beta}{c}x_B) < (t_A - \frac{\beta}{c}x_A),$$

il che equivale, per le trasformazioni di Lorentz, a

$$t'_B < t'_A,$$

perciò esiste almeno un riferimento galileiano $\mathcal{R}'$ dove è possibile l'inversione dell'ordine temporale. ■

Osservazione: Il teorema precedente dimostra che la disuguaglianza

$$t_B - t_A < \tau_{AB}$$

supposta verificata per due dati eventi non equilocati in un riferimento galileiano $\mathcal{R}$, è *C.N.E.S* affinché esista almeno un riferimento $\mathcal{R}'$ dove è possibile l'inversione dell'ordine temporale fra

gli stessi eventi osservati in $\mathcal{R}$, ma non dice se di tali riferimenti ne esista uno solo o più di uno, né stabilisce con quale velocità dovrebbe muoversi $\mathcal{R}'$ rispetto ad $\mathcal{R}$. Ma è subito visto che dovendosi verificare la (4.12), cioè

$$0 < \frac{t_B - t_A}{\tau_{AB}} < \frac{v}{c} < 1,$$

che equivale alla

$$0 < c\frac{t_B - t_A}{\tau_{AB}} < v < c$$

ovvero, moltiplicando membro a membro per c, si avrà

$$0 < c^2\frac{t_B - t_A}{x_B - x_A} < v < c,$$

di tali riferimenti galileiani $\mathcal{R}'$ ne esistono perciò infiniti, mobili di moto traslatorio rettilineo uniforme rispetto ad $\mathcal{R}$ con velocità v soddisfacente alla precedente disuguaglianza.

4.1.9 Relatività delle lunghezze di regoli rigidi (contrazione di Lorentz-Fitzgerald)

Consideriamo, come al solito, due sistemi galileiani $\mathcal{R}$ ed $\mathcal{R}'$ in configurazione x-standard, con $\mathcal{R}'$ che si muova, verso destra, con velocità $\mathbf{v} = cost.$ rispetto ad $\mathcal{R}$.
Fissiamo l'attenzione su un regolo rigido (sbarra rigida) AB che sia in quiete nel riferimento $\mathcal{R}'$ e disposto sull'asse x', i cui estremi abbiano coordinate: $A = (x'_A, 0, 0)$, $B = (x'_B, 0, 0)$, con $x'_B > x'_A$. La sua lunghezza misurata in $\mathcal{R}'$ (*chiamata lunghezza di quiete o lunghezza propria*) sarà quindi $\ell' = x'_B - x'_A = \Delta x'$. Per valutare la lunghezza ℓ della medesima sbarra in un altro riferimento galileiano $\mathcal{R}$ (supposto in configurazione x-standard con $\mathcal{R}'$), gli osservatori di $\mathcal{R}$ rilevano ad un certo istante t, cioè simultaneamente ($\Delta t = 0$), le tracce lasciate in $\mathcal{R}$ dalle estremità A e B della sbarra, le quali staranno ovviamente sull'asse x. Denotata con $\ell = x_B - x_A = \Delta x$ la lunghezza della sbarra AB vista da $\mathcal{R}$, per la 1ª delle (3.35a) si avrà:

$$\Delta x' = \gamma \Delta x, \tag{4.13}$$

cioè

$$\ell = \frac{\ell'}{\gamma} = \ell'\sqrt{1 - \frac{v^2}{c^2}} = \ell'\sqrt{1 - \beta^2}, \qquad (\ell < \ell'). \tag{4.14}$$

Il risultato, noto come **contrazione delle lunghezze**, si può cos ì enunciare:

> *Un regolo rigido in moto traslatorio rettilineo uniforme con velocità v rispetto a un dato riferimento galileiano, risulta contratto, nella direzione del moto, per il fattore $1/\gamma = \sqrt{1 - \beta^2}$, inverso del fattore di Lorentz.*

Osserviamo che **la lunghezza propria** di un'asta rigida, cioè la lunghezza misurata in un riferimento in cui l'asta si trova in quiete, è la più grande fra tutte le lunghezze osservate da sistemi di riferimento in moto rispetto all'asta. Si noti, invece, che le sbarre disposte perpendicolarmente alla direzione del moto non subiscono nessuna contrazione. Le sbarre disposte in modo qualsiasi, poi, subiscono una contrazione compresa tra 1 e $1/\gamma$. Anche il volume di un solido viene alterato per il fattore $1/\gamma = \sqrt{1 - \beta^2}$, cioè :

$$V = V'\sqrt{1 - \beta^2}.$$

Per tale effetto un corpo in quiete rispetto ad $\mathcal{R}'$, che ivi abbia forma sferica di raggio a, in $\mathcal{R}$ ha la forma di un ellissoide rotondo con asse di rotazione parallelo alla direzione del moto, quindi

coincidente con l'asse x, il cui semiasse (*semiasse polare*) essendo contratto rispetto al raggio della sfera, misura $a\sqrt{1-\beta^2}$; il semiasse ortogonale alla direzione del moto (*semiasse equatoriale*), non subendo alcuna contrazione, ha misura a uguale al raggio della sfera. Perciò anche le nozioni di lunghezza, di volume, ecc., perdono in relatività, il loro carattere assoluto.
Chiudiamo questo paragrafo proponendo il seguente:
Esercizio (Paradosso dell'asta e del ripostiglio).[3]

Un atleta entra correndo in un ripostiglio lungo 1 m, trasportando un'asta orizzontale lunga 2 m ad una velocità $v = \dfrac{\sqrt{3}}{2}c \approx 0,866\,c$ e chiude la porta. Si spieghi come ciò sia possibile sia nel sistema di riferimento solidale col ripostiglio, sia nel sistema di riferimento solidale con l'atleta. Naturalmente si deve supporre che nel retro del ripostiglio ci sia un blocco di cemento armato talmente resistente da sopportare l'urto e non ci sia alcuna questione su come l'asta si ferma nel riferimento inerziale del ripostiglio.

Precisiamo un pò meglio i termini della questione: indichiamo con AB l'asta a riposo sull'asse x del riferimento inerziale $\mathcal{R} \equiv Oxyz$ solidale con l'atleta, come evidenziato nella Fig.4.3. Supponiamo inoltre che il secondo estremo B dell'asta coincida con l'origine del sistema di riferimento $B \equiv O$.
Sia $l_A = 2$ m la lunghezza dell'asta nel riferimento $\mathcal{R}$ solidale all'atleta la cui velocità rispetto al riferimento inerziale $\mathcal{R}'$ solidale al ripostiglio è $v = \dfrac{\sqrt{3}}{2}c$.
Sia $A'B'C'D'$ il ripostiglio di forma rettangolare, a riposo nel riferimento $\mathcal{R}'$. Tale ripostiglio (vedi Fig.4.3) ha, rispetto a un osservatore solidale con $\mathcal{R}'$, larghezza $A'B' = l'_R = 1$ m, e dispone di una porta P sul muro $A'D'$ che quando è aperta consente ad AB di entrare al suo interno.
All'istante $t = t' = 0$ supponiamo che $\mathcal{R} \equiv \mathcal{R}'$ (cioè l'estremo B dell'asta urta il muro $B'C'$ nell'origine O').

Figura 4.3: Paradosso dell'asta e del ripostiglio.

Si chiede di:

a) Stabilire se per l' osservatore (atleta) solidale a $\mathcal{R}$ all'istante $t = t' = 0$ l'asta si trova, oppure no, all'interno del ripostiglio (basta determinare la lunghezza dell'asta e la larghezza del ripostiglio nel riferimento $\mathcal{R}$ e confrontarle);

b) Analogamente, stabilire se per l'osservatore solidale ad $\mathcal{R}'$ all'istante $t = t' = 0$ l'asta si trova, oppure no, all'interno del ripostiglio (basta determinare la lunghezza dell'asta e la larghezza del ripostiglio nel riferimento $\mathcal{R}'$ e confrontarle);

[3]Sull'origine di questo paradosso il lettore può consultare utilmente: W. Rindler, *Relativity, Special, General and Cosmological*, Second Edition-Oxford University Press, 2006, paragrafo 3.4, pag. 63. Un paradosso analogo, ma concettualmente diverso, è il cosiddetto *paradosso dell'asta nel fienile* che lo studioso può trovare nel bel libro (fortemente raccomandato per un principiante) di Daniel F. Styer, *Capire davvero la Relatività*, Zanichelli, Bologna, 2012, pagg. 181-190.

c) Spiegare la situazione paradossale che deriva dai risultati apparentemente contraddittori che si trovano per le situazioni relative ai punti a) e b).

Soluzione:

a) Per l'atleta solidale ad $\mathcal{R}$ la lunghezza dell'asta è $l_A = 2\,\mathrm{m}$ poiché questa si trova a riposo rispetto ad esso. Si noti che per un osservatore solidale ad $\mathcal{R}$ il riferimento $\mathcal{R}'$ appare in moto con velocità $-v$ e perciò la larghezza del ripostiglio, a tale osservatore, appare contratta e sarà data da $l_R = \dfrac{l'_R}{\gamma}$. Tenuto conto che la velocità dell'atleta è $v = \dfrac{\sqrt{3}}{2}c$ si ottiene subito che il corrispondente fattore di Lorentz è $\gamma = 2$ ed essendo $l'_R = 1\,\mathrm{m}$ si avrà $l_R = 0,5\,\mathrm{m}$, perciò $l_A > l_R$. Possiamo pertanto concludere che, nell'istante $t = t' = 0$ in cui i due riferimenti coincidono ($\mathcal{R} \equiv \mathcal{R}'$), **nel riferimento $\mathcal{R}$ solidale all'atleta l'asta non si trova completamente all'interno del ripostiglio.**

b) Per un osservatore solidale ad $\mathcal{R}'$ la larghezza del ripostiglio è $l'_R = 1\,\mathrm{m}$ poiché questo si trova a riposo in R'. Per tale osservatore il riferimento $\mathcal{R}$ appare in moto con velocità v e quindi l'asta subisce la contrazione di Lorentz e la sua lunghezza sarà data da $l'_A = \dfrac{l_A}{\gamma} = 1\,\mathrm{m}$, ed essendo $\gamma = 2$, avremo $l'_A = l'_R$. Possiamo pertanto concludere che, nell'istante $t = t' = 0$ in cui i due riferimenti coincidono ($\mathcal{R} \equiv \mathcal{R}'$), per **un osservatore solidale a $\mathcal{R}'$ l'asta si trova completamente all'interno del ripostiglio**.
Dunque i pareri espressi dai due osservatori, il primo solidale con $\mathcal{R}$, il secondo con $\mathcal{R}'$, sono, nell'istante $t = t' = 0$, apparentemente contraddittori; vediamo allora come è possibile spiegare questo risultato paradossale.

c) Mostriamo che anche per un osservatore solidale ad $\mathcal{R}$ l'asta, ad un certo istante, si trova contenuta nel ripostiglio. Ricordiamo che all'istante $t = t' = 0$ i due riferimenti coincidono e che l'estremo B dell'asta tocca la parete $B'C'$ in O. Indichiamo, all'istante $t = 0$, con $x_A = -2\,\mathrm{m}$ l'ascissa dell'estremità A dell'asta e con $x_P = l_R = -0.5\,\mathrm{m}$ l'ascissa della porta del ripostiglio sul suo lato anteriore $A'D'$.
Occorre osservare come l'estremità A, dopo che si è verificato l'urto fra l'estremità B e la parete $B'C'$, continui nel suo moto, rettilineo e uniforme, verso destra (vedi Fig.4.3) finché l'onda d'urto elastica - che si origina nell'istante dell'urto e si propaga nell'asta a **velocità finita** non superiore alla velocità c della luce - non arrivi in A.
Per un osservatore solidale a $\mathcal{R}$ anche il ripostiglio si muove verso sinistra, andando cioè incontro ad A, con velocità $-\mathbf{v}$.
Per verificare se l'asta, anche nel riferimento $\mathcal{R}$, è, ad un certo istante, dentro il ripostiglio, indichiamo con $\bar{t}$ il tempo necessario all'onda d'urto (supposto che si muova alla velocità della luce) per raggiungere l'estremità A. Basta allora verificare che si abbia

$$\overline{x}_A \geq \overline{x}_P$$

dove $\overline{x}_A$ rappresenta l'ascissa dell'estremità A all'istante $\bar{t}$ mentre $\overline{x}_P$ denota l'ascissa della porta allo stesso istante.
Per determinare il tempo $\bar{t}$ che occorre all'onda d'urto per raggiungere l'estremo[4] A, basta risolvere la seguente equazione rispetto al tempo $\bar{t}$:

$$x_A + v\bar{t} = -c\bar{t},$$

si ottiene così:

$$\bar{t} = \frac{-x_A}{c+v} \approx \frac{1.071}{c}\,s,$$

[4]Si tenga conto che sia l'estremo A dell'asta sia il ripostiglio si muovono di moto rettilineo uniforme.

e, conseguentemente, si ha

$$\overline{x}_A = x_A + v\overline{t} \approx -2 + \frac{\sqrt{3}}{2}c \cdot \frac{1.071}{c} \approx -1.07\,\mathrm{m}\,, \quad \overline{x}_P = x_P - v\overline{t} \approx -1.42\,\mathrm{m}.$$

Risulta quindi che $\overline{x}_A > \overline{x}_P$ e quindi l'asta anche in $\mathcal{R}$ si trova, all'istante $\overline{t}$, all'interno del ripostiglio.

Osservazione 1:
In via di principio noi non dubitiamo della fattibilità di questo esperimento, cioè sulla "realtà" della contrazione delle lunghezze. Quando l'asta si ferma nel sistema di riferimento del ripostiglio, essa tende ad assumere la sua lunghezza originaria rispetto al ripostiglio. Allora, se essa sopravvive all'impatto, deve ora incurvarsi o sfondare la porta, o rimanere compressa.
Osservazione 2:
Val la pena osservare che in ciò che precede gioca anche il fatto che la **simultaneità è un concetto relativo** al sistema di riferimento, infatti in $\mathcal{R}'$ l'asta si trova nel ripostiglio all'istante $t = 0$, mentre in $\mathcal{R}$ vi si trova all'istante $\overline{t} > 0$.
Chiudiamo la spiegazione di questo utile esercizio rimarcando il fatto che la condizione affinchè l'asta continui a muoversi di moto rettilineo uniforme fintanto che l'estremo A non venga raggiunto dall'onda d'urto elastica proveniente da B è che l'asta non sia un corpo rigido! Quindi questo semplice esercizio mette anche in evidenza come il concetto di rigidità, fondamentale in meccanica classica, non possa essere accettato in Relatività Ristretta.
Lo studente curioso di cimentarsi con altri simili paradossi può, con vantaggio, consultare il seguente libro: J. Freund, *Special Relativity for Beginners*, World Scientific Publishing Co. Pte. Ltd. Singapore, 2008.
Ci occuperemo del più famoso dei paradossi: **il paradosso dei gemelli** nel Cap.5, paragrafo 5.7.

4.1.10 Legge relativistica di composizione delle velocità

Stabiliamo ora quali siano le formule di composizione delle velocità nell'ambito relativistico. Considerati i due soliti riferimenti galileiani $\mathcal{R}$ ed $\mathcal{R}'$ in configurazione x-standard, siano $x = x(t)$, $y = y(t)$, $z = z(t)$, le equazioni del moto di una particella mobile $\mathcal{P}$ nel riferimento $\mathcal{R}$, e siano $u_x = dx/dt$, $u_y = dy/dt$, $u_z = dz/dt$ le componenti cartesiane della sua velocità $\mathbf{u}$ ad un istante generico t.
Nel riferimento $\mathcal{R}'$, invece, le equazioni del moto di $\mathcal{P}$ saranno del tipo $x' = x'(t')$, $y' = y'(t')$, $z' = z'(t')$, e le componenti della sua velocità $\mathbf{u}'$ sugli assi x', y', z', ad un generico istante t', sono espresse dalle seguenti formule[5]

$$u'_x = \frac{dx'}{dt'} = \frac{dx'}{dt}\frac{dt}{dt'} = \frac{\frac{dx'}{dt}}{\frac{dt'}{dt}}, \quad u'_y = \frac{dy'}{dt'} = \frac{dy'}{dt}\frac{dt}{dt'} = \frac{\frac{dy'}{dt}}{\frac{dt'}{dt}}, \quad u'_z = \frac{dz'}{dt'} = \frac{dz'}{dt}\frac{dt}{dt'} = \frac{\frac{dz'}{dt}}{\frac{dt'}{dt}}. \tag{4.15}$$

Per ricavare le relazioni tra le componenti della velocità di $\mathcal{P}$ in $\mathcal{R}$ e in $\mathcal{R}'$ facciamo riferimento alla trasformazione di Lorentz (3.32a): derivando membro a membro rispetto a t le equazioni

[5]Le componenti della velocità $\mathbf{u}'$ sugli assi coordinati del riferimento $\mathcal{R}'$ dovrebbero essere indicate con $u'_{x'}, u'_{y'}, u'_{z'}$, ma, per non appesantire troppo le notazioni, spesso tali componenti vengono indicate più semplicemente con u'_x, u'_y, u'_z come abbiamo fatto anche noi.

che ivi figurano e sostituendo le derivate così calcolate nelle formule precedenti si ottiene:

$$\begin{cases} u'_x = \dfrac{u_x - v}{1 - \dfrac{v}{c^2}u_x} \\[2ex] u'_y = \dfrac{u_y\sqrt{1 - v^2/c^2}}{1 - \dfrac{v}{c^2}u_x} = \dfrac{u_y}{\gamma(v)\left(1 - \dfrac{v}{c^2}u_x\right)} \\[2ex] u'_z = \dfrac{u_z\sqrt{1 - v^2/c^2}}{1 - \dfrac{v}{c^2}u_x} = \dfrac{u_z}{\gamma(v)\left(1 - \dfrac{v}{c^2}u_x\right)}. \end{cases} \tag{4.16}$$

Per quanto riguarda le trasformazioni inverse basta seguire lo stesso procedimento di sopra utilizzando la trasformazione inversa di Lorentz (3.32b), si avrà:

$$\begin{cases} u_x = \dfrac{u'_x + v}{1 + \dfrac{v}{c^2}u'_x} \\[2ex] u_y = \dfrac{u'_y\sqrt{1 - \beta^2}}{1 + \dfrac{v}{c^2}u'_x} = \dfrac{u'_y}{\gamma(v)\left(1 + \dfrac{v}{c^2}u'_x\right)} \\[2ex] u_z = \dfrac{u'_z\sqrt{1 - \beta^2}}{1 + \dfrac{v}{c^2}u'_x} = \dfrac{u'_z}{\gamma(v)\left(1 + \dfrac{v}{c^2}u'_x\right)} \end{cases} \tag{4.17}$$

queste ultime si ottengono più semplicemente cambiando le lettere accentate con quelle non accentate e viceversa, e cambiando v in $-v$, come si è detto più volte.
Facciamo osservare che per i moduli delle velocità $\mathbf{u}$, $\mathbf{u}'$ si ha:

$$|\mathbf{u}| = (u_x^2 + u_y^2 + u_z^2)^{1/2}, \qquad |\mathbf{u}'| = (u'^2_x + u'^2_y + u'^2_z)^{1/2}. \tag{4.18}$$

Facendo tendere $c \to \infty$ nelle (4.16), (4.17) si ottiene la legge classica di composizione delle velocità, e ciò conferma ancora una volta che le formule classiche sono un caso limite delle corrispondenti formule relativistiche.
Nel caso particolare in cui la particella si muove in $\mathcal{R}$ con velocità $\mathbf{u}$ diretta lungo l'asse x e nello stesso verso del moto di $\mathcal{R}'$ rispetto ad $\mathcal{R}$, nel sistema $\mathcal{R}'$ essa si muoverà lungo l'asse x' con velocità data dalla formula di Einstein:

$$u' = \frac{u - v}{1 - \dfrac{v}{c^2}u}, \tag{4.19}$$

che per $c \to \infty$ si riduce alla legge galileiana di composizione delle velocità:

$$u' = u - v. \tag{4.20}$$

Le formule relativistiche di composizione delle velocità (4.16), (4.17) ed in particolare la (4.19) rispettano ovviamente il postulato dell'invarianza della velocità della luce, nel senso che se in $\mathcal{R}$ un raggio luminoso si propaga con velocità c lungo l'asse delle x, in virtù della (4.19) esso avrà in $\mathcal{R}'$ la stessa identica velocità

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

Se vogliamo la legge di composizione delle velocità, in forma vettoriale, per una trasformazione di Loretnz fra sistemi di riferimento che si muovano di moto traslatorio in una direzione arbitraria basta derivare rispetto al tempo t le due formule che compaiono in (3.53) ovvero in (3.57) per ottenere con facili calcoli:

$$\mathbf{u}' = \frac{\mathbf{u} + \left[(\gamma - 1)\dfrac{\mathbf{u}\cdot\mathbf{v}}{v^2} - \gamma\right]\mathbf{v}}{\gamma\left(1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}\right)}, \qquad (\mathbf{v} \neq 0), \tag{4.21}$$

oppure

$$\mathbf{u}' = \frac{\mathbf{u} + \gamma\left[\dfrac{\gamma}{\gamma+1}\dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2} - 1\right]\mathbf{v}}{\gamma\left(1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}\right)}, \quad (\textit{valida qualunque sia } \mathbf{v}). \tag{4.22}$$

Una conseguenza interessante delle (4.16), (4.17) è il seguente:

Teorema 4.2 *A seconda che la particella $\mathcal{P}$ abbia velocità minore, uguale o maggiore di c in un riferimento galileiano, essa avrà velocità minore, uguale o maggiore di c in ogni altro riferimento galileiano.*

Dimostrazione. Supponiamo che nel riferimento $\mathcal{R}$ un punto materiale abbia una velocità $\mathbf{u}$ il cui modulo sia $u \lesseqgtr c$, vogliamo dimostrare che in $\mathcal{R}'$ si avrà $u' \lesseqgtr c$. Dire che $u \lesseqgtr c$ equivale a dire che $u_x^2 + u_y^2 + u_z^2 = u^2 \lesseqgtr c^2$. Utilizzando le (4.17), (4.18) avremo:

$$\frac{\gamma^2\left(u'_x + v\right)^2}{\gamma^2\left(1 + \dfrac{v}{c^2}u'_x\right)^2} + \frac{u'^2_y}{\gamma^2\left(1 + \dfrac{v}{c^2}u'_x\right)^2} + \frac{u'^2_z}{\gamma^2\left(1 + \dfrac{v}{c^2}u'_x\right)^2} \lesseqgtr c^2,$$

da cui:

$$u'^2_x\gamma^2 + 2u'_x v\gamma^2 + v^2\gamma^2 + u'^2_y + u'^2_z \lesseqgtr c^2\gamma^2 + 2u'_x v\gamma^2 + u'^2_x\frac{v^2}{c^2}\gamma^2,$$

$$u'^2_x\gamma^2\left(1 - \frac{v^2}{c^2}\right) + u'^2_y + u'^2_z \lesseqgtr \gamma^2(c^2 - v^2);$$

da quest'ultima, semplificando:

$$u'^2_x + u'^2_y + u'^2_z = u'^2 \lesseqgtr c^2,$$

ovvero $u' \lesseqgtr c$, c.v.d. ■

La velocità della luce c, dunque, costituisce un confine tra due classi di velocità, una costituita da velocità minori di c, l'altra da velocità maggiori di c: l'appartenenza all'una o all'altra di queste classi rappresenta un carattere assoluto di un dato movimento.

Vedremo nel paragrafo 4.1.12 che la classe delle velocità maggiori della velocità c della luce non è fisicamente raggiungibile da particelle che in precedenza abbiano velocità inferiori a quella della luce.

Osservazione: Le formule relative alla legge relativistica di composizione delle accelerazioni sono molto più complicate delle (4.16) e poiché non avremo occasione di usarle non ci soffermeremo su di esse.[6]

[6]Tali formule si possono trovare nel libro di Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), pag. 108.

4.1.11 Legge di trasformazione del fattore di Lorentz

Come applicazione della legge relativistica di composizione delle velocità, in questo paragrafo, ci occupiamo di un'altra utile formula di cui avremo bisogno nel seguito, e cioè la legge di trasformazione del fattore di Lorentz nel passare da un riferimento $\mathcal{R}$ ad un altro $\mathcal{R}'$ che si trovino, come al solito, in relazione x-standard: si tratta dunque di trovare il legame fra $\gamma(u)$, $\gamma(u')$ e $\gamma(v)$.
A tal fine utilizzando le (4.16) otteniamo

$$u'^2 = u_x'^2 + u_y'^2 + u_z'^2 = \frac{(u_x - v)^2 + u_y^2(1 - \frac{v^2}{c^2}) + u_z^2(1 - \frac{v^2}{c^2})}{(1 - \frac{v}{c^2}u_x)^2} =$$
$$= \frac{(u_x - v)^2 + (u^2 - u_x^2)(1 - \frac{v^2}{c^2})}{(1 - \frac{v}{c^2}u_x)^2}, \tag{4.23}$$

da cui, con semplici calcoli avremo

$$1 - \frac{u'^2}{c^2} = \frac{(1 - \frac{v^2}{c^2})(1 - \frac{u^2}{c^2})}{(1 - \frac{v}{c^2}u_x)^2}. \tag{4.24}$$

In termini dei fattori di Lorentz, questo risultato può scriversi

$$\gamma(u') = \gamma(u)\gamma(v)\big(1 - \frac{v}{c^2}u_x\big), \tag{4.25}$$

la cui relazione inversa si ottiene, come al solito, cambiando le lettere accentate con le analoghe senza accento e cambiando v in $-v$:

$$\gamma(u) = \gamma(u')\gamma(v)\big(1 + \frac{v}{c^2}u_x'\big); \tag{4.26}$$

queste formule ci saranno utili in seguito quando ricaveremo le formule di trasformazione delle quantità di moto e delle forze.

4.1.12 Insuperabilità della velocità della luce nel vuoto

Nessun riferimento galileiano può muoversi rispetto ad un altro con velocità uguale o maggiore di c, tale limitazione, come abbiamo già visto, è conseguenza diretta del fatto che noi vogliamo che le trasformazioni di Lorentz (3.32a), (3.32b) siano reali. Ciò succede anche per la velocità di una qualsiasi particella, anche se la teoria della relatività di per sé non esclude l'esistenza di particelle che viaggino a velocità superiore alla velocità della luce, **i cosiddetti tachioni** (tali particelle dovrebbero però nascere già con velocità superiore a c). Ciò che la teoria esclude è che una particella materiale che parta dalla quiete o si trovi già in moto con velocità inferiore alla velocità della luce, possa essere accelerata fino a raggiungere o superare la velocità c, come vedremo in sede dinamica. Tale limitazione viene imposta non dalla teoria della relatività ma dal *principio di causalità (o principio di antecedenza)* che abbiamo accettato, all'inizio di queste lezioni, come principio metodologico (Cap 0, paragrafo 0.2).
In altre parole, si deve ragionevolmente presumere che **la velocità della luce non possa essere superata in alcun "moto fisico", cioè tale che ogni stato istantaneo di moto sia "causa" degli stati successivi**, come può essere il moto di una particella o la trasmissione di un generico segnale ovvero l'interazione fra due particelle o ancora l'interazione fra una particella ed un

campo. Si riconosce infatti che se non si pongono limiti alla velocità di una particella o di un segnale, si possono trarre conclusioni contraddicenti, come abbiamo già detto poco fa, il **principio di causalità** che la Fisica non può non accettare.

Teorema 4.3 Insuperabilità della velocità della luce nel vuoto:
La velocità della luce nel vuoto è una velocità limite non superabile in alcun moto fisico.

Dimostrazione. Siano $\mathcal{R}$ ed $\mathcal{R}'$ i due soliti riferimenti galileiani in correlazione *x-standard*, e consideriamo due eventi distinti $\mathcal{A}$, $\mathcal{B}$ che in $\mathcal{R}$ accadano sull'asse x in due punti di ascisse x_A, x_B rispettivamente, con $x_B > x_A$. Supponiamo inoltre che l'evento $\mathcal{B}$ (effetto) sia provocato dall'evento $\mathcal{A}$ (causa) (e che perciò sia $t_B > t_A$) per mezzo della trasmissione di un *agente fisico*. Se questo agente si trasmettesse con velocità u superiore alla velocità c della luce nel vuoto, cioè $u > c$, ne seguirebbe che

$$\frac{1}{u} < \frac{1}{c},$$

e moltiplicando primo e secondo membro per $x_B - x_A > 0$ si avrebbe

$$\frac{x_B - x_A}{u} < \frac{x_B - x_A}{c},$$

e ciò equivale a dire che, in $\mathcal{R}$, sarebbe verificata la disuguaglianza

$$t_B - t_A < \tau_{AB.}$$

Ma questa, per il teorema 4.1 del paragrafo 4.1.8, è $C.N.E.S.$ affinché esista almeno un riferimento galileiano $\mathcal{R}'$ nel quale si ha l'inversione dell'ordine temporale fra i due eventi considerati, cioè in $\mathcal{R}'$ l'evento $\mathcal{B}$ (effetto) sarebbe precedente all'evento $\mathcal{A}$ (causa), il che è manifestamente assurdo per il principio di causalità che abbiamo citato più sopra. È perciò assurdo supporre che esista un agente fisico che si propaghi con velocità superiore alla velocità della luce nel vuoto $(u > c)$. c.d.d.
Analogo risultato assurdo si può ottenere con un altro ragionamento un pò diverso dal precedente. Supponiamo che, nel riferimento $\mathcal{R}$, venga lanciato, all'istante $t_1 = 0$, un segnale da O nel verso positivo dell'asse x, con velocità $u > c$. Questo evento, diciamolo $\mathcal{E}_1$, nel riferimento $\mathcal{R}$, ha coordinate $x_1 = 0$, $y_1 = 0$, $z_1 = 0$, $t_1 = 0$ e, nel riferimento $\mathcal{R}'$, coordinate $x'_1 = 0$, $y'_1 = 0$, $z'_1 = 0$, $t'_1 = 0$, in quanto all'istante iniziale per entrambi i riferimenti questi coincidono per ipotesi. Nel riferimento $\mathcal{R}$, il segnale, dopo un tempo $t_2 > 0$, avrà percorso una distanza $x_2 = ut_2 > 0$; l'evento $\mathcal{E}_2$ consistente nella ricezione di questo segnale nel punto di ascissa x_2 ha quindi coordinate $x_2 = ut_2$, $y_2 = 0$, $z_2 = 0$, $t_2 > 0$, in $\mathcal{R}$. In $\mathcal{R}'$, invece, tale evento avviene all'istante

$$t'_2 = \gamma\left(t_2 - \frac{v}{c^2}x_2\right) = \gamma t_2\left(1 - \frac{vu}{c^2}\right), \tag{4.27}$$

Ora se teniamo conto dell'ipotesi $u > c$, possiamo osservare che il termine $\frac{vu}{c^2} = \frac{v}{c}\frac{u}{c}$, che compare nella (4.27), si presenta come prodotto di un numero $\frac{v}{c} < 1$ per un numero $\frac{u}{c} > 1$ per cui il prodotto, a priori, potrebbe assumere valori minori, maggiori o anche uguali ad 1; ma essendo sia v che u a nostra disposizione, purché si rispettino le disuguaglianze $v < c$ e $u > c$, è sempre possibile scegliere tali valori in modo che sia $\frac{vu}{c^2} > 1$, in tal caso si avrebbe $t'_2 < 0$, perciò il segnale partito da O all'istante $t'_1 = 0$ verrebbe percepito, in $\mathcal{R}'$ prima dell'evento che lo ha generato, in manifesta contraddizione con il principio di causalità!
Concludiamo osservando che: se è fissata a priori la velocità $v < c$ con cui si muove il sistema $\mathcal{R}'$ rispetto ad $\mathcal{R}$, per poter verificare la disuguaglianza $\frac{vu}{c^2} > 1$ basta scegliere $u > \frac{c^2}{v}$, se invece

si fissa a priori la velocità $u > c$ del segnale, allora bisogna prendere un $\mathcal{R}'$ che si muova con velocità $v > \dfrac{c^2}{u}$ rispetto ad $\mathcal{R}$. ■

Osservazione 1: Val la pena di sottolineare, come già detto all'inizio di questo paragrafo, che l'ipotesi ammessa più sopra di velocità superiori a c non è in contraddizione logica con i postulati di Einstein, essa però, come si è appena visto, contraddice tuttavia in modo evidente il **"principio di causalità", secondo il quale l'effetto non può mai precedere la causa**.
Osservazione 2: Il teorema ora dimostrato *non esclude che una velocità maggiore di c sia possibile in moti aventi carattere puramente cinematico, senza relazione causale tra i suoi singoli eventi*. Per esempio, supponiamo che sul semiasse positivo x sia disposta una distribuzione continua di lampadine puntiformi ed altrettanti operatori. In base ad un accordo preso in precedenza, l'operatore di ascissa x all'istante $t = \dfrac{x}{u}$, u essendo una costante maggiore di c ($u > c$), accende la propria lampadina, che poi rimane accesa. Analogamente per ogni altro operatore. Si viene in tal modo a formare sulla semiretta Ox un segmento luminoso OL il cui estremo L si muove con velocità $u > c$. Questo moto è effettivamente possibile; non si tratta però di un *"moto fisico"*. Infatti tra due accensioni successive di lampadine non intercorre nessuna relazione di causa e di effetto. E nessuna contraddizione col principio di causalità vi sarebbe se in qualche riferimento galileiano due lampadine apparissero accendersi in ordine diverso di quello che appare nel riferimento $\mathcal{R}$.
Osservazione 3: È interessante osservare che la legge di composizione delle velocità sussiste anche per velocità superiori a c, come quelle degli esempi sopra indicati.

4.1.13 Invarianza della legge di propagazione delle onde elettromagnetiche per trasformazioni di Lorentz

Una ulteriore prova della validità della teoria di Einstein viene dal fatto che le leggi dell'elettromagnetismo sono invarianti per trasformazioni di Lorentz. Come esempio di ciò, dimostriamo ora che la legge di propagazione delle onde elettromagnetiche è invariante rispetto alla trasformazione speciale di Lorentz.
Per fare questo consideriamo i due soliti riferimenti galileiani $\mathcal{R}$ e $\mathcal{R}'$ in configurazione x-standard. Supponiamo che un'onda elettromagnetica si propaghi nello spazio vuoto e che nel riferimento $\mathcal{R}$ sia governata da un'equazione alle derivate parziali del 2° ordine del tipo:

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = 0, \tag{4.28}$$

dove $\varphi(x, y, z, t) = 0$ è l'equazione cartesiana del fronte d'onda all'istante t e c è la sua velocità di avanzamento, che nel caso delle onde elettromagnetiche che si propagano nel vuoto coincide con la velocità della luce.
Dimostriamo che applicando la trasformazione speciale di Lorentz alla (4.28) questa si trasforma per invarianza, si ottiene cioè un'equazione avente la stessa struttura della (4.28), espressa però nelle variabili x', y', z', t' del riferimento $\mathcal{R}'$.
Infatti, se calcoliamo le derivate parziali della funzione $\varphi(x, y, z, t)$ rispetto alle variabili x, y, z, t, tenendo conto che esse sono funzioni derivabili delle analoghe variabili x', y', z', t' tramite le

trasformazioni speciali di Lorentz (3.32a) si ottengono dapprima le derivate prime:

$$
\begin{aligned}
\frac{\partial\varphi}{\partial x} &= \frac{\partial\varphi}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial\varphi}{\partial t'}\frac{\partial t'}{\partial x} = \gamma\frac{\partial\varphi}{\partial x'} - \frac{\gamma\beta}{c}\frac{\partial\varphi}{\partial t'},\\
\frac{\partial\varphi}{\partial y} &= \frac{\partial\varphi}{\partial y'},\\
\frac{\partial\varphi}{\partial z} &= \frac{\partial\varphi}{\partial z'},\\
\frac{\partial\varphi}{\partial t} &= \frac{\partial\varphi}{\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial\varphi}{\partial t'}\frac{\partial t'}{\partial t} = -\gamma\beta c\frac{\partial\varphi}{\partial x'} + \gamma\frac{\partial\varphi}{\partial t'},
\end{aligned}
$$

dalle quali è facile calcolare anche le derivate seconde

$$
\begin{aligned}
\frac{\partial^2\varphi}{\partial x^2} &= \gamma\left[\frac{\partial^2\varphi}{\partial x'^2}\frac{\partial x'}{\partial x} + \frac{\partial^2\varphi}{\partial x'\partial t'}\frac{\partial t'}{\partial x}\right] - \frac{\gamma\beta}{c}\left[\frac{\partial^2\varphi}{\partial t'\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial^2\varphi}{\partial t'^2}\frac{\partial t'}{\partial x}\right] =\\
&= \gamma^2\frac{\partial^2\varphi}{\partial x'^2} - 2\frac{\gamma^2\beta}{c}\frac{\partial^2\varphi}{\partial x'\partial t'} + \frac{\gamma^2\beta^2}{c^2}\frac{\partial^2\varphi}{\partial t'^2},\\
\frac{\partial^2\varphi}{\partial y^2} &= \frac{\partial^2\varphi}{\partial y'^2},\\
\frac{\partial^2\varphi}{\partial z^2} &= \frac{\partial^2\varphi}{\partial z'^2},\\
\frac{\partial^2\varphi}{\partial t^2} &= -\gamma\beta c\left[\frac{\partial^2\varphi}{\partial x'^2}\frac{\partial x'}{\partial t} + \frac{\partial^2\varphi}{\partial x'\partial t'}\frac{\partial t'}{\partial t}\right] + \gamma\left[\frac{\partial^2\varphi}{\partial t'\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial^2\varphi}{\partial t'^2}\frac{\partial t'}{\partial t}\right] =\\
&= \gamma^2\beta^2c^2\frac{\partial^2\varphi}{\partial x'^2} - 2\gamma^2\beta c\frac{\partial^2\varphi}{\partial x'\partial t'} + \gamma^2\frac{\partial^2\varphi}{\partial t'^2}.
\end{aligned}
$$

Sostituendo infine queste ultime nella (4.28) si ottiene:

$$
\begin{aligned}
&\gamma^2\frac{\partial^2\varphi}{\partial x'^2} - 2\frac{\gamma^2\beta}{c}\frac{\partial^2\varphi}{\partial x'\partial t'} + \frac{\gamma^2\beta^2}{c^2}\frac{\partial^2\varphi}{\partial t'^2} + \frac{\partial^2\varphi}{\partial y'^2} + \frac{\partial^2\varphi}{\partial z'^2} -\\
&- \gamma^2\beta^2\frac{\partial^2\varphi}{\partial x'^2} + 2\frac{\gamma^2\beta}{c}\frac{\partial^2\varphi}{\partial x'\partial t'} - \frac{\gamma^2}{c^2}\frac{\partial^2\varphi}{\partial t'^2} = 0,
\end{aligned}
$$

ovvero:

$$
\gamma^2\left(1-\beta^2\right)\frac{\partial^2\varphi}{\partial x'^2} + \frac{\partial^2\varphi}{\partial y'^2} + \frac{\partial^2\varphi}{\partial z'^2} - \frac{\gamma^2}{c^2}\left(1-\beta^2\right)\frac{\partial^2\varphi}{\partial t'^2} = 0;
$$

quindi, poiché $\gamma^2 = \left(1-\beta^2\right)^{-1}$, avremo

$$
\frac{\partial^2\varphi}{\partial x'^2} + \frac{\partial^2\varphi}{\partial y'^2} + \frac{\partial^2\varphi}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t'^2} = 0. \tag{4.29}
$$

che ha evidentemente la stessa struttura della (4.28), c.v.d.
La legge di propagazione di un'onda elettromagnetica, come è noto, è conseguenza delle equazioni di Maxwell, ed il fatto che essa sia invariante per trasformazioni di Lorentz non implica, ben inteso, che lo siano anche le equazioni di Maxwell. L'invarianza delle equazioni di Maxwell deve perciò essere dimostrata in modo indipendente: la dimostrazione è però un po' più complicata della precedente data la presenza, in tali equazioni, del campo elettrico **E** e del campo magnetico **H**, perciò per ragioni di brevità, per il momento, la tralasciamo (La dimostrazione verrà fatta nel capitolo sull'elettrodinamica).
Possiamo quindi concludere che, contrariamente a quanto si pensava, o si presupponeva, nell'ambito classico, cioè che l'elettromagnetismo fosse, almeno apparentemente, in conflitto con

il principio di relatività, tale presunzione si è rivelata falsa per il semplice motivo che le trasformazioni corrette, che fanno passare da un sistema galileiano, ad un altro, non sono le trasformazioni di Galileo bensì le trasformazioni di Lorentz. Perciò, siccome le equazioni della meccanica sono invarianti per trasformazioni di Galileo, ne segue che esse non possono essere allo stesso tempo invarianti per trasformazioni di Lorentz. Tutto ciò implica che, per ironia della sorte, le equazioni della meccanica di Galileo-Newton che sembravano inattaccabili perché ormai ben consolidate da almeno due secoli di esperienze, alla luce del principio einsteiniano di relatività, devono essere corrette per renderle Lorentz-invarianti. Invece, l'elettromagnetismo, nonostante sia stato introdotto nella Fisica molto più tardi della meccanica, non necessita di nessuna modifica in quanto le sue equazioni, scritte nella forma classica fornita da Maxwell, sono già di per sé stesse invarianti per trasformazioni di Lorentz e quindi soddisfacenti al principio di relatività di Einstein esteso a tutti i fenomeni fisici.

4.2 Conferme sperimentali della Cinematica Relativistica

L'esperienza, come ora vedremo, sembra offrire significative conferme delle formule relativistiche già in ambito puramente cinematico. In particolare, le esperienze che avevano una spiegazione nell'ambito della fisica classica vengono spiegate altrettanto bene dalla teoria relativistica, ed inoltre, quelle non spiegate dalla teoria classica, hanno valide spiegazioni in relatività.
Fra le tante esperienze che sono state fatte a sostegno della validità della teoria della relatività speciale, per il momento ne illustreremo soltanto alcune nell'ambito della cinematica relativistica, le prime tre delle quali sono già state discusse nel Cap.2 nell'ambito della fisica classica; esse sono:

1) **L'aberrazione della luce stellare**;

2) **L'effetto Doppler longitudinale e trasversale**;

3) **L'esperienza di Michelson-Morley**;

4) **L'allungamento della vita media del mesone μ in virtù della dilatazione del tempo**.

Vedremo ulteriori conferme nell'ambito della dinamica relativistica.

4.2.1 L'aberrazione della luce stellare: spiegazione relativistica

Di questo fenomeno astronomico abbiamo già dato, nel Cap.2, paragrafo 2.4, una spiegazione soddisfacente nell'ambito della meccanica classica. Ci proponiamo ora di spiegarlo dal punto di vista relativistico.
Consideriamo, come nel caso classico, un sistema di riferimento galileiano $\mathcal{R} \equiv (Oxyz, t)$ nella cui origine O sia situato l'occhio dell'osservatore terrestre nella posizione occupata sulla propria orbita attorno al Sole ad un determinato istante che consideriamo come iniziale $t = 0$ per gli orologi di $\mathcal{R}$. Prendiamo l'asse x, uscente da O, parallelo ed equiorientato come la velocità orbitale $\mathbf{v}$ della Terra, l'asse z ortogonale ad x ed al piano dell'orbita terrestre ed orientato verso l'emisfero celeste in cui si trova la stella da osservare, l'asse y giacente su tale piano ed orientato in modo che $Oxyz$ sia una terna destra (Fig.4.4).

Supponiamo però che il riferimento $\mathcal{R}$ non venga trascinato dalla Terra ma sia in quiete nello spazio ovvero sia **solidale con l'etere cioè con le stelle fisse**. Consideriamo inoltre un secondo sistema $\mathcal{R}' \equiv (O'x'y'z', t)$ che all'istante $t = 0$, precedentemente fissato, sia coincidente con $\mathcal{R} \equiv (Oxyz, t)$ ma trascinato dalla Terra con orientamento invariabile, e che perciò possiamo ritenere, almeno per un breve periodo di tempo, in moto traslatorio uniforme con velocità

Figura 4.4: Aberrazione della luce stellare: spiegazione relativistica.

costante **v** rispetto ad $\mathcal{R}$ e quindi anch'esso galileiano. I due sistemi $\mathcal{R}$ ed $\mathcal{R}'$ sono dunque in configurazione x-standard. Il riferimento $\mathcal{R}'$ non lo possiamo prendere solidale alla Terra perché questa, oltre che essere animata dal moto annuo attorno al Sole, è animata dal moto rotatorio diurno attorno al proprio asse e perciò se $\mathcal{R}'$ fosse ad essa solidale non potrebbe essere inerziale anche per un breve intervallo di tempo.
Supponiamo inoltre, per semplicità, di effettuare l'osservazione di una stella nell'istante t in cui i due sistemi di riferimento sono coincidenti e che tale stella giaccia sul piano xz e quindi anche sul piano $x'z'$ ad esso sovrapposto. L'osservatore terrestre O' vedrà la stella in una certa posizione apparente S' sulla sfera celeste; sia S la posizione vera della stessa stella come si vedrebbe dal sistema $\mathcal{R}$ solidale con l'etere.
Indichiamo con **c** il vettore velocità della luce proveniente dalla stella all'occhio dell'osservatore nel sistema $\mathcal{R}$ e con $\mathbf{c}'$ l'analogo vettore velocità della luce proveniente dalla stessa stella nel riferimento $\mathcal{R}'$, inoltre indichiamo con θ e θ' gli angoli, convessi compresi fra 0 e π, formati rispettivamente dai vettori **c** e $\mathbf{c}'$ con gli assi equiorientati e sovrapposti $x \equiv x'$. I versori degli assi della terna $Oxyz$ siano $\mathbf{i}, \mathbf{j}, \mathbf{k}$ ed analogamente $\mathbf{i}', \mathbf{j}', \mathbf{k}'$ rispettivamente quelli della terna $O'x'y'z'$ e scriviamo le componenti cartesiane di **c** e $\mathbf{c}'$ secondo gli assi di $\mathcal{R}$ ed $\mathcal{R}'$, avremo:

$$\begin{cases} c_x = \mathbf{c}\cdot\mathbf{i} = |\mathbf{c}||\mathbf{i}|\cos\theta = c\,\cos\theta, \quad c_y = 0,, \quad c_z = \mathbf{c}\cdot\mathbf{k} = |\mathbf{c}||\mathbf{k}|\cos(\frac{3}{2}\pi - \theta) = -c\,\sin\theta \\ c'_x = \mathbf{c}'\cdot\mathbf{i}' = |\mathbf{c}'||\mathbf{i}'|\cos\theta' = c\,\cos\theta', \; c'_y = 0, \quad c'_z = \mathbf{c}'\cdot\mathbf{k}' = |\mathbf{c}'||\mathbf{k}'|\cos(\frac{3}{2}\pi - \theta') = -c\,\sin\theta', \end{cases} \tag{4.30}$$

essendo, per il 2° assioma della relatività speciale, $|\mathbf{c}| = |\mathbf{c}'| = c$.
Sostituiamo ora le (4.30) nelle formule (4.16) che ci danno la legge relativistica di composizione delle velocità, avremo:

$$\begin{cases} \cos\theta' = \dfrac{\cos\theta - \beta}{1 - \beta\cos\theta}, \\ \\ \sin\theta' = \dfrac{\sin\theta\sqrt{1-\beta^2}}{1-\beta\cos\theta}, \end{cases} \tag{4.31}$$

la seconda equazione non è stata scritta riducendosi all'identità $c'_y = c_y = 0$. **Le** (4.31) **costituiscono le formule relativistiche esatte di aberrazione.**

Se dividiamo membro a membro la seconda per la prima delle (4.31) otteniamo.

$$\tan\theta' = \frac{\sin\theta\sqrt{1-\beta^2}}{\cos\theta - \beta}. \tag{4.32}$$

Se nella formula (4.32) scambiamo θ' con θ e β con $-\beta$ si ottiene la formula inversa che fornisce θ in funzione dell'angolo θ' osservato dalla Terra:

$$\tan\theta = \frac{\sin\theta'\sqrt{1-\beta^2}}{\cos\theta' + \beta}. \tag{4.33}$$

La (4.32) o la sua inversa (4.33) è una delle diverse forme in cui si può esprimere **la formula relativistica esatta di aberrazione.**[7]
Notiamo che, in virtù del moto della Terra, il raggio di luce proveniente dalla stella nel riferimento $\mathcal{R}'$ trascinato dalla Terra forma con l'asse $x' \equiv x$ un angolo θ' che è sempre maggiore dell'analogo angolo θ ($\theta' > \theta$) formato dal raggio di luce proveniente dalla stessa stella visto però nel riferimento $\mathcal{R}$ supposto solidale con le stelle fisse. La differenza $\Delta\theta = \theta' - \theta$ di tali angoli è il cosìddetto **angolo di aberrazione** ed è sempre positivo ($\Delta\theta > 0$).
Se ora teniamo conto che la velocità orbitale della Terra è $v \approx 30\,\mathrm{km/s}$ mentre la velocità della luce $c \approx 3 \times 10^5\,\mathrm{km/s}$ abbiamo $\beta = 10^{-4}$ che è già abbastanza piccolo ma non trascurabile, mentre è del tutto trascurabile il suo quadrato $\beta^2 = 10^{-8}$, la (4.32) ci da allora

$$\tan\theta' \approx \frac{\sin\theta}{\cos\theta - \beta}, \tag{4.34}$$

che coincide con la formula classica esatta di aberrazione (2.23) data nel Cap.2, paragrafo 2.4.
La (4.34), con facili calcoli, può essere scritta in una forma più conveniente, basta sostituire $\tan\theta'$ con $\dfrac{\sin\theta'}{\cos\theta'}$ e ridurre l'espressione a forma intera, si avrà:

$$\sin\Delta\theta \approx \beta\sin\theta' \tag{4.35}$$

che coincide con la formula classica (2.24) già stabilita; essa fornisce immediatamente l'angolo di aberrazione $\Delta\theta$ non appena si sia misurato θ', ed è proprio quest'ultima forma quella che comunemente viene indicata come *formula approssimata di aberrazione della luce stellare*.
Essendo $\sin\theta' \leq 1$ si ha per l'angolo di aberrazione $\Delta\theta \approx \dfrac{v}{c} = 10^{-4} radianti \approx 20'',50$ in ottimo accordo con i risultati sperimentali.
Entro i limiti della precisione strumentale odierna, quindi, il fenomeno dell'aberrazione della luce stellare è giustificato ugualmente bene in fisica classica come in fisica relativistica. **Soltanto tecniche sperimentali capaci di valutare differenze del secondo ordine in β potrebbero far preferire la spiegazione relativistica a quella classica.**

4.2.2 Effetto Doppler relativistico longitudinale e trasversale

Applicheremo ora le trasformazioni di Lorentz all'effetto Doppler per la luce, di notevole importanza teorica ed applicativa specialmente in Astronomia e Astrofisica. Consideriamo, come al solito, i due sistemi di riferimento inerziali $\mathcal{R}$ ed $\mathcal{R}'$ in configurazione x-standard, che si muovono l'uno rispetto all'altro con velocità di modulo v. La figura di riferimento è la stessa del caso classico, cioè la (Fig.2.5), che qui riproduciamo per comodità del lettore (vedi Fig.4.5).

[7] Un'altra interessante espressione della formula di aberrazione si può ottenere utilizzando in maniera opportuna la seguente identità trigonometrica: $\tan\dfrac{\theta'}{2} = \dfrac{\sin\theta'}{1+\cos\theta'}$. Infatti da questa indentità e dalle formule (4.31) è facile

Figura 4.5: Effetto Doppler relativistico.

Supponiamo che un'onda elettromagnetica piana monocromatica sia emessa da una **sorgente luminosa** $\mathcal{S}$ posta nell'origine O di $\mathcal{R}$.[8] Per semplicità supponiamo che l'onda si propaghi nel piano xy (e quindi anche nel piano $x'y'$) nella direzione e verso individuato dal vettore d'onda $\mathbf{k} = (k_x, k_y, 0) = k\mathbf{n}$, essendo k il suo modulo ed $\mathbf{n}$ il suo versore. Inoltre supponiamo che $\mathbf{k}$ formi un angolo θ con l'asse x e perciò tale che le sue componenti cartesiane, in $\mathcal{R}$, siano $(k_x = k\cos\theta, k_y = k\sin\theta, k_z = 0)$. Supponiamo ancora che l'onda, in $\mathcal{R}$, abbia frequenza ν, pulsazione $\omega = 2\pi\nu$, e si muova con velocità c e lunghezza d'onda $\lambda = \dfrac{2\pi}{k}$ con $\lambda\nu = c$. Per quel che abbiamo visto nel paragrafo 2.5, in ambito classico, l'equazione dell'onda sarà del tipo (2.25), cioè:

$$\Phi = a\,\cos\left(\omega t - \mathbf{k}\cdot\mathbf{x}\right) = a\,\cos\,\omega\Big(t - \frac{\mathbf{n}\cdot\mathbf{x}}{c}\Big), \tag{4.36}$$

nella quale $\mathbf{x} = (x, y, 0)$ è il raggio vettore di un generico punto P del fronte d'onda e dove si è tenuto conto che $\mathbf{k} = k\mathbf{n}$ e $k = \dfrac{\omega}{c}$. Nella (4.36) la quantità

$$\phi = \omega\Big(t - \frac{\mathbf{n}\cdot\mathbf{x}}{c}\Big) = 2\pi\nu\Big(t - \frac{x\cos\theta + y\sin\theta}{c}\Big), \tag{4.37}$$

rappresenta **la fase** dell'onda.
Nel riferimento $\mathcal{R}'$ i fronti d'onda saranno ancora piani, poiché la trasformazione di Lorentz è lineare e trasforma un piano in un piano, perciò la propagazione dell'onda - in virtù dei due assiomi einsteiniani - è rappresentata da un'equazione dello stesso tipo della (4.36), la cui fase sarà espressa, analogamente alla (4.37), dalla

$$\phi' = 2\pi\nu'\Big(t' - \frac{x'\cos\theta' + y'\sin\theta'}{c}\Big), \tag{4.38}$$

con $\lambda'\nu' = c$ e θ' l'angolo formato da $\mathbf{k}'$ con l'asse x'.
Ora noi già sappiamo, dal caso classico, che, tenuto conto del suo significato fisico: **la fase di un'onda piana è una quantità invariante** *che rimane la stessa nel passare dal riferimento $\mathcal{R}$ al riferimento $\mathcal{R}'$, cioè $\phi = \phi'$.*
Se perciò applichiamo alla fase ϕ', espressa in termini delle variabili x', y', z', t', la trasformazione speciale di Lorentz (3.32a), essa verrà espressa in termini delle variabili x, y, z, t e dovrà

ottenere (esercizio) la seguente notevole formula esatta di aberrazione:

$$\tan\frac{\theta'}{2} = \sqrt{\frac{1+\beta}{1-\beta}}\tan\frac{\theta}{2} = \sqrt{\frac{c+v}{c-v}}\tan\frac{\theta}{2}.$$

[8]In virtù dei due assiomi einsteiniani, si perverrebbe agli stessi risultati se anziché considerare la sorgente in O si considerasse situata in O'.

coincidere con la (4.37), cioè

$$2\pi\nu'\Big[\gamma(t-\frac{\beta}{c}x)-\frac{\gamma(x-\beta ct)\cos\theta'+y\sin\theta'}{c}\Big]=2\pi\nu\Big(t-\frac{x\cos\theta+y\sin\theta}{c}\Big). \tag{4.39}$$

Dividendo primo e secondo membro della (4.39) per 2π e raccogliendo opportunamente i termini nelle variabili t, x, y, si ottiene

$$\nu'\Big[\gamma(1+\beta\cos\theta')t-\frac{\gamma}{c}(\beta+\cos\theta')x-\frac{\sin\theta'}{c}y\Big]=\nu\Big(t-\frac{\cos\theta}{c}x-\frac{\sin\theta}{c}y\Big). \tag{4.40}$$

Quest'ultima equazione deve essere soddisfatta per valori arbitrari delle variabili t, x, y, perciò i loro coefficienti, al 1° e 2° membro, devono essere uguali, otteniamo allora le seguenti tre relazioni:

$$\nu=\nu'\frac{1+\beta\cos\theta'}{\sqrt{1-\beta^2}}, \tag{4.41}$$

$$\nu\cos\theta=\nu'\frac{\beta+\cos\theta'}{\sqrt{1-\beta^2}}, \tag{4.42}$$

$$\nu\sin\theta=\nu'\sin\theta'. \tag{4.43}$$

Se dividiamo membro a membro la terza per la seconda di queste relazioni riotteniamo **la formula relativistica dell'aberrazione** (4.33) ricavata per altra via nel paragrafo 4.2.1 e cioè

$$\tan\theta=\frac{\sin\theta'\sqrt{1-\beta^2}}{\beta+\cos\theta'}. \tag{4.44}$$

La (4.41) è invece **la formula relativistica dell'effetto Doppler**, cioè la relazione, che ci eravamo proposti di ricavare, che lega le frequenze ν e ν' dell'onda elettromagnetica in considerazione, misurate rispettivamente nei riferimenti $\mathcal{R}$ ed $\mathcal{R}'$. Essa può essere invertita col solito procedimento di scambiare le lettere accentate con quelle non accentate e β in $-\beta$, ottenendo così:

$$\nu'=\nu\frac{1-\beta\cos\theta}{\sqrt{1-\beta^2}}. \tag{4.45}$$

a) **Effetto Doppler longitudinale**.
Se ora ragioniamo dal punto di vista dell'osservatore $\mathcal{O}'$, solidale con $\mathcal{R}'$, che riceve l'onda partita dalla sorgente luminosa $\mathcal{S}$, in quiete in $\mathcal{R}$, e la cui frequenza ν è la cosiddetta **frequenza propria**[9] (talvolta indicata con ν_0), la formula da usarsi sarà la (4.41) invertita, cioè

$$\nu'=\nu_0\frac{\sqrt{1-\beta^2}}{1+\beta\cos\theta'}. \tag{4.46}$$

Verifichiamo ora, che in prima approssimazione, trascurando termini del 2° ordine in β cioè per $v \ll c$, la formula relativistica si riduce a quella classica. Infatti, trascurando nella (4.46) i termini del 2° ordine in β si ottiene

$$\nu'\approx\frac{\nu_0}{1+\beta\cos\theta'}\simeq\nu_0(1-\beta\cos\theta'), \tag{4.47}$$

che è il risultato classico (2.36). Ciò si capisce meglio considerando due casi particolari più familiari. Infatti, se l'osservatore (ricevitore) si allontana dalla sorgente lungo l'asse x'

[9]Per **frequenza propria** si intende la frequenza di un'onda elettromagnetica la cui sorgente sia in quiete in un dato sistema di riferimento.

(ovvero sorgente ed osservatore sono in reciproco allontanamento lungo l'asse x') si avrà $\theta' = 0°$ e la (4.47) si riduce alla

$$\nu' \simeq \nu_0(1-\beta) = \nu_0\Big(1-\frac{v}{c}\Big), \tag{4.48}$$

che dimostra che **la frequenza osservata ν' è più piccola della frequenza propria ν_0** (spostamento delle righe spettrali verso il rosso: "red shift").
Se invece l'osservatore si avvicina alla sorgente sempre lungo l'asse x' (ovvero sorgente ed osservatore sono in reciproco avvicinamento) si ha $\theta' = 180°$ e la (4.47) si riduce alla

$$\nu' \approx \nu_0(1+\beta) = \nu_0\Big(1+\frac{v}{c}\Big), \tag{4.49}$$

che dimostra che **la frequenza osservata ν' è più grande della frequenza propria ν_0** (spostamento delle righe spettrali verso il blu: "blu shift"). Infine se $\theta' = 90°$ oppure $\theta' = 270°$, nel qual caso la direzione di propagazione della luce è perpendicolare a quella del moto relativo tra sorgente e ricevitore, non c'è classicamente alcun effetto Doppler, cioè $\nu' = \nu_0$. Osserviamo che questi effetti classici sono del 1° ordine in $\beta = \frac{v}{c}$.
Tornando alla formula relativistica esatta (4.46), se $\theta' = 0°$ **(sorgente ed osservatore in allontanamento)** avremo

$$\nu' = \nu_0\sqrt{\frac{1-\beta}{1+\beta}} = \nu_0\sqrt{\frac{c-v}{c+v}} \qquad (\nu' < \nu_0), \tag{4.50}$$

se invece è $\theta' = 180°$ **(sorgente ed osservatore in avvicinamento)** avremo

$$\nu' = \nu_0\sqrt{\frac{1+\beta}{1-\beta}} = \nu_0\sqrt{\frac{c+v}{c-v}} \qquad (\nu' > \nu_0). \tag{4.51}$$

In questi due ultimi casi si parla di **effetto Doppler longitudinale**.

2) **Effetto Doppler trasversale**.
La conseguenza più notevole della formula relativistica esatta (4.46) è l'esistenza di un **effetto Doppler trasversale** che come abbiamo visto poco fa non è previsto dalla fisica classica. Tale effetto corrisponde al caso in cui è $\theta' = 90°$ oppure $\theta' = 270°$, nel qual caso la direzione di propagazione della luce è perpendicolare a quella del moto relativo dei due sistemi di riferimento $\mathcal{R}$ e $\mathcal{R}'$, e la (4.46) ci dà

$$\nu' = \nu_0\sqrt{1-\beta^2} \qquad (\nu' < \nu_0), \tag{4.52}$$

che è un effetto del 2° ordine in β ed è perciò molto più difficile da osservare dell'effetto Doppler longitudinale, ma genuinamente relativistico nel senso che se si vogliono verificare le predizioni della teoria della relatività si deve essere in grado di rivelare effetti del 2° ordine.

Le previsioni teoriche riguardanti l'effetto Doppler longitudinale furono confermate sperimentalmente per la prima volta da Ives e Stilwell[10] nel 1938 e ripetute nel 1941, che seguendo un suggerimento fatto per primo da Einstein nel 1907, usarono come sorgente di radiazione un fascio di atomi di idrogeno eccitati di velocità e direzione ben definita. L'esperimento fu ripetuto nel 1961 da Mandelberg e Witten[11], confermando di nuovo l'effetto relativistico.

[10] H. E. Ives e G. R. Stillwell, *J. Opt. Soc. Am*, **28**, p.215 (1938); e **31**, p.369 (1941).
[11] Mandelberg e Witten *J. Opt. Soc. Am*, **52**, p.529 (1962).

Per quanto riguarda l'effetto Doppler trasversale, esso fu confermato sperimentalmente nel 1938 e 1941 sempre da Ives e Stilwell e da Otting[12] nel 1939. Più recentemente Kundig[13] ha ottenuto eccellenti risultati confermando la formula relativistica entro l'errore sperimentale dell'1.1%. Infine, nel 1985 Kaivola, Poulsen, Riis e Lee[14] hanno verificato la validità della formula relativistica dell'effetto Doppler trasversale con un'accuratezza di 4 parti su 10^5.
È interessante osservare che l'effetto Doppler trasversale è sia una conseguenza diretta della dilatazione del tempo sia una sua conferma. L'osservatore (ricevitore) $\mathcal{O}$ solidale ad $\mathcal{R}'$ è in moto rispetto alla sorgente $\mathcal{S}$ solidale ad $\mathcal{R}$ e può perciò essere considerato come un orologio in moto che rallenta rispetto all'orologio della sorgente, sicché il periodo T' dell'onda ricevuta ed il **periodo proprio** T_0 dell'onda emessa sono legati dalla relazione

$$T' = \frac{T_0}{\sqrt{1-\beta^2}}, \tag{4.53}$$

ed essendo $T' = \dfrac{1}{\nu'}$ e $T_0 = \dfrac{1}{\nu_0}$, avremo

$$\nu' = \nu_0\sqrt{1-\beta^2}, \tag{4.54}$$

che coincide appunto con la formula (4.52) dell'effetto Doppler trasversale. Osserviamo infine che la teoria della relatività, sia per la spiegazione dell'effetto Doppler sia per quella riguardante il fenomeno dell'aberrazione della luce stellare, fornisce un'intrinseca semplificazione rispetto alla interpretazione classica di questi fenomeni. Infatti essa, in accordo con le osservazioni sperimentali, non fa alcuna distinzione fra i due casi: sorgente a riposo ed osservatore in moto, oppure sorgente in moto ed osservatore a riposo, trattati invece in modi diversi nella fisica classica. Inoltre è notevole il fatto che con una sola dimostrazione si riesce a spiegare contemporaneamente tre effetti: l'effetto Doppler longitudinale, l'aberrazione e l'effetto Doppler trasversale.
Tenuto conto delle molteplici conferme sperimentali sull'effetto Doppler, di cui abbiamo riferito più sopra, specialmente dell'effetto Doppler trasversale, possiamo concludere che a differenza del fenomeno dell'aberrazione, quest'ultimo offre un argomento preferenziale tra cinematica classica e cinematica relativistica, ed é con la seconda che si trova d'accordo.

4.2.3 L'esperimento di Michelson-Morley: spiegazione relativistica

Come già visto, il risultato dell'esperienza di Michelson-Morley risultava in netto contrasto con le previsioni della fisica classica. Il risultato negativo dell'esperimento è invece spiegato in modo immediato in ambito relativistico, in primo luogo perché la validità del principio einsteiniano di relatività implica che nessun esperimento svolto in un ambiente galileiano (la Terra in prima approssimazione) può mettere in evidenza l'eventuale moto dell'ambiente stesso rispetto ad un altro riferimento galileiano (Sole o stelle fisse). Inoltre, il risultato negativo di tale esperimento si spiega anche tenendo conto del 2° principio di Einstein cioè l'invarianza della velocità della luce in tutti i riferimenti galileiani, ciò che implica in particolare un comportamento della luce perfettamente isotropo in ciascuno di essi.
D'altra parte, possiamo constatare che se indichiamo, come già abbiamo fatto nel caso classico, con $\mathcal{R}$ il sistema galileiano solidale all'interferometro e con $\mathcal{R}'$ il sistema solidale con la presunta corrente d'etere, e con $\mathbf{u}$ e $\mathbf{u}'$ le velocità della luce rispetto a tali sistemi inerziali, avremo (Fig.4.6):

[12] G. Otting *Phys. Z.*, **40**, p.681 (1939).
[13] W. Kundig *Phys. Rev.*, **129**, p.2371 (1963).
[14] M. Kaivola, O. Poulsen, E. Riis e S. A. Lee, *Phys. Rev. Lett.*, **54**, p.255 (1985).

Figura 4.6: Esperimento di Michelson-Morley: spiegazione relativistica.

a) per il percorso AB, di andata e ritorno, lungo l'asse x, le componenti cartesiane della velocità della luce sono:

$$u_x = \pm c, \qquad u_y = 0, \qquad u_z = 0, \tag{4.55}$$

che sostituite nelle formule (4.16) di composizione delle velocità, ci danno

$$u'_x = \pm c, \qquad u'_y = 0, \qquad u'_z = 0, \tag{4.56}$$

dalle quali si deduce immediatamente che i moduli delle due velocità in $\mathcal{R}$ e in $\mathcal{R}'$ sono identici

$$|\mathbf{u}| = \sqrt{u_x^2 + u_x^2 + u_x^2} = c, \qquad |\mathbf{u}'| = \sqrt{u_x'^2 + u_x'^2 + u_x'^2} = c,$$

come ci si aspetta in base al 2° principio.

b) per il percorso AC, di andata e ritorno, lungo l'asse y, le componenti cartesiane della velocità della luce sono invece:

$$u_x = 0, \qquad u_y = \pm c, \qquad u_z = 0, \tag{4.57}$$

che sostituite nelle (4.16) ci danno

$$u'_x = -v, \qquad u'_y = \pm\sqrt{c^2 - v^2}, \qquad u'_z = 0, \tag{4.58}$$

dalle quali si deduce che i moduli delle due velocità in $\mathcal{R}$ e in $\mathcal{R}'$ sono ancora entrambi identici a c. Dunque la luce, nel riferimento $\mathcal{R}$, impiega lo stesso tempo di andata e ritorno sia lungo l'asse x, sia lungo l'asse y e perciò, nell'interferometro di Michelson usato per l'esperimento, non si formano frange di interferenza. Si può verificare che la stessa cosa succede anche nel riferimento $\mathcal{R}'$: la dimostrazione in questo caso è però leggermente più complicata.

Possiamo concludere affermando che la corretta interpretazione dell'esperimento di Michelson-Morley, ancor più dell'effetto Doppler longitudinale e trasversale, è completamente a favore della Teoria della Relatività di Einstein.

4.2.4 L'allungamento della vita media del mesone μ causata dalla dilatazione del tempo

Un altro importante esperimento a conferma della cinematica relativistica è la constatazione che certe particelle presenti nella radiazione cosmica come i mesoni μ (detti anche muoni) viaggiando a velocità prossima alla velocità della luce hanno una vita media notevolmente più lunga di quando si trovano a riposo.[15]
È noto che queste particelle hanno la stessa carica dell'elettrone (negativo o positivo) e massa circa 200 volte maggiore. Tali particelle sono radioattive, nel senso che dal momento in cui sono state originate, ciascuna di esse si disintegra spontaneamente in un elettrone dello stesso segno ed in due neutrini non direttamente osservabili, ma la cui esistenza viene indirettamente congetturata per la necessità di soddisfare al principio di conservazione dell'energia e della quantità di moto. Utilizzando in maniera opportuna questa disintegrazione è possibile, in laboratorio, mediante una tecnica elettronica raffinata, valutare *la vita media a riposo* τ_0 dei mesoni incidenti su di una lastra metallica e da questa frenati. Le misure forniscono, con l'approssimazione di pochi centesimi, il valore

$$\tau_0 = 2,15 \times 10^{-6}\,\mathrm{s}.$$

Durante una vita così breve il mesone, anche se fosse animato da una velocità pari alla velocità della luce $c = 3 \times 10^8\,\mathrm{m/s}$, non potrebbe percorrere che un cammino piuttosto breve

$$l_0 = \tau_0 c = 645\,\mathrm{m}.$$

Questo risultato è però in contraddizione con i risultati di altre osservazioni, concernenti la sopravvivenza di mesoni a diverse quote negli alti strati dell'atmosfera, secondo le quali è certo che molti mesoni osservati ad una certa quota, hanno avuto origine a quote diversi km più alte. La contraddizione è facilmente eliminabile se si interpreta il valore sperimentale τ_0 come *vita media di riposo* del mesone e tenendo conto che se esso viaggia alla velocità v allora la sua vita media, secondo la relatività ristretta, sarà

$$\tau = \gamma(v)\tau_0 = \frac{\tau_0}{\sqrt{1 - \dfrac{v^2}{c^2}}}.$$

Se supponiamo che il mesone viaggi ad una velocità $v = 0,995\,c$, il fattore di Lorentz vale $\gamma = \frac{1}{\sqrt{1-(0,995)^2}} \approx 10$, sicché anche il percorso medio del mesone risulta circa decuplicato, cioè $l = v\tau = 0,995 c\gamma\tau_0 = 0,995\gamma l_0 = 0,995 \times 10 \times 645\,\mathrm{m} = 6.418\,\mathrm{m}$, in accordo con le osservazioni sperimentali.
A conclusione di questa rapida rassegna di fatti sperimentali, possiamo dire che, già dal punto di vista cinematico, i risultati ottenuti parlano in favore della validità della teoria della relatività ristretta rispetto alla fisica classica. Più numerose ed altrettanto soddisfacenti conferme si avranno nell'ambito della dinamica relativistica.

[15] Si vedano gli esperimenti di D. H. Frisch e J. H. Smith, in *Am. J. Phys.*, 31, 1963, pag. 342. Il primo esperimento sulla dilatazione della vita media delle particelle costituenti i raggi cosmici è quello di B. Rossi e D. B. Hall, in *Phys. Rev.*, 59, 1941, pag. 223.

Capitolo 5

SPAZIOTEMPO DI MINKOWSKI

Mi dispiace di avervi somministrato una dose così grande
di geometria quadri-dimensionale. Non mi scuso perché
non sono responsabile del fatto che la Natura
è fondamentalmente quadri-dimensionale.
(Alfred North Whitehead, *The Concept of Nature*, 1920)

In questo capitolo ci occuperemo del concetto di **spaziotempo 4-dimensionale** introdotto nel 1908 dal matematico lituano Hermann Minkowski[1] appena tre anni dopo la pubblicazione del lavoro fondamentale di Einstein (1905). Il nuovo concetto di spaziotempo soppianterà i vecchi concetti newtoniani di spazio e tempo considerati separatamente; infatti, la parte iniziale del discorso di Minkowski pronunciato in occasione dell'ottantesima Assemblea degli Scienziati della Natura e dei Medici Tedeschi (21 settembre, 1908) è divenuta famosa:
"I concetti di spazio e di tempo che desidero esporvi traggono origine dal terreno della fisica sperimentale, e in ciò risiede la loro forza. Sono radicali. D'ora in avanti lo spazio singolarmente inteso, ed il tempo singolarmente inteso, sono destinati a svanire in nient'altro che ombre, e solo una connessione dei due potrà preservare una realtà indipendente".[2]
Per comprendere il nuovo punto di vista sullo spazio e sul tempo ed introdurre il concetto di **spaziotempo 4-dimensionale di Minkowski**, utilizzeremo uno strumento matematico particolarmente adatto a tali questioni: il calcolo tensoriale sui cosiddetti *4-vettori* e *4-tensori*. I fondamenti di questo Calcolo in tutta la sua generalità - assolutamente necessario per formulare la Relatività Generale - è presentato nel Cap.8 del libro. Tuttavia, ai fini dello studio della Relatività Ristretta bastano poche nozioni elementari che via via presenteremo nei prossimi paragrafi.

5.1 Invariante fondamentale della Relatività Ristretta

Abbiamo visto in precedenza che alcune grandezze (ed alcuni concetti) aventi un senso assoluto in fisica classica, cioè indipendente dal sistema di riferimento in cui sono definite, nella teoria einsteiniana assumono invece carattere relativo, dipendente perciò dal sistema di riferimento galileiano nel quale vengono considerate. Tali sono per esempio, il concetto di simultaneità, l'intervallo temporale fra due eventi, la lunghezza di una sbarra rigida, ecc. Come vedremo in

[1]Hermann Minkowski (1864-1909), matematico di origine ebraica, nato in Lituania e naturalizzato tedesco, fu professore di Einstein al Politecnico di Zurigo.
[2]Hermann Minkowski, *Raum und Zeit*, trad. inglese in H. A. Lorentz, A. Einstein, H. Minkowski e H. Weyl, *The Principle of Relativity,*, a cura di A. Sommerfeld, Dover, New York, 1952, pp. 75-91.

seguito, anche altre grandezze fisiche, quali la frequenza di una radiazione luminosa, la massa di una particella, ecc., sono da considerarsi, in relatività, come grandezze relative.
Tuttavia, anche nella teoria della relatività ci sono **grandezze assolute**, invarianti cioè al variare del riferimento galileiano: prima fra tutte la velocità c della luce nel vuoto[3] come è stabilito nel 2° postulato einsteiniano, che assieme al 1° postulato, cioè al principio di relatività, porta alla nozione di **invariante fondamentale della relatività ristretta,** che abbiamo già introdotto nel paragrafo 3.2.7, come conseguenza diretta dei due postulati di Einstein, e che ora riprenderemo deducendolo dalle trasformazioni di Lorentz.
Siano $\mathcal{R}$ ed $\mathcal{R}'$ due riferimenti galileiani in mutua relazione *x-standard* e quindi legati dalle trasformazioni speciali di Lorentz (3.32a). Si considerino due eventi $\mathcal{A}$, $\mathcal{B}$ infinitamente vicini, le cui coordinate spazio-temporali in $\mathcal{R}$ differiscano per delle quantità infinitesime dx, dy, dz, dt:

$$\mathcal{A} \equiv (x, y, z, t), \qquad \mathcal{B} \equiv (x + dx, y + dy, z + dz, t + dt). \tag{5.1}$$

Osserviamo che i due dati eventi $\mathcal{A}$, $\mathcal{B}$ sono separati dalla *distanza spaziale* (euclidea) $d\ell \equiv |\mathbf{dx}| = \sqrt{dx^2 + dy^2 + dz^2}$, e dalla *distanza temporale* dt.[4] Gli stessi eventi avranno in $\mathcal{R}'$ altre coordinate spazio- temporali le cui differenze saranno dx', dy', dz', dt'. Le relazioni fra le due quaterne dx, dy, dz, dt e dx', dy', dz', dt' sono date dalle trasformazioni di Lorentz (3.36a), (3.36b) scritte in forma differenziale, sicché un calcolo diretto ed immediato permette di constatare, che la forma quadratica differenziale

$$\mathbf{ds}^2 \equiv c^2dt^2 - dx^2 - dy^2 - dz^2, \tag{5.2}$$

applicando le (3.36b), si trasforma nell'identica forma quadratica nelle nuove variabili dx', dy', dz', dt':

$$\mathbf{ds}^2 \equiv c^2dt'^2 - dx'^2 - dy'^2 - dz'^2, \tag{5.3}$$

e viceversa, se si applicano le trasformazioni di Lorentz (3.36a) alla (5.3) si ottiene la (5.2). Infatti, con una sostituzione diretta nella (5.2), avremo

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = c^2\gamma^2(dt' + \frac{\beta}{c}dx')^2 - \gamma^2(dx' + \beta c dt')^2 - dy'^2 - dz'^2,$$

e sviluppando i quadrati al secondo membro si ottiene

$$c^2dt^2 - dx^2 - dy^2 - dz^2 =$$

$$= c^2\gamma^2dt'^2 + 2\beta c\gamma^2dt'dx' + \beta^2\gamma^2dx'^2 - \gamma^2dx'^2 - 2\beta c\gamma^2dt'dx' - \beta^2c^2\gamma^2dt'^2 - dy'^2 - dz'^2,$$

con facili semplificazioni ricaviamo l'importantissimo risultato

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = c^2dt'^2 - dx'^2 - dy'^2 - dz'^2, \tag{5.4}$$

che possiamo scrivere più concisamente così

$$c^2dt^2 - \mathbf{dx}^2 \equiv c^2dt'^2 - \mathbf{dx}'^2 = \text{Invariante di Lorentz.} \tag{5.5}$$

[3]Oltre alla velocità della luce nel vuoto, che sappiamo essere una *velocità limite*, sono invarianti tutte le *interazioni* fra particelle o fra particelle e campi aventi anch'esse la stessa velocità limite. Sono inoltre invarianti le cosiddette *grandezze proprie*: per es. il *tempo proprio* e la *lunghezza propria* incontrate nel capitolo precedente ed altre che incontreremo in seguito.

[4]Siccome, per le considerazioni che faremo, è molto scomodo usare distanze spaziali espresse con radici quadrate, è più conveniente utilizzare il quadrato della distanza spaziale fra i due eventi, cioè $d\ell^2 \equiv \mathbf{dx}^2 = dx^2 + dy^2 + dz^2$, ed in maniera analoga utilizzare la quantità c^2dt^2 - avente le stesse dimensioni di $\mathbf{dx}^2$ cioè del quadrato di una lunghezza - per valutare l'intervallo temporale fra gli stessi eventi. La differenza fra il quadrato della distanza temporale ed il quadrato della distanza spaziale fra gli eventi, cioè la forma quadratica differenziale

$$\mathbf{ds}^2 \equiv c^2dt^2 - \mathbf{dx}^2 \equiv c^2dt^2 - (dx^2 + dy^2 + dz^2) = c^2dt^2 - dx^2 - dy^2 - dz^2$$

misurerà, in qualche modo, la separazione spaziotemporale fra gli eventi stessi.

La forma quadratica (5.2) e l'uguaglianza (5.5) possono essere scritte in una forma più concisa, e più conveniente per le considerazioni che faremo in seguito, con un semplice cambio di notazione sulle coordinate, ponendo $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, considerando perciò come prima coordinata la variabile temporale e successivamente le variabili spaziali.[5] Un evento puntiforme sarà allora indicato in $\mathcal{R}$ con $(x^i) \equiv (x^0, x^1, x^2, x^3)$, $(i = 0, 1, 2, 3)$ ed in $\mathcal{R}'$ in modo analogo $(x'^i) \equiv (x'^0, x'^1, x'^2, x'^3)$, o anche, con un leggero abuso di notazione, $(x^{i'}) \equiv (x'^i)$. Infatti, se introduciamo il simbolo η_{ik}, $(i, k = 0, 1, 2, 3)$ per i coefficienti della forma quadratica (5.2) immaginandola scritta in forma completa nelle variabili dx^0, dx^1, dx^2, dx^3, includendovi cioè anche i termini misti nelle stesse variabili (che ovviamente sono tutti nulli), avremo la seguente matrice diagonale simmetrica

$$(\eta_{ik}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad (i, k = 0, 1, 2, 3) \tag{5.6}$$

associata a tale forma quadratica, e l'uguaglianza (5.4) o (5.5) si potrà perciò scrivere, utilizzando la convenzione di sommazione di Einstein, in forma tensoriale così:

$$\eta_{ik} dx^i dx^k = \eta_{i'k'}\, dx^{i'} dx^{k'} = \text{Invariante di Lorentz.} \tag{5.7}$$

La quaterna di differenziali $(cdt, dx, dy, dz) \equiv (dx^0, dx^1, dx^2, dx^3) \equiv (dx^i)$, come vedremo, con più dettagli, nel paragrafo 5.1.1, si dice che definisce nello spaziotempo degli eventi (*spazio di Minkowski*)(vedi paragrafo 5.2) *un 4-vettore infinitesimo (o 4-tensore del 1° ordine)*, che unisce i due eventi puntiformi $\mathcal{A}$, $\mathcal{B}$, e che diremo *spostamento infinitesimo* da $\mathcal{A}$ a $\mathcal{B}$ e indicheremo con la notazione

$$\mathbf{ds} \equiv (dx^i), \tag{5.8}$$

la cui *norma o quadrato* (prodotto scalare di **ds** per sé stesso):[6]

$$\mathbf{ds} \cdot \mathbf{ds} \equiv \mathbf{ds}^2 \equiv ||\mathbf{ds}|| = \eta_{ik} dx^i dx^k, \tag{5.9}$$

si identifica proprio con la forma quadratica (5.2). Perciò la (5.5) o (5.7) potrà scriversi in maniera più sintetica così

$$\mathbf{ds}^2 = \mathbf{ds}'^2. \tag{5.10}$$

L'invariante fondamentale della relatività ristretta (5.9) verrà chiamato, come abbiamo già detto, anche **divario o intervallo spaziotemporale fra i due eventi $\mathcal{A}$, $\mathcal{B}$.** *Esso costituisce il vero cardine della teoria einsteiniana ed il punto di partenza di tutti i suoi ulteriori sviluppi.*[7]

[5]Con questa nuova notazione la coordinata temporale $x^0 = ct$ ha il vantaggio di avere le dimensioni di una lunghezza come le tre coordinate spaziali x^1, x^2, x^3.

[6]Adotteremo la convenzione di indicare i 4-vettori dello spaziotempo di Minkowski, dei quali si parlerà più avanti, generalmente con lettere maiuscole in grassetto, per es. **U**, **V**, unica eccezione, nell'ambito della RR, sarà il 4-vettore spostamento infinitesimo che indicheremo, com'è d'uso, con lettere minuscole così: **ds**. Per evitare ogni confusione, i 3-vettori dello spazio euclideo ordinario saranno indicati con lettere minuscole in grassetto, es. **u**, **v**, **x**, ecc. Nei manoscritti converremo di indicare i 3-vettori con lettere sottolineate, es. $\underline{u}$, $\underline{v}$, $\underline{x}$, ecc. mentre i 4-vettori possono essere indicati, con lettere per es., sormontate da una freccia es. $\vec{U}$, $\vec{V}$, ecc.

[7]Osserviamo che la quasi totalità degli autori indica la forma quadratica fondamentale con ds^2 e la chiama *quadrato dell'intervallo spazio-temporale ds*. Tale notazione benché universalmente adottata presenta l'inconveniente di apparire come il quadrato di una quantità scalare ds, il che non è vero: si tratta invece, come ha osservato correttamente Rindler e come vedremo con precisione nel seguito, del quadrato (o norma) del 4-vettore spostamento elementare $\mathbf{ds} \equiv (dx^i)$ che unisce due eventi infinitamente vicini dello spazio di Minkowski e può assumere valori positivi, nulli o negativi.

Osserviamo infine, che se indichiamo le coordinate spazio temporali con x^0, x^1, x^2, x^3, considerando come prima coordinata la variabile temporale x^0, le trasformazioni speciali di Lorentz (3.32a) e (3.32b) possono scriversi in maniera più compatta nel seguente modo:

$$\begin{cases} x'^0 = \gamma(v)(x^0 - \beta x^1) \\ x'^1 = \gamma(v)(x^1 - \beta x^0) \\ x'^2 = x^2 \\ x'^3 = x^3 \end{cases} \tag{5.11}$$

$$\begin{cases} x^0 = \gamma(v)(x'^0 + \beta x'^1) \\ x^1 = \gamma(v)(x'^1 + \beta x'^0) \\ x^2 = x'^2 \\ x^3 = x'^3. \end{cases} \tag{5.12}$$

Osservazione:
Nelle considerazioni precedenti abbiamo considerato due eventi $\mathcal{A}$, $\mathcal{B}$, infinitamente vicini perché ciò torna comodo in diverse circostanze e per diverse considerazioni che faremo nel seguito; è ovvio però che si possono considerare eventi le cui coordinate differiscono per quantità finite

$$\mathcal{P} \equiv (x, y, z, t), \qquad \mathcal{Q} \equiv (x + \Delta x, y + \Delta y, z + \Delta z, t + \Delta t), \tag{5.13}$$

e per i quali si possono ripetere le stesse considerazioni svolte per $\mathcal{A}$, $\mathcal{B}$. Per essi **l'invariante fondamentale o divario spaziotemporale** sarà finito e quindi si scriverà

$$\mathbf{\Delta s^2} = \eta_{ik}\Delta x^i \Delta x^k = c^2\Delta t^2 - (\Delta x^2 + \Delta y^2 + \Delta z^2), \tag{5.14}$$

o, se si vuole:

$$\mathbf{\Delta s^2} = c^2(t_Q - t_P)^2 - [(x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2].$$

5.1.1 Cenno alle trasformazioni generali di Lorentz in notazione tensoriale

Ciò che si è detto nel paragrafo 5.1 riguardo all'invarianza lorentziana può essere generalizzato definendo dapprima **le trasformazioni generali di Lorentz omogenee (TLO) come le trasformazioni lineari invertibili**

$$x^{i'} = L^{i'}_j x^j, \quad (i', j = 0, 1, 2, 3) \tag{5.15}$$

che lasciano invariato, in valore e forma, il divario spaziotemporale fra due eventi arbitrari

$$\mathbf{ds}^2 = \eta_{ik} dx^i dx^k \quad (i, k = 0, 1, 2, 3).$$

La trasformazione (5.15) non è altro che la trasformazione speciale di Lorentz (5.11) scritta per mezzo delle nuove variabili x^i. I coefficienti $L^{i'}_j$ sono 16 quantità reali costanti cioè non dipendenti dalle coordinate. Secondo questa definizione, sono da considerarsi trasformazioni di Lorentz non solo *le trasformazioni speciali*, dette anche *trasformazioni di velocità* o con termine anglosassone *boost*, che considerano riferimenti inerziali in moto traslatorio rettilineo uniforme l'uno rispetto all'altro, ma anche combinazioni di trasformazioni speciali con *rotazioni*.
La più generale trasformazione di Lorentz non omogenea detta anche trasformazione di Poincaré (TP) è una trasformazione lineare invertibile del tipo

$$x^{i'} = L^{i'}_j x^j + a^{i'} \quad (a^{i'} \text{ costanti}), \tag{5.16}$$

che lascia invariato il ds^2.
Per una trasformazione di Lorentz, omogenea o non omogenea, le componenti dx^i, $dx^{i'}$, del 4-vettore spostamento infinitesimo **ds**, in due diversi sistemi di riferimento $\mathcal{R}$, $\mathcal{R}'$, sono legate, dalle relazioni

$$dx^{i'} = L^{i'}_j dx^j, \tag{5.17}$$

che si ottengono differenziando le (5.15) e (5.16).
E poiché si ha[8]

$$dx^{i'} = \frac{\partial x^{i'}}{\partial x^j} dx^j, \tag{5.18}$$

se questa si confronta con la (5.17) si ottiene

$$L^{i'}_j = \frac{\partial x^{i'}}{\partial x^j}. \tag{5.19}$$

Esaminiamo ora a quali condizioni devono soddisfare i coefficienti $L^{i'}_j$ affinché le (5.16) rappresentino una trasformazione generale di Lorentz. A tal fine, teniamo conto del fatto che tale trasformazione deve lasciare invariato il divario spaziotemporale **ds**^{2}, cioè

$$\eta_{i'k'} dx^{i'} dx^{k'} = \eta_{ik} dx^i dx^k. \tag{5.20}$$

Sostituendo due volte la (5.17) nella (5.20), e cambiando in maniera opportuna gli indici, avremo

$$\eta_{i'k'} L^{i'}_j L^{k'}_h dx^j dx^h = \eta_{jh} dx^j dx^h, \tag{5.21}$$

valevole per ogni scelta delle dx^j. Ne conseguono le relazioni

$$\eta_{jh} = \eta_{i'k'} L^{i'}_j L^{k'}_h. \tag{5.22}$$

Queste sono le condizioni generali a cui devono soddisfare i coefficienti di una trasformazione di coordinate del tipo (5.15) o (5.16) per essere una trasformazione di Lorentz. Si tratta, per la simmetria rispetto agli indici j, h, di 10 relazioni algebriche tra i 16 coefficienti $L^{i'}_j$, perciò questi ultimi dipendono di fatto soltanto da $16 - 10 = 6$ parametri indipendenti.
Un'importante proprietà delle trasformazioni di Lorentz omogenee (TL) è che esse formano un *gruppo* detto **gruppo di Lorentz**. L'insieme delle TL soddisfa infatti le seguenti condizioni:

1. *(Chiusura)*: Componendo due trasformazioni di Lorentz, cioè effettuandole di seguito, si ottiene ancora una trasformazione di Lorentz;

2. *(Associatività)*: La composizione di tre o più trasformazioni di Lorentz è associativa;

3. *(Identità)*: Esiste la trasformazione identità;

4. *(Inversa)*: Per ogni trasformazione esiste la trasformazione inversa.

Anche le trasformazioni di Lorentz non omogenee, o trasformazioni di Poincaré (TP), formano un gruppo, **il gruppo di Poincaré**, che contiene il gruppo di Lorentz come sottogruppo.
Esce fuori dai compiti che ci siamo proposti nel redigere questo libro fare uno studio approfondito dei gruppi di Lorentz e Poincaré, per i quali esistono in letteratura trattati appositi.[9]

[8]Osserviamo che la (5.18) è *la più semplice legge tensoriale di controvarianza*. Ciò significa, che in un generico cambiamento di coordinate $x^i \to x^{i'}$, la n-pla (dx^j), dei differenziali delle variabili x^j, costituisce un **tensore controvariante del 1° ordine**: anzi (dx^j) viene considerato come *il prototipo dei tensori controvarianti del 1° ordine* chiamati anche **vettori controvarianti**. Tutto ciò sarà visto con maggiori dettagli nel Cap.8 dove svolgeremo in tutta generalità il Calcolo Tensoriale.

[9]Per un approfondimento di questi argomenti si consiglia lo studente di consultare il più volte citato ottimo trattato di Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), ristampa 2009, pp. 170-179 e 193-205.

5.1.2 Digressione sugli spazi affini, euclidei e pseudo-euclidei

Il concetto di **spazio vettoriale reale 3-dimensionale** nasce in modo naturale quando si descrivono le proprietà principali dello **spazio fisico ordinario**, nel quale è stato scelto preventivamente un punto O da cui vengono spiccati tutti i vettori dello spazio e considerato come origine di un sistema arbitrario di coordinate cartesiane. Poiché tutti i punti dello spazio fisico si suppongono equivalenti, conviene avere, insieme al concetto di **spazio vettoriale** - che supponiamo sia noto al lettore - anche un modello matematico corrispondente al concetto di spazio nel quale non si fissi in anticipo alcun punto come origine. Esempi di questi modelli sono lo **spazio affine** e lo **spazio euclideo puntuale**. Per comodità del lettore, richiamiamo qui di seguito alcune nozioni intuitive di base di questi spazi ed anche il concetto di **spazio pseudo-euclideo**.
Cominciamo col ricordare alcune proprietà dell'ordinario spazio euclideo 3- dimensionale E^3. Come è ben noto, esso è dotato anzitutto di **una struttura affine**, caratterizzata dagli **assiomi: di appartenenza, dell'ordine, di continuità e di parallelismo.**
Accanto alla struttura affine si introduce **una struttura metrica**, fondata sulle nozioni quantitative di distanza fra due punti, area, volume, ampiezza angolare, ortogonalità ecc. La struttura affine dà senso al concetto di **vettore libero** in E^3, alle operazioni di somma e differenza di vettori, moltiplicazione di un vettore per un numero reale, alla nozione di componenti affini di un vettore rispetto ad un prefissato **riferimento affine** $\mathcal{RA} \equiv (O, \mathbf{e_i})$, ecc. La struttura metrica di E^3 dà senso all'operazione di prodotto scalare fra due vettori, alla nozione di norma e di modulo di un vettore, di angolo di due vettori, di prodotto vettoriale, ecc.
La struttura affine di E^3 implica l'esistenza di infiniti sistemi di coordinate privilegiate (x^i), $(i = 1, 2, 3)$ (coordinate cartesiane affini), dall'uno all'altro di questi sistemi si passa mediante una trasformazione lineare invertibile

$$x^{i'} = A^{i'}_k x^k + A^{i'}, \quad (i', k = 1, 2, 3)$$

essendo $A^{i'}_k \equiv \dfrac{\partial x^{i'}}{\partial x^k}$ una matrice regolare, costituita da elementi costanti come sono costanti anche le $A^{i'}$. L'insieme di tali trasformazioni costituisce un gruppo detto **il gruppo affine**.
La struttura metrica di E^3 restringe la classe dei sistemi di coordinate privilegiati alle sole coordinate cartesiane ortogonali; in tale classe si passa da un sistema ad un altro mediante trasformazioni di tipo particolare **(gruppo delle trasformazioni ortogonali).**
Queste ultime, com'è noto, sono tutte e sole quelle trasformazioni lineari che lasciano inalterata la distanza di un punto arbitrario $P(x^1, x^2, x^3)$ dello spazio dall'origine O, cioè che lasciano invariata, in valore e forma, la forma quadratica definita positiva

$$\delta_{ik} x^i x^k \equiv (x^1)^2 + (x^2)^2 + (x^3)^2,$$

cioè tali che passando dal sistema di coordinate (x^i), $(i = 1, 2, 3)$ al sistema $(x^{i'})$, $(i' = 1, 2, 3)$ si abbia

$$\delta_{i'k'} x^{i'} x^{k'} = \delta_{ik} x^i x^k = \text{Invariante},$$

dove con δ_{ik} si denota il ben noto delta di Kronecker ($\delta_{ik} = 1, se\ i = k$, $\delta_{ik} = 0, se\ i \neq k$). Si tratta di una forma quadratica definita positiva, avente il significato di quadrato della distanza del punto (x^i), $(i = 1, 2, 3)$ dall'origine $O \equiv (0, 0, 0)$ del sistema di coordinate. A tale forma quadratica invariante si dà il nome di **forma quadratica fondamentale** di E^3.
Le proprietà ora richiamate dello spazio ordinario, hanno suggerito l'introduzione di spazi astratti che sono una spontanea generalizzazione dello spazio euclideo 3-dimensionale. La generalizzazione consiste in due punti:

1) la dimensione dello spazio viene portata da 3 ad un numero intero n qualsiasi;

2) la forma quadratica fondamentale, non soltanto passa da 3 ad n variabili indipendenti, ma non conserva necessariamente il carattere di forma definita positiva.

La generalizzazione 1) dà luogo alla nozione di **spazio affine ad n dimensioni**, nel quale esistono infiniti sistemi di coordinate affini, dall'uno all'altro dei quali si passa con una trasformazione lineare invertibile (cioè a determinante diverso da zero, $det(A^{i'}_k) \neq 0$):

$$x^{i'} = A^{i'}_k x^k + A^{i'}, \quad (i', k = 1, 2, ..., n).$$

Anche in tali spazi generalizzati l'introduzione di una struttura metrica (prodotto scalare, norma, misure angolari, ecc.) restringe la classe dei sistemi coordinati ammissibili. Questi sono caratterizzati dal fatto che le trasformazioni che fanno passare dall'uno all'altro di essi sono tutte e sole quelle che lasciano invariata una certa forma quadratica

$$g_{ik} x^i x^k \equiv g_{i'k'} x'^i x'^k \quad (i, k, i', k' = 1, 2, ..., n),$$

che chiameremo ancora **forma quadratica fondamentale**.
Ai coefficienti g_{ik} si chiede ancora di costituire una matrice diagonale ($g_{ik} = 0$ *per* $i \neq k$), con termini diagonali g_{ii} unitari ma non necessariamente tutti positivi. Alla forma fondamentale, che ancora viene assunta a definire la norma di un vettore, non si richiede più, di essere definita positiva, potendo assumere per certi vettori un valore positivo, per altri un valore negativo, e per altri ancora un valore nullo, anche se tutte le componenti del vettore sono non nulle.
Un tale spazio si dirà **spazio euclideo ad n dimensioni**, se la forma quadratica fondamentale è definita positiva; si dirà invece uno **spazio pseudo-euclideo ad n dimensioni**, se la forma quadratica fondamentale è indefinita.[10]
Di quest'ultimo tipo è lo spaziotempo 4-dimensionale di Minkowski del quale ci occuperemo nel prossimo paragrafo.

5.2 Lo spaziotempo di Minkowski come modello di Universo

Il carattere assoluto degli eventi, la cui esistenza è completamente indipendente da ogni riferimento galileiano e **l'invarianza della forma quadratica** (5.9) di fronte ad ogni trasformazione di Lorentz, suggerì a Minkowski di interpretare l'insieme di tutti i possibili eventi che accadono nell'Universo, come uno spazio affine $\mathcal{E}^4$ a 4 dimensioni, nel quale le coordinate $(x^i) \equiv (x^0, x^1, x^2, x^3)$ fungono da coordinate cartesiane ortogonali di punto, munito di *una metrica* costituita precisamente dall'invariante fondamentale[11]

$$\mathbf{ds}^2 \equiv \eta_{ik} dx^i dx^k \gtreqless 0. \tag{5.23}$$

[10]Per uno studio più approfondito degli spazi pseudoeuclidei, ed in particolare dello spazio di MInkowski, si rimanda il lettore ai seguenti testi: N. V. Efimov, *Géométrie Supérieure*, pagg. 456-505, Edition Mir Moscou (1981) and N. V. Efimov and E. R. Rozendorn, *Linear Algebra and Multi- Dimensional Geometry*, pagg. 258-259, Mir Publishers - Moscow (1975).

[11]Il concetto di **spazio affine** viene introdotto in Geometria Analitica generalmente in due modi diversi equivalenti fra di loro come ora spiegheremo (*continua nella pagina successiva*):

- **Def.1**: Dicesi **spazio affine** $\mathcal{A}^n$, ad n dimensioni, un insieme di elementi (o "punti": $A, B, C, ...$), posti in corrispondenza biunivoca con le *ennuple di numeri reali* (x^i) $(i = 1, 2, ..., n)$, definite a meno di trasformazioni lineari invertibili: $x^{i'} = A^{i'}_k x^k + A^{i'}$, essendo $A^{i'}_k \equiv \dfrac{\partial x^{i'}}{\partial x^k}$ una matrice regolare, costante come le $A^{i'}$. L'insieme di tali trasformazioni costituisce un gruppo detto **il gruppo affine**.
- **Def. 2**: Uno **spazio affine** $\mathcal{A}^n$, ad n dimensioni è un insieme di elementi (o "punti": A, B, C,....), associato ad uno **spazio vettoriale** $\mathcal{V}^n$, in modo che esista un'applicazione suriettiva α tra le coppie ordinate di punti di $\mathcal{A}^n$ ed i vettori di $\mathcal{V}^n$: $\mathcal{A}^n \times \mathcal{A}^n \to \mathcal{V}^n$, cioè $(A, B) \to \mathbf{V} \equiv \alpha(A, B) \equiv \overrightarrow{AB}$, che verifichi le condizioni:
  - i) $\forall A \in \mathcal{A}^n$ e $\mathbf{V} \in \mathcal{V}^n$, $\exists!\, B \in \mathcal{A}^n$ tale che $\overrightarrow{AB} = \mathbf{V}$;
  - ii) $\forall A, B, C \in \mathcal{A}^n$ si ha $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$, (relazione triangolare).

Tale spazio è uno **spazio pseudo-euclideo**, visto che la forma (5.23) non è definita positiva bensì indefinita.[12] Osserviamo che, essendo (dx^i) un 4- vettore ed $\eta_{ik}dx^i dx^k$ un invariante, per un noto criterio di tensorialità, η_{ik} è un tensore doppio simmetrico chiamato *tensore metrico* (cfr. Cap. 8, paragrafo 8.15, Teorema 8.5).
Anche se alla forma medesima si dà spesso il nome di *quadrato della distanza, o dell'intervallo, tra due punti infinitamente vicini dello spazio quadridimensionale* in questione, è chiaro che si tratta soltanto di una espressione puramente convenzionale, alla quale non si deve attribuire il significato fisico ordinario. Infatti, come vedremo nei prossimi paragrafi, il $\mathbf{ds}^2$ di Minkowski può essere interpretato, a meno del segno, come quadrato di una effettiva distanza (in un particolare riferimento) soltanto quando esso ha un valore negativo. Se il $\mathbf{ds}^2$ è positivo, invece, è piuttosto interpretabile, a meno del fattore c^2, come il quadrato di un intervallo temporale (sempre in un particolare riferimento).
È preferibile comunque interpretare il $\mathbf{ds}^2$ di Minkowski come una nuova grandezza, avente carattere assoluto, capace di misurare in qualche modo il divario spazio-temporale tra due eventi. Ciò non toglie che l'analogia formale con la forma quadratica fondamentale dell'ordinario spazio euclideo tridimensionale

$$\mathbf{dx}^2 \equiv \delta_{ik}dx^i dx^k = dx^2 + dy^2 + dz^2,$$

dove δ_{ik} è il delta di Kronecker ($\delta_{ik} = 1$ *se* $i = k$, $\delta_{ik} = 0$ *se* $i \neq k$), può suggerire di adottare, nello spazio 4-dimensionale degli eventi, un comodo ed espressivo linguaggio geometrico. Possiamo perciò dare la seguente
Definizione:
Dicesi spaziotempo di Minkowski *l'insieme* $\mathcal{E}^4$ *di tutti i possibili eventi che accadono nel nostro universo e caratterizzati ciascuno - in un prefissato sistema galileiano - dalla quaterna di coordinate spazio-temporali* $(x^i) \equiv (x^0, x^1, x^2, x^3)$, *munito della metrica*

$$\mathbf{ds}^2 \equiv \eta_{ik}dx^i dx^k = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \gtrless 0.$$

Tale spazio, che indicheremo con la notazione $\mathcal{M}^4 \equiv (\mathcal{E}^4, \eta_{ik})$, non porta soltanto il vantaggio di un comodo e sintetico linguaggio geometrico, ma fornisce un potente strumento per costruire teorie fisiche soddisfacenti ai due postulati della Relatività Ristretta.
Infatti, come sarà evidenziato meglio nel seguito, **una teoria fisica che sia formulata in termini geometrici intrinseci, vettoriali o tensoriali di $\mathcal{M}^4$ gode automaticamente della proprietà di essere espressa in una medesima forma in tutti i riferimenti galileiani, soddisfacendo in tal modo automaticamente al principio einsteiniano di relatività.**
Torneremo ad approfondire la struttura geometrica dello spaziotempo 4-dimensionale pseudo-euclideo di Minkowski nei paragrafi 5.4 e 5.5 di questo capitolo.

5.3 Classificazione degli intervalli spaziotemporali

Abbiamo visto, in ciò che precede che, anche nel caso della più generale trasformazione di Lorentz, *il divario spazio-temporale è una quantità invariante* in valore e forma nel passare da un riferimento inerziale ad un altro arbitrario, cioè:

$$\mathbf{ds}^2 = \mathbf{ds}'^2. \tag{5.24}$$

Osserviamo che in $\mathcal{A}^n$ si possono scegliere, in infiniti modi, insiemi ordinati di $n+1$ punti: $\Omega, U_1, U_2, ..., U_n$ (diciamo pure *una origine ed n punti unità*), tali che i vettori $\mathbf{E}_i = \overrightarrow{\Omega U_i}$ $(i = 1, 2, ..., n)$ siano linearmente indipendenti in $\mathcal{A}^n$ (formino dunque una base per questo spazio) e definiscano ivi un *riferimento affine*: $\mathcal{RA} \equiv (\Omega, \mathbf{E}_i)$.

[12] Ricordiamo che, se $\mathcal{V}^n$ è **lo spazio vettoriale associato ad uno spazio affine** $\mathcal{A}^n$ e φ una forma quadratica atta a definire **un prodotto scalare** fra gli elementi di $\mathcal{V}^n$, la coppia $(\mathcal{A}^n, \varphi)$ viene chiamata **"spazio euclideo"** se φ è definita positiva (o anche definita negativa), **"spazio pseudo-euclideo"** se φ è indefinita.

La forma quadratica fondamentale $\mathbf{ds}^2$ costituisce la *metrica dello spazio di Minkowski*; essa è indefinita e di segnatura $(+\ -\ -\ -)$, per cui può assumere valori sia positivi sia negativi che nulli.[13] Quindi anche eventi *distinti* possono avere tra loro divario spazio-temporale nullo. Come ora vedremo, il segno e anche il valore numerico del divario tra due eventi hanno importanti interpretazioni fisiche.

5.3.1 Divario spazio-temporale del genere tempo

Se la distanza spaziale che separa due eventi è minore della distanza che la luce percorre nell'intervallo temporale che separa gli eventi stessi, cioè se

$$dl^2 < c^2 dt^2, \tag{5.25}$$

e con la convenzione da noi adottata risulta $\mathbf{ds}^2 > 0$.[14] I due eventi sono allora collegabili con un effettivo segnale, in altri termini, possono essere posti l'uno con l'altro in relazione di causa ed effetto. Infatti, poiché il $\mathbf{ds}^2 > 0$ è invariante nel passare dal sistema galileiano $\mathcal{R}$ al sistema $\mathcal{R}'$, sarà anche $\mathbf{ds}'^2 > 0$, ed essendo $\mathbf{ds}'^2 \equiv c^2 dt'^2 - d\ell'^2$, è allora possibile soddisfare a questa condizione con $dl'^2 = 0$, *esiste dunque un riferimento galileiano nel quale i due eventi risultano equilocati* e il $\mathbf{ds}'^2$ si riduce al solo termine $c^2 dt'^2$, cioè, a meno del fattore c^2, al quadrato di un intervallo temporale. Se ne trae perciò la seguente conclusione:

> *Se il divario spazio-temporale tra due eventi è positivo, esso si può interpretare, a meno del fattore c^2, come* **il quadrato dell'intervallo temporale** *che intercorre tra i due eventi nel riferimento galileiano, certamente esistente, in cui essi sono equilocati.*

In tal caso diremo che **il divario spazio-temporale è del genere tempo**.

5.3.2 Divario spazio-temporale del genere spazio

Se la distanza spaziale che separa i due eventi è invece maggiore della distanza che la luce percorre nell'intervallo temporale che separa gli eventi stessi, cioè se

$$dl^2 > c^2 dt^2, \tag{5.26}$$

risulta $\mathbf{ds}^2 < 0$. Si può constatare subito che i due eventi ora non sono collegabili con alcun segnale fisico né con alcuna azione causale. Infatti, poiché il $\mathbf{ds}^2 < 0$ è invariante nel passare dal sistema galileiano $\mathcal{R}$ al sistema $\mathcal{R}'$ sarà anche $\mathbf{ds}'^2 < 0$ ed essendo $\mathbf{ds}'^2 \equiv c^2 dt'^2 - d\ell'^2$, è allora possibile soddisfare a questa condizione con $dt'^2 = 0$. Non esiste, in questo caso, un riferimento galileiano nel quale i due eventi avvengono nello stesso luogo; *esiste in compenso un riferimento galileiano nel quale i due eventi risultano simultanei* e il $\mathbf{ds}'^2$ si riduce al solo termine $-dl'^2$, cioè, a meno del segno, si riduce al quadrato di un intervallo spaziale. Si trae allora la seguente interpretazione:

> *Se il divario spazio-temporale tra due eventi è negativo, esso si può interpretare, a meno del segno, come* **il quadrato dell'intervallo spaziale** *che separa i due eventi nel riferimento galileiano, certamente esistente, in cui essi appaiono simultanei.*

In questo caso diremo che **il divario spazio-temporale è del genere spazio**.

[13] Non tutti gli autori definiscono il $\mathbf{ds}^2$ come abbiamo fatto noi, alcuni lo definiscono col segno cambiato, cioè $\mathbf{ds}^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2$ oppure $\mathbf{ds}^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2$; entrambe le notazioni presentano pregi e difetti, perciò non c'è nessuna seria motivazione per preferire l'una piuttosto che l'altra, è soltanto questione di gusto o di abitudine. Osserviamo comunque che la notazione da noi introdotta è quella originale usata da Einstein nel lavoro del 1916 sui *Fondamenti della Relatività Generale* e non c'è nessuna buona ragione per cambiarla. Essa è stata adottata anche da Landau e Lifschitz e, dopo di loro, da moltissimi altri relativisti.

[14] Si osservi che, se si adotta per il $\mathbf{ds}^2$, la convenzione di segnatura $(+,+,+,-)$, cioè opposta a quella da noi adottata, allora gli intervalli spazi-temporali del genere tempo saranno caratterizzati dal segno negativo, cioè $\mathbf{ds}^2 < 0$; in modo analogo si avrà cambio di segno anche per gli intervalli del genere spazio per i quali sarà $\mathbf{ds}^2 > 0$.

5.3.3 Divario spazio-temporale del genere luce

Se infine, la distanza spaziale fra i due eventi è esattamente uguale alla distanza che la luce percorre nell'intervallo temporale che separa gli eventi, cioè se

$$dl^2 = c^2 dt^2, \tag{5.27}$$

risulta $\mathbf{ds}^2 = 0$, e perciò anche $\mathbf{ds}'^2 = 0$. Possiamo dunque dire che

> *Se i due eventi, pur essendo distinti, hanno un divario spazio- temporale nullo, essi possono essere collegati causalmente, ma soltanto mediante azioni fisiche propagantisi con la velocità della luce nel vuoto.*

Diremo che i due eventi sono separati da un **divario spazio- temporale del genere luce**.
Il caso ora considerato, benché intermedio tra i due precedenti, può concettualmente farsi rientrare, come caso limite, nella classe del genere tempo, visto che, in ogni caso, gli eventi sono collegabili da una azione causale.
Possiamo perciò concludere col seguente schema

$$\begin{cases} \mathbf{ds}^2 > 0, & \text{divario del genere tempo} \\ \mathbf{ds}^2 < 0, & \text{divario del genere spazio} \\ \mathbf{ds}^2 = 0, & \text{divario del genere luce} \end{cases} \tag{5.28}$$

5.4 Quadrivettori dello spazio di Minkowski

5.4.1 Proprietà geometriche di $\mathcal{M}^4$

Come già detto, un *punto* di $\mathcal{M}^4 \equiv (\mathcal{E}^4, \eta_{ik})$ corrisponde ad un *evento* (puntiforme). In un riferimento fisico galileiano ogni evento $\mathcal{E}$ è individuato da quattro coordinate $(x^0 \equiv ct,\ x^1,\ x^2,\ x^3)$; la prima coordinata ci dà l'*istante* $t = \dfrac{x^0}{c}$ in cui l'evento ha luogo, e dice *quando* è avvenuto, le altre tre coordinate individuano invece la sua *posizione* spaziale e perciò dicono *dove* è avvenuto.
Le medesime quattro coordinate possono essere interpretate come coordinate cartesiane ortogonali di punto nello spazio pseudo-euclideo $\mathcal{M}^4$. Ne consegue che scegliere un riferimento fisico galileiano $\mathcal{R} \equiv (Oxyz,\ t)$ equivale a scegliere un riferimento cartesiano ortogonale in $\mathcal{M}^4$ formato da un asse temporale x^0 e da tre assi spaziali x^1, x^2, x^3 che corrispondono rispettivamente agli assi x, y, z del riferimento galileiano.
Il punto $\mathcal{O} \in \mathcal{M}^4$ nel quale hanno origine i quattro assi rappresenta l'evento che in tutti i riferimenti galileiani considerati viene assunto come origine spaziale all'istante iniziale. I punti dell'asse x^0 rappresentano i successivi eventi dei quali è composta la storia dell'origine O del riferimento galileiano considerato. Cambiando riferimento galileiano cambiano gli assi, ed in particolare l'asse temporale, che viene a rappresentare la storia di un altro oggetto: l'origine O' del nuovo riferimento galileiano.
Sappiamo che ad ogni spazio affine $\mathcal{A}$ è associato uno spazio vettoriale i cui elementi sono i vettori individuati dalle coppie ordinate di punti di $\mathcal{A}$. Perciò lo spazio di Minkowski $\mathcal{M}^4$ può essere anche interpretato come uno spazio vettoriale a 4 dimensioni: due suoi punti qualsiasi $X \equiv (x^i)$ e $Y \equiv (y^i)$, $(i = 0, 1, 2, 3)$ individuano un *vettore* a quattro componenti $\overrightarrow{XY} \equiv (y^i - x^i)$ che chiameremo **quadrivettore** o più sinteticamente **4- vettore**. E siccome le coordinate che individuano gli eventi, nel passare da un riferimento inerziale $\mathcal{R}$ ad un altro $\mathcal{R}'$, in configurazione

$x - standard$, sono legate dalle trasformazioni speciali di Lorentz, cioè $x^{i'} = L^{i'}_j x^j$ con $L^{i'}_j = \dfrac{\partial x^{i'}}{\partial x^j}$, possiamo definire un **4-vettore** $\overrightarrow{V}$ di $\mathcal{M}^4$ come un oggetto le cui componenti controvarianti $V^i \equiv (V^0, V^1, V^2, V^3)$, nel passare da $\mathcal{R}$ ad $\mathcal{R}'$, si trasformano secondo la legge tensoriale

$$V^{i'} = \frac{\partial x^{i'}}{\partial x^j} V^j = L^{i'}_j V^j. \tag{5.29}$$

Tale legge scritta esplicitamente ci dà

$$\begin{cases} V'^0 = \gamma(v)(V^0 - \beta V^1) \\ V'^1 = \gamma(v)(V^1 - \beta V^0) \\ V'^2 = V^2 \\ V'^3 = V^3 \end{cases} \tag{5.30}$$

che non è altro che la trasformazione speciale di Lorentz applicata alle componenti controvarianti del 4-vettore V^i.

5.4.2 Prodotto scalare, norma e modulo dei vettori di $\mathcal{M}^4$

Vediamo adesso di definire una **norma** per i 4-vettori. Prima di fare questo notiamo che la metrica di Minkowski definisce automaticamente un **prodotto scalare** sui 4-vettori di $\mathcal{M}^4$. Infatti, se indichiamo con $\{\mathbf{E_i}\}, i = 0, 1, 2, 3$ la base canonica di $\mathcal{M}^4$ in quanto spazio vettoriale, cioè la base individuata dai versori degli assi x^0, x^1, x^2, x^3, ogni 4-vettore $\mathbf{V}$ di $\mathcal{M}^4$ si potrà rappresentare come combinazione lineare dei vettori di base, cioè $\mathbf{V} = V^i \mathbf{E}_i$ con (V^i) le cosidette *componenti controvarianti* di $\mathbf{V}$. Se consideriamo un qualsiasi altro 4-vettore $\mathbf{W} = W^k \mathbf{E}_k$ di $\mathcal{M}^4$, possiamo definire il loro prodotto scalare utilizzando il tensore metrico η_{ik} nel seguente modo:

$$\mathbf{V} \cdot \mathbf{W} \equiv (V^i \mathbf{E}_i) \cdot (W^k \mathbf{E}_k) = V^i W^k \mathbf{E}_i \cdot \mathbf{E}_k \stackrel{\text{def}}{=} \eta_{ik} V^i W^k = V^0 W^0 - V^1 W^1 - V^2 W^2 - V^3 W^3, \tag{5.31}$$

avendo indentificato i coefficienti η_{ik} del tensore metrico con i prodotti scalari dei vettori di base, cioè:

$$\eta_{ik} = \mathbf{E}_i \cdot \mathbf{E}_k = diag(+1, -1, -1, -1). \tag{5.32}$$

Come si può notare, quello definito dalla (5.31) non è un prodotto scalare definito positivo ma indefinito. A questo punto si può dare una definizione di **norma**[15] di un quadrivettore $\mathbf{V}$:

$$||\mathbf{V}|| \stackrel{\text{def}}{=} \mathbf{V} \cdot \mathbf{V} = \mathbf{V^2} = \left(V^0\right)^2 - \left(V^1\right)^2 - \left(V^2\right)^2 - \left(V^3\right)^2 \gtrless 0. \tag{5.33}$$

Inoltre si chiama **modulo** del 4-vettore $\mathbf{V}$ la radice quadrata del modulo della sua norma, cioè

$$|\mathbf{V}| \equiv V = \sqrt{|\mathbf{V} \cdot \mathbf{V}|} \geq 0. \tag{5.34}$$

Due quadrivettori, diciamo $\mathbf{V}$ e $\mathbf{W}$, si dicono **ortogonali** se risulta:

$$\mathbf{V} \cdot \mathbf{W} = 0; \tag{5.35}$$

val la pena osservare che la definizione è formalmente uguale a quella che si dà in uno spazio euclideo. Osserviamo che la base canonica $\{\mathbf{E}_i\}$ è **una base ortonormale** in quanto i vettori

[15] Avvertiamo il lettore che la definizione di **norma di un vettore** così come l'abbiamo data noi, è un po' diversa da quella che usualmente viene data negli spazi euclidei. In questi spazi per **norma** generalmente si intende il **modulo** del vettore, cioè $||\mathbf{x}|| \equiv |\mathbf{x}| = \sqrt{(x^1)^2 + (x^2)^2 + ...(x^n)^2}$.

che la compongono sono a due a due ortogonali e di modulo unitario, cioè $\mathbf{E}_i \cdot \mathbf{E}_k = 0, i \neq k$ e $|\mathbf{E}_i| = +1$, inoltre $||\mathbf{E}_0|| = +1$ mentre $||\mathbf{E}_\alpha|| = -1, \ \alpha = 1, 2, 3$.[16]
Si può constatare facilmente che le componenti cartesiane della base canonica sono

$$\mathbf{E}_0 \equiv (1, 0, 0, 0), \ \mathbf{E}_1 \equiv (0, 1, 0, 0), \ \mathbf{E}_2 \equiv (0, 0, 1, 0), \ \mathbf{E}_3 \equiv (0, 0, 0, 1).$$

Con le definizioni date più sopra, il 4-vettore infinitesimo $\mathbf{ds} = (dx^i)$ deve essere considerato come uno spostamento da un dato evento ad un altro infinitamente vicino in $\mathcal{M}^4$ e la sua norma, per definizione sarà

$$||\mathbf{ds}|| \equiv \mathbf{ds} \cdot \mathbf{ds} \equiv \mathbf{ds}^2,$$

mentre il suo modulo potrà scriversi semplicemente come

$$ds = \sqrt{|\mathbf{ds}^2|} \geq 0.$$

5.5 Proprietà dei quadrivettori

5.5.1 Classificazione dei quadrivettori e definizione del cono di luce

Una conseguenza immediata della (5.33), è che la norma di un quadrivettore può essere positiva, nulla o anche negativa; in particolare è negativa la norma dei 4-vettori $\{\mathbf{E}_\alpha\}, \ \alpha = 1, 2, 3$, della base canonica, mentre è positiva la norma del 4-vettore $\mathbf{E}_0$, come abbiamo già osservato nel paragrafo precedente. L'insieme dei 4-vettori si può quindi dividere nelle tre classi:

1) *Vettori del genere tempo*, o vettori a norma positiva;

2) *Vettori del genere spazio*, o vettori a norma negativa;

3) *Vettori del genere luce o isotropi*, o vettori a norma nulla.

Questi nomi derivano dal fatto che i vettori di $\mathcal{M}^4$ rappresentano intervalli spazio-temporali fra due eventi, i quali costituiscono gli estremi del quadrivettore stesso; infatti, un vettore a norma positiva in $\mathcal{M}^4$ corrisponde ad un divario spazio-temporale del genere tempo e così via.
Se per un punto $\overline{P} \equiv (\overline{x}^i)$ di $\mathcal{M}^4$ spicchiamo tutti i possibili vettori del genere luce, il luogo dei loro estremi $P \equiv (x^i)$ costituisce un cono indefinito $\mathcal{C}$, 3-dimensionale a due falde, di vertice $\overline{P}$ che si suole chiamare *cono di luce o cono isotropo* relativo al punto evento $\overline{P}$. Che si tratti di un cono è subito visto dalla sua definizione in quanto deve essere:[17]

$$||P - \overline{P}|| = 0,$$

cioè

$$(x^0 - \overline{x}^0)^2 - (x^1 - \overline{x}^1)^2 - (x^2 - \overline{x}^2)^2 - (x^3 - \overline{x}^3)^2 = 0,$$

che è appunto l'equazione cartesiana di un cono 3-dimensionale, avente il vertice nel punto $\overline{P} \equiv (\overline{x}^0, \overline{x}^1, \overline{x}^2, \overline{x}^3)$, in uno spazio affine a quattro dimensioni. Anziché prendere un punto

[16] D'ora in avanti, seguendo Landau e Lifschitz, adotteremo la seguente convenzione: con lettere latine minuscole $a, b, c, ...i, j, h, k, ...$ ecc., si indicheranno gli indici spazio-temporali che assumono i valori $0, 1, 2, 3$, mentre con lettere greche minuscole $\alpha, \beta, ..., \mu, \nu, ...$, ecc., saranno indicati gli indici puramente spaziali aventi i valori $1, 2, 3$. È bene avvertire che, questa convenzione è adottata da circa la metà degli autori, l'altra metà adotta esattamente la convenzione opposta.

[17] Ricordiamo al lettore che nello spazio euclideo ordinario E^3 riferito ad un sistema di coordinate cartesiane ortogonali $Oxyz$, l'equazione di un generico cono circolare retto, ad asse parallelo all'asse z ed avente il vertice nell'origine O delle coordinate, è $x^2 + y^2 - z^2 = 0$ (o equivalentemente $z^2 - x^2 - y^2 = 0$). Se invece il cono è ancora ad asse parallelo all'asse z ma col vertice in un generico punto $V(x_0, y_0, z_0)$ dello spazio, la sua equazione sarà: $(x - x_0)^2 + (y - y_0)^2 - (z - z_0)^2 = 0$.

arbitrario di $\mathcal{M}^4$ possiamo, per semplicità, scegliere l'origine $\mathcal{O}$ del sistema di coordinate e classificare tutti gli eventi di $\mathcal{M}^4$ rispetto ad $\mathcal{O}$, in tal caso l'equazione cartesiana del relativo cono $\mathcal{C}_0$ si semplifica nella

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = 0.$$

Il cono $\mathcal{C}_0$ divide gli eventi di $\mathcal{M}^4$ in tre categorie, a seconda della loro relazione rispetto ad $\mathcal{O}$ (vedi Fig.5.1). La prima è costituita dagli eventi appartenenti alla falda inferiore $\mathcal{C}_0^-$ del cono e rappresenta il **passato** dell'evento $\mathcal{O}$; gli eventi in essa contenuti sono quelli che possono aver avuto un'influenza fisica sull'evento $\mathcal{O}$ ma che non possono essere influenzati in nessun modo da $\mathcal{O}$. La seconda, costituita dagli eventi appartenenti alla falda superiore $\mathcal{C}_0^+$ del cono, rappresenta il **futuro** dell'evento $\mathcal{O}$; essa è costituita da tutti gli eventi sui quali l'evento $\mathcal{O}$ può avere una influenza fisica ma dai quali non può subire nessuna influenza. La terza categoria è l'insieme dei punti di $\mathcal{M}^4$ che sono esterni al cono di luce: sono tutti gli eventi che non possono né influire fisicamente sull'evento $\mathcal{O}$, né subirne l'influenza; l'insieme può essere denominato, con una accezione che è soltanto in parte diversa da quella abituale, il **presente indefinito** dell'evento $\mathcal{O}$.
Risulta evidente che *la linea oraria* di una particella (ovvero la sua storia, intesa come evoluzione spazio-temporale) ha la caratteristica di avere, in ogni punto, la tangente contenuta nel cono di luce relativo a quel punto.
L'asse dei tempi x^0 rappresenta la storia di un osservatore che si trova in quiete nell'origine di un riferimento galileiano ed è asse di simmetria del cono di luce relativo ad $\mathcal{O}$; esso viene naturalmente orientato verso il futuro.
Ogni iperpiano $x^0 = cost.$, luogo degli eventi simultanei rispetto all'osservatore galileiano che ha l'asse temporale come linea oraria, costituisce lo spazio fisico tridimensionale Σ_t visto dall'osservatore stesso, all'istante $t = \dfrac{x^0}{c}$.
Ai risultati ottenuti si può dare una rappresentazione grafica; naturalmente, non potendo raffigurare graficamente lo spazio di Minkowski 4-dimensionale $\mathcal{M}^4$, ci si limita a considerare spazi di Minkowski a 3 o a 2 dimensioni dei quali è possibile una semplice rappresentazione grafica. Un $\mathcal{M}^3$, per esempio, potrà essere rappresentativo, di tutti gli eventi che avvengono sul piano xy, o sui piani xz, o yz di un dato riferimento galileiano $\mathcal{R}$, essendo in ogni caso x^0 l'asse temporale di $\mathcal{M}^3$ (vedi Fig.5.1).

Un caso più semplice è uno spazio di Minkowski 2-dimensionale $\mathcal{M}^2$ in quanto può essere rappresentato su di un piano; di questo spazio ci occuperemo dettagliatamente nei prossimi due paragrafi.
Concludiamo, dimostrando il seguente interessante teorema.

Teorema 5.1 *Se* $\mathbf{V}$ *è un vettore del genere tempo, ogni vettore* $\mathbf{W} \neq 0$ *ortogonale a* $\mathbf{V}$ *è del genere spazio.*

Dimostrazione. Essendo per ipotesi $\mathbf{V}$ del genere tempo possiamo scegliere in $\mathcal{M}^4$ l'asse temporale x^0 parallelo a $\mathbf{V}$ per cui sarà $\mathbf{V} = \lambda \mathbf{E}_0$ con $\lambda \neq 0$, ma $\mathbf{V} = V^0\mathbf{E}_0 + V^1\mathbf{E}_1 + V^2\mathbf{E}_2 + V^3\mathbf{E}_3$ perciò deve essere $V^0 = \lambda \neq 0,\ V^1 = V^2 = V^3 = 0$. Tenuto conto di ciò e dell'altra ipotesi che $\mathbf{V}$ e $\mathbf{W}$ sono ortogonali, cioè $\mathbf{V} \cdot \mathbf{W} = 0$, si ottiene $V^0W^0 = 0$, ed essendo $V^0 \neq 0$ sarà $W^0 = 0$, dunque $||\mathbf{W}|| = -\sum_\alpha (W^\alpha)^2 < 0$, cioè $\mathbf{W}$ è un 4-vettore del genere spazio. ■

Diamo qui di seguito una dimostrazione alternativa, del tutto generale, del precedente teorema:

Dimostrazione. Essendo per ipotesi $\mathbf{V}$ del genere tempo possiamo scrivere: $||\mathbf{V}|| = (V^0)^2 - \sum_\alpha (V^\alpha)^2 > 0$, con $V^0 \neq 0$. La quantità al primo membro essendo positiva possiamo scriverla così $(V^0)^2 - \sum_\alpha (V^\alpha)^2 = \mu^2$ con μ reale. Avremo allora

$$(V^0)^2 = \sum_\alpha (V^\alpha)^2 + \mu^2 > 0. \tag{5.36}$$

Figura 5.1: Diagramma spazio-temporale di Minkowski in $\mathcal{M}^3$; cono di luce relativo ad $\mathcal{O}$ e linea oraria di una particella.

Utilizziamo ora l'ipotesi che $\mathbf{W}$ è un qualunque vettore ortogonale a $\mathbf{V}$, cioè $\mathbf{V} \cdot \mathbf{W} = V^0 W^0 - \sum_\alpha V^\alpha W^\alpha = 0$, dalla quale si ricava

$$W^0 = \frac{\sum_\alpha V^\alpha W^\alpha}{V^0}. \tag{5.37}$$

D'altra parte la norma di $\mathbf{W}$ è

$$||\mathbf{W}|| = (W^0)^2 - \sum_\alpha (W^\alpha)^2,$$

sostituendo in quest'ultima la (5.37) otteniamo

$$||\mathbf{W}|| = \frac{(\sum_\alpha V^\alpha W^\alpha)^2}{(V^0)^2} - \sum_\alpha (W^\alpha)^2 = \frac{1}{(V^0)^2}\Big[\Big(\sum_\alpha V^\alpha W^\alpha\Big)^2 - (V^0)^2 \sum_\alpha (W^\alpha)^2\Big]. \tag{5.38}$$

Tenendo ora conto della (5.36) la (5.38) si scriverà

$$||\mathbf{W}|| = \frac{1}{(V^0)^2}\Big[(\mathbf{v} \cdot \mathbf{w})^2 - v^2\, w^2 - \mu^2 w^2\Big];$$

avendo indicato con $\mathbf{v} = (V^\alpha)$ e $\mathbf{w} = (W^\alpha)$ i due 3- vettori aventi per componenti le componenti spaziali dei 4-vettori $\mathbf{V}$, $\mathbf{W}$ rispettivamente. Per la disuguaglianza di Cauchy-Schwartz $(\mathbf{v} \cdot \mathbf{w})^2 \leq v^2\, w^2$, l'espressione entro parentesi quadre è sicuramente negativa, cioè $||\mathbf{W}|| < 0$, il che significa che il 4-vettore $\mathbf{W}$ è del genere spazio. ■

L'ortogonalità tra quadrivettori di cui uno del genere tempo e l'altro del genere spazio costituisce dunque una sorta di **coniugio** rispetto al cono di luce nel senso che uno è interno e l'altro esterno a tale cono.

5.5.2 Diagrammi spazio-temporali e linee di universo in $\mathcal{M}^2$.

Il caso più semplice di spazio di Minkowski è naturalmente il caso in cui gli eventi sono caratterizzati soltanto da due coordinate $x^0 = ct$ ed x^1, una temporale ed una spaziale e perciò si riferiscono ad avvenimenti che accadono sopra una retta arbitraria dello spazio fisico, che possiamo identificare con l'asse delle x (o con uno degli assi y o z), di un dato sistema inerziale $\mathcal{R}$.[18] In questo caso semplicissimo lo spaziotempo 2-dimensionale sarà indicato con $\mathcal{M}^2$. Inoltre, per semplificare un pò le notazioni, conveniamo di indicare la variabile spaziale x^1 con x ($x^1 \equiv x$) e la variabile temporale x^0 con T ($x^0 \equiv T = ct$) come abbiamo già fatto nel paragrafo 3.3, formule (3.33a) e (3.33b). Lo spazio di Minkowski $\mathcal{M}^2$ può allora essere rappresentato su di un piano euclideo ordinario, scegliendo in maniera arbitraria l'origine[19], che indicheremo con $\mathcal{O}$, ed i cui assi cartesiani ortogonali siano x, assunto come asse delle ascisse, ed T preso come asse delle ordinate (vedi Fig.5.2).

Figura 5.2: Diagramma spazio-temporale di Minkowski in $\mathcal{M}^2$; eventi simultanei ed eventi equilocati.

Tale rappresentazione, si chiama **diagramma spazio-temporale**: ogni evento è individuato allora, su tale diagramma, dal punto $\mathcal{P} \equiv (x, T)$.
I punti $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, ... appartenenti ad una retta parallela all'asse x, avendo tutti la stessa ordinata temporale T, rappresenteranno eventi **simultanei**, ed i punti $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$, ... appartenenti ad una retta parallela all'asse dei tempi T, avendo tutti la stessa ascissa, rappresenteranno eventi che avvengono nello stesso luogo, cioè sono **equilocati** sul corrispondente asse x del riferimento inerziale $\mathcal{R}$ dello spazio fisico (Fig.5.2).
Il moto di una particella nello spazio fisico, che si muova sopra l'asse x del riferimento $\mathcal{R}$, avrà un'equazione oraria del tipo $x = x(t)$, che nello spazio $\mathcal{M}^2$ si scriverà, in maniera più opportuna, $x = x(T)$ (con $T = ct$) e rappresenta ovviamente una linea λ del piano (x, T) che diremo **linea oraria o linea di universo o storia** della particella mobile, e si tratta di una linea orientata verso il futuro (Fig. 5.2).
A scanso di equivoci, è opportuno notare che, nei diagrammi spazio-temporali 2-dimensionali, la variabile indipendente T non è riportata, come si fa usualmente, sull'asse delle ascisse (asse orizzontale) ma sull'asse delle ordinate (asse verticale). Ciò comporta, per esempio, che la pendenza di una linea di universo in un suo punto generico sarà ancora data dal coefficiente angolare della retta tangente alla linea in quel punto ma riferito all'angolo che la tangente forma,

[18] Più in generale si potrebbero considerare eventi che avvengono su una curva generica dello spazio fisico su cui venga prefissato un sistema di ascisse curvilinee in modo che ad ogni punto P della curva corrisponda un evento caratterizzato dalle coordinate $(x^0 = ct, x^1)$ dove x^1 rappresenta l'ascissa curvilinea di P e t l'istante che le compete.

[19] Conveniamo di indicare con lettere maiuscole dell'alfabeto latino A, B, C,.... ecc., i punti o particelle dello spazio fisico ordinario e con le analoghe lettere in corsivo $\mathcal{A}, \mathcal{B}, \mathcal{C}$,.... ecc., i corrispondenti punti-evento dello spaziotempo.

non con l'asse delle ascisse x, ma con l'asse delle ordinate T. Insomma la Geometria Analitica svolta nel piano (x, y) con x ascissa ed y ordinata continua a valere scambiando il ruolo delle ascisse con quello delle ordinate.
Osserviamo esplicitamente che, anche se una particella è in quiete nello spazio fisico in una posizione individuata dall'ascissa $\overline{x}$, la sua linea di universo, non si riduce mai ad un punto, ma è rappresentata dalla retta parallela all'asse T di equazione $x = \overline{x}$.
Perciò, nello spaziotempo il punto rappresentativo di una successione di eventi non è mai in quiete perché il tempo non può fermarsi.
Per esempio, la linea di universo di una particella che si trova in quiete nell'origine O del sistema inerziale $\mathcal{R}$ coincide con l'asse dei tempi T, uscente dall'origine $\mathcal{O}$, in quanto tutti i suoi punti sono caratterizzati dall'avere coordinate $(0, T)$ con T arbitrario. Così, se consideriamo l'intero asse x del sistema galileiano $\mathcal{R}$, esso è rappresentato in $\mathcal{M}^2$ dal fascio di rette parallele all'asse dei tempi.
Consideriamo ora il più semplice moto di una particella che si muova sull'asse x di $\mathcal{R}$ di moto uniforme con velocità costante $u = cost.$; tale moto è rappresentato da un'equazione lineare in t: $x = ut + x_0$, ovvero dall'equivalente equazione lineare in T: $x = \beta T + x_0$ con $\beta = \dfrac{u}{c}$ la solita velocità römeriana (vedi Fig.5.3).

Figura 5.3: Linee orarie in $\mathcal{M}^2$ del moto uniforme e del moto vario di una particella.

Poiché, nelle notazioni adottate, l'asse T è l'asse delle ordinate, la precedente equazione conviene scriverla nella forma $T = \dfrac{1}{\beta}x - \dfrac{x_0}{\beta}$ il cui coefficiente angolare è $\tan\varphi = \dfrac{1}{\beta}$ con φ angolo d'inclinazione della retta sull'asse x. Perciò tale retta sarà inclinata di un angolo $\psi = \dfrac{\pi}{2} - \varphi$ sull'asse T, ne segue allora che

$$\beta = \tan\psi, \qquad \text{ovvero,} \quad \psi = \arctan\beta.$$

Da queste ultime relazioni si deduce che, essendo sempre $0 \leq \beta \leq 1$, perché la velocità c della luce non è superabile in nessun moto fisico $(u \leq c)$, dovrà aversi sempre $0 \leq \psi \leq 45°$.
Le precedenti considerazioni possono essere generalizzate al caso di una particella che si muova di moto arbitrario sopra l'asse x del sistema galileiano $\mathcal{R}$. Se il moto della particella è rappresentato dall'equazione $x = x(T)$, la condizione che essa si muova con una velocità inferiore alla velocità della luce è espressa dalla relazione

$$\frac{dx}{dT} = \tan\psi = \frac{u}{c} = \beta < 1,$$

dove u è la velocità istantanea della particella e ψ l'angolo formato dalla tangente geometrica alla linea di universo in un suo punto generico con l'asse dei tempi T. Possiamo perciò concludere

dicendo che: **le linee orarie di qualsiasi particella, che si muova di moto vario sull'asse x del sistema inerziale $\mathcal{R}$, godono della proprietà che la tangente geometrica in ogni loro punto-evento forma con l'asse T un angolo sempre minore di $45°$; soltanto un segnale luminoso ha per linea oraria una retta inclinata di $45°$ sull'asse T.**
Tenendo conto di ciò che abbiamo detto più sopra, se fissiamo l'attenzione sull'origine $\mathcal{O} \equiv (0,0)$ del piano (x,T), le bisettrici del I° e II° quadrante essendo inclinate di $45°$ sull'asse T ed avendo equazioni $x - T = 0$ e $x + T = 0$ rispettivamente, rappresentano due raggi di luce che si propagano dall'origine O del sistema galileiano $\mathcal{R}$, lungo l'asse delle x ed in direzioni opposte (Fig.5.4).

Figura 5.4: Raggi di luce e cono di luce relativo all'origine $\mathcal{O}$, in $\mathcal{M}^2$.

Esse delimitano una regione piana, quella contenente l'asse temporale T, che può considerarsi come **un cono circolare retto degenere** $\mathcal{C}$, di apertura $45°$, rappresentato dall'equazione omogenea di 2° grado nelle variabili x, T:[20]

$$x^2 - T^2 = 0. \tag{5.39}$$

Tale cono degenere dicesi **cono di luce** in quanto è delimitato da due rette che rappresentano le linee orarie dei due raggi di luce di cui sopra. Tutte le linee orarie uscenti da $\mathcal{O}$, rappresentanti moti di particelle, devono necessariamente giacere entro il cono di luce ed avere la proprietà che le tangenti nei loro punti devono formare con l'asse T angoli $\psi < 45°$. I punti che stanno nella **falda inferiore** $\mathcal{C}^-$ del cono di luce rappresentano eventi che si trovano nel **passato** di $\mathcal{O}$, quelli che giacciono nella sua **falda superiore** $\mathcal{C}^+$, appartengono al futuro di $\mathcal{O}$. I punti che si trovano all'esterno del cono di luce rappresentano eventi che appartengono al **presente** di $\mathcal{O}$ e che non possono comunicare con $\mathcal{O}$ con nessun segnale perché questo dovrebbe viaggiare a velocità superiore alla velocità della luce e ciò è impossibile.
Analoga costruzione del cono di luce si può fare considerando come suo vertice un generico punto $\mathcal{P}_0 \equiv (x_0, T_0)$ del piano (x,T): esistono pertanto tanti coni di luce quanti sono gli eventi di $\mathcal{M}^2$.

5.5.3 Interpretazione geometrica della trasformazione di Lorentz in $\mathcal{M}^2$

Usando le variabili (x,T), introdotte nel paragrafo precedente, relative al riferimento galileiano $\mathcal{R}$, e le (x', T'), relative al riferimento $\mathcal{R}'$, in configurazione *x-standard* col primo, le trasformazioni speciali di Lorentz si scriveranno così:

$$\begin{cases} x' = \gamma(x - \beta T) \\ T' = \gamma(T - \beta x) \end{cases} \tag{5.40}$$

[20]Si ricordi che si chiama **apertura** di un cono circolare retto l'angolo ψ formato fra l'asse del cono ed una sua generatrice: esso è sempre inferiore a $90°(\psi < 90°)$. Nel caso limite che $\psi = 90°$ il cono degenera in un piano.

e le relative formule inverse:

$$\begin{cases} x = \gamma(x' + \beta T') \\ T = \gamma(T' + \beta x'). \end{cases} \tag{5.41}$$

Osserviamo che i due sistemi di coordinate in $\mathcal{M}^2$, per le ipotesi fatte sui sistemi inerziali in configurazione *x-standard*, devono avere la stessa origine $\mathcal{O} \equiv \mathcal{O}'$ all'istante $t = t' = 0$ e consideriamo ortogonali i due assi x e T di equazioni cartesiane $T = 0$, $x = 0$ rispettivamente. Analogamente gli assi x' e T', nel nuovo sistema di riferimento, avranno equazioni $T' = 0$, $x' = 0$, ma non possiamo pretendere che siano anch'essi ortogonali in quanto ci troviamo in uno spazio pseudo-euclideo. Tenuto conto della trasformazione di Lorentz (5.40), i nuovi assi nel vecchio sistema di riferimento hanno equazione, rispettivamente, $T = \beta x$ e $x = \beta T$. Quindi, posto $\psi = \arctan \beta$, l'asse x' forma un angolo ψ con l'asse x e l'asse T' forma un angolo $\frac{\pi}{2} - \psi$ con lo stesso asse e perciò forma ancora un angolo ψ con l'asse T (vedi Fig.5.5).

Figura 5.5: Interpretazione geometrica della trasformazione di Lorentz in $\mathcal{M}^2$: eventi equilocati.

Dunque il sistema cartesiano $(\mathcal{O}'x'T')$, relativo al riferimento $\mathcal{R}'$, non è ortogonale ma ha gli assi che si ottengono dagli analoghi di $(\mathcal{O}xT)$ ruotandoli entrambi di un angolo ψ verso la bisettrice del 1° quadrante. Poiché da $\beta < 1$ discende che deve essere $\psi < \pi/4$, la posizione degli assi x' e T' può dunque variare dalla posizione degli assi x e T fin verso la bisettrice del 1° quadrante del piano $(\mathcal{O}xT)$ che rappresenta una posizione limite non raggiungibile perchè in tal caso $\mathcal{R}'$ dovrebbe muoversi con velocità $v = c$ rispetto ad $\mathcal{R}$, ciò che sappiamo essere impossibile.
Per determinare le coordinate di un dato evento nel riferimento $\mathcal{R}$ è sufficiente mandare dal punto $\mathcal{A}$ rappresentativo dell'evento le parallele agli assi x e T e individuarne le intersezioni con gli assi stessi, avremo le coordinate $(x_{\mathcal{A}}, T_{\mathcal{A}})$; la stessa cosa per quanto riguarda il riferimento $\mathcal{R}'$, solo che in questo caso bisogna mandare le parallele agli assi x' e T' ottenendo così $(x'_{\mathcal{A}}, T'_{\mathcal{A}})$. In questo modo si può notare, per esempio, che eventi equilocati in $\mathcal{R}$ (ovvero giacenti su una retta parallela all'asse T), come $\mathcal{A}$ e $\mathcal{B}$, non sono equilocati in $\mathcal{R}'$ (tale retta non è infatti parallela all'asse T') e viceversa.

Analogamente, eventi simultanei in $\mathcal{R}$ (ovvero giacenti su una retta parallela all'asse x), come $\mathcal{C}$ e $\mathcal{D}$ non sono simultanei in $\mathcal{R}'$ (tale retta non è infatti parallela all'asse x') e viceversa (vedi Fig.5.6).
Osserviamo ancora che nella presente interpretazione dei diagrammi spazio-temporali di Minkowski, nel passare dal sistema (x, T) al sistema (x', T') oltre che un cambio di assi si ha anche un cambiamento delle unità di misura. Infatti se consideriamo i due punti unità, $\mathcal{J}'(0, 1)$ sull'asse x' e $\mathcal{K}'(1, 0)$ sull'asse T', applicando le trasformazioni di Lorentz (5.41) possiamo calcolare

Figura 5.6: Interpretazione geometrica della trasformazione di Lorentz in $\mathcal{M}^2$: eventi simultanei.

le coordinate degli stessi punti nel riferimento (x, T), avremo (vedi Fig.5.7):

$$\begin{cases} x_{\mathcal{J}'} = \gamma \\ T_{\mathcal{J}'} = \beta\gamma \end{cases} \tag{5.42}$$

e

$$\begin{cases} x_{\mathcal{K}'} = \beta\gamma \\ T_{\mathcal{K}'} = \gamma. \end{cases} \tag{5.43}$$

Figura 5.7: Cambiamento d'assi e cambiamento delle unità di misura

Perciò, valutando le lunghezze nel piano (x, T) avente gli assi ortogonali, otteniamo:

$$\mathcal{O}\mathcal{J}' = \sqrt{x_{\mathcal{J}'}^2 + T_{\mathcal{J}'}^2} = \gamma\sqrt{1+\beta^2} = \sqrt{\frac{1+\beta^2}{1-\beta^2}} > 1 \tag{5.44}$$

e

$$\mathcal{O}\mathcal{K}' = \sqrt{x_{\mathcal{K}'}^2 + T_{\mathcal{K}'}^2} = \gamma\sqrt{1+\beta^2} = \sqrt{\frac{1+\beta^2}{1-\beta^2}} > 1. \tag{5.45}$$

Riassumendo: *ammettere per il sistema galileiano $\mathcal{R}'$, in moto T.R.U. rispetto all'analogo sistema $\mathcal{R}$, con velocità costante* **v**, *la trasformazione di Lorentz* (5.40) *e la sua inversa* (5.41), *equivale a riferire il piano (x, T) ad un nuovo sistema di assi (x', T') ruotati (come in Fig.5.7) di uno stesso angolo ψ verso la bisettrice del 1° quadrante, ed adottare nuove unità di misura, eguali fra loro, sui nuovi assi.*
Osserviamo infine che utilizzando i diagrammi spazio-temporali di Minkowski, si potrebbero ritrovare, per via puramente geometrica, molte proprietà della Cinematica Relativistica. Nel prossimo paragrafo, come esempio, faremo vedere come è possibile ritrovare la contrazione delle lunghezze.

5.5.4 Deduzione geometrica della contrazione di Lorentz in $\mathcal{M}^2$

Cominciamo col considerare la rappresentazione grafica sul piano (x, T) di una sbarra AB disposta lungo l'asse x del riferimento $\mathcal{R}$ ed inizialmente in quiete. Essa sarà rappresentata dal segmento $\mathcal{A}_0\mathcal{B}_0$ dell'asse x del piano (x, T) (vedi Fig.5.8).

Figura 5.8: Diagramma orario di una sbarra in quiete in $\mathcal{R}$ e giacente sull'asse x.

Ogni suo punto ha come linea oraria la retta passante per esso e parallela all'asse dei tempi T, perciò il suo diagramma orario è la striscia di piano, diciamola $\mathcal{S}$, le cui rette che la delimitano sono le rette passanti per $\mathcal{A}_0$ e $\mathcal{B}_0$ e parallele all'asse T.
Consideriamo ora la sbarra AB disposta lungo l'asse x' del riferimento $\mathcal{R}'$ e perciò in moto uniforme con velocità costante v rispetto ad $\mathcal{R}$. Essa si muove ovviamente lungo l'asse $x \equiv x'$ ed il suo moto è rappresentato nel piano (x, T) da una striscia, diciamola $\mathcal{S}'$, tanto più inclinata sull'asse x quanto è più grande v (vedi Fig.5.9). Tale striscia, nel piano (x', T'), sarà dunque parallela all'asse T' dovendo rappresentare una sbarra in quiete nel riferimento $\mathcal{R}'$.

Figura 5.9: Sbarra in moto e deduzione geometrica della contrazione di Lorentz in $\mathcal{M}^2$.

Per l'osservatore del sistema inerziale $\mathcal{R}'$ che accompagna la sbarra nel suo movimento questa rappresentazione dovrà essere riferita al piano (x', T') i cui assi x' e T' sono inclinati di un angolo $\psi = \arctan\beta < 45°$ sugli assi x ed T. Per l'osservatore di $\mathcal{R}$ la lunghezza ℓ della sbarra (distanza fra gli estremi A e B considerati *in uno stesso istante valutato in* $\mathcal{R}$) è data da una sezione della striscia $\mathcal{S}'$ con una parallela all'asse x misurata con l'unità di misura $\mathcal{O}\mathcal{J} = 1$ adottata su tale asse. Possiamo considerare, per semplicità, la sezione $\mathcal{A}_0\mathcal{B}_0$, della striscia $\mathcal{S}'$ con l'asse x stesso, si ha perciò

$$\ell = \frac{\mathcal{A}_0\mathcal{B}_0}{\mathcal{O}\mathcal{J}}.$$

Per l'osservatore di $\mathcal{R}'$ la lunghezza ℓ' della sbarra sarà determinata, analogamente, misurando con l'unità $\mathcal{O}\mathcal{J}'$ (si tenga presente che $\mathcal{O} \equiv \mathcal{O}'$) la sezione $\mathcal{A}_0'\mathcal{B}_0'$ della striscia $\mathcal{S}'$ con l'asse x', si avrà

$$\ell' = \frac{\mathcal{A}_0'\mathcal{B}_0'}{\mathcal{O}\mathcal{J}'}.$$

Otterremo per il rapporto di tali lunghezze:

$$\frac{\ell}{\ell'} = \frac{\mathcal{A}_0\mathcal{B}_0}{\mathcal{A}_0'\mathcal{B}_0'}\frac{\mathcal{O}\mathcal{J}'}{\mathcal{O}\mathcal{J}}.$$

Conduciamo ora da $\mathcal{A}_0$ la parallela all'asse x', che intercetta sulla striscia $\mathcal{S}'$ il segmento $\mathcal{A}_0\mathcal{C}$ (vedi Fig.5.9), e dallo stesso punto $\mathcal{A}_0$ la perpendicolare $\mathcal{A}_0\mathcal{H}$ alla retta $\mathcal{B}_0\mathcal{B}_0'$. Si ottiene così il triangolo $\mathcal{A}_0\mathcal{H}\mathcal{C}$, rettangolo in $\mathcal{H}$, nel quale è facile riconoscere che $\mathcal{A}_0\mathcal{B}_0$ è bisettrice dell'angolo in $\mathcal{A}_0$ che misura 2ψ. Dalla considerazione dei due triangoli rettangoli $\mathcal{A}_0\mathcal{H}\mathcal{B}_0$ e $\mathcal{A}_0\mathcal{H}\mathcal{C}$, tenuto conto che è $\mathcal{A}_0\mathcal{C} = \mathcal{A}_0'\mathcal{B}_0'$, si ottiene

$$\mathcal{A}_0\mathcal{B}_0 \cos\psi = \mathcal{A}_0'\mathcal{B}_0' \cos 2\psi = \mathcal{A}_0\mathcal{H} \quad \text{(altezza della striscia)}.$$

Esprimiamo ora $\cos\psi$ e $\cos 2\psi$ in funzione di $\tan\psi = \beta$, si ha

$$\cos^2\psi = \frac{\cos^2\psi}{\cos^2\psi + \sin^2\psi} = \frac{1}{1+\tan^2\psi} = \frac{1}{1+\beta^2},$$

ed essendo $\psi < 45°$, è $\cos\psi > 0$, perciò avremo

$$\cos\psi = \frac{1}{\sqrt{1+\beta^2}}.$$

D'altra parte è

$$\cos 2\psi = \cos^2\psi - \sin^2\psi = \cos^2\psi(1-\tan^2\psi) = \frac{1-\beta^2}{1+\beta^2}.$$

Tenendo conto di tutto ciò avremo

$$\frac{\ell}{\ell'} = \frac{\mathcal{A}_0\mathcal{B}_0}{\mathcal{A}_0'\mathcal{B}_0'}\frac{\mathcal{O}\mathcal{J}'}{\mathcal{O}\mathcal{J}} = \frac{\cos 2\psi}{\cos\psi}\frac{\mathcal{O}\mathcal{J}'}{\mathcal{O}\mathcal{J}}.$$

Basta infine ricordare che $\mathcal{O}\mathcal{J} = 1$ e che, per la (5.44), $\mathcal{O}\mathcal{J}' = \sqrt{\dfrac{1+\beta^2}{1-\beta^2}}$ per ottenere, con facili calcoli, il risultato voluto

$$\ell = \ell'\sqrt{1-\beta^2}.$$

Osserviamo che con ragionamento poco diverso dal precedente si potrebbe, pure, ritrovare la dilatazione dei tempi, ma su ciò crediamo sia superfluo insistere ulteriormente.

Osservazione:

Alcuni autori per dimostrare, sui diagrammi di Minkowski, la contrazione delle lunghezze e la dilatazione dei tempi, utilizzano le due iperboli equilatere coniugate, di equazioni

$$x^2 - T^2 = 1, \qquad e \qquad x^2 - T^2 = -1,$$

i cui asintoti comuni sono le bisettrici dei quadranti del piano iperbolico (x, T), aventi per equazioni $x = \pm T$. Perciò, nella metrica di Minkowski, il cosiddetto **"cerchio unitario"**, cioè il luogo dei punti estremi di tutti i possibili vettori uscenti dall'origine ed aventi norma di modulo unitario, consiste di due iperboli euclidee sulla prima delle quali giacciono i punti estremi

dei vettori unitari, aventi la norma $||\mathbf{x}|| = 1$, e sull'altra giacciono i punti estremi dei vettori unitari immaginari di norma $||\mathbf{x}|| = -1$. Per seguire, con profitto, queste considerazioni è però necessario conoscere i fondamenti della geometria iperbolica piana, ma ciò esula dagli scopi di questo libro.[21]

5.6 Grandezze cinematiche assolute

Nello spaziotempo $\mathcal{M}^4$ pensiamo scelto un generico sistema di coordinate cartesiane (x^i), $(i = 0, 1, 2, 3)$, ciò che corrisponde alla scelta di un determinato riferimento galileiano $\mathcal{R} \equiv (Oxyz, t)$ $(x^\alpha \equiv (x, y, z)$, $(\alpha = 1, 2, 3)$ coordinate spaziali e $x^0 \equiv ct$ coordinata temporale, nel riferimento medesimo).
Come in meccanica classica, anche in meccanica relativistica il più semplice schema rappresentativo di un corpo naturale è il **punto materiale o particella**. Geometricamente esso non ha estensione; da un punto di vista cinematico la sua evoluzione spaziotemporale è rappresentata da una successione continua di ∞^1 eventi costituenti, com'è noto, la sua **linea di universo o linea oraria della particella**, orientata verso il futuro, e sempre del genere tempo perchè essendo il modulo della sua velocità $u < c$ si ha $\mathbf{ds}^2 \equiv c^2dt^2 - dl^2 = dt^2(c^2 - \dfrac{dl^2}{dl^2}) = dt^2(c^2 - u^2) > 0$. Questa linea, che ha anche il significato di **storia della particella** si può rappresentare con equazioni parametriche del tipo:

$$x^i = x^i(\lambda) \qquad (i = 0, 1, 2, 3) \tag{5.46}$$

con λ parametro arbitrario, non avente necessariamente un significato fisico, od anche del tipo più particolare

$$x^\alpha = x^\alpha(x^0) \qquad (\alpha = 1, 2, 3) \tag{5.47}$$

esprimenti direttamente le coordinate spaziali in funzione della variabile temporale x^0 (equazioni orarie). La condizione $\mathbf{ds}^2 > 0$ è imposta, come abbiamo visto poco fa, dalla impossibilità fisica, per una arbitraria particella materiale, di raggiungere la velocità della luce.
Le grandezze cinematiche fondamentali che ora dovremo definire sono: **il tempo proprio, la 4-velocità e la 4-accelerazione di una particella.**

5.6.1 Tempo proprio

Consideriamo un riferimento galileiano $\mathcal{R} \equiv Oxyz$ e sia t il **tempo coordinato o tempo relativo** segnato da un orologio posto in O. Sia poi P una particella mobile rispetto ad $\mathcal{R}$ che consideriamo coincidente con l'origine O' di un altro sistema di riferimento $\mathcal{R}' \equiv O'x'y'z'$, che si muove insieme alla particella ed avente gli assi equiorientati con quelli del sistema $\mathcal{R}$. Sia t' il tempo segnato da un orologio solidale con l'origine O' di $\mathcal{R}'$ e quindi con la particella P (vedi Fig.5.10).

Durante un intervallo infinitesimo di tempo dt, misurato dall'osservatore di $\mathcal{R}$, il riferimento $\mathcal{R}'$ può essere considerato inerziale rispetto ad $\mathcal{R}$. Nello stesso intervallo di tempo, P avrà percorso lo spazio dl rispetto ad $\mathcal{R}$ e $dl' = 0$ rispetto ad $\mathcal{R}'$, essendo in quiete in $\mathcal{R}'$. Per l'invarianza del divario spazio-temporale avremo

$$\mathbf{ds}'^2 = c^2dt'^2 - dl'^2 = c^2dt^2 - dl^2 = \mathbf{ds}^2,$$

[21] Per i primi elementi della geometria iperbolica nel piano di Minkowski si consiglia lo studente di consultare: N. V. Efimov e E. R. Rozendorn, *Linear Algebra and multidimensional Geometry*, Mir Publishers - Moscow, 1975, pagg. 278-287. Un buon libro dove si studia, in maniera elementare, la contrazione delle lunghezze e la dilatazione dei tempi con l'uso delle due iperboli coniugate $x^2 - T^2 = 1$, e $x^2 - T^2 = -1$, è: Piero Caldirola, *Istituzioni alla Fisica teorica*, pag. 204-207, Casa Editrice Viscontea - Milano, 1968.

Figura 5.10: Riferimento galileiano di moto incipiente e tempo proprio.

ed essendo $dl' = 0$, si ottiene

$$\mathbf{ds}'^2 = c^2 dt'^2 = c^2 dt^2 - dl^2 = \mathbf{ds}^2,$$

da cui

$$dt'^2 = dt^2 - \frac{dl^2}{c^2} = dt^2 \left(1 - \frac{1}{c^2}\frac{dl^2}{dt^2}\right) = \frac{1}{c^2}\mathbf{ds}^2 > 0. \tag{5.48}$$

La quantità dl/dt non è altro che la velocità $u(t)$ della particella rispetto ad $\mathcal{R}$, quindi, posto $\beta(t) = u(t)/c$, dalla (5.48) si ha:

$$dt' = dt\sqrt{1 - \beta^2(t)} = \frac{dt}{\gamma(t)}; \tag{5.49}$$

dalla stessa (5.48) si ha anche[22]:

$$dt' = \frac{1}{c}\sqrt{\mathbf{ds}^2} = \frac{ds}{c}. \tag{5.50}$$

La quantità dt' data dalle (5.49) e (5.50) è universalmente indicata con $d\tau$

$$d\tau \stackrel{\text{def}}{=} \frac{ds}{c} = dt\sqrt{1 - \beta^2(t)} = \frac{dt}{\gamma(t)}. \tag{5.51}$$

e viene chiamata **intervallo elementare di tempo proprio** relativo alla particella considerata. L'intervallo elementare di tempo proprio è *l'intervallo di tempo misurato da un orologio* **comobile (in inglese: comoving)** *con la particella, cioè mobile solidalmente con la particella, nel particolare riferimento galileiano nel quale essa si trova in quiete*; il corrispondente **intervallo di tempo coordinato (o relativo)** dt, valutato in qualsiasi altro riferimento galileiano, sarà sempre maggiore di $d\tau$ visto che $\gamma > 1$. Si osservi che il riferimento $\mathcal{R}' \equiv O'x'y'z'$, pur muovendosi di moto traslatorio rettilineo uniforme nell'intervallo infinitesimo dt, è chiaramente non inerziale durante un intervallo finito di tempo, ma possiamo considerarlo inerziale istante per istante, pur rappresentando nei diversi istanti riferimenti inerziali diversi: esso viene chiamato **riferimento di quiete istantanea o riferimento di moto incipiente**.
La definizione di intervallo elementare di tempo proprio poteva essere data a partire dalla linea di universo $\mathcal{L}$ della particella in $\mathcal{M}^4$, visto che $\mathbf{ds}^2 = c^2dt^2 - dl^2$ è il divario spazio-temporale tra due eventi infinitamente prossimi di $\mathcal{L}$.
Dalla definizione di $d\tau$ è possibile passare, mediante integrazione lungo $\mathcal{L}$ alla variabile temporale τ, che verrà chiamata **tempo proprio** della particella:

$$\tau = \int d\tau. \tag{5.52}$$

[22]Si noti che, poich é $\mathbf{ds}^2 > 0$, il radicando è positivo, perciò $\sqrt{\mathbf{ds}^2} = \sqrt{|\mathbf{ds}^2|} = ds$.

Se l'integrazione avviene tra due punti $\mathcal{P}_1$ e $\mathcal{P}_2$, avremo:

$$\tau_2 - \tau_1 = \int_{\mathcal{P}_1}^{\mathcal{P}_2} d\tau = \int_{t_1}^{t_2} \sqrt{1-\beta^2(t)}\, dt \tag{5.53}$$

Per il suo carattere assoluto e per il suo significato fisico, la variabile τ può essere utilmente assunta quale parametro per individuare i singoli eventi di $\mathcal{L}$. Infatti, a partire dal tempo proprio, definiremo ora, nello spaziotempo $\mathcal{M}^4$ di Minkowski, due grandezze cinematiche assolute associate alla particella: la quadrivelocità e la quadriaccelerazione.

5.6.2 Quadrivelocità

In analogia con la definizione di 3-velocità di una particella nello spazio fisico ordinario, chiamiamo **4-velocità** della stessa particella il 4-vettore $\mathbf{U}$ di $\mathcal{M}^4$ derivato del 4-vettore di posizione $\overrightarrow{\mathcal{OP}} = \mathbf{X} = x^i\mathbf{E}_i$ rispetto al tempo proprio:

$$\mathbf{U} \stackrel{\text{def}}{=} \frac{d\mathbf{X}}{d\tau} = U^i\mathbf{E}_i \Longleftrightarrow U^i \stackrel{\text{def}}{=} \frac{dx^i}{d\tau}, \tag{5.54}$$

che ha per **componenti controvarianti** U^i (con indici in alto), le derivate delle coordinate spazio-temporali della particella rispetto al suo tempo proprio.
Se si tiene conto della relazione (5.51) tra tempo proprio e tempo coordinato (o tempo relativo), si ha:

$$\begin{cases} U^0 &= \dfrac{dx^0}{dt}\dfrac{dt}{d\tau} = \gamma(u)c, \\ \\ U^\alpha &= \dfrac{dx^\alpha}{dt}\dfrac{dt}{d\tau} = \gamma(u)u^\alpha, \quad \alpha = 1,2,3 \end{cases} \tag{5.55}$$

dove le $u^\alpha = dx^\alpha/dt$ sono le componenti della 3-velocità ordinaria $\mathbf{u}$ della particella nel riferimento galileiano associato alle coordinate x, y, z.
Il vettore 4-velocità si scrive talvolta nel modo seguente:

$$U^i = \frac{dx^i}{d\tau} = \frac{dx^i}{dt}\frac{dt}{d\tau} = \gamma(u)\frac{dxi}{dt} = (\gamma(u)c, \gamma(u)\mathbf{u}) = \gamma(u)(c, \mathbf{u}) \tag{5.56}$$

dove nelle ultime due espressioni in parentesi si è separato, in virtù delle (5.55), la parte temporale da quella spaziale del vettore 4- velocità.
È facile riconoscere che $\mathbf{U}$ è un **vettore del genere tempo**, tangente alla linea d'universo $\mathcal{L}$ della particella, in quanto si constata facilmente che la sua norma è positiva, infatti:

$$\|\mathbf{U}\| = (U^0)^2 - \sum_{\alpha=1}^{3}(U^\alpha)^2 = (\gamma c)^2 - \sum_{\alpha=1}^{3}(\gamma u^\alpha)^2 = \gamma^2(c^2-u^2) = c^2 > 0. \tag{5.57}$$

Stante il legame (5.57), tre soltanto delle quattro componenti U^i della 4- velocità sono indipendenti.
Osservazione 1:
Talvolta è utile usare anziché le **componenti controvarianti** U^i di $\mathbf{U}$, le cosiddette **componenti covarianti** definite utilizzando il tensore metrico η_{ik} dello spazio di Minkowski:

$$U_i \stackrel{\text{def}}{=} \eta_{ik}U^k = \eta_{i0}U^0 + \eta_{i\alpha}U^\alpha = \eta_{i0}U^0 + \eta_{i1}U^1 + \eta_{i2}U^2 + \eta_{i3}U^3. \tag{5.58}$$

Se si tiene conto che la matrice associata al tensore metrico η_{ik} è una matrice diagonale i cui elementi con indici diversi sono tutti nulli, mentre gli elementi della diagonale principale sono

$$\eta_{00} = 1, \quad \eta_{11} = \eta_{22} = \eta_{33} = -1, \quad \eta_{ik} = 0 \ (per\ i \neq k)$$

avremo

$$U_0 = U^0 = c\gamma(u), \quad U_\alpha = -U^\alpha = -u^\alpha\gamma(u) \quad (\alpha = 1, 2, 3). \tag{5.59}$$

È ovvio che i concetti di *componenti controvarianti e di componenti covarianti* hanno senso, non solo per il 4-vettore velocità, ma per un qualsiasi 4-vettore **V** (*tensore del 1° ordine*) dello spazio di Minkowski, per cui possiamo affermare che hanno validità generale le seguenti relazioni:

$$V_0 = V^0, \qquad V_\alpha = -V^\alpha \quad (\alpha = 1, 2, 3). \tag{5.60}$$

Osservazione 2:
Avvertiamo il lettore che circa la metà degli autori, per es. Møller, Rindler, Fock, Stephani, Cattaneo, ecc., definiscono la 4-velocità come abbiamo fatto noi, come derivata delle coordinate spazio-temporali rispetto al tempo proprio, cioè:

$$U^i \stackrel{\text{def}}{=} \frac{dx^i}{d\tau}. \tag{5.61}$$

L'altra metà degli autori, per es. Landau-Lifschitz, Adler-Bazin- Schiffer, Barone, ecc., definiscono la 4-velocità in maniera leggermente diversa:

$$u^i \stackrel{\text{def}}{=} \frac{dx^i}{ds}, \tag{5.62}$$

facendo la derivata delle coordinate spazio-temporali rispetto ad s e indicando le sue componenti con la lettera minuscola u^i.
Poiché $ds = cd\tau$, la relazione tra la definizione da noi adottata e quest'ultima è data dalla formula

$$U^i = cu^i, \tag{5.63}$$

cioè le due definizioni differiscono per una costante che è la velocità c della luce nel vuoto. D'altronde le due definizioni formalmente coincidono se si conviene, come talvolta si fa per semplificare i calcoli, di usare le cosiddette *unità relativistiche* che assumono come unitaria la velocità della luce nel vuoto: $c = 1$.
Naturalmente se si adotta la notazione u^i per la 4-velocità, indicare la 3-velocità con la stessa lettera minuscola $\mathbf{u} \equiv (u^\alpha)$ potrebbe generare confusione, ma dal contesto si capisce sempre di che cosa stiamo parlando. Ovviamente la 3-velocità di una particella si può indicare con una lettera diversa da **u**, per esempio con $\mathbf{v} \equiv (v^\alpha)$ come fanno molti autori; in tal caso, per non confondere la velocità **v** della particella con la velocità relativa di traslazione di due riferimenti galileiani, in configurazione x-standard, quest'ultima viene spesso indicata con una lettera maiuscola **V**.
Osserviamo infine che la 4-velocità U^i, come l'abbiamo definita noi, ha le dimensioni di una velocità ordinaria, cioè spazio diviso tempo, mentre la u^i è adimensionale. Inoltre mentre U^i ha norma uguale al quadrato della velocità della luce, cioè $\|U^i\| = c^2$, il suo modulo è per definizione $|\mathbf{U}| = \sqrt{|\|\mathbf{U}\||} = c$. Osserviamo ancora che il 4-vettore **u** non è altro che il versore di **U**:

$$\mathbf{u} = vers\mathbf{U} = \frac{\mathbf{U}}{|\mathbf{U}|},$$

infine è facile verificare che **u** ha norma e modulo unitari: $\|\mathbf{u}\| = |u^i| = 1$.

5.6.3 Quadriaccelerazione

Si chiama **4-accelerazione** di una particella il quadrivettore **A** derivato della quadrivelocità rispetto al suo tempo proprio:

$$\mathbf{A} \stackrel{\text{def}}{=} \frac{d\mathbf{U}}{d\tau} = A^i\mathbf{E_i} \Longleftrightarrow A^i \stackrel{\text{def}}{=} \frac{dU^i}{d\tau}. \tag{5.64}$$

Tenendo conto delle (5.55) si trova:

$$\begin{cases} A^0 &= \dfrac{dU^0}{dt}\dfrac{dt}{d\tau} = \dfrac{d(\gamma c)}{dt}\gamma = \dfrac{d\gamma}{dt}c\gamma \\ \\ A^\alpha &= \dfrac{dU^\alpha}{dt}\dfrac{dt}{d\tau} = \dfrac{d(\gamma u^\alpha)}{dt}\gamma = \left(\dfrac{d\gamma}{dt}u^\alpha + \gamma\dfrac{du^\alpha}{dt}\right)\gamma; \quad \alpha = 1,2,3 \end{cases}$$

inoltre, poiché risulta

$$\frac{d\gamma}{dt} = \frac{d}{dt}\frac{1}{\sqrt{1-\dfrac{\mathbf{u}\cdot\mathbf{u}}{c^2}}} = -\frac{1}{2\left(\sqrt{1-\dfrac{u^2}{c^2}}\right)^3}\left(-\frac{1}{c^2}2\mathbf{u}\cdot\mathbf{a}\right) = \gamma^3\frac{\mathbf{u}\cdot\mathbf{a}}{c^2}, \tag{5.65}$$

si ha, in definitiva:

$$\begin{cases} A^0 &= \gamma^4\dfrac{\mathbf{u}\cdot\mathbf{a}}{c} \\ \\ A^\alpha &= \gamma^4\dfrac{\mathbf{u}\cdot\mathbf{a}}{c^2}u^\alpha + \gamma^2 a^\alpha, \quad \alpha = 1,2,3 \end{cases} \tag{5.66}$$

essendo le $a^\alpha = du^\alpha/dt$ le componenti della 3-accelerazione ordinaria $\mathbf{a}$ della particella.
A differenza della quadrivelocità, la quadriaccelerazione è un vettore del genere spazio e perciò ortogonale alla linea d'universo $\mathcal{L}$, la cui norma è negativa:

$$\|\mathbf{A}\| < 0. \tag{5.67}$$

Infatti, derivando la (5.57) rispetto a τ si ha:

$$\frac{d}{d\tau}(\mathbf{U}\cdot\mathbf{U}) = \frac{d\mathbf{U}}{d\tau}\cdot\mathbf{U} + \mathbf{U}\cdot\frac{d\mathbf{U}}{d\tau} = 2\ \mathbf{U}\cdot\mathbf{A} = 0,$$

e dall'ultima uguaglianza, essendo $\mathbf{U}$ un vettore del genere tempo, ne segue, per il Teorema 5.1, che $\mathbf{A}$ è del genere spazio. Inoltre, l'uguaglianza

$$\mathbf{U}\cdot\mathbf{A} = \eta_{ik}U^iA^k = 0, \tag{5.68}$$

esprime l'ortogonalità dei due 4-vettori di componenti U^i, A^k. Il legame (5.68) mostra che anche le componenti della quadriaccelerazione non sono tra loro indipendenti.
Osserviamo ancora che le singole componenti $A^i = (A^0, A^\alpha)$ non hanno un'immediata interpretazione fisica. Se però si assume in $\mathcal{M}^4$ un sistema cartesiano tale che l'asse temporale x^0 abbia la direzione della tangente alla linea di universo $\mathcal{L}$ in un suo punto, essendo in tale punto la 4-velocità $\mathbf{U}$ parallela all'asse x^0, avrà diversa da zero soltanto la componente temporale U^0 mentre le componenti spaziali saranno tutte nulle, cioè $U^\alpha = 0$. Ma $U^\alpha = \gamma(u)u^\alpha$ ed essendo $\gamma(u) \neq 0$ ne segue che $u^\alpha = 0$, $\alpha = 1,2,3$, ciò implica l'annullarsi della 3-velocità della particella, $\mathbf{u} = 0$, e di conseguenza sarà anche $\gamma(u) = 1$. Pertanto, in virtù delle (5.66) avremo $A^0 = 0, A^\alpha = a^\alpha$, cioè la 4-accelerazione della particella coincide sostanzialmente con la sua accelerazione ordinaria.
Il riferimento galileiano corrispondente a tale scelta dell'asse temporale nello spazio di Minkowski $\mathcal{M}^4$, verrà detto *riferimento galileiano di quiete istantanea o di moto incipiente*. Nel riferimento galileiano di moto incipiente, e soltanto in esso, la 4- accelerazione si identifica con l'accelerazione ordinaria. L'adozione di un riferimento galileiano di quiete istantanea nel quale si ha $U^0 = c, U^\alpha = 0, A^0 = 0, A^\alpha = a^\alpha$ permette di confermare immediatamente, in termini di componenti, l'annullarsi dell'invariante $\mathbf{U}\cdot\mathbf{A} = \eta_{ik}U^iA^k$, nonché il valore negativo della norma

di **A**.
L'eventuale annullarsi della 4-accelerazione **A** *per tutta la durata del moto caratterizza la costanza della 4-velocità* **U** *non soltanto in norma ma anche in direzione, e cioè* **il carattere rettilineo della linea di universo in** $\mathcal{M}^4$. *Quest'ultima proprietà si traduce a sua volta, in termini relativi, nella dipendenza lineare delle* x^α *da* x^0 *e cioè* **nel carattere rettilineo ed uniforme del moto** *rispetto ad ogni riferimento galileiano. Come in cinematica classica,* **così anche in cinematica relativistica, la rettilinearità e l'uniformità del moto costituisce dunque una proprietà invariantiva nel passaggio da un riferimento galileiano ad un altro**, *pur cambiando in senso relativistico, il valore della velocità.*

5.7 Cenno al Paradosso dei Gemelli o Paradosso degli orologi

In un libro sulla Relatività non si può non citare il più popolare dei paradossi: quello dei Gemelli. Ci piace perciò chiudere questo capitolo dando almeno un cenno di questo famosissimo paradosso noto anche, con più precisione, come **Paradosso degli orologi**. Per semplicità, riferiamoci dapprima a due orologi identici O, O' in quiete nello stesso punto dell'asse x di un dato sistema inerziale $\mathcal{R}$ e sincronizzati; per orologio intendiamo un dispositivo qualunque la cui evoluzione ci consente di misurare il tempo t che figura nella trasformazione di Lorentz.
Diamo ad uno di essi, per esempio ad O', un'accelerazione che ci consente di portarlo dalla quiete ad una certa velocità v e poi lo lasciamo muoversi (solidalmente ad un riferimento inerziale $\mathcal{R}'$), rispetto all'orologio O (solidale con $\mathcal{R}$) di moto rettilineo uniforme con la velocità v raggiunta. Dopo un certo tempo, con un'opportuna decelerazione, lo riportiamo alla quiete, nello stesso punto di partenza, nel sistema di riferimento $\mathcal{R}$ dal quale siamo partiti e nel quale è rimasto l'orologio gemello O senza muoversi. Supponiamo trascurabili le durate dei periodi di accelerazione e decelerazione rispetto ai periodi di moto uniforme.
Nella fisica prerelativistica si ammette che, allorquando l'orologio O' è ritornato a riposo nel suo punto di partenza, esso segna lo stesso tempo di quello O che è rimasto fisso.
Nella Teoria della Relatività, al contrario, i periodi di accelerazione e decelerazione, producono sulla marcia dell'orologio O', un effetto *che persiste durante i periodi di moto uniforme.*
L'orologio O' che si è mosso **ritarda costantemente rispetto all'orologio fisso** O; quando i due orologi sono di nuovo in quiete relativa (nello stesso luogo), **quello che si è mosso rispetto al sistema di riferimento** $\mathcal{R}$**, cioè** O'**, è in ritardo sull'orologio fisso** O**.** Se indichiamo con Δt_O l'intervallo di tempo misurato in $\mathcal{R}$ fra la partenza e l'arrivo dell'orologio O', e con $\Delta t'_O$ l'analogo intervallo di tempo valutato in $\mathcal{R}'$ - per la (4.6) del paragrafo 4.1.7 in cui i ruoli di $\mathcal{R}$ ed $\mathcal{R}'$ sono fra di loro scambiati - è valutato dalla relazione: $\Delta t_{O'} = (1/\gamma(v))\Delta t_O = \sqrt{1-\beta^2}\Delta t_O$. Il divario fra i due orologi sarà dunque dato da:

$$\Delta t_O - \Delta t_{O'} = \left(1 - \sqrt{1-\beta^2}\right)\Delta t_O, \qquad \left(\beta = \frac{v}{c}\right). \tag{5.69}$$

Per una velocità v abbastanza piccola avremo:[23]

$$\Delta t_O - \Delta t_{O'} \approx \frac{\beta^2}{2}\Delta t_O = \frac{vl}{c^2}, \tag{5.70}$$

essendo $2l$ la distanza percorsa, misurata nel sistema $\mathcal{R}$ in cui O è rimasto in quiete.
La conseguenza precedente è molto interessante. In effetti, finora, abbiamo confrontato sempre e soltanto le indicazioni di un orologio O di $\mathcal{R}$ con le indicazioni di tutti gli orologi di $\mathcal{R}'$ che sfilavano davanti ad esso, essendo i due sistemi in moto reciproco ed in configurazione x-standard.
Nel caso attuale le cose vanno diversamente. Due orologi identici sono confrontati, in epoche

[23]Si è fatto uso del seguente sviluppo in serie $\sqrt{1-\beta^2} = 1 - \frac{1}{2}\beta^2 - \frac{1}{2\cdot 4}\beta^4 - \frac{1\cdot 3}{2\cdot 4\cdot 6}\beta^6 - .... \approx 1 - \frac{1}{2}\beta^2$.

differenti, in uno stesso sistema di riferimento. Alle epoche di confronto, essi sono entrambi a riposo. Noi verifichiamo che il loro comportamento è stato differente (non si ha reciprocità come nelle considerazioni del paragrafo 4.1.7). **Due eventi che si verificano in uno stesso punto** (*confronto degli orologi prima della partenza di uno di essi e confronto al loro incontro*) **sono separati da durate diverse.** Sembra che siamo giunti ad un risultato in contraddizione col Principio di Relatività; in effetti sembra che l'esperienza permetta di decidere quale dei due orologi è stato in movimento (*è qui che si ravvisa il paradosso*).
In realtà, l'esperienza permette soltanto di sapere quale dei due orologi è sempre stato a riposo in uno stesso riferimento inerziale (riferimento proprio) e non chi dei due sistemi di riferimento fosse in moto e l'altro fermo.
A questo punto siamo tentati di applicare le conseguenze precedentemente raggiunte a degli esseri viventi. Questi esseri possono costituire oggetto di studio alla stessa stregua di qualsiasi altro fenomeno. Conviene tuttavia attirare l'attenzione sul fatto che le formule usate si rapportano unicamente a degli orologi la cui marcia è uniforme in conformità al tempo della trasformazione di Lorentz. L'estensione ai fenomeni biologici o psichici richiede dunque preliminarmente la verifica dell'uniformità di questi fenomeni.
Le conclusioni precedenti, incontestabili nell'ambito della Teoria della Relatività, conducono, con le riserve più sopra formulate, a delle previsioni sorprendenti consistenti nel famosissimo **Paradosso dei gemelli**: *Immaginiamo due gemelli, Al e Bert, dei quali uno, per esempio Bert, supposto astronauta, effettua un viaggio nelle condizioni descritte all'inizio del paragrafo. Al ritorno del suo viaggio, Bert è più giovane di Al.*
Il paradosso sorge perché le situazioni dei due gemelli si ritengono simmetriche ed interscambiabili, ipotesi questa del tutto errata. Infatti, il primo dei due gemelli se ne sta in quiete nel suo riferimento inerziale, mentre il secondo gemello viaggia in un riferimento che per ritornare, dopo un lungo viaggio, al punto di partenza non può essere, per tutta la durata del viaggio, in un sistema inerziale, ma deve necessariamente subire delle accelerazioni e delle decelerazioni. La situazione dei due gemelli dunque non è simmetrica. Essendo necessario l'uso, oltre che di sistemi inerziali, anche di sistemi accelerati, una spiegazione corretta e rigorosa del paradosso non può ottenersi semplicemente nell'ambito della sola Relatività Ristretta, ma deve necessariamente coinvolgere la Relatività Generale, perciò si tratta di una spiegazione piuttosto complessa che esula dai limiti della nostra trattazione. Tuttavia possiamo dare una semplice spiegazione, necessariamente incompleta ma efficace, utilizzando un diagramma spazio-temporale nell'ambito della RR.[24]
Sia (x, ct) un diagramma spazio-temporale 2-dimensionale dove due eventi A e B possono essere collegati con diverse linee di universo. Non ci stupisce che la distanza spaziale fra A e B dipenda dal percorso considerato. È altrettanto vero che anche il tempo misurato dagli orologi in movimento dipende dal percorso considerato. Il tempo registrato da un orologio, trasportato da un corpo puntiforme che descrive una linea di universo è il suo *il tempo proprio*.

Si è visto che la relazione fra il tempo proprio τ ed il tempo coordinato t misurato in un dato riferimento inerziale $\mathcal{R}$, è $d\tau = dt\sqrt{1 - \frac{u^2}{c^2}}$, essendo u la velocità della particella rispetto ad $\mathcal{R}$. Per il moto in una sola dimensione spaziale x, se indichiamo con dx lo spostamento elementare della particella, la formula precedente si può scrivere nella forma $d\tau = \sqrt{dt^2 - \frac{dx^2}{c^2}}$. Perciò il tempo proprio che separa i due eventi A e B si ottiene integrando la precedente espressione lungo la linea di universo $\mathcal{L}_{AB}$ che unisce A e B:

$$\Delta\tau = \int_{\mathcal{L}_{AB}} \sqrt{dt^2 - \frac{dx^2}{c^2}}. \tag{5.71}$$

[24]Un libro esclusivamente dedicato al Paradosso dei Gemelli è quello di: Leslie Marder, *Time and Space-Traveller*, University of Pensylvania Press, Inc.; New Edition, 1972.

Figura 5.11: Diagramma orario del paradosso dei gemelli.

Si consideri ora il caso particolare, (vedi Fig.5.11), in cui una linea di universo congiungente A con B rappresenti un orologio in quiete sull'asse x in una determinata posizione: ad esso corrisponde una linea retta parallela all'asse ct. Si consideri ora un secondo percorso lungo il quale un secondo orologio, identico al primo, viene portato da A a B. Tale orologio si allontana dal primo e poi , dopo un certo percorso, ritorna ad esso, in modo che i due orologi coincidano in A e in B. Il tempo proprio trascorso lungo la prima linea di universo è:

$$\Delta\tau = \int_A^B \sqrt{dt^2 - \frac{dx^2}{c^2}} = \int_A^B \sqrt{dt^2 - 0} = \int_A^B dt = t_B - t_A, \tag{5.72}$$

dato che $dx = 0$ lungo questo cammino ed il tempo proprio coincide con l'intervallo di tempo coordinato, $t_B - t_A$, registrato dall'orologio in riposo nel punto di ascissa x.
Lungo la seconda linea di universo, tuttavia il tempo proprio trascorso è:

$$\Delta\tau' = \int_A^B \sqrt{dt^2 - \frac{dx^2}{c^2}}, \tag{5.73}$$

ed essendo sempre $dx^2 > 0$ si ha che:

$$\Delta\tau' < \Delta\tau. \tag{5.74}$$

Quando gli orologi si incontreranno di nuovo segneranno tempi diversi: l'orologio che ha viaggiato sarà indietro, registrando una differenza di tempi più piccola, rispetto a quello rimasto in quiete nell'originario riferimento inerziale.
È appena necessario notare che il riferimento (x, ct) è un riferimento inerziale, perciò il moto dell'orologio viaggiante è rappresentato in questo riferimento da una linea di universo curva, il che significa che questo orologio, per poter tornare nel luogo di partenza e fermarsi, subisce sicuramente delle accelerazioni e delle decelerazioni, cioè il suo sistema di riferimento non può essere, per tutto il viaggio, inerziale. La Teoria della Relatività Ristretta può prevedere anche il comportamento di oggetti accelerati, a patto che, nella formulazione delle leggi fisiche, si consideri sempre il punto di vista degli osservatori inerziali, ed è ciò che abbiamo fatto costantemente nei nostri ragionamenti.
Concludendo, ribadiamo ancora una volta, che la situazione dei due orologi (o gemelli) *non è per niente simmetrica*: **uno si trova sempre in un riferimento inerziale, l'altro no!**

Capitolo 6

MECCANICA RELATIVISTICA DI UNA PARTICELLA

La Meccanica è il paradiso delle Matematiche.
(Leonardo da Vinci)

Il progresso della scienza consiste nell'osservare le interazioni e nel mostrare con paziente ingegnosità che gli eventi di questo mondo sottoposto a continue mutazioni non sono altro che casi particolari di poche e generali relazioni dette leggi. Scoprire il generale nel particolare e il permanente nel transitorio è lo scopo del pensiero scientifico.
(Alfred North Whitehead, *Introduzione alla Matematica)*

6.1 Introduzione

In questa introduzione alla Meccanica relativistica seguiremo da vicino l'esposizione di Rindler[1]. Finora i nostri studi sono stati essenzialmente una elaborazione dei due postulati fondamentali di Einstein, senza l'aggiunta di ulteriori ipotesi. Abbiamo così scoperto alcuni fenomeni ottici riguardanti la propagazione della luce, come l'aberrazione annua delle stelle, l'esperimento di Michelson-Morley e l'effetto Doppler sia longitudinale che trasversale; infine abbiamo scoperto la struttura dello spaziotempo implicita negli assiomi einsteiniani.
Veniamo ora al prossimo punto del programma della Relatività Ristretta, cioè all'esame delle leggi della meccanica newtoniana, invarianti per trasformazioni di Galileo, ed alla loro modifica in modo da renderle invarianti rispetto al gruppo di Lorentz. A questo scopo *sono necessarie nuove ipotesi* poichè **non esiste alcun procedimento logico od empirico per dimostrare una legge fondamentale della fisica**. Possiamo costruire soltanto un modello matematico della realtà - una invezione umana - che deve essere consistente con la nostra esperienza limitata, ma deve anche permetterci di fare nuove previsioni che potrebbero, in via di principio essere *falsificate* da futuri esperimenti (secondo il pensiero del filosofo austriaco Karl Raimund Popper). **Tali previsioni ed i tentativi per falsificarle (usualmente chiamati tentativi per verificare le previsioni) sono in se stessi il maggior stimolo per il progresso della fisica**. Le leggi che non possono essere falsificate sono, secondo Popper, delle **tautologie**, *cioè delle affermazioni vere per qualsiasi valore di verità degli elementi che le compongono e perciò prive di valore informativo: le tautologie logiche ragionano circolarmente attorno agli argomenti e alle affermazioni*.

[1]W. Rindler, *Introduction to Special Relativity*, Oxford University Press, 1991, reprinted in 1996, pagg. 69-73.

Concludiamo questo preambolo rilevando alcuni disaccordi tra Meccanica Newtoniana e Principi della Relatività Ristretta:

1) In meccanica classica si può accelerare una particella ad una qualsiasi velocità applicando con continuità una forza costante, invece nell'ambito relativistico la velocità di una particella non può mai superare la velocità della luce nel vuoto;

2) In meccanica classica vale il **Principio di azione e reazione (3° Principio della dinamica)** in cui due corpi, posti ad una distanza qualunque l'uno dall'altro, si scambiano **simultaneamente** forze di uguale intensità e di verso opposto: ciò è generalmente privo di senso in meccanica relativistica perché il concetto di simultaneità in luoghi diversi non è assoluto ma relativo al sistema di riferimento, a meno che i due corpi vengano a contatto (come nei processi d'urto): soltanto in questo caso è valido il 3° Principio perché il concetto di simultaneità diventa assoluto se i due eventi avvengono nello stesso luogo che si identifica con il punto di contatto dei due corpi.

3) Le leggi della meccanica classica sono invarianti per trasformazioni di Galileo e perciò non possono essere invarianti per trasformazioni di Lorentz, il che significa che esse ubbidiscono al principio galileiano di relatività e non a quello einsteiniano.

L'evidenza sperimentale, soprattutto nel campo delle alte velocità, è ampiamente in accordo con la Teoria della Relatività Ristretta e non con la Meccanica Classica. Tuttavia nella costruzione della nuova meccanica molti dei concetti fondamentali saranno introdotti, fin dove è possibile, in analogia con i concetti classici. **Deve essere però ben chiaro che la Meccanica Classica non essendo Lorentz-invariante deve essere necessariamente modificata**. La nuova meccanica è nota come **Meccanica relativistica.** Questo nome non è del tutto appropriato perché, come sappiamo, anche la meccanica newtoniana è relativistica in quanto assume fra i suoi principi il **Principio di Relatività di Galileo**, essa però è relativistica sotto l'azione *"dell'errato" gruppo di trasformazioni di Galileo.* La meccanica newtoniana è utilizzata in modo eccellente dall'Astronomia e dalla Meccanica Celeste, per esempio nel predire con estrema precisione le eclissi ed i moti planetari in generale. Essa è stata usata come teoria di base per operazioni incredibilmente delicate, per inviare sonde spaziali sulla Luna e su alcuni pianeti, ed è di applicazione costante e quotidiana in innumerevoli applicazioni terrestri. Perciò essa non può essere interamente sbagliata. Prima dell'avvento della Relatività Generale, cioè fino al ventesimo secolo, è noto che la teoria newtoniana ha avuto un enorme successo ed ha fallito in un solo caso: non è riuscita a spiegare il famoso avanzamento del perielio del pianeta Mercurio di circa $43''$ per secolo, scoperto dagli astronomi. La spiegazione di quest'ultimo è stata invece un vero e grande successo della Relatività Generale.
Dopo l'entrata in funzione dei moderni acceleratori di particelle, tuttavia, sono state scoperte notevoli discrepanze con la teoria newtoniana, mentre si è constatato che la meccanica relativistica fornisce descrizioni corrette. Dobbiamo anche avvertire che la meccanica classica ha subito non una ma *due correzioni* importanti: una dovuta alla relatività e l'altra alla teoria quantistica. Noi, in questo libro, ci occuperemo solo della prima.
La meccanica classica, in pratica, fornisce con ottima approssimazione gli stessi risultati della meccanica relativistica in un vasto dominio di applicazioni e cioè quando si ha a che fare con moti lenti rispetto alla velocità della luce. **La vecchia meccanica però è in larga misura in errore quando il fattore di Lorentz γ è molto maggiore dell'unità.** In esperienze di laboratorio riguardanti collisioni di particelle è abbastanza usuale che $\gamma \approx 10^4$ e fattori γ dell'ordine di 10^{11} sono stati calcolati per urti fra protoni appartenenti ai raggi cosmici nell'alta atmosfera. Applicata a tali situazioni, la meccanica newtoniana, non è leggermente errata ma *totalmente errata.* Ovviamente nel dominio dei moti lenti la meccanica newtoniana sarà sempre utilizzata sia per ragioni di convenienza concettuale, essendo più semplice, sia per ragioni pratiche in

quanto fornisce risultati sostanzialmente coincidenti con quelli della nuova meccanica ed in perfetto accordo con l'evidenza sperimentale. **Come costruzione logica la meccanica newtoniana rimarrà perfetta ed inviolabile quanto la geometria euclidea, invece come modello della Natura essa è meno aderente alla realtà della nuova teoria einsteiniana**.

6.2 Grandezze dinamiche assolute e relative

In questo paragafo illustreremo brevemente il concetto fondamentale di **massa propria o massa a riposo di una particella**, sostanzialmente noto dalla meccanica newtoniana, corredandolo di alcune utili osservazioni, ed introdurremo i nuovi concetti di **massa relativistica, di 4-impulso, 4-forza**, e loro analoghi 3- dimensionali, il concetto di **energia cinetica**, quello di **energia totale ed energia di riposo** di una particella. Discuteremo, inoltre, i teoremi dinamici relativistici che legano fra di loro queste grandezze e le loro leggi di conservazione, sia nello spazio di Minkowski $\mathcal{M}^4$, 4-dimensionale, sia nello spazio fisico ordinario 3-dimensionale E^3. Ci soffermeremo, ovviamente, su quello che è considerato come uno dei risultati più importanti, se non il più importante, della Relatività Ristretta e cioè la notevole relazione trovata da Einstein $E = \dfrac{m_0 c^2}{\sqrt{1 - u^2/c^2}}$ che fornisce l'energia totale di una particella in funzione della sua massa propria m_0 e della sua velocità u e che contiene, come caso particolare, la sconvolgente, famosissima relazione

$$E_0 = m_0 c^2$$

esprimente **l'energia posseduta dalla particella quando si trova in quiete**.
Concluderemo il capitolo con qualche applicazione della teoria svolta a problemi d'urto fra due particelle ed allo studio del moto di una particella carica in un campo magnetico costante.

6.2.1 Massa propria di una particella

Analogamente a quello che si fa nella meccanica newtoniana, anche in meccanica relativistica la costituzione fisica di un punto materiale si suole riassumere in uno **scalare invariante** m_0, detto **massa inerziale o massa propria o massa a riposo**, generalmente positivo, tranne che per alcune particelle relativistiche, *come i fotoni* per le quali la massa a riposo è nulla: $m_0 = 0$ o trascurabile[2].
La massa a riposo è una grandezza che rimane sempre la stessa sia che la particella si trova in quiete in un dato sistema galileiano, sia che essa sia in movimento rispetto ad esso, inoltre essa rimane invariata nel passare da un sistema inerziale ad un altro. Se si considera la massa propria come una **costante caratteristica della particella** non suscettibile di variare durante la sua storia, si dice che **la particella è priva di struttura interna** $(m_0 = cost.)$. La massa propria dovrà invece pensarsi suscettibile di variare, se il punto materiale viene assunto quale schema macroscopico di un **sistema dotato di una struttura interna**, più o meno complessa. L'aggiunta del concetto di massa propria ai concetti cinematici di tempo proprio, di 4-velocità e 4-accelerazione, introdotti nel Cap.4, permette ora di definire le grandezze dinamiche fondamentali nello spaziotempo di Minkowski.
Osservazione: In tutte le figure di questo capitolo in cui si confrontano grandezze classiche con grandezze relativistiche si è scelto di porre nei grafici $m_0 = 2$ e $c = 1$.

[2] I **neutrini** ipotizzati da Wolfgang Pauli nel 1931, per risolvere un'apparente anomalia nel decadimento beta dei nuclei, e scoperti sperimentalmente, molto più tardi, nel 1956 da Cowan e Reines, erano fino a poco tempo fa considerati anch'essi particelle di massa nulla. Vi sono oggi alcune evidenze sperimentali a favore di una loro massa piccolissima ma diversa da zero. La prima indicazione inequivocabile in questo senso si è avuta dall'esperimento fatto in Giappone e chiamato *SuperKamiokande*, si veda Y. Fukuda e altri, in Phys. Rev. Lett., 81, 1988, p. 1562. Esistono anche altre particelle di massa nulla, per es. i cosiddetti **gluoni**.

6.2.2 Quadri-quantità di moto, massa relativistica e tri-quantità di moto

Introdurremo ora, in maniera piuttosto formale, una grandezza fondamentale dello spaziotempo di Minkowski, la 4-quantità di moto, che contiene in sé il concetto più familiare di 3-quantità di moto di una particella. *Questo approccio formale per introdurre concetti di base della meccanica relativistica è, rispetto ad alcuni altri approcci, che utilizzano argomenti di natura più specificatamente fisica, certamente il più rapido ed economico ed ha l'enorme pregio che*: **ogni proposizione fisica che sia espressa in termini geometrici intrinseci di $\mathcal{M}^4$, consista cioè in un'equazione tensoriale fra 4-tensori, soddisfa automaticamente al principio einsteiniano di relatività, cioè è Lorentz-invariante**[3]. In effetti **lo spaziotempo di Minkowski rappresenta un potentissimo mezzo euristico per la formulazione delle leggi relativistiche**: la dinamica einsteiniana, come vedremo, ne costituirà un primo importantissimo esempio.
In analogia con il concetto di quantità di moto della meccanica newtoniana, definiamo **quadrimpulso** o **quadrimomento** o **quadri-quantità di moto** di una particella il prodotto della sua massa propria per la sua quadrivelocità[4]:

$$\mathbf{P} \stackrel{\text{def}}{=} m_0 \mathbf{U}. \tag{6.1}$$

Come $\mathbf{U}$, anche $\mathbf{P}$ è tangente alla linea di universo della particella, con orientamento verso il futuro. Le sue componenti, per la (6.1) e le (5.55), risultano essere:

$$\begin{cases} P^0 &= m_0 U^0 = m_0 \gamma(u) c, \\ P^\alpha &= m_0 U^\alpha = m_0 \gamma(u) u^\alpha \quad (\alpha = 1,2,3), \end{cases} \tag{6.2}$$

dove le u^α $(\alpha = 1,2,3)$ sono le componenti cartesiane della 3- velocità $\mathbf{u}$ della particella.
Del significato della componente temporale P^0 del 4-impulso ci occuperemo in seguito. Per il momento prendiamo in considerazione le sole componenti spaziali P^α e notiamo che esse costituiscono un trivettore

$$\mathbf{p} = m_0 \gamma(u) \mathbf{u} = \frac{m_0}{\sqrt{1 - \dfrac{u^2}{c^2}}} \mathbf{u}, \tag{6.3}$$

di componenti $p^\alpha = m_0 \gamma(u) u^\alpha$ $(\alpha = 1,2,3)$, sicché il 4-vettore $\mathbf{P}$, in virtù delle (5.56), potrà esprimersi anche nella forma $\mathbf{P} = (m_0 \gamma(u) c, \mathbf{p})$ dove entro parentesi si è separata la parte temporale da quella spaziale del 4-impulso.
A questo punto, per proseguire nella nostra discussione, il formalismo ci suggerisce, in modo spontaneo, due possibili procedimenti da seguire che corrispondono a due diverse correnti di pensiero entrambe sostenute da fisici eminenti, e cioè:

1) interpretare la quantità

$$m(u) \equiv m_0 \gamma(u) = \frac{m_0}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \frac{m_0}{\sqrt{1 - \beta^2}}, \tag{6.4}$$

come *massa della particella*, **relativa al riferimento galileiano considerato, funzione del modulo u della sua velocità**. Tale massa sarà chiamarla *massa inerziale relativa* o anche

[3]Per rendersi conto di questa affermazione fondamentale si veda la successiva nota 4 a piè di pagina e per maggiori dettagli si legga **l'Osservazione importantissima** riportata alla fine del paragrafo 8.12.

[4]Osserviamo che l'uguaglianza (6.1) fra 4-vettori di $\mathcal{M}^4$ si può scrivere anche in termini di componenti così: $P^i = m_0 U^i$ $(i = 0,1,2,3)$. Si può constatare immediatamente che **essa è invariante per trasformazioni di Lorentz**: infatti, come sappiamo, si ha $P^i = L^i_{j'} P^{j'}$, $U^i = L^i_{j'} U^{j'}$, che sostituite nella formula precedente ci danno $L^i_{j'}(P^i_{j'} - m_0 U^{j'}) = 0$, che è un sistema lineare omogeneo di 4 equazioni in 4 incognite, con determinante dei coefficienti $det(L^i_{j'}) \neq 0$, dunque esso ammette la sola soluzione nulla $P^i_{j'} - m_0 U^{j'} = 0$, che equivale a $P^{j'} = m_0 U^{j'}$ $(j' = 0,1,2,3)$ c.d.d.

massa inerziale relativistica, $m_{rel}(u) \equiv m(u)$, nel qual caso il 3-vettore $\mathbf{p}$, può interpretarsi come **quantità di moto relativa** al riferimento galileiano associato alle coordinate $\{x^i\}$, definito dalla

$$\mathbf{p} \stackrel{\text{def}}{=} m(u)\mathbf{u}, \tag{6.5}$$

cioè come il prodotto della *massa relativistica* $m(u)$ della particella per il 3-vettore velocità $\mathbf{u}$, in perfetta analogia formale con quello che si fa nella meccanica classica.

Figura 6.1: Confronto fra massa classica e massa relativistica.

Qualunque possa essere il significato di $m = m(u)$, osserviamo che essa è funzione della velocità u e cresce con essa: quando $u = 0$, $\gamma(0) = 1$ ed $m = m_0$ che si identifica con la **massa a riposo o massa propria** della particella già definita nel paragrafo 6.2.1. È interessante dare uno sguardo al grafico della massa relativistica in funzione di β, cioè al grafico della funzione $m(\beta) = \dfrac{m_0}{\sqrt{1-\beta^2}}$ che, a meno del fattore costante m_0, si identifica col grafico del fattore di Lorentz (Fig.6.1).

2) Oppure, definire semplicemente la **3-quantità di moto relativistica** della particella così come si presenta spontaneamente nella formula (6.3), cioè come prodotto della massa propria, identificata ora con la massa inerziale newtoniana m, per il fattore di Lorentz e per la sua velocità, quindi in maniera diversa dal caso newtoniano per la presenza del fattore $\gamma(u)$. In questo secondo caso perciò, per la particella, esiste una sola massa: quella newtoniana.
La quantità di moto relativistica sarà allora definita dalla

$$\mathbf{p} \stackrel{\text{def}}{=} m\gamma(u)\mathbf{u}, \tag{6.6}$$

con $m = cost.$ indipendente dalla velocità.

Nel primo caso si ritiene che la massa possa variare al variare del modulo della velocità e nel passaggio da un sistema inerziale ad un altro - ciò che non sarebbe per niente scandaloso dato che in relatività altre grandezze come le lunghezze, gli intervalli temporali, le frequenze, ecc. dipendono dalla velocità relativa dei sistemi di riferimento inerziali nei quali vengono considerate.
Nel secondo caso si ritiene invece che la massa m della particella sia uno scalare invariante e che non esista altro concetto di massa se non quello introdotto in ambito classico come massa inerziale. Il primo approccio è stato, storicamente, quello seguito da Einstein - anche se sembra che lui stesso avesse delle forti perplessità al riguardo - e dalla stragrande maggioranza dei

fisici e matematici di tutto il mondo (Born, Sommerfeld, Pauli, Møller, Synge, Rindler, ecc.), perlomeno fino all'avvento dell'impetuoso sviluppo della fisica delle particelle e della fisica delle alte energie. Successivamente molti fisici, anche illustri, come Landau e Lifschitz, Bergman, Taylor e Wheeler, Weinberg, Adler, Okun, ecc., ma anche fisici matematici come Lanczos hanno cominciato ad optare per la seconda interpretazione ritenendo che la prima è sostanzialmente formale, non utile nelle applicazioni ed anche, per certi aspetti, fuorviante.[5] Forse è questione di gusti scegliere l'uno o l'altro dei due approcci, in ogni caso è ovvio che entrambi, pur utilizzando una terminologia un pò diversa, conducono sostanzialmente agli stessi risultati significativi espressi in un linguaggio leggermente diverso.
In questo libro noi aderiamo al 1°punto di vista, cioè quello che introduce il concetto di massa relativistica, e che è stato sostenuto recentemente, per esempio, da T. R. Sandin, Q. ter Spill e R. Baierlein, nei lavori citati nella nota 5 che è poi quello adottato tradizionalmente da moltissimi autori compreso il più volte citato Wolfgang Rindler del quale abbiamo seguito più da vicino il pensiero. Secondo Sandin, il punto di forza di questo approccio che anche noi condividiamo, lungi dall'essere una reminiscenza storica, consiste - come la pratica dell'insegnamento ha dimostrato - *"nell'essere didatticamente più semplice, più simmetrico, più bello, per niente fuorviante, e soprattutto di maggiore comprensibilità da parte degli studenti che intraprendono per la prima volta lo studio della relatività".*
Nell'adottare questo punto di vista ci è di conforto il seguente pensiero estetico, a dire il vero, abbastanza sconcertante, di Paul Adrian Maurice Dirac[6]: **"È più importante che le equazioni siano belle piuttosto che in accordo con gli esperimenti.[....] Sembra che, se si lavora con il proposito di ottenere equazioni dotate di bellezza, e si possiede un'intuizione davvero solida, si è sicuramente sulla strada del progresso".**
Osservazione:
Riguardo al secondo procedimento, al quale abbiamo accennato prima, consistente nel definire la quantità di moto per mezzo della (6.6) senza introdurre il concetto di massa relativistica ma usando solo il concetto di massa inerziale newtoniana, daremo un cenno nell'APPENDICE alla fine del capitolo.
In ciò che segue ci atterremo esclusivamente al primo procedimento consistente nel definire **la massa relativistica** con la:

$$m(u) \stackrel{\text{def}}{=} m_0\gamma(u),$$

dipendente dunque dalla velocità tramite il fattore di Lorentz, ed in conseguenza definire **la quantità di moto o impulso** o, con denominazione proveniente dalla lingua inglese, **momento lineare** o semplicemente **momento** con la:

$$\mathbf{p} \stackrel{\text{def}}{=} m(u)\mathbf{u}.$$

[5]Sul concetto di massa in relatività è consigliabile la lettura dei seguenti articoli:

1) in difesa del secondo punto di vista di cui abbiamo parlato più sopra, cioè contro il concetto di massa relativistica si può vedere: C.G. Adler, *Does mass really depend on velocity, dad ?*, in "Am. J. Phys.", bf 55, p. 739-743 (1987) e Lev B. Okun, *The Concept of Mass*, "Physics Today" **42** (6), p.31-36 (July 1989).

2) In difesa del primo punto di vista, cioè il concetto di massa relativistica dipendente dalla velocità, oltre gli autori classici, ci sono anche autori moderni che lo sostengono, per es. T. R. Sandin, *In Defense of Relativistic Mass*, "American Journal of Physics" **59** (11), (Nov. 1991); Q. ter Spill, *Mass and Energy*, 's Gravesande, Institute of Physics Education, Leiden - Netherlands, p. 1-70; Ralph Baierlein, *Teaching $E = mc^2$: An exploration of some issues*, "Phys. Teach. **29**, 170-175 (1991).

[6]Si raccomanda vivamente al lettore di consultare il bellissimo libro di P.A.M. Dirac, *La bellezza come metodo*, Indiana Editore srl, Milano 2013, a cura di Vincenzo Barone; la frase sopra riportata si trova a pag.104. A proposito di bellezza e simmetria delle leggi fisiche, si raccomanda anche la lettura dell'interessante libro di Vincenzo Barone, *L'ordine del mondo, le simmetrie in Fisica da Aristotele a Higgs*, Bollati Boringhieri, 2014. In questo libro di Barone, a pag.14, si ritrova la frase di Dirac sopra citata ed altre ancora che ribadiscono la *bellezza come metodo* per la costruzione delle teorie fisiche.

Figura 6.2: Confronto fra la quantità di moto classica e quella relativistica.

Concludendo, è molto istruttivo fare un confronto fra i grafici della quantità di moto classica e relativistica, Fig.6.2, e renderci conto che, per velocità piccole rispetto alla velocità della luce, le due definizioni praticamente sono coincidenti, mentre per velocità che si approssimano a quella della luce la divergenza è notevole. L'esperienza dimostra che la definizione corretta è quella relativistica.

6.2.3 La conservazione del 4-impulso come base della Meccanica Relativistica

Ci sono diversi modi di introdurre la meccanica relativistica di una o più particelle e tutti conducono alla stessa, finora, ben provata teoria. Storicamente il primo approccio fu attraverso la teoria di Maxwell, ed in particolare attraverso la forza di Lorentz come prototipo di forza relativistica. Ma ciò dà un'indebita preminenza logica alla teoria elettromagnetica, e di fatto al concetto di forza come si fa anche nella teoria newtoniana. Se invece seguiamo l'approccio fondato sulla meccanica delle collisioni (o urti) fra particelle, non c'è bisogno d'introdurre subito il concetto di forza che sarà invece introdotto in un secondo tempo.
Com'è noto in meccanica classica è valida la **legge di conservazione della quantità di moto nelle collisioni fra corpi puntiformi:** *In un processo d'urto (***elastico o anelastico***)*[7] *fra particelle, la quantità di moto totale prima dell'urto è uguale alla quantità di moto totale dopo l'urto, quale che siano le variazioni delle velocità delle particelle coinvolte nella collisione.*
Per il modo stesso con cui è formulata questa legge, è plausibile che essa mantenga la sua validità nell'ambito della nuova dinamica (che stiamo per costruire); essa concerne infatti solo azioni reciproche che si esercitano fra le varie particelle che partecipano alla collisione, senza influenze esterne e quindi senza che vi sia un riferimento ad altre particelle, che non sono direttamente implicate nella collisione, e senza riferimento ad alcun sistema di coordinate.
D'altra parte noi vogliamo che le nuove leggi siano Lorentz-invarianti, cioè che ubbidiscano al principio einsteiniano di relatività, ciò può essere realizzato, abbastanza facilmente, esprimendole in forma tensoriale cioè formulando la nuova meccanica con l'uso di grandezze definite nello spaziotempo $\mathcal{M}^4$ di Minkowski. Ora noi conosciamo già qualche 4-vettore come la 4-velocità **U** ed il 4-momento **P** di una particella, perciò assumiamo - *come tentativo, i cui risultati dovranno essere verificati in seguito sperimentalmente* - che la legge classica di conservazione della

[7] Un urto fra particelle, o fra corpi, si dice **elastico** se si conserva l'energia cinetica totale delle particelle (cioè la somma delle energie cinetiche delle particelle prima dell'urto è uguale alla somma delle analoghe energie cinetiche dopo l'urto). L'urto si dice invece **anelastico** se non c'è conservazione dell'energia cinetica, e se le particelle che collidono rimangono fuse una sull'altra formando un'unica particella. **Sia negli urti elastici sia in quelli anelastici vale invece sempre il principio di conservazione della quantità di moto e dell'energia**. Per la teoria degli urti si consulti un qualunque libro di *Fisica I* del 1° biennio, per es. D. Halliday, R. Resnick e K. S. Krane, *Fisica I*, 4ª Edizione, Casa Editrice Ambrosiana, Milano (1997), Cap. 10, pag. 215.

quantità di moto si possa enunciare in termini di 4-vettori, così:
Legge di conservazione della 4-quantità di moto:
la somma delle 4-quantità di moto di tutte le particelle che confluiscono in una collisione è uguale alla somma delle 4-quantità di moto di tutte le particelle che emergono da essa.
La collisione può essere **elastica o anelastica**, e il numero delle particelle che confluiscono nella collisione può essere minore, uguale o maggiore del numero di quelle che emergono dalla collisione. Se il numero totale delle particelle prima e dopo la collisione è n ed il numero delle particelle prima della collisione è h, il numero delle particelle che emergono dalla collisione sarà $k = n - h$, perciò la legge di conservazione della 4-quantità di moto si potrà scrivere così:

$$\sum_{i=1}^{h} \mathbf{P}_i = \sum_{j=h+1}^{n} \mathbf{P}_j \qquad (conserv.\ della\ 4-quant.\ di\ moto). \tag{6.7}$$

Possiamo esprimere ciò, passando tutti i termini al 1° membro, anche nella forma

$$\sum_a{}^* \mathbf{P}_a = 0, \quad (a = 1, 2, 3, ..., n) \qquad (conserv.\ della\ 4-quant.\ di\ moto), \tag{6.8}$$

dove l'indice a (*non tensoriale*) numera tutte le particelle prima e dopo la collisione e dove $\sum_a{}^*$ è una somma che conta i termini prima della collisione positivamente e quelli dopo la collisione negativamente.
Questa legge fra 4-vettori, nella forma (6.7) o (6.8), si dimostrerà molto soddisfacente sia dal punto di vista teorico sia da quello empirico e sperimentale.
Adesso, tenendo conto delle (6.2), (6.3), (6.4), la legge di conservazione della 4-quantità di moto (6.8) potrà scriversi così

$$\sum_a{}^* \mathbf{P}_a \equiv \sum_a{}^* (m_a c, \mathbf{p}_a) = 0, \tag{6.9}$$

dove $m_a = m_{0\,a}\gamma(u_a)$ sono le masse relativistiche di tutte le particelle coinvolte nella collisione e $\mathbf{p}_a = m_a \mathbf{u}_a$ le loro 3-quantità di moto.
Perciò il 4-vettore $\sum_a{}^*(m_a c, \mathbf{p}_a)$ quantità di moto totale essendo nullo avrà ciascuna delle sue quattro componenti nulle separatamente, ed essendo c una costante, avremo allora

$$\sum_a{}^* m_a = 0, \quad (conserv.\ della\ massa). \tag{6.10}$$

e

$$\sum_a{}^* \mathbf{p}_a \equiv \sum_a{}^* m_a \mathbf{u}_a = 0 \quad (conserv.\ della\ 3-quant.\ di\ moto). \tag{6.11}$$

Per il momento lasciamo da parte la (6.10) riservandoci di discuterla in seguito.
La (6.11) ha una diretta interpretazione fisica in quanto essa è formalmente analoga alla **legge di conservazione della quantità di moto della meccanica newtoniana**, pur di assumere come definizione del 3-momento la (6.6), ed anzi tende proprio ad essa per $\gamma(u) \to 1$ cioè per $c \to \infty$, ovvero praticamente per $u \ll c$. È ciò ragionevole? Ragionevole o no, è la sola legge Lorentz-invariante della forma (6.11).
Insomma, la conclusione del ragionamento precedente è che se vogliamo che anche in relatività, in una collisione fra più particelle, valga **la legge di conservazione della 3-quantità di moto**, cioè la

$$\sum_a{}^* m_a(u_a)\mathbf{u}_a = 0 \quad (conserv.\ della\ 3-quant.\ di\ moto). \tag{6.12}$$

dobbiamo ammettere necessariamente che, in un generico riferimento inerziale $\mathcal{R}$, sia valida la relazione (6.4) cioè: *la massa relativistica deve essere uguale al prodotto della massa propria per il fattore di Lorentz.*

$$m(u) = m_0\gamma(u). \tag{6.13}$$

Questa relazione fondamentale è stata da noi introdotta con considerazioni piuttosto formali, vogliamo ora darne una dimostrazione di natura fisica supponendo di eseguire un esperimento ideale. Consideriamo l'urto fra due particelle identiche $\mathcal{P}_1$, $\mathcal{P}_2$, entrambe di massa propria m_0,

Figura 6.3: Urto anelastico fra due particelle identiche $\mathcal{P}_1$, $\mathcal{P}_2$ (1^a interpretazione della massa).

mobili, in direzioni opposte ma con velocità di ugual modulo u (Fig.6.3). La prima si muova sul semiasse positivo delle z verso il basso cioè verso l'origine O di un dato sistema galileiano $\mathcal{R}$, mentre la seconda si muova sul semiasse negativo delle z' verso l'alto cioè verso l'origine O' di un altro sistema inerziale $\mathcal{R}'$. I due sistemi $\mathcal{R}$, $\mathcal{R}'$, siano in configurazione x-standard e mobili l'uno rispetto all'altro con velocità relativa v, inoltre supponiamo che prima della collisione fra $\mathcal{P}_1$ e $\mathcal{P}_2$, $\mathcal{R}'$ si trovi a sinistra di $\mathcal{R}$, rispetto ad un osservatore in quiete nell'origine O di $\mathcal{R}$ stesso, in piedi sul piano xy ed orientato dai piedi alla testa come il semiasse positivo delle z, e che abbia lo sguardo rivolto verso il semiasse positivo delle y. Immaginiamo che la collisione sia **anelastica**, cioè che le particelle dopo la collisione rimangano fuse una sull'altra e che ciò avvenga, nell'origine $O \equiv O'$ dei due sistemi di coordinate, nell'istante in cui essi vengono a coincidere. Ammettiamo ancora che, per ragioni di simmetria, la particella composta dopo l'urto rimanga in quiete in $\mathcal{R}$. Utilizzando la legge relativistica di composizione delle velocità nella forma (4.16) o (4.17), cioè

$$\begin{cases} u'_x = \dfrac{u_x - v}{1 - \dfrac{v}{c^2}u_x}, & u_x = \dfrac{u'_x + v}{1 + \dfrac{v}{c^2}u'_x} \\[2ex] u'_y = \dfrac{u_y\gamma^{-1}(v)}{1 - \dfrac{v}{c^2}u_x}, & u_y = \dfrac{u'_y\gamma^{-1}(v)}{1 + \dfrac{v}{c^2}u_x} \\[2ex] u'_z = \dfrac{u_z\gamma^{-1}(v)}{1 - \dfrac{v}{c^2}u_x}, & u_z = \dfrac{u'_z\gamma^{-1}(v)}{1 + \dfrac{v}{c^2}u_x} \end{cases}$$

avremo

PRIMA DELL'URTO

Particella $\mathcal{P}_1$, $in\,\mathcal{R}$: $u_{1x} = 0,\ u_{1y} = 0,\ u_{1z} = -u,$ (componenti assegnate);

Particella $\mathcal{P}_1$, $in\ \mathcal{R}'$: $u'_{1x} = -v,\ u'_{1y} = 0,\ u'_{1z} = -u\gamma^{-1}(v),$ (componenti calcolate);

Particella $\mathcal{P}_2$, $in\,\mathcal{R}'$: $u'_{2x} = 0,\ u'_{2y} = 0,\ u'_{2z} = u,$ (componenti assegnate);

Particella $\mathcal{P}_2$, $in\,\mathcal{R}$: $u_{2x} = v,\ u_{2y} = 0,\ u_{2z} = u\gamma^{-1}(v),$ (componenti calcolate);

DOPO L'URTO (istante dell'urto)

Particelle fuse $\mathcal{P}_1 + \mathcal{P}_2$, *in* $\mathcal{R}$: $U_x = U_y = U_z = 0$, (componenti assegnate);
Particelle fuse $\mathcal{P}_1 + \mathcal{P}_2$, *in* $\mathcal{R}'$: $U'_x = -v$, $U'_y = U'_z = 0$, (componenti calcolate);

avendo indicato con lettere minuscole le componenti delle velocità prima dell'urto e con maiuscole le analoghe componenti dopo l'urto.
Applichiamo ora **la legge di conservazione della quantità di moto** in $\mathcal{R}$ lungo l'asse z, in accordo con la (6.12), proiettata sull'asse z, avremo

$$m(u_1)u_{1z} + m(u_2)u_{2z} = M(U)U_z,$$

dove

$$u_1 = \sqrt{(u_{1x})^2 + (u_{1y})^2 + (u_{1z})^2} = u,$$

e

$$u_2 = \sqrt{(u_{2x})^2 + (u_{2y})^2 + (u_{2z})^2} = \sqrt{v^2 + u^2\gamma^{-2}(v)},$$

sono rispettivamente i moduli delle velocità delle due particelle $\mathcal{P}_1$ e $\mathcal{P}_2$ prima dell'urto ed U il modulo della velocità delle particelle fuse dopo l'urto, nel riferimento $\mathcal{R}$.
Tenuto conto dei valori delle componenti delle velocità $u_{1z} = -u$, $u_{2z} = u\gamma^{-1}(v)$ e $U_z = 0$, riportate nello schema precedente che descrive l'urto, otteniamo

$$m(u)(-u) + m(u_2)u\gamma^{-1}(v) = M(U) \cdot 0 = 0$$

cioè

$$m(u_2) = m(u)\gamma(v). \tag{6.14}$$

Se ora facciamo tendere a zero u ($u \to 0$), considerando una sequenza continua di esperimenti con u sempre più piccola, avremo che $u_2 \to v$, ciò porta a concludere che la (6.14) diventa

$$m(v) = m(0)\gamma(v), \tag{6.15}$$

con v velocità della particella $\mathcal{P}_2$ nel riferimento $\mathcal{R}$; ma quest'ultimo è un generico riferimento inerziale e $\mathcal{P}_2$ è una generica particella che si muove con velocità $|\mathbf{v}| = v$, perciò la (6.15) - a meno dello scambio di v con u - dimostra la nostra tesi e cioè che: **se vogliamo salvare la forma (6.12), analoga formalmente alla legge newtoniana della conservazione della 3-quantità di moto, dobbiamo permettere ad m di variare precisamente secondo la formula (6.15) che abbiamo accettato in precedenza in via del tutto formale.**
Osserviamo inoltre che con la definizione (6.13) di m, le due leggi Lorentz-invarianti (6.10) (*conservazione della massa*) e (6.11) (*conservazione della quantità di moto*) si riducono alle ben note analoghe della meccanica newtoniana (con masse indipendenti dalle velocità) quando tutte le velocità in gioco sono così piccole che $m \approx m_0$ per ogni particella. Ma per tali situazioni le leggi di Newton sono estremamente ben convalidate. Cosicché se consideriamo una collisione che avvenga con moto lento (newtoniano) in un sistema galileiano $\mathcal{R}_0$, in un altro sistema $\mathcal{R}$ ad alta velocità rispetto ad $\mathcal{R}_0$, la stessa collisione soddisfa automaticamente le nuove leggi relativistiche per la loro Lorentz-invarianza. Questa è sicuramente un'altra raccomandazione per tali leggi.
Dopo le considerazioni precedenti, ciascuno di noi può adesso essere pronto ad accettare con più convinzione, su base puramente teorica, la legge tensoriale (6.8) che abbiamo proposto all'inizio della nostra discussione! *È bene avvertire che tale legge è universalmente accettata come base della meccanica relativistica, ed è stata ampiamente confermata dalle osservazioni sulle collisioni*

delle particelle elementari per ormai più di un secolo.
La massa propria m_0 di una particella è equivalente alla sua *massa inerziale newtoniana*, cioè quella che è determinata nelle collisioni lente. Essa è un **invariante**, cioè tutti gli osservatori concordano col suo valore ad ogni istante. Ma non c'è nessuna ragione per cui essa debba rimanere **costante**: la massa propria di una particella può essere alterata in una collisione se il suo stato interno cambia, e può essere anche alterata nel passare attraverso certi campi di forza. Ma per una particella che si muove di **moto libero** *(cioè una* **particella isolata** *non soggetta all'azione di alcun agente esterno)*, anche se essa subisce cambiamenti interni di stato, m_0 deve essere costante, in virtù della prima legge di Newton (legge d'inerzia) e per la conservazione della quantità di moto.

6.2.4 Energia totale, energia di riposo ed energia cinetica di una particella

È giunto il momento di occuparci della componente temporale della 4-quantità di moto **P** che si identifica con la prima delle (6.2), cioè $P^0 = m_0\gamma(u)c$ e che finora abbiamo tralasciato. Tale componente, a parte il fattore costante c, dipende soltanto dal prodotto della massa propria della particella per il fattore di Lorentz, mentre le componenti spaziali $P^\alpha = m_0\gamma(u)u^\alpha$ oltre che dipendere da tale prodotto dipendono anche dalle componenti u^α della 3-velocità della particella.
Il significato e l'importanza della componente P^0, deriva specialmente dalla sua stretta connessione con la nozione di energia. Consideriamo infatti il prodotto di P^0 per c, cioè

$$cP^0 = m_0\gamma(u)c^2 = mc^2, \tag{6.16}$$

avendo indicato con $m = m_0\gamma(u)$ la massa relativistica introdotta con la (6.4). La quantità $mc^2 = m_0\gamma(u)c^2$, essendo $m_0 > 0$ e $\gamma(u) > 0$ è **sempre positiva ed ha le dimensioni di una energia**, in quanto il fattore di Lorentz è adimensionale. La indicheremo con la lettera E ponendo

$$E \stackrel{\text{def}}{=} mc^2 = m_0\gamma(u)c^2 = \frac{m_0c^2}{\sqrt{1-\dfrac{u^2}{c^2}}}. \tag{6.17}$$

Sviluppiamo ora la funzione $\gamma(u)$ in serie di Taylor relativa al punto $u = 0$,[8] la formula precedente diventa

$$E = m_0c^2 + \frac{1}{2}m_0u^2 + \frac{3}{8}m_0\frac{u^4}{c^2} + \frac{5}{16}m_0\frac{u^6}{c^4} + \dots \tag{6.18}$$

dalla quale si riconosce che, per valori abbastanza piccoli di u rispetto a c, cioè $u \ll c$, nello sviluppo precedente si possono trascurare i termini dal terzo in poi, e perciò si avrà

$$E \approx m_0c^2 + \frac{1}{2}m_0u^2 \tag{6.19}$$

in tal caso dunque l'energia totale della particella differisce per la costante additiva m_0c^2 dall'energia cinetica classica $\frac{1}{2}m_0u^2$.
Lo scalare E è quindi una grandezza relativa, che per basse velocità, quando cioè la meccanica relativistica deve, per programma, coincidere con la meccanica classica, si identifica, a meno di

[8] Per sviluppare il fattore di Lorentz $\gamma(u)$ conviene considerarlo come funzione di β scrivendolo nella forma $\gamma(\beta) = (1-\beta^2)^{-1/2}$ ed utilizzare la serie binomiale $(1+x)^\alpha = \sum_{k=0}^{+\infty}\binom{\alpha}{k}x^k$, con $x, \alpha \in \mathbb{R}$, e $\binom{\alpha}{k} = \dfrac{\alpha(\alpha-1)(\alpha-2).....(\alpha-k+1)}{k!}$. Nel nostro caso specifico è $\alpha = -1/2$, $x = -\beta^2$, ed essendo $|x| < 1$ la serie è senz'altro assolutamente convergente; si ottiene allora $\gamma(\beta) = 1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \frac{5}{16}\beta^6 + ...$

una costante additiva, m_0c^2, con l'energia cinetica (ordinaria) della particella. Ciò suggerisce di assumere la (6.17) come definizione della **energia totale** della particella ed inoltre ci induce ad identificare, nel secondo membro della (6.18), la somma dei termini dal secondo in poi come **l'energia cinetica relativistica della particella**, cioè

$$T_{rel} = \frac{1}{2}m_0u^2 + \frac{3}{8}m_0\frac{u^4}{c^2} + \frac{5}{16}m_0\frac{u^6}{c^4} + \dots , \tag{6.20}$$

la quale, tenuto conto dello sviluppo di $\gamma(u)$ dato nella nota 8 a piè di pagina, si può riscrivere

Figura 6.4: Confronto tra energia cinetica classica e relativistica.

in maniera più sintetica così:

$$T_{rel} \stackrel{\text{def}}{=} m_0c^2\Big[\gamma(u) - 1\Big]. \tag{6.21}$$

Nella Fig.6.4 si confrontano i grafici dell'energia cinetica classica con quella relativistica e si vede chiaramente che $T_{class.} \to +\infty$ per $u \to +\infty$, mentre $T_{rel} \to +\infty$ per $u \to c$. L'opportunità delle precedenti definizioni di energia totale e di energia cinetica verranno del resto confermate un pò più avanti, in sede dinamica, quando queste grandezze verranno poste in relazione con le forze agenti sulla particella.
La definizione (6.17) attribuisce un'energia non nulla anche ad una particella in quiete ($u = 0$) chiamata **energia di riposo** o anche **energia di massa** della particella:

$$E_0 \stackrel{\text{def}}{=} m_0c^2. \tag{6.22}$$

La relazione (6.17) *fra massa relativistica ed energia totale di una particella e la* (6.22) *fra massa a riposo ed energia di massa della stessa particella, sono le formule emblematiche di tutta la relatività e costituiscono un risultato sorprendente e di enorme importanza. La* (6.22) *ci dice che, per il semplice fatto di avere una massa inerziale m_0, un corpo possiede anche una quantità enorme di energia pari ad m_0c^2.* **Massa ed energia sono "equivalenti": la massa può convertirsi in energia, e, viceversa l'energia può convertirsi in massa**, un fatto che non ha nessun riscontro nella meccanica newtoniana. È opportuno rilevare che tale equivalenza non ha importanza nella esperienza quotidiana ma è importante a livello nucleare e subnucleare.
Questa energia, peraltro, nel caso di una particella priva di struttura interna non è un'energia disponibile, cioè tale da poter essere trasformata in altre forme di energia e quindi utilizzabile. Diverso invece sarà il caso di una particella dotata di struttura interna; in tal caso anche **l'energia di riposo può divenire, in parte, energia disponibile**.
La (6.18) può essere ora riscritta in maniera più sintetica

$$E = E_0 + T_{rel}, \tag{6.23}$$

Figura 6.5: Energia totale ed energia propria.

cioè **l'energia totale di una particella è la somma della sua energia di riposo più l'energia cinetica dovuta alla sua velocità**. Nella Fig.6.5 sono riportati i grafici di queste grandezze fondamentali E ed E_0.
Siamo ora finalmente in grado di dare un'interpretazione fisica alla (6.10), che avevamo lasciato da parte nel paragrafo precedente, e per far ciò basta moltiplicarla per c, avremo allora

$$\sum_a{}^* m_a c^2 \equiv \sum_a{}^* E_a = 0, \quad (a = 1, 2, 3, ...) \tag{6.24}$$

che possiamo riscrivere in maniera più espressiva così:

$$\left(\sum_{i=1}^{h} E_i\right)_{prima\ dell'urto} = \left(\sum_{j=h+1}^{n} E_j\right)_{dopo\ l'urto} \quad (conserv.\ massa - energia) \tag{6.25}$$

si tratta dunque della fondamentale **legge di conservazione dell'energia in una collisione fra particelle in numero arbitrario** ovvero essa può essere interpretata anche come **legge di conservazione della massa relativistica in una collisione fra particelle in numero arbitrario**.
Se nella (6.25) teniamo conto della (6.23), cioè scomponiamo l'energia totale di ciascuna particella nella somma dell'energia a riposo e dell'energia cinetica, avremo

$$\left(\sum m_0 c^2\right)_{prima\ urto} + \left(\sum T\right)_{prima\ urto} = \left(\sum m_0 c^2\right)_{dopo\ urto} + \left(\sum T\right)_{dopo\ urto}. \tag{6.26}$$

I due termini, di massa e cinetico, non sono separatamente conservati, salvo che in qualche caso particolare come un urto elastico, in cui le particelle rimangono inalterate e si ha

$$\left(\sum m_0 c^2\right)_{prima\ urto} = \left(\sum m_0 c^2\right)_{dopo\ urto} \quad e \quad \left(\sum T\right)_{prima\ urto} = \left(\sum T\right)_{dopo\ urto}.$$

In generale, massa ed energia cinetica si convertono l'una nell'altra, e solo l'energia totale rimane costante.
Ritorneremo di nuovo, nel paragrafo 6.4, con più dettagli sul concetto fondamentale di equivalenza fra massa ed energia.

6.3 Le leggi della dinamica relativistica

6.3.1 Quadriforza e leggi assolute del moto per una particella di massa propria costante

Vediamo ora di introdurre le leggi della dinamica relativistica operando nello spazio di Minkowski. Conformemente alle considerazioni esposte nell'introduzione a questo capitolo, le leggi del moto che ora dobbiamo introdurre in sostituzione delle leggi newtoniane, debbono a priori soddisfare alla duplice condizione di essere invarianti per trasformazioni di Lorentz e di ridursi alle ordinarie leggi newtoniane per $c \to \infty$. Alla prima condizione si soddisfa automaticamente se le leggi si esprimono in termini geometrici intrinseci dello spaziotempo di Minkowski $\mathcal{M}^4$. *Naturalmente queste due condizioni, pur necessarie, non bastano ad individuare le leggi del moto; soltanto l'accordo delle loro conseguenze osservabili con l'evidenza sperimentale potrà, a posteriori, giustificare la loro scelta.*

Possiamo anzitutto osservare che il **principio newtoniano d'inerzia (1ª legge della Dinamica)** soddisfa senza modificazioni, al principio d'invarianza lorentziana. E infatti, come abbiamo già osservato alla fine del paragrafo 5.6.3, la rettilineità e uniformità del moto, si traduce nella rettilineità della linea di universo della particella, e questo è un carattere invariante per trasformazioni di Lorentz come lo è anche per trasformazioni di Galileo. La legge d'inerzia può dunque assumersi, anche in relatività, come indice dell'assenza di forze agenti. Assumeremo pertanto come primo postulato dinamico il seguente:

I - Principio d'inerzia:

Se un punto materiale privo di struttura interna è esente da forze esso permane in quiete o in moto rettilineo uniforme rispetto ad ogni riferimento galileiano.

La rettilineità della linea di universo nello spazio di MInkowski si esprime nella costanza del vettore 4-velocità, cioè $\mathbf{U} = cost.$ ovvero è costante il quadrimpulso della particella: $\mathbf{P} = cost.$ e ciò si può esprimere così

$$\frac{d\mathbf{P}}{d\tau} = 0, \tag{6.27}$$

e cioè: *è nulla la derivata del 4-impulso* $\mathbf{P}$ *di una particella rispetto al suo tempo proprio* τ.

Se invece il punto materiale si muove di moto vario (e ciò accade allora in tutti i riferimenti inerziali), noi diremo che **esso è soggetto ad una forza**. Ebbene, in analogia con la legge newtoniana $\mathbf{f} = \frac{d\mathbf{p}}{dt}$, operando ora però nello spazio di Minkowski, diremo **quadriforza o forza di Minkowski** applicata ad una particella, il quadrivettore $\mathbf{F}$ uguale alla derivata del suo quadrimpulso rispetto al suo tempo proprio:

$$\mathbf{F} \stackrel{\text{def}}{=} \frac{d\mathbf{P}}{d\tau}. \tag{6.28}$$

L'uguaglianza ora scritta si può naturalmente interpretare come una definizione della *quadriforza* $\mathbf{F}$, ma se, analogamente a quanto si fa per la forza in meccanica **si assume la quadriforza come una grandezza primitiva, rappresentante l'azione dell'ambiente sul punto materiale**, la (6.28) si può veramente interpretare come una relazione causale tra *quadriforza (causa) e variazione del moto (effetto)*. Enunceremo pertanto la seguente legge che gode della proprietà di essere invariante in forma nel passaggio da un sistema inerziale ad un altro.[9]

II - Legge del moto della dinamica relativistica:

[9]Che la legge (6.28), scritta nella forma (6.29), espressa in termini di 4-vettori conservi la stessa forma nel passaggio da un sistema galileiano $\mathcal{R}$ ad un altro $\mathcal{R}'$ in configurazione x-standard, si può vedere in maniera molto semplice. Riscriviamola in termini di componenti $\frac{dP^i}{d\tau} = F^i$ nel riferimento $\mathcal{R}$ e nelle coordinate x^i. Se si esegue una trasformazione speciale di Lorentz, consistente in una trasformazione lineare di coordinate $x^i \to x^{i'}$, i due 4-vettori

Se su un punto materiale di massa propria m_0 agisce una quadriforza **F**, *questa induce in esso una variazione della sua quadriquantità di moto* **P**, *e si ha precisamente*

$$\frac{d\mathbf{P}}{d\tau} = \mathbf{F}. \tag{6.29}$$

Se la particella è priva di struttura interna avrà, per definizione, massa propria costante $m_0 = cost.$, in tal caso la 2^a legge della dinamica relativistica si scriverà nella forma

$$m_0\mathbf{A} = \mathbf{F}, \tag{6.30}$$

formalmente analoga a quella classica newtoniana, essendo **A** la 4-accelerazione.
Naturalmente perché la (6.29) o la (6.30) possa avere utilizzazione pratica per l'effettiva determinazione del moto occorre ammettere che sia possibile conoscere a priori la legge della quadriforza e cioè la sua dipendenza dalle coordinate x^i che individuano la posizione della particella, dalla sua quadrivelocità U^i e dal suo tempo proprio τ.
Per quanto riguarda **la terza legge della dinamica classica**, cioè **il principio di azione e reazione**, esso non ha generalmente senso in meccanica relativistica sostanzialmente per due motivi:

1) le mutue azioni che due corpi si scambiano devono essere, per la meccanica classica, **simultanee**, mentre in relatività la simultaneità non è un concetto assoluto ma relativo ad un determinato sistema di riferimento galileiano;

2) ogni azione fisica si propaga con velocità finita, non superiore alla velocità della luce nel vuoto, perciò non sono ammesse azioni immediate a distanza.

Dunque:
III - 3ª Legge della dinamica relativistica:
la terza legge della dinamica, nella meccanica relativistica, conserva la sua validità soltanto nel caso in cui si ha a che fare con **azioni dirette o mediate**, *tanto nel caso di contatto diretto di particelle, quanto nel caso di interazione tra una particella ed un campo di forze, ovvero tra campo e campo.*
Essa è dunque certamente valida nell'istante in cui due particelle vengono in collisione.
Tornando alla definizione data di quadriforza, è opportuno precisare che in relatività, questa contiene in sé tutte le possibili azioni (meccaniche, elettromagnetiche, termiche, ecc.) che possono agire sulla particella, e non solo quelle puramente meccaniche come si assume nella meccanica newtoniana.
Si noti inoltre, fatto questo molto importante, che la (6.29) non è equivalente alla (6.30); l'equivalenza fra queste due forme in meccanica classica, quindi, non si trasporta in meccanica relativistica. L'esperienza mostra, in ogni caso, che la definizione corretta è la (6.29)[10]. Concludiamo questo paragrafo con le seguenti due utili Osservazioni.
Osservazione 1: Si noti che, in relatività, la quadriforza **F** agente su di una particella priva di struttura interna ($m_0 = cost.$) non può essere qualsiasi ma, come **A**, è vincolata alla condizione di essere ortogonale ad **U**:

$$\mathbf{F}\cdot\mathbf{U} = \eta_{ik}F^iU^k = 0. \tag{6.31}$$

P^i ed F^i si trasformeranno così $P^i = \frac{\partial x^i}{\partial x^{i'}}P^{i'}$ ed $F^i = \frac{\partial x^i}{\partial x^{i'}}F^{i'}$. Perciò, tenuto conto che gli elementi della matrice $\left(\frac{\partial x^i}{\partial x^{i'}}\right)$ sono costanti e perciò indipendenti da τ, avremo $\frac{d}{d\tau}\left(\frac{\partial x^i}{\partial x^{i'}}P^{i'}\right) = \frac{\partial x^i}{\partial x^{i'}}F^{i'}$, da cui $\frac{\partial x^i}{\partial x^{i'}}\left(\frac{dP^{i'}}{d\tau} - F^{i'}\right) = 0$, ed essendo le trasformazioni di Lorentz non degeneri, cioè $det\left(\frac{\partial x^i}{\partial x^{i'}}\right) \neq 0$, otteniamo $\frac{dP^{i'}}{d\tau} = F^{i'}$, *c.d.d.*

[10] D'altra parte, anche in meccanica classica, quando entrano in gioco masse variabili col tempo, è valida la $\mathbf{f} = d\mathbf{p}/dt$ e non la $\mathbf{f} = m_0\mathbf{a}$.

ne segue che le quattro componenti di **F** non sono tra loro indipendenti e che la "legge" della quadriforza non può a priori essere qualunque ma sarà comunque vincolata al rispetto della condizione (6.31).
Osservazione 2: Nella legge del moto (6.29) o (6.30) il 4-vettore **P** o il 4-vettore **A** hanno carattere intrinseco in $\mathcal{M}^4$. Se si vuole che la legge medesima soddisfi al principio di relatività occorre naturalmente che anche la legge della 4-forza **F** (cioè la sua dipendenza dal punto di universo, dalla 4-velocità della particella e da τ) sia espressa mediante elementi intrinseci di $\mathcal{M}^4$, inoltre essa dovrà soddisfare identicamente alla condizione (6.31).

6.3.2 Formulazione relativa 3-dimensionale della dinamica di una particella

L'equazione (6.29) costituisce la formulazione assoluta 4-dimensionale della legge del moto. Si pone ora il problema di formulare la legge fondamentale della dinamica di una particella in senso relativo 3-dimensionale cioè in ogni riferimento galileiano. Ciò non solo allo scopo fondamentale di ottenere relazioni che siano controllabili in un fissato riferimento galileiano, ma anche per confrontare direttamente le equazioni relativistiche con quelle classiche. Per il principio einsteiniano di relatività, **tali equazioni dovranno essere formalmente invarianti nel passaggio da un sistema di riferimento galileiano ad un altro**; cambieranno invece i valori delle singole grandezze coinvolte, diversamente da quanto accade nella formulazione 4-dimensionale.
Ai fini della formulazione 3-dimensionale della dinamica, scriviamo separatamente la componente temporale e le componenti spaziali della legge di moto (6.29):

$$\begin{cases} F^0 &= \dfrac{dP^0}{dt}\dfrac{dt}{d\tau} = \dfrac{d(mc)}{dt}\gamma(u), \\ F^\alpha &= \dfrac{dP^\alpha}{dt}\dfrac{dt}{d\tau} = \dfrac{dp^\alpha}{dt}\gamma(u), \qquad (\alpha = 1,2,3) \end{cases} \tag{6.32}$$

dove la variabile temporale t rappresenta il tempo relativo (detto anche *tempo coordinato*) ad un determinato sistema inerziale $\mathcal{R}$.
Osserviamo che la seconda delle (6.32) può essere riscritta così

$$\frac{dp^\alpha}{dt} = \frac{F^\alpha}{\gamma(u)}, \qquad (\alpha = 1,2,3) \tag{6.33}$$

e perciò interpretabile come una relazione 3-dimensionale. Introducendo allora il 3-vettore **f**, che chiameremo **3-forza relativa** al riferimento galileiano considerato, di componenti

$$f^\alpha = \frac{F^\alpha}{\gamma(u)}, \qquad (\alpha = 1,2,3) \tag{6.34}$$

tenendo conto che $p^\alpha = m_0\gamma(u)u^\alpha = mu^\alpha$ sono le componenti della 3-quantità di moto **p** della particella, possiamo interpretare la (6.33) come sua equazione di moto nel riferimento galileiano in cui tale particella si muove con velocità **u**, cioè

$$\frac{d(m\mathbf{u})}{dt} = \mathbf{f}. \tag{6.35}$$

Si noti che la (6.35) esprime, in forma *relativa* al riferimento galileiano considerato, **la seconda legge della dinamica relativistica** da alcuni autori chiamata *legge o equazione di Minkowski* sebbene sia più corretto chiamare **equazione di Minkowski l'equazione 4-dimensionale** (6.29). Si badi bene che la (6.35) è soltanto formalmente identica alla sua analoga newtoniana, in quanto la quantità di moto relativistica non è definita come quella newtoniana e cioè come il prodotto

della massa propria per la velocità della particella ma differisce da essa per la presenza del fattore di Lorentz: $\mathbf{p} = m_0\gamma(u)\mathbf{u}$. Inoltre, c'è un'altra differenza, meno evidente, con la legge newtoniana. Tale differenza riguarda il modo con cui si trasforma la forza nel passaggio da un sistema galileiano ad un altro. Affinché la (6.35) sia consistente col principio di relatività, cioè sia valida in tutti i sistemi galileiani, è necessario che, rispetto ad una trasformazione di Lorentz, $\mathbf{f}$ cambi nello stesso modo con cui cambia il primo membro $\frac{d(m\mathbf{u})}{dt}$. Le proprietà di trasformazione dei due membri della (6.35) verranno studiate un pò più avanti. Vedremo che il principio di relatività einsteiniano implica che le componenti della forza nell'equazione di moto (6.35) si trasformano in modo alquanto diverso dalle componenti della forza che compare nella legge di Newton, che secondo la relatività galileiana non cambiano affatto.
La legge fondamentale (6.35), così introdotta in maniera formale, è suffragata da una mole imponente di dati e conferme sperimentali, dalle misure di Bucherer del 1908-09 (A. H. Bucherer, *Ann.Physik*, **28**, 513 (1909)) alla moderna fisica degli acceleratori.
Occupiamoci ora della interpretazione della 1ª delle (6.32), cioè della componente temporale della 4-forza. **Nel caso di una particella priva di struttura interna** (m_0 = costante), la (6.27) può essere riscritta come:

$$\mathbf{F} = m_0\mathbf{A}. \tag{6.36}$$

In questo caso le quattro componenti della quadriforza non sono tra loro indipendenti; infatti, stante la (6.31), risulta che $\mathbf{F}$ è ortogonale ad $\mathbf{U}$, dato che lo è $\mathbf{A}$. In termini espliciti, quindi:

$$\mathbf{F}\cdot\mathbf{U} = F^0U^0 - \sum_{\alpha=1}^{3} F^\alpha U^\alpha = \frac{d(mc)}{dt}\gamma_u(\gamma_u c) - \sum_{\alpha=1}^{3} f^\alpha\gamma_u(\gamma_u u^\alpha) = 0.$$

Dall'ultima uguaglianza, dividendo per γ_u^2 e portando la costante c dentro il segno di derivata si ottiene:

$$\frac{d(mc^2)}{dt} = \mathbf{f}\cdot\mathbf{u},$$

dove $\mathbf{f}\cdot\mathbf{u}$ non è altro che **la potenza istantanea della 3- forza f, cioè il lavoro che essa compie nell'unità di tempo**. Tenendo conto della definizione (6.17) che abbiamo dato di energia totale della particella, possiamo scrivere la relazione precedente così:

$$\frac{dE}{dt} = \mathbf{f}\cdot\mathbf{u}. \tag{6.37}$$

La (6.37) rappresenta il cosiddetto **teorema dell'energia**:
La derivata temporale dell'energia totale di una particella è uguale alla potenza istantanea della forza ad essa applicata.
Possiamo dunque concludere che: a partire dalla formulazione 4-dimensionale della dinamica, con le definizioni introdotte e con la legge di moto (6.29) adottata, è resa possibile, la ricostituzione formale delle equazioni classiche della quantità di moto (o 2^a equazione fondamentale della dinamica relativistica) (la (6.35)) e dell'energia (la (6.37))

$$\begin{cases} \dfrac{d\mathbf{p}}{dt} = \mathbf{f} & (\textit{Teorema della quantità di moto}) \\[2ex] \dfrac{dE}{dt} = \mathbf{f}\cdot\mathbf{u} & (\textit{Teorema dell'energia}). \end{cases} \tag{6.38}$$

Naturalmente, quando sono in gioco piccole velocità tali equazioni si identificano completamente, e non soltanto formalmente, con le equazioni classiche. Infatti, in tal caso, tanto le p^α quanto E assumono (la seconda a meno di una costante additiva che scompare nella derivazione rispetto al tempo del teorema dell'energia) la forma classica.

6.3.3 Teorema dell'energia

Come già in meccanica classica, così anche in meccanica relativistica il teorema dell'energia è una conseguenza della 2ª legge fondamentale della dinamica relativistica, almeno per le particelle prive di struttura interna $m_0 = cost.$. Se infatti si moltiplicano entrambi i membri della prima delle (6.38) scalarmente per $\mathbf{u}$, si ottiene:

$$\frac{d\mathbf{p}}{dt} \cdot \mathbf{u} = \mathbf{f} \cdot \mathbf{u},$$

cioè

$$\sum_{\alpha=1}^{3} \frac{d(mu^\alpha)}{dt} u^\alpha = \mathbf{f} \cdot \mathbf{u},$$

e sviluppando la derivata al 1° membro, otteniamo

$$\sum_{\alpha=1}^{3} \left[m_0 \frac{d\gamma}{dt} u^\alpha + m_0 \gamma \frac{du^\alpha}{dt} \right] u^\alpha = \mathbf{f} \cdot \mathbf{u}.$$

Trasformiamo ora la sommatoria che sta al primo membro della precedente uguaglianza, tenendo conto della (5.65), cioè della $\frac{d\gamma}{dt} = \gamma^3 \frac{\mathbf{u} \cdot \mathbf{a}}{c^2}$, e della identità $\gamma^2 \frac{u^2}{c^2} + 1 = \gamma^2$, avremo:

$$m_0 \gamma^3 \frac{\mathbf{u} \cdot \mathbf{a}}{c^2} \sum_{\alpha=1}^{3} (u^\alpha)^2 + m_0 \gamma \sum_{\alpha=1}^{3} u^\alpha a^\alpha = m_0 \gamma^3 \frac{\mathbf{u} \cdot \mathbf{a}}{c^2} u^2 + m_0 \gamma \, \mathbf{u} \cdot \mathbf{a} =$$
$$= m_0 \gamma \, \mathbf{u} \cdot \mathbf{a} \left[\gamma^2 \frac{u^2}{c^2} + 1 \right] = m_0 \gamma^3 \, \mathbf{u} \cdot \mathbf{a} = m_0 c^2 \frac{d\gamma}{dt} = \frac{d(m_0 c^2 \gamma)}{dt} = \frac{d(mc^2)}{dt} = \frac{dE}{dt},$$

ciò dimostra il teorema dell'energia c.d.d.
Val la pena sottolineare ancora che questa dipendenza del teorema dell'energia dalle equazioni del moto è subordinata essenzialmente all'ipotesi che la massa propria m_0 sia una costante e cioè che la particella considerata sia priva di struttura interna.
Nel caso che la particella sia dotata di struttura interna e perciò $m_0 \neq cost.$, sono ancora valide le due equazioni (6.38) ma esse saranno indipendenti[11].
Osservazione:
Facciamo ora una osservazione tanto semplice quanto importante. Se integriamo, rispetto al tempo, primo e secondo membro dell'equazione che esprime il teorema dell'energia, questo può scriversi anche in termini finiti; supponendo che all'istante $t = 0$ la particella sia in quiete, avremo

$$E - E_0 = \int_0^t \mathbf{f} \cdot \mathbf{u} \, dt. \tag{6.39}$$

Da questa relazione è facile dedurre che se $u \to c$, nel qual caso $\gamma(u) \to \infty$ e di conseguenza anche $E \to \infty$, il secondo membro deve tendere anch'esso all'infinito. In altre parole, affinché la particella possa raggiungere la velocità della luce occorrerebbe esercitare su di essa un lavoro infinito, o il che è lo stesso, somministrargli un'energia infinita, ma ciò non è fisicamente possibile. **Questa semplice osservazione costituisce una spiegazione dinamica dell'irraggiungibilità della velocità della luce per una particella materiale**.

6.3.4 Energia cinetica relativistica ed energia totale dedotte dal teorema dell'energia

In questo paragrafo desideriamo reintrodurre il concetto di energia cinetica relativistica e di energia totale di una particella, in maniera diversa da quel che abbiamo fatto, in modo piuttosto

[11] Delle **particelle con struttura interna** daremo un cenno nell'ultimo paragrafo di questo capitolo.

formale, nel paragrafo 6.2.4. Precisamente ci faremo guidare, *"in modo sensato alla Galilei"*, dall'analogia newtoniana e constateremo di ottenere gli stessi risultati del paragrafo 6.2.4. Ciò costituirà un utile confronto fra i due approcci e, nello stesso tempo, avremo una comprensione più profonda dei concetti fondamentali di energia cinetica e di energia totale di una particella in ambito relativistico.
Nella meccanica classica di una particella, com'è noto, vale il **teorema dell'energia cinetica (o del lavoro)**:

$$dT = dL \iff \frac{dT}{dt} = \mathbf{f} \cdot \mathbf{u}, \tag{6.40}$$

dove dT rappresenta l'incremento che subisce l'energia cinetica $T = \frac{1}{2}m_0u^2$ di una particella in un intervallo infinitesimo di tempo dt, e $dL = \mathbf{f} \cdot \mathbf{u}\, dt$ è il lavoro corrispondente compiuto dalla forza $\mathbf{f}$ agente su di essa per un suo generico spostamento elementare $dP = \mathbf{u}\, dt$. Con $\mathbf{u}$ abbiamo indicato la velocità della particella e con $\mathbf{f} \cdot \mathbf{u}$ la **potenza istantanea** della forza $\mathbf{f}$ ad essa applicata. La (6.40) può essere scritta in termini finiti integrando 1° e 2° membro così:

$$T - T_0 = \int_{t_0}^{t} \mathbf{f} \cdot \mathbf{u}\, dt = L(t_0, t), \tag{6.41}$$

dove con $L(t_0, t)$ abbiamo indicato il lavoro compiuto dalla forza $\mathbf{f}$.
Se all'istante iniziale $t_0 = 0$ si suppone che la particella sia a riposo (cioè con velocità nulla), la sua energia cinetica è nulla per definizione ($T_0 = 0$), sicché la (6.41) può interpretarsi così:
Teorema: *L'energia cinetica T di una particella, in un generico istante t, è uguale al lavoro che deve compiere la forza $\mathbf{f}$ per portare la particella dalla velocità nulla alla velocità $\mathbf{u}$, all'istante t.*
Possiamo allora definire l'energia cinetica, anziché come al solito con la $T = \frac{1}{2}m_0u^2$, utilizzando il teorema precedente, così:[12]

$$T_{class} \stackrel{\text{def}}{=} \int_0^t \mathbf{f} \cdot \mathbf{u}\, dt. \tag{6.42}$$

Per motivi che ora vedremo, in meccanica relativistica conviene definire **l'energia cinetica** a partire proprio dal teorema del lavoro, cioè in analogia con la (6.42).
Definiamo allora "energia cinetica relativistica" di una particella il lavoro che la 3-forza $\mathbf{f}$, applicata alla particella, deve compiere per portarla da velocità nulla alla velocità $\mathbf{u}$, cioè:

$$T_{rel} \stackrel{\text{def}}{=} \int_0^t \mathbf{f} \cdot \mathbf{u}\, dt \iff \frac{dT_{rel}}{dt} = \mathbf{f} \cdot \mathbf{u}. \tag{6.43}$$

Per ottenere un'espressione esplicita di T_{rel} in funzione del modulo u della velocità della particella, basta tener conto dell'equazione di moto relativistica (6.35), cioè della $\frac{d(m\mathbf{u})}{dt} = \mathbf{f}$ che inserita nel secondo membro della seconda formula di (6.43), ci da:

$$\frac{dT_{rel}}{dt} = \frac{d(m\mathbf{u})}{dt} \cdot \mathbf{u} = m_0\Big(\frac{d\gamma}{dt}\mathbf{u} \cdot \mathbf{u} + \gamma\frac{d\mathbf{u}}{dt} \cdot \mathbf{u}\Big) = m_0\Big(\gamma^3\frac{\mathbf{a} \cdot \mathbf{u}}{c^2}u^2 + \gamma\mathbf{a} \cdot \mathbf{u}\Big) = m_0\gamma\mathbf{a} \cdot \mathbf{u}\Big(\gamma^2\frac{u^2}{c^2} + 1\Big)$$
$$= m_0\gamma^3\mathbf{a} \cdot \mathbf{u} = m_0c^2\frac{d\gamma}{dt} = \frac{d}{dt}(mc^2).$$

Abbiamo perciò ottenuto l'importantissima formula

$$\frac{dT_{rel}}{dt} = \frac{d}{dt}(mc^2) \tag{6.44}$$

[12]Infatti se consideriamo la (6.42) come definizione dell'energia cinetica, avremo: $T_{class} = \int_0^t \mathbf{f} \cdot \mathbf{u}\, dt = \int_0^t m_0\mathbf{a} \cdot \mathbf{u}\, dt = \int_0^t \frac{d}{dt}\big(\frac{1}{2}m_0\mathbf{u} \cdot \mathbf{u}\big)\, dt = \int_0^t \frac{d}{dt}\big(\frac{1}{2}m_0u^2\big)\, dt = \frac{1}{2}m_0u^2 - \frac{1}{2}m_0u_0^2$, e se $u_0 = 0$ si avrà l'espressione consueta dell'energia cinetica $T_{class} = \frac{1}{2}m_0u^2$.

nella quale la quantità mc^2 ha le dimensioni di un'energia e la indicheremo perciò, come abbiamo già fatto in precedenza, con la lettera E, ponendo

$$E = mc^2 = m_0\gamma(u)c^2. \tag{6.45}$$

La (6.44) scritta nella forma

$$\frac{dT_{rel}}{dt} = \frac{dE}{dt} \tag{6.46}$$

può essere integrata, fra l'istante $t_0 = 0$ e l'istante generico t, ottenedosi

$$\int_0^t \frac{dT_{rel}}{dt}dt = \int_0^t \frac{dE}{dt}dt,$$

cioè, tenuto conto della (6.45):

$$T_{rel} = E - E_0 = m_0c^2\,[\gamma(u) - 1\,] = m_0c^2\left[\frac{1}{\sqrt{1-\dfrac{u^2}{c^2}}} - 1\right], \tag{6.47}$$

che è la formula cercata.
Osserviamo esplicitamente che, come per tutte le formule relativistiche, quando $c \to \infty$ essa tende al valore classico $T_{class} = \dfrac{1}{2}m_0u^2$. Infatti, per $c \to \infty$ si ha $\gamma(u) \to 1$ e perciò $(\gamma(u) - 1) \to 0$, dunque $T_{rel} \to \infty \cdot 0$, assume cioè forma indeterminata, ma è subito visto che:

$$c^2(\gamma - 1) = c^2\Big(1 + \frac{1}{2}\frac{u^2}{c^2} + \frac{3}{8}\frac{u^4}{c^4} + \frac{5}{16}\frac{u^6}{c^6} + \ldots\ldots - 1\Big) = \frac{1}{2}u^2 + \frac{3}{8}\frac{u^4}{c^2} + \frac{5}{16}\frac{u^6}{c^4} + \ldots\ldots$$

Perciò se $c \to \infty$, $c^2(\gamma - 1) \to \dfrac{1}{2}u^2$ e $T_{rel} \to T_{class} = \dfrac{1}{2}m_0u^2$, siamo dunque sulla buona strada!
Per la (6.47) l'energia cinetica relativistica di una particella si presenta come differenza fra la quantità $E = mc^2$ e la quantità $E_0 = m_0c^2$, ed inoltre quando $u \to 0$, $E \to E_0$ e perciò $T \to 0$ come ci aspettiamo. Queste considerazioni suggeriscono di riguardare

$$E_0 = m_0c^2, \tag{6.48}$$

come **energia di riposo o energia di massa** della particella; cioè in relatività si presenta spontaneo questo fatto nuovo, che non ha riscontro nella meccanica classica newtoniana: **una particella in quiete, per il solo fatto di avere una massa m_0, possiede anche una energia $E_0 = m_0c^2$.** Inoltre quando la particella è in moto con velocità di modulo u essa possiede anche una energia cinetica T_{rel} e per la (6.47) la quantità

$$E = mc^2 = \frac{m_0c^2}{\sqrt{1-\dfrac{u^2}{c^2}}}, \tag{6.49}$$

sarà interpretata come **energia totale** della particella:

$$E = E_0 + T_{rel}. \tag{6.50}$$

La (6.48) dimostra la perfetta equivalenza, a meno del fattore c^2, fra la massa propria e l'energia di riposo di una particella in quiete, mentre la (6.49) dimostra l'equivalenza fra la massa relativistica e l'energia totale di una particella in moto. Le formule (6.48), (6.49), (6.50) sono considerate fra le più importanti di tutta la fisica e specialmente la (6.49) - formula emblematica universalmente nota ormai anche a livello popolare - venne considerata da

Einstein come il risultato più profondo della sua Teoria della Relatività Ristretta!
Un'importante differenza tra l'energia cinetica classica

$$T_{class} = \frac{1}{2} m_0 u^2$$

e quella relativistica

$$T_{rel} = m_0 c^2 \left[\gamma(u) - 1 \right]$$

è che la prima diverge (cioè tende ad infinito) solo se $u \to \infty$, mentre la seconda tende ad infinito per $u \to c$, il che evidenzia ancora una volta il ruolo di velocità limite svolto da c. In un esperimento condotto dal fisico italo-americano William Bertozzi al Massachussets Institute of Technology (M.I.T) nei primi anni sessanta, la relazione (6.47) che fornisce l'energia cinetica relativistica in funzione della velocità u della particella, venne sottoposta a verifica sperimentale (vedi Fig.6.6).[13] Bertozzi misurò il tempo che elettroni di diversa energia (da $T = 0.5$ Mev a $T = 15$ Mev) impiegavano a percorrere la distanza tra due rivelatori distanti circa 8.4 m, determinando in tal modo la loro velocità. I dati sperimentali rappresentati nel grafico di Fig.6.6 sono in perfetto accordo con la previsione relativistica (6.47). I quattro punti segnati sul grafico sono i dati sperimentali di Bertozzi, la curva continua rappresenta la previsione relativistica, la curva tratteggiata è la previsione non relativistica. Si noti come i dati evidenziano l'esistenza di una velocità limite pari alla velocità della luce c. Vale la pena osservare, a questo proposito, che

Figura 6.6: Relazione tra energia cinetica T e velocità u di elettroni relativistici: esperimento di Bertozzi (1964)

tutta la moderna ingegneria degli acceleratori di particelle è basata sulla dinamica relativistica. Il corretto funzionamento di queste macchine rappresenta dunque una continua verifica delle leggi einsteiniane.
Un'ultima osservazione: combinando le definizioni di impulso (6.3) e di energia totale (6.49) di una particella si ricava la sua velocità $\mathbf{u}$ in funzione dell'energia E e dell'impulso $\mathbf{p}$:

$$\mathbf{u} = \frac{c^2}{E} \mathbf{p}, \tag{6.51}$$

ovvero possiamo esprimere la quantità di moto della particella in funzione della sua energia totale e della sua velocità:

$$\mathbf{p} = \frac{E}{c^2} \mathbf{u}, \tag{6.52}$$

relazioni che si rivelano molto utili in svariate applicazioni.

[13] I risultati dell'esperimento vennero pubblicati in: W. Bertozzi, *"Speed and Kinetic Energy of Relativistic Electrons"*, **Am. Journ. Phys., 32**, 1964, p.551.

6.3.5 Relazione tra energia e momento: relazione di mass shell

La precedente equazione (6.52) costituisce ovviamente una relazione fra il momento **p** e l'energia totale E di una particella libera con l'intervento della sua velocità **u**. Esiste però, fra tali grandezze, un'altra un'importantissima relazione, molto utile nelle applicazioni, detta **relazione di mass shell** che si può ottenere facilmente nel modo seguente. Riscriviamo la (6.49) così $\frac{E}{c} = m_0\gamma c$; se ora ne eleviamo ambo i membri al quadrato e la stessa cosa la facciamo sulla relazione $p = m_0\gamma u$ che definisce il momento, e poi sottraiamo membro a membro le relazioni ottenute, si avrà:

$$\frac{E^2}{c^2} - p^2 = m_0^2\gamma^2c^2 - m_0^2\gamma^2u^2 = m_0^2\gamma^2(c^2 - u^2) = m_0^2c^2, \tag{6.53}$$

ovvero

$$E^2 = p^2c^2 + m_0^2c^4 \iff E^2 = (pc)^2 + E_0^2, \tag{6.54}$$

da cui

$$E = c\sqrt{p^2 + m_0^2c^2}. \tag{6.55}$$

Una qualunque delle formule equivalenti (6.54) e (6.55) viene chiamata **relazione di mass shell** ed ha una grandissima importanza nelle applicazioni alla fisica delle particelle di alta energia[14]
Se ora e nel seguito, l'energia cinetica relativistica T_{rel}, la indicheremo semplicemente con T, la seconda delle relazioni (6.54) può essere messa in funzione dell'energia cinetica e del momento sotto la forma $(T + E_0)^2 = (pc)^2 + E_0^2$ che può semplificarsi nella $\frac{T^2}{c^2} + 2m_0T = p^2$; per $c \to \infty$ il primo termine della somma tende a zero, perciò l'espressione stessa si riduce, come già abbiamo visto in precedenza, alla forma newtoniana $T_{class} = \frac{p^2}{2m_0} = \frac{1}{2}m_0u^2$.
Otteniamo infine un'altra utile relazione derivando la (6.55) rispetto a p, avremo:

$$\frac{dE}{dp} = \frac{cp}{\sqrt{m_0^2c^2 + p^2}} = \frac{c^2p}{c\sqrt{m_0^2c^2 + p^2}} = \frac{c^2p}{E},$$

che, essendo $E = m_0\gamma c^2$ e $p = m_0\gamma u$, in definitiva si può scrivere così:

$$\frac{dE}{dp} = u, \tag{6.56}$$

cioè *la velocità di una particella si può ottenere come derivata della sua energia totale rispetto al suo momento*.
Osserviamo ancora che la relazione di mass shell (6.54) ha la stessa forma in tutti i sistemi di riferimento inerziali: ciò è ovvio se la si scrive nella forma $E^2 - c^2p^2 = E_0^2$ in quanto E_0^2 è palesemente un **invariante** nel passaggio da un sistema di riferimento inerziale ad un altro. Ciò può vedersi anche in maniera più formale così: dati due sistemi inerziali $\mathcal{R}$ e $\mathcal{R}'$, nei quali una particella di massa propria m_0 ha momento ed energia rispettivamente, **p**, E e **p**$'$, E', si ha

$$E^2 - c^2p^2 = E'^2 - c^2p'^2 = E_0^2 = m_0^2c^4 \quad (\textbf{Invariante di Lorentz}). \tag{6.57}$$

Dalla (6.55) possiamo esprimere il momento p in funzione dell'energia totale E,

$$p = \frac{1}{c}\sqrt{E^2 - m_0^2c^4}, \tag{6.58}$$

[14]Per ricordare **la relazione di mass shell** si può ricorrere all'ausilio mnemonico di un triangolo rettangolo la cui ipotenusa sia rappresentata dall'energia totale $E = mc^2$ della particella, ed i due cateti siano rappresentati rispettivamente da pc e da $E_0 = m_0c^2$.

o dell'energia cinetica $T = E - m_0c^2$,

$$p = \frac{1}{c}\sqrt{T(T + 2m_0c^2)}. \tag{6.59}$$

Per grandi valori di E, cioè per $E \gg m_0c^2$, risulta $m_0^2c^4 = E_0^2$ trascurabile rispetto ad E^2, e dalla (6.58) si vede che il momento cresce linearmente con l'energia.
Osservazione:
La relazione di mass shell (6.54) si può ottenere in maniera ancora più semplice ragionando su grandezze 4-dimensionali, precisamente basta calcolare il quadrato del 4-momento $\mathbf{P} = m_0\mathbf{U}$ in due modi diversi. Infatti, tenuto conto che $\mathbf{U} \cdot \mathbf{U} = c^2$ si ha

$$\mathbf{P}^2 = \mathbf{P} \cdot \mathbf{P} = (m_0\mathbf{U}) \cdot (m_0\mathbf{U}) = m_0^2\mathbf{U}^2 = m_0^2c^2. \tag{6.60}$$

D'altra parte, essendo l'energia totale $E = mc^2$ è $\dfrac{E}{c} = mc$, e ricordando che $P^0 = m_0U^0$, avremo che il 4-momento è esprimibile anche così

$$\mathbf{P} = \left(P^0, P^\alpha\right) = \left(m_0U^0, p^\alpha\right) = \left(m_0c\gamma, p^\alpha\right) = \left(mc, \mathbf{p}\right) = \left(\frac{E}{c}, \mathbf{p}\right).$$

Dunque il quadrato di $\mathbf{P}$ si può scrivere anche nella forma

$$\mathbf{P}^2 = \frac{E^2}{c^2} - p^2. \tag{6.61}$$

Dal confronto delle due relazioni (6.60), (6.61) avremo

$$m_0^2c^2 = \frac{E^2}{c^2} - p^2,$$

da cui si ottiene appunto la relazione di mass shell:

$$E^2 = (pc)^2 + E_0^2.$$

6.3.6 Particelle di massa propria nulla

Spesso si sente parlare di particelle di massa nulla, ma è veramente possibile che esistano tali particelle? E com'è possibile, almeno dal punto di vista teorico e concettuale, rendersi conto di tale possibilità? In effetti è un'importante peculiarità della teoria della relatività ristretta il fatto che essa permette l'esistenza di **particelle di massa propria nulla**. Ora se noi consideriamo le due formule fondamentali che ci danno il momento e l'energia di una particella

$$\mathbf{p} = \frac{m_0\mathbf{u}}{\sqrt{1 - \dfrac{u^2}{c^2}}}, \qquad E = \frac{m_0c^2}{\sqrt{1 - \dfrac{u^2}{c^2}}}, \tag{6.62}$$

e poniamo in esse $m_0 = 0$ sembrerebbe che debba essere $\mathbf{p} = 0$ ed $E = 0$ almeno finché $u < c$, ma se contemporaneamente ad $m_0 = 0$ si prende anche $u = c$ le due espressioni si presentano sotto forma indeterminata, è allora possibile, almeno dal punto di vista matematico, che se si calcola il doppio limite per $m_0 \to 0$ e $u \to c$ esse possano tendere a valori finiti. Ciò è proprio quello che si verifica, infatti se consideriamo la relazione di mass shell $E = c\sqrt{p^2 + m_0^2c^2}$ stabilita nel paragrafo precedente e poniamo in essa $m_0 = 0$ avremo un'espressione del tutto sensata

$$E = pc. \tag{6.63}$$

La (6.63) *è la relazione energia-momento per una particella di massa propria nulla.* Combinando quest'ultima relazione con la (6.51), di cui si prendano i valori assoluti del 1° e 2° membro, cioè $u = \frac{c^2}{E}p$, ne segue che *deve essere necessariamente* $u = c$. Se tentassimo di porre $m_0 = 0$ nella relazione non relativistica tra l'energia E, che in tal caso si identifica con T, ed il momento p, cioè nella $T = p^2/2m_0$, quest'ultima relazione perderebbe senso.
Dal punto di vista sperimentale possiamo dire che in natura esistono effettivamente delle particelle di massa nulla, un esempio notevole è rappresentato dai **fotoni** che sono i quanti della radiazione elettromagnetica, infatti la (6.63), che abbiamo dedotto in ambito relativistico, è esattamente la stessa che si ottiene nell'ambito dell'elettromagnetismo classico[15].
Rigorosamente parlando, tuttavia, la locuzione *"massa a riposo nulla"* è un po' equivoca in quanto è impossibile trovare un sistema di riferimento in cui i fotoni (o qualunque cosa che viaggi alla velocità della luce) siano in riposo. Comunque, se m_0 è determinata da misure di energia ed impulso secondo la relazione $m_0 = \sqrt{(E/c^2)^2 - (p/c)^2}$, si avrà $m_0 = 0$ quando, come per il fotone, $p = E/c$.
Il risultato, che una particella di massa propria nulla possa avere un'energia ed una quantità di moto finite e che tali particelle debbano muoversi alla velocità della luce è anche consistente con il significato che abbiamo dato alla massa propria come energia interna. Poiché se la massa propria è energia interna, che esiste quando un corpo è in riposo, un "corpo" con massa a riposo nulla non ha energia interna. La sua energia è tutta esterna, essendo legata al suo moto nello spazio.
Se un corpo di questo tipo si muovesse ad una velocità inferiore alla velocità c della luce in un sistema di riferimento, potremmo sempre trovare un altro sistema di riferimento in cui esso è in riposo. Ma se esso si muove alla velocità c in un sistema di riferimento, si muoverà a questa stessa velocità in tutti i sistemi di riferimento. È quindi coerente con le trasformazioni di Lorentz che un corpo di massa propria nulla si debba muovere alla velocità della luce e non ammetta perciò alcun sistema di riposo.
Il concetto di fotone, per spiegare **l'effetto fotoelettrico**, fu introdotto da Einstein nel 1905 (il cosidetto *"Annus Mirabilis"* cioè l'anno mirabile in cui il grande genio **formulò la teoria della relatività ristretta, scoprì l'equivalenza fra massa ed energia e fece anche la scoperta del moto browniano**). L'effetto fotoelettrico consiste nel fatto che l'energia degli elettroni emessi da una superficie metallica irradiata con raggi X dipende dalla frequenza e non dall'intensità della radiazione incidente, secondo la celebre formula di Einstein

$$E = h\nu, \tag{6.64}$$

dove ν è la frequenza della radiazione incidente ed $h = 6,626 \times 10^{-34} J \cdot s$ la **costante di Planck**. Per le (6.63), (6.64) il momento di un fotone di frequenza ν vale

$$p = \frac{E}{c} = \frac{h\nu}{c}. \tag{6.65}$$

Qualche anno dopo la scoperta dell'effetto fotoelettrico, Compton fece ricorso all'ipotesi dei fotoni per interpretare i dati della diffusione di raggi X da parte di elettroni liberi. Il fatto che i fotoni abbiano massa propria nulla non è fortuito ma legato ad un'importante simmetria

[15]Per chi non dovesse ricordare che in elettromagnetismo si dimostra che la quantità di moto di un fotone è legata alla sua energia dalla relazione $p = E/c$, può ottenere questo risultato ragionando, come abbiamo fatto più sopra, direttamente sull'energia totale relativistica $E = mc^2/\sqrt{1 - u^2/c^2}$. Questa implica che $E \to \infty$ se $u \to c$ a meno che contemporaneamente si abbia $m_0 \to 0$. Perciò i fotoni, che per definizione hanno $u = c$, devono avere $m_0 = 0$. Quindi dalla relazione di mass shell $E = c\sqrt{p^2 + m_0^2c^2}$ segue che i fotoni devono soddisfare alla relazione $p = E/c$. Il fatto che questo risultato sia stato trovato in modo indipendente nell'elettromagnetismo classico dimostra la compatibilità fra quest'ultimo e la relatività ristretta.

dell'elettrodinamica, la cosiddetta *invarianza di gauge*.
Ci sono in fisica altre particelle di massa propria nulla: fino a poco tempo fa erano considerati tali anche i **neutrini**. Essi furono ipotizzati da Wolfgang Pauli nel 1931 per risolvere un'apparente anomalia nel decadimento beta dei nuclei e scoperti sperimentalmente molto più tardi nel 1956 da Cowan e Reines. Vi sono oggi alcune evidenze sperimentali a favore di una loro massa propria piccolissima ma diversa da zero. Ciò si è potuto sperimentare in maniera inequivocabile con l'esperimento *Super-Kamiokande*, si veda Y. Fukuda e altri, in Phys. Rev. Lett., 81, 1998, pag. 1562. Nell'ambito del cosiddetto **modello standard** delle particelle elementari anche i **gluoni** sono considerate particelle di massa propria nulla.

6.3.7 Relazione tra forza, accelerazione e velocità di una particella nello spazio fisico

Esaminiamo, in questo paragrafo, il legame esistente tra la 3-accelerazione di una particella e la 3-forza che la sollecita, nell'ipotesi consueta che la particella sia priva di struttura interna, cioè con $m_0 = cost$. La 3-forza, come già visto, è data da

$$\mathbf{f} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m_0\gamma\mathbf{u}) = m_0\frac{d\gamma}{dt}\mathbf{u} + m_0\gamma\frac{d\mathbf{u}}{dt}. \tag{6.66}$$

Il termine $\frac{d\gamma}{dt}$, che figura nella (6.66), è già stato calcolato, alla fine del Cap.5 con la formula (5.65), cioè $\frac{d\gamma}{dt} = \gamma^3\frac{\mathbf{a}\cdot\mathbf{u}}{c^2}$, ma ora intendiamo esprimere tale derivata del fattore di Lorentz rispetto al tempo in maniera diversa, nel modo seguente.
Ricordando il teorema dell'energia (6.37) cioè $\frac{dE}{dt} = \mathbf{f}\cdot\mathbf{u}$ e che $E = m_0\gamma c^2$, avremo

$$\frac{d(m_0\gamma c^2)}{dt} = \mathbf{f}\cdot\mathbf{u}$$

che ci dà

$$\frac{d\gamma}{dt} = \frac{\mathbf{f}\cdot\mathbf{u}}{m_0c^2}. \tag{6.67}$$

Sostituendo quest'ultimo risultato nella (6.66), otteniamo

$$\mathbf{f} = m\mathbf{a} + \frac{\mathbf{f}\cdot\mathbf{u}}{c^2}\mathbf{u}, \tag{6.68}$$

avendo indicato con $\mathbf{a} = d\mathbf{u}/dt$ l'accelerazione della particella. La (6.68) può essere riscritta isolando il termine contenente l'accelerazione così

$$m\mathbf{a} = \mathbf{f} - \frac{\mathbf{f}\cdot\mathbf{u}}{c^2}\mathbf{u}, \tag{6.69}$$

nella quale $m = m_0\gamma$ è la massa relativistica della particella.
Le (6.68) e (6.69) sono forme alternative dell'equazione fondamentale della dinamica relativistica o *equazione di Minkowski*.
Ciò che le (6.68), (6.69) ci dicono immediatamente è che, in generale, in relatività, l'accelerazione *non* è parallela e concorde alla forza, poiché l'ultimo termine in tali espressioni è nella direzione della velocità.
Ci sono due semplici, ma interessanti casi, tuttavia, in cui l'accelerazione è parallela alla forza.
1) Un caso è quello in cui la forza $\mathbf{f}$ è parallela e concorde alla velocità $\mathbf{u}$, cioè $\mathbf{f} \parallel \mathbf{u}$, di modo che, in virtù della (6.69), $\mathbf{a}$ è parallela e concorde sia ad $\mathbf{u}$ che ad $\mathbf{f}$. In questo caso la particella

si muove lungo una linea retta. Supponiamo dunque che sia $\mathbf{u} = u\widehat{\mathbf{u}}$, con $\widehat{\mathbf{u}} = \text{vers}\,\mathbf{u}$, essendo $\mathbf{f} \parallel \mathbf{u}$ ed aventi lo stesso verso sarà $\mathbf{f} = f\widehat{\mathbf{u}}$. Sostituendo nella (6.68) avremo

$$f\widehat{\mathbf{u}} = m_0\gamma a\widehat{\mathbf{u}} + \frac{(f\widehat{\mathbf{u}} \cdot u\widehat{\mathbf{u}})}{c^2}u\widehat{\mathbf{u}}$$

che può riscriversi più semplicemente così

$$\left[f(1 - \frac{u^2}{c^2}) - m_0\gamma a\right]\widehat{\mathbf{u}} = 0$$

relazione che può mettersi nella forma più semplice

$$\mathbf{f}_{\parallel} = m_0\gamma^3\,\mathbf{a}_{\parallel}, \tag{6.70}$$

nella quale i simboli $\mathbf{f}_{\parallel}$ e $\mathbf{a}_{\parallel}$ significano che si tratta di vettori paralleli alla velocità $\mathbf{u}$ della particella.
2) Un altro caso in cui $\mathbf{a}$ è parallela e concorde ad $\mathbf{f}$ è quello in cui la forza $\mathbf{f}$ è perpendicolare alla velocità $\mathbf{u}$, $(\mathbf{f} \perp \mathbf{u})$, di modo che $\mathbf{f} \cdot \mathbf{u} = 0$ e la (6.69) diventa

$$m_0\gamma\mathbf{a} = \mathbf{f}.$$

Per ricordare che questo risultato si applica quando $\mathbf{f}$ ed $\mathbf{a}$ *sono perpendicolari alla velocità* $\mathbf{u}$, l'ultima formula la riscriveremo così:

$$\mathbf{f}_{\perp} = m_0\gamma\,\mathbf{a}_{\perp}. \tag{6.71}$$

Esempio 1: Si verifica la prima circostanza, per esempio, nel caso che una particella carica si muova in un campo elettrico uniforme ($\mathbf{E} = cost.$) e perciò la sua traiettoria sarà rettilinea; in tal caso la particella è soggetta alla forza $\mathbf{f} = q\mathbf{E}$, e questa è parallela sia alla velocità $\mathbf{u}$ che alla accelerazione $\mathbf{a}$.
Esempio 2: Si verifica il secondo caso quando una particella carica si muove in un campo magnetico uniforme, in assenza di campo elettrico ($\mathbf{E} = 0, \mathbf{B} = cost.$). In tal caso essa è soggetta alla forza di Lorentz $\mathbf{f} = q\dfrac{\mathbf{u} \wedge \mathbf{B}}{c}$ ed allora è certamente $(\mathbf{f} \perp \mathbf{u})$.

Osservazione:
la quantità $m_L = m_0\gamma^3(u)$ che figura nella (6.70), veniva chiamata, in passato, *massa longitudinale*. Mentre la quantità $m_T = m_0\gamma(u)$ che compare nella (6.71) veniva chiamata *massa trasversale*; queste denominazioni sono però attualmente cadute in disuso.

6.3.8 Trasformazione di Lorentz della 3-quantità di moto, dell'energia e della massa

Nei paragrafi precedenti abbiamo introdotto i concetti fondamentali di **massa relativistica e quantità di moto relativistica** (paragrafo 6.2.2), di **energia totale, energia di riposo ed energia cinetica** (paragrafi 6.2.4 e 6.3.4) ed abbiamo studiato **la formulazione relativa 3-dimensionale della dinamica di una particella** (paragrafo 6.3.2). Tuttavia tutto ciò che abbiamo detto in tali paragrafi vale esclusivamente in un dato, seppur generico, sistema galileiano $\mathcal{R}$.
Occupiamoci ora di vedere come si trasformano le componenti della 3-quantità di moto $\mathbf{p} = m_0\gamma(u)\mathbf{u}$ e dell'energia totale $E = m_0\gamma(u)c^2$ di una particella nel passare da un sistema galileiano $\mathcal{R}$ ad un altro $\mathcal{R}'$ in reciproca configurazione x-standard.
In un dato sistema inerziale $\mathcal{R}$, le componenti della 3-quantità di moto di una particella di massa

propria m_0 e velocità $\mathbf{u}$ e la sua energia totale sono espresse dalle relazioni:

$$\begin{cases} p_x &= m_0\gamma(u)u_x = \dfrac{m_0u_x}{\sqrt{1-u^2/c^2}} = \dfrac{m_0u_x}{\sqrt{1-\beta_u^2}}, \\ p_y &= m_0\gamma(u)u_y = \dfrac{m_0u_y}{\sqrt{1-u^2/c^2}} = \dfrac{m_0u_y}{\sqrt{1-\beta_u^2}}, \\ p_z &= m_0\gamma(u)u_z = \dfrac{m_0u_z}{\sqrt{1-u^2/c^2}} = \dfrac{m_0u_z}{\sqrt{1-\beta_u^2}}, \\ E &= m_0\gamma(u)c^2 = \dfrac{m_0c^2}{\sqrt{1-u^2/c^2}} = \dfrac{m_0c^2}{\sqrt{1-\beta_u^2}}. \end{cases} \tag{6.72}$$

dove $u^2 = u_x^2 + u_y^2 + u_z^2$ e dove, per semplificare un pò le notazioni, abbiamo scritto $\beta_u = u/c$ in luogo di $\beta(u) = u/c$.
Nel riferimento $\mathcal{R}'$ le quantità corrispondenti sono espresse da relazioni analoghe alle precedenti:

$$\begin{cases} p'_x &= m_0\gamma(u')u'_x = \dfrac{m_0u'_x}{\sqrt{1-u'^2/c^2}} = \dfrac{m_0u'_x}{\sqrt{1-\beta_{u'}^2}}, \\ p'_y &= m_0\gamma(u')u'_y = \dfrac{m_0u'_y}{\sqrt{1-u'^2/c^2}} = \dfrac{m_0u'_y}{\sqrt{1-\beta_{u'}^2}}, \\ p'_z &= m_0\gamma(u')u'_z = \dfrac{m_0u'_z}{\sqrt{1-u'^2/c^2}} = \dfrac{m_0u'_z}{\sqrt{1-\beta_{u'}^2}}, \\ E' &= m_0\gamma(u')c^2 = \dfrac{m_0c^2}{\sqrt{1-u'^2/c^2}} = \dfrac{m_0c^2}{\sqrt{1-\beta_{u'}^2}}. \end{cases} \tag{6.73}$$

dove $u'^2 = u_x'^2 + u_y'^2 + u_z'^2$ e $\beta_{u'} = u'/c$.
Se ora ricordiamo le formule relativistiche di trasformazione delle componenti della velocità $\mathbf{u}$, (4.16), (4.17), l'espressione (6.17) dell'energia totale e le formule di trasformazione del fattore di Lorentz (4.25), (4.26) e le inseriamo nelle formule precedenti, dopo facili calcoli otteniamo:

$$\begin{cases} p_x &= \gamma(v)\left(p'_x + v\dfrac{E'}{c^2}\right), \\ p_y &= p'_y, \\ p_z &= p'_z, \\ E &= \gamma(v)\left(E' + vp'_x\right). \end{cases} \tag{6.74}$$

Le relazioni inverse si ottengono, come al solito nel modo più semplice, cambiando v in $-v$, scambiando le quantità accentate con quelle non accentate e ricordando che il fattore di Lorentz

è una funzione pari, cioè $\gamma(-v) = \gamma(v)$, otteniamo:

$$\begin{cases} p'_x &= \gamma(v)\Big(p_x - v\dfrac{E}{c^2}\Big), \\ \\ p'_y &= p_y, \\ \\ p'_z &= p_z, \\ \\ E' &= \gamma(v)\Big(E - vp_x\Big). \end{cases} \tag{6.75}$$

Se si confrontano queste formule con le trasformazioni di Lorentz (3.32a), (3.32b) si rileva subito una notevole analogia. Le quantità $p_x, p_y, p_z, E/c^2$ si trasformano esattamente come le coordinate spazio-temporali x, y, z, t di una particella, o, se si vuole, $p_x, p_y, p_z, E/c$ si trasformano come le coordinate spazio-temporali x, y, z, ct. Per le formule di trasformazione della quantità di moto ciò è immediato riconoscerlo sulle formule così come sono scritte più sopra; invece per le formule di trasformazione dell'energia totale l'analogia è meno evidente, ma basta riscriverle sotto la forma equivalente

$$\frac{E}{c^2} = \gamma(v)\Big(\frac{E'}{c^2} + \frac{v}{c^2}p'_x\Big), \tag{6.76}$$

$$\frac{E'}{c^2} = \gamma(v)\Big(\frac{E}{c^2} - \frac{v}{c^2}p_x\Big), \tag{6.77}$$

perché l'analogia diventi palese. Questa analogia con le trasformazioni di Lorentz permette di ricordare facilmente sia le formule di trasformazione della quantità di moto, sia le formule di trasformazione dell'energia totale.
Osserviamo che il legame fra l'energia totale E di una particella e la sua quantità di moto $\mathbf{p}$ è un fatto che abbiamo già constatato in diverse formule fondamentali, per esempio: nella (6.51), nella relazione di mass shell (6.54), nella (6.56). Il motivo fondamentale di tale legame deriva dal fatto che il 4-vettore quantità di moto, così come è stato definito, ha la sua componente temporale che, a meno del fattore c, coincide proprio con l'energia totale della particella.
Forse la connessione più profonda fra l'energia totale e la quantità di moto, nella Relatività Ristretta, è questa: *se l'energia e la quantità di moto sono conservati in una interazione per un osservatore inerziale, essi sono necessariamente conservati, in questa interazione, per qualunque altro osservatore inerziale; inoltre se la quantità di moto è conservata, anche l'energia deve essere conservata.*
Per concludere questo paragrafo ci rimane da vedere come si trasforma la massa nel passare da un sistema inerziale $\mathcal{R}$ ad un'altro $\mathcal{R}'$ in configurazione x-standard col primo. È pressoché evidente che le equazioni di trasformazione della massa seguono direttamente dalle trasformazioni per l'energia. Cioè, essendo $E = mc^2$, con $m = m_0\gamma(u)$, l'energia di una particella di massa propria m_0 che si muove con velocità di modulo u in $\mathcal{R}$ ed essendo $E' = m'c^2$, con $m' = m_0\gamma(u')$, l'analoga espressione dell'energia della particella in $\mathcal{R}'$, è piuttosto semplice trovare la formula di trasformazione che esprime m' in funzione di m e viceversa. Tuttavia è più semplice ricavare tali espressioni, ricordando le formule di trasformazione del fattore di Lorentz. Infatti utilizzando la (4.25) avremo

$$m' = m_0\gamma(u') = m_0\gamma(u)\gamma(v)\big(1 - \frac{v}{c^2}u_x\big),$$

perciò otteniamo

$$m' = m\gamma(v)\big(1 - \frac{v}{c^2}u_x\big). \tag{6.78}$$

Analogamente, utilizzando la (4.26), si ottiene la formula inversa della precedente

$$m = m'\gamma(v)\big(1 + \frac{v}{c^2}u'_x\big). \tag{6.79}$$

Le (6.78), (6.79) sono le formule volute di trasformazione della massa.

6.3.9 Trasformazione di Lorentz della 3-forza

Occupiamoci ora di come deve trasformarsi nel passaggio da $\mathcal{R}$ ad $\mathcal{R}'$ la 3-forza $\mathbf{f}$ che figura nella *seconda legge* (6.35) *della dinamica relativistica* di una particella di massa propria m_0, legge che ora riscriviamo così

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}, \tag{6.80}$$

dove $\mathbf{p} = m\mathbf{u} = m_0\gamma(u)\mathbf{u}$ è la quantità di moto della particella.
In accordo con il Principio di Relatività, richiediamo che la (6.80) abbia la stessa forma in tutti i sistemi galileiani, perciò in $\mathcal{R}'$ si avrà una relazione analoga:

$$\frac{d\mathbf{p}'}{dt'} = \mathbf{f}', \tag{6.81}$$

ciò implica che la forza $\mathbf{f}$ dovrà trasformarsi come il vettore $\frac{d\mathbf{p}}{dt}$. Consideriamo la componente f'_x, essa, per la prima delle (6.75) è data da

$$f'_x = \frac{dp'_x}{dt'} = \gamma(v)\frac{d}{dt'}\Big(p_x - v\frac{E}{c^2}\Big). \tag{6.82}$$

Se teniamo conto della relazione che lega dt' a dt nella trasformazione di Lorentz (3.36a), scritta in forma differenziale, la derivata temporale $\frac{d(...)}{dt'}$ si trasforma nel modo seguente

$$\frac{d(...)}{dt'} = \frac{d(...)}{dt}\frac{dt}{dt'} = \frac{d(...)}{dt}\Big(\frac{dt'}{dt}\Big)^{-1} = \frac{d(...)}{dt}\Big[\gamma(v)\Big(1 - \frac{v}{c^2}u_x\Big)\Big]^{-1}. \tag{6.83}$$

In virtù della relazione precedente, la (6.82) diventa allora

$$f'_x = \frac{1}{\Big(1 - \frac{v}{c^2}u_x\Big)}\Big(\frac{dp_x}{dt} - \frac{v}{c^2}\frac{dE}{dt}\Big). \tag{6.84}$$

Considerando la componente secondo l'asse x del 1° e 2° membro della (6.80) si ha $\frac{dp_x}{dt} = f_x$ e ricordando il *teorema dell'energia* (6.37), la (6.84) diventa

$$f'_x = \frac{f_x - \frac{v}{c^2}\mathbf{f}\cdot\mathbf{u}}{1 - \frac{v}{c^2}u_x}, \tag{6.85}$$

dove $\mathbf{u}$ è la velocità in $\mathcal{R}$ della particella soggetta alla forza $\mathbf{f}$.
Per la componente f'_y, in virtù della seconda delle (6.74), si ha

$$f'_y = \frac{dp'_y}{dt'} = \frac{dp_y}{dt'} = \frac{dp_y}{dt}\frac{dt}{dt'}, \tag{6.86}$$

e per la (6.83) avremo in definitiva

$$f'_y = \frac{f_y}{\gamma(v)\Big(1 - \frac{v}{c^2}u_x\Big)}. \tag{6.87}$$

In modo analogo si ricava

$$f'_z = \frac{f_z}{\gamma(v)\left(1 - \frac{v}{c^2}u_x\right)}. \tag{6.88}$$

Raccogliendo i risultati ottenuti, le leggi di trasformazione delle componenti della forza $\mathbf{f}$ applicata ad una particella, nel passaggio da un sistema inerziale $\mathcal{R}$ ad un'altro $\mathcal{R}'$ mobile rispetto al primo con velocità costante $\mathbf{v}$ lungo l'asse $x \equiv x'$, e perciò in configurazione x-standard, sono:

$$\begin{cases} f'_x &= \dfrac{f_x - \frac{v}{c^2}\mathbf{f}\cdot\mathbf{u}}{1 - \frac{v}{c^2}u_x}, \\[2ex] f'_y &= \dfrac{f_y}{\gamma(v)\left(1 - \frac{v}{c^2}u_x\right)}, \\[2ex] f'_z &= \dfrac{f_z}{\gamma(v)\left(1 - \frac{v}{c^2}u_x\right)}. \end{cases} \tag{6.89}$$

Le relazioni inverse si ottengono, come al solito, cambiando v in $-v$, scambiando le quantità accentate con quelle non accentate e ricordando che il fattore di Lorentz è una funzione pari, cioè $\gamma(-v) = \gamma(v)$, otteniamo perciò:

$$\begin{cases} f_x &= \dfrac{f'_x + \frac{v}{c^2}\mathbf{f}'\cdot\mathbf{u}'}{1 + \frac{v}{c^2}u'_x}, \\[2ex] f_y &= \dfrac{f'_y}{\gamma(v)\left(1 + \frac{v}{c^2}u'_x\right)}, \\[2ex] f_z &= \dfrac{f'_z}{\gamma(v)\left(1 + \frac{v}{c^2}u'_x\right)}. \end{cases} \tag{6.90}$$

Le equazioni (6.89) e (6.90) sono le equazioni di trasformazione per le componenti della forza. Si può verificare molto semplicemente che queste relazioni, nel limite newtoniano $c \to \infty$, si riducono alla relazione classica $\mathbf{f} = \mathbf{f}'$, come deve essere.
Un aspetto interessante delle equazioni (6.89), (6.90) è che la forza in un riferimento $\mathcal{R}$ *è legata alla potenza sviluppata dalla sua analoga* in un altro riferimento $\mathcal{R}'$; ciò si vede bene dall'espressione di f'_x che contiene la potenza $\mathbf{f}\cdot\mathbf{u}$ di $\mathbf{f}$ in $\mathcal{R}$ e dall'analoga espressione di f_x che contiene la potenza $\mathbf{f}'\cdot\mathbf{u}'$ di $\mathbf{f}'$ in $\mathcal{R}'$. Questo fatto è analogo alla dipendenza di p_x da E', ovvero di p'_x da E come si vede dalle (6.74) e (6.75), le quali suggeriscono che come l'energia e la quantità di moto sono collegate, quando sono espresse in forma 4-dimensionale, analogo legame si ha tra forza e potenza. Infatti, non è difficile dimostrare (esercizio) che una quarta relazione, che lega la potenza alle componenti della forza, si può aggiungere alle (6.89), precisamente

$$\mathbf{f}'\cdot\mathbf{u}' = \frac{\mathbf{f}\cdot\mathbf{u} - vf_x}{1 - \frac{v}{c^2}u_x}, \tag{6.91}$$

ed analogamente si può aggiungere la relazione inversa

$$\mathbf{f} \cdot \mathbf{u} = \frac{\mathbf{f}' \cdot \mathbf{u}' + v f'_x}{1 + \frac{v}{c^2} u'_x}, \tag{6.92}$$

alle (6.90).
I ragionamenti precedenti mettono in luce il grande potere di generalizzazione e di unificazione della Teoria della Relatività. Infatti si può notare che durante lo sviluppo storico della meccanica classica, alcuni tra i fondatori *consideravano come misura significativa di una forza il suo effetto nel tempo* $f dt = dp$, mentre per altri *la quantità significativa era l'effetto della forza per lo spostamento* $f dx = dE$. *Cioè* f *può essere considerata come la derivata rispetto al tempo dell'impulso* $f = \frac{dp}{dt}$, *o come derivata rispetto alla coordinata spaziale dell'energia* $f = \frac{dE}{dx}$.
Esaminiamo infine un caso particolare delle equazioni (6.90) che ci sarà utile nel prossimo capitolo sull'elettrodinamica quando studieremo in che modo il valore numerico dei campi elettrici e magnetici dipenda dal riferimento dell'osservatore. Si consideri una particella istantaneamente in riposo nel riferimento $\mathcal{R}'$, dove è soggetta ad una forza di componenti f'_x, f'_y, f'_z. Poiché $u' = 0$, in questo riferimento proprio, le trasformazioni (6.90) per la forza diventano

$$\begin{cases} f_x &= f'_x, \\ f_y &= \dfrac{f'_y}{\gamma(v)} = f'_y \sqrt{1 - \beta_v^2}, \\ f_z &= \dfrac{f'_z}{\gamma(v)} = f'_z \sqrt{1 - \beta_v^2}, \end{cases} \tag{6.93}$$

dove $\beta_v = v/c$, ed f_x, f_y, f_z sono le componenti della forza in qualche altro riferimento $\mathcal{R}$, ovviamente non proprio. È facile constatare, per le (6.93), che *la forza nel riferimento in cui la particella è istantaneamente a riposo, è, in modulo, più grande della forza corrispondente in qualunque altro riferimento.*

6.4 Equivalenza tra massa ed energia

6.4.1 Leggi di conservazione nei fenomeni d'urto

Per velocità relativamente piccole, rispetto alla velocità della luce, la legge classica di conservazione della quantità di moto per i fenomeni d'urto si accorda ottimamente con l'esperienza. Per velocità non tanto piccole, invece, l'esperienza mostra che la legge classica di conservazione non è più valida, se si considera la definizione newtoniana di quantità di moto. Tuttavia, se la quantità di moto viene definita in modo relativistico cioè con la (6.3), l'esperienza mostra che il principio di conservazione negli urti è ancora valido.
Prendiamo ora in esame un urto anelastico. In particolare consideriamo due particelle identiche A e B, aventi entrambe massa a riposo m_0, e supponiamo che, rispetto ad un certo riferimento galileiano $\mathcal{R}'$, la situazione prima e dopo l'urto sia rappresentata nella Fig.6.7a. Prima dell'urto le particelle A e B hanno ciascuna velocità di modulo u', e sono dirette in senso opposto lungo l'asse x'. Siccome abbiamo supposto che l'urto sia anelastico, esse dopo l'urto formeranno un unico corpo C che rimarrà in quiete, come è richiesto dalla conservazione della quantità di moto, con una certa massa a riposo M_0.
In un altro sistema di riferimento $\mathcal{R}$, che si muove rispetto ad $\mathcal{R}'$ alla velocità $v = u'$ verso sinistra lungo l'asse comune x-x', il corpo composto C avrà una velocità di modulo v diretta verso

Figura 6.7: Un secondo esempio di urto anelastico fra due particelle.

destra lungo l'asse x poiché in $\mathcal{R}'$ è stazionario. In $\mathcal{R}$, il corpo A sarà stazionario prima dell'urto e il corpo B avrà una velocità u_B da calcolarsi con le formule relativistiche di composizione delle velocità. La situazione nel riferimento $\mathcal{R}$ è presentata nella Fig.6.7b.
La velocità u_B nel riferimento $\mathcal{R}$, come abbiamo detto, può essere ricavata dalle equazioni relativistiche di composizione delle velocità (4.17):

$$u_B = \frac{u' + v}{1 + \frac{v}{c^2}u'} = \frac{2u'}{1 + \frac{u'^2}{c^2}}. \tag{6.94}$$

Per quanto riguarda la massa relativistica di B, nel riferimento $\mathcal{R}$, si ha, per le (6.4), (6.94):

$$m_B = \frac{m_0}{\sqrt{1 - \frac{u_B^2}{c^2}}} = \frac{m_0}{\sqrt{1 - \frac{1}{c^2}\frac{4u'^2}{\left(1 + \frac{u'^2}{c^2}\right)^2}}} = \frac{m_0\left(1 + \frac{u'^2}{c^2}\right)}{\sqrt{\left(1 + \frac{u'^2}{c^2}\right)^2 - 4\frac{u'^2}{c^2}}},$$

da cui, in definitiva:

$$m_B = m_0 \frac{1 + \frac{u'^2}{c^2}}{1 - \frac{u'^2}{c^2}}. \tag{6.95}$$

Dopo l'urto, la massa composta C viaggia, in $\mathcal{R}$, ad una velocità $v = u'$, poiché in $\mathcal{R}'$ era stazionaria. Quindi, applicando nel riferimento $\mathcal{R}$ la legge relativistica di conservazione della quantità di moto nella direzione x (le componenti y e z sono automaticamente conservate), si ha:

$$m_B u_B + m_0 \cdot 0 = M_C v,$$

dove con $M_C = \dfrac{M_0}{\sqrt{1-\dfrac{u'^2}{c^2}}}$ si è indicata la massa relativistica del corpo C nel riferimento $\mathcal{R}$ dopo l'urto. Con $v = u'$, u_B ed m_B date, rispettivamente, dalle (6.94) e (6.95), la precedente relazione diventa:

$$m_0 \frac{1+\dfrac{u'^2}{c^2}}{1-\dfrac{u'^2}{c^2}} \cdot \frac{2u'}{1+\dfrac{u'^2}{c^2}} = \frac{M_0 u'}{\sqrt{1-\dfrac{u'^2}{c^2}}},$$

perciò:

$$M_0 = \frac{2m_0}{\sqrt{1-\dfrac{u'^2}{c^2}}} = 2m_0\gamma(u'). \tag{6.96}$$

La (6.96) ci dice che la massa a riposo del corpo formato nell'urto non è uguale alla somma delle masse a riposo dei corpi originali ($2m_0$), ma è alterata dal fattore di Lorentz $\gamma(u')$; in particolare, la differenza tra le due risulta essere:

$$\Delta m_0 = M_0 - 2m_0 = 2m_0[\gamma(u') - 1] > 0. \tag{6.97}$$

D'altra parte, prima dell'urto, in $\mathcal{R}'$, l'energia cinetica totale dei due corpi era

$$T'_A + T'_B = 2T' = 2m_0c^2[\gamma(u') - 1] = \Delta m_0 c^2, \tag{6.98}$$

dopo l'urto invece, essendo il corpo C in quiete, la sua energia cinetica è manifestamente nulla, e perciò tutta l'energia cinetica iniziale $2T'$ è scomparsa. Al suo posto, dopo l'urto, appare una qualche forma di energia interna, quale l'energia termica o l'energia di eccitazione. Vediamo ora che questa energia interna, apparsa dopo l'urto, dà luogo ad una massa a riposo del corpo C più grande della massa a riposo totale dei due corpi A e B prima dell'urto. In altre parole, la massa a riposo è equivalente all'energia (quella che abbiamo chiamato energia a riposo) e deve essere inclusa nell'applicazione del principio di conservazione dell'energia.
Da un esame delle equazioni (6.97) e (6.98) si vede chiaramente, in questo caso, che l'energia associata all'aumento della massa a riposo dopo l'urto, $\Delta m_0 c^2$, è uguale all'energia cinetica presente prima dell'urto. Possiamo dire quindi che, sebbene in un urto anelastico l'energia cinetica da sola non sia conservata, **l'energia totale è conservata**. L'energia totale comprende l'energia (di massa) a riposo più l'energia cinetica:

$$E = E_0 + T, \tag{6.99}$$

come deve essere in generale.

6.4.2 Equivalenza tra massa ed energia

Vediamo adesso di dimostrare che, **nell'urto anelastico la conservazione dell'energia totale è equivalente alla conservazione della massa relativistica**. Cominciamo col valutare l'energia totale.
Consideriamo dapprima il riferimento $\mathcal{R}'$. Prima dell'urto, l'energia totale è:

$$(E'_{tot})_{prima\ dell'urto} = E'_A + E'_B = (m_0c^2 + T') + (m_0c^2 + T') = 2m_0c^2 + 2T' =$$
$$= 2m_0c^2 + 2m_0c^2[\gamma(u') - 1],$$

cioè

$$(E'_{tot})_{prima\ dell'urto} = 2m_0c^2\gamma(u'). \tag{6.100}$$

Dopo l'urto, tenuto conto della (6.96), l'energia totale è:

$$(E'_{tot})_{dopo\ l'urto} = M_0c^2 = 2m_0c^2\gamma(u'). \tag{6.101}$$

Quindi, come si vede, nel riferimento $\mathcal{R}'$, l'energia totale è conservata nell'urto.
Consideriamo ora il riferimento $\mathcal{R}$. Prima dell'urto, l'energia totale è:

$$(E_{tot})_{prima\ dell'urto} = E_A + E_B = m_0c^2 + m_Bc^2 = m_0c^2\left[1 + \frac{1+\frac{u'^2}{c^2}}{1-\frac{u'^2}{c^2}}\right] = \frac{2m_0c^2}{1-\frac{u'^2}{c^2}},$$

cioè

$$(E_{tot})_{prima\ dell'urto} = 2m_0c^2\gamma^2(u'). \tag{6.102}$$

Dopo l'urto, tenuto conto del valore di M_C, l'energia totale è:

$$(E_{tot})_{dopo\ l'urto} = M_Cc^2 = M_0c^2\gamma(u') = 2m_0c^2\gamma^2(u'). \tag{6.103}$$

Quindi l'energia totale è conservata anche nel riferimento $\mathcal{R}$.
Valutiamo adesso la massa relativistica sia in $\mathcal{R}$ che in $\mathcal{R}'$.
Consideriamo dapprima il riferimento $\mathcal{R}'$. Prima dell'urto, la massa è:

$$m_0\gamma(u') + m_0\gamma(u') = 2m_0\gamma(u'). \tag{6.104}$$

Dopo l'urto la massa è:

$$M_0 = 2m_0\gamma(u'). \tag{6.105}$$

Quindi, la massa relativistica è conservata nel riferimento $\mathcal{R}'$. Consideriamo ora il riferimento $\mathcal{R}$. Prima dell'urto, la massa è:

$$m_0 + m_B = m_0\left[1 + \frac{1+\frac{u'^2}{c^2}}{1-\frac{u'^2}{c^2}}\right] = 2m_0\gamma^2(u'). \tag{6.106}$$

Dopo l'urto, la massa è:

$$M_C = M_0\gamma(u') = 2m_0\gamma^2(u'). \tag{6.107}$$

Quindi la massa relativistica è conservata anche nel riferimento $\mathcal{R}$.
Abbiamo così stabilito che **nei fenomeni d'urto la massa relativistica si conserva, così come l'energia totale, inoltre le due leggi di conservazione coincidono, a meno di un fattore** c^2.
c.v.d.
Dai risultati precedenti abbiamo visto che l'invarianza dell'energia implica l'invarianza della massa e viceversa. Massa ed energia sono equivalenti: esse sono una sola grandezza invariante che possiamo chiamare **massa-energia**. Moltiplicando per la costante universale c^2 relazioni che involgono solo la massa si ottengono le corrispondenti relazioni per l'energia e viceversa.
La relazione

$$E = mc^2 \tag{6.108}$$

esprime il fatto che la massa-energia può essere espressa in unità di energia E, o equivalentemente in unità di massa $m = E/c^2$. Cioè, la massa e l'energia sono equivalenti, e le loro unità differiscono di un fattore c^2. Una conveniente identità è $c^2 = (3 \times 10^8\text{m/s})^2 = 931$ Mev/u.m.a., dove u.m.a significa unità di massa atomica.

Nelle pratiche applicazioni è diventato ormai abituale esprimere le masse in termini di elettronvolt e suoi multipli, dicendo, per esempio, che la massa di un elettrone è $0,51$ Mev; ciò è molto utile nei calcoli energetici[16].
La precedente equazione è forse la legge più conosciuta di tutta la fisica. Essa è stata confermata da numerose applicazioni pratiche e conseguenze teoriche: *Einstein la considerò come la conseguenza più significativa della sua teoria della Relatività Ristretta.*
Derivando, rispetto al tempo, 1° e 2° membro della (6.108) avremo

$$\frac{dE}{dt} \equiv \frac{dT}{dt} = c^2 \frac{dm}{dt}, \tag{6.109}$$

già vista in precedenza sotto la forma leggermente diversa (6.37). Essa afferma: **una variazione dell'energia totale di una particella, coincide con l'analoga variazione della sua energia cinetica e produce una variazione ad essa proporzionale della sua massa relativistica**.
Se l'energia cinetica di un corpo è considerata come una forma di energia esterna, l'energia a riposo può essere considerata come l'energia interna del corpo. Questa energia interna consiste, in parte, da entità quali il moto molecolare, che varia quando dal corpo viene assorbita o liberata energia termica, o l'energia potenziale intermolecolare, che varia quando avvengono reazioni chimiche. Oppure l'energia interna si può identificare con l'energia potenziale atomica, che può variare quando un atomo assorbe della radiazione e si eccita o emette radiazione diseccitandosi, o di energia potenziale nucleare, che può essere variata da reazioni nucleari. Il contributo più grande all'energia interna è, comunque, l'energia totale a riposo posseduta dalle particelle "fondamentali", che è considerata come la sorgente primaria dell'energia interna. Anche questa può variare, come, per esempio, nella creazione o annichilazione di coppie elettrone-positrone. La massa a riposo o massa propria di un corpo, non è, quindi, in generale costante. L'approfondimento di queste interessantissime tematiche esula però dai nostri scopi.
Osservazione:
Ritornando ancora alle collisioni o interazioni fra corpi, ricordiamo che si è visto che, indipendentemente dalla natura della collisione, l'energia totale si conserva e che tale conservazione è equivalente alla conservazione della massa relativistica. Nella fisica classica avevamo due distinti principi di conservazione:

1) la conservazione della massa classica, come nelle reazioni chimiche;

2) la conservazione dell'energia.

In Relatività, questi due principi si fondono in un unico principio quello di conservazione della massa-energia.

6.4.3 Moto di una particella carica in un campo magnetico uniforme

Un problema importante che si incontra sia nell'elettrodinamica classica che in quella relativistica è il moto di una particella carica positivamente o negativamente $(q^+ \; o \; q^-)$ in un campo elettromagnetico assegnato. Si tratta del moto di particelle cariche che si realizza, per opportune condizioni iniziali, in un *sincrotrone*, cioè in una ordinaria macchina acceleratrice di particelle.
Il problema si inquadra nella dinamica di una particella in un prefissato riferimento galileiano $\mathcal{R} \equiv Oxyz$, che in prima approssimazione si può ritenere coincidente con il laboratorio in cui è situato il sincrotrone. Il moto di tale particella è governato, in ambito relativistico, dalle seguenti

[16] È opportuno sottolineare il fatto che la massa non è numericamente uguale all'energia, in quanto le loro unità sono differenti. Tuttavia, esse sono quantità fisicamente equivalenti e connesse fra di loro. C'è una corrispondenza analoga a quella fra l'altezza di una colonna di mercurio e la pressione dell'aria.

equazioni

$$\begin{cases} \dfrac{d(m\mathbf{u})}{dt} = \mathbf{f}, \qquad m \equiv m_0\gamma(u) = \dfrac{m_0}{\sqrt{1-\dfrac{u^2}{c^2}}} \\ \\ \dfrac{dE}{dt} = \mathbf{f}\cdot\mathbf{u}, \qquad (E = mc^2)\ (\textit{Energia totale della particella}). \end{cases} \tag{6.110}$$

Se la particella è immersa in un campo elettromagnetico assegnato $(\mathbf{E}, \mathbf{H})$ e se q è la sua carica, la quale si suppone non modifichi in alcun modo il campo elettromagnetico (particella di prova), la forza su di essa agente sarà la forza di Lorentz

$$\mathbf{f} = q(\mathbf{E} + \frac{1}{c}\mathbf{u}\wedge\mathbf{H}). \tag{6.111}$$

Ciò premesso consideriamo il problema dinamico nelle seguenti ipotesi (vedi Fig.6.8):

I. assenza di campo elettrico $(\mathbf{E} = 0)$;

II. campo magnetico uniforme $(\mathbf{H} = cost.)$;

III. velocità iniziale $\mathbf{u}_0$ della particella perpendicolare ad $\mathbf{H}$ $(\mathbf{u}_0 \perp \mathbf{H})$.

Figura 6.8: Moto di una particella carica in un campo magnetico uniforme.

L'ipotesi I, per la (6.111), implica che le equazioni dinamiche (6.110) si riducano alle

$$\begin{cases} \dfrac{d(m\mathbf{u})}{dt} = \dfrac{q}{c}\mathbf{u}\wedge\mathbf{H} \\ \\ \dfrac{d(mc^2)}{dt} = \mathbf{f}\cdot\mathbf{u} = 0, \end{cases} \tag{6.112}$$

dalla seconda delle quali si deduce l'integrale primo dell'energia $mc^2 = cost.$ il quale implica che è costante sia la massa relativistica $m(u) = m_0\gamma(u)$ della particella sia il modulo u della sua velocità. Pertanto l'ipotesi I implica che **il moto della particella è uniforme.** La sua traiettoria è determinata dalla prima delle equazioni (6.112) che costituisce la sua equazione di moto, cioè[17]

$$m\mathbf{a} = \frac{q}{c}\mathbf{u}\wedge\mathbf{H}. \tag{6.113}$$

[17]Ricordiamo che l'equazione di moto relativistica è $\mathbf{f} = m(u)\mathbf{a} + \dfrac{\mathbf{f}\cdot\mathbf{u}}{c^2}\mathbf{u}$, ma nel nostro caso essendo $\mathbf{f} \perp \mathbf{u}$ si avrà $\mathbf{f} = m(u)\mathbf{a}$ come nel caso classico.

Teniamo ora conto dell'ipotesi II, cioè che sia costante il campo magnetico ($\mathbf{H} = cost.$), il quale, scegliendo l'asse delle z del riferimento $\mathcal{R}$ parallelo ed equiorientato con $\mathbf{H}$, possiamo esprimerlo così $\mathbf{H} = H\mathbf{k}$ con $H > 0$, essendo $\mathbf{k}$ il versore dell'asse z. Ne consegue l'esistenza di un altro integrale primo del moto, e precisamente la conservazione della quantità di moto rispetto all'asse z perché $\mathbf{f} \perp z$. Infatti, in tale ipotesi, si ha $f_z = 0$ e perciò la prima delle (6.112), proiettata sull'asse z, ci dà $\dfrac{d(m\dot{z})}{dt} = 0$, ovvero $\dot{z} = cost. = \dot{z}_0$ per ogni $t \geq t_0$. Pertanto, se teniamo conto della condizione III, cioè $\mathbf{u}_0 \perp \mathbf{H}$ ciò equivale a dire che $\dot{z}_0 = 0$ (condizione che si realizza praticamente in una macchina acceleratrice) e ciò equivale a dire che **il moto della particella è piano** essendo xy $(z = 0)$ il piano del moto, anzi possiamo affermare con più precisione che **la traiettoria è una circonferenza**.
Infatti, se deriviamo rispetto al tempo l'equazione di moto (6.113) e teniamo conto dell'ipotesi II, otteniamo

$$m\dot{\mathbf{a}} = \frac{q}{c}\mathbf{a} \wedge \mathbf{H}, \tag{6.114}$$

la quale implica che i vettori $\mathbf{a}$ e $\dot{\mathbf{a}}$ sono ortogonali fra di loro (oltre che ad $\mathbf{H}$), cioè $\mathbf{a} \cdot \dot{\mathbf{a}} = 0$ che equivale a $\dfrac{d(\mathbf{a} \cdot \mathbf{a})}{dt} = 0$, ovvero ad $a = cost.$. Dunque le ipotesi I, II, III, implicano, non soltanto che il moto è uniforme e piano, ma che è costante anche il modulo dell'accelerazione della particella. Infatti, essendo il modulo dell'accelerazione esprimibile nella forma

$$a^2 = \ddot{s}^2 + \left(\frac{\dot{s}^2}{\rho}\right)^2 = cost.,$$

dove s è l'ascissa curvilinea sulla traiettoria della particella, $\dot{s}$ e $\ddot{s}$ le sue derivate temporali prima e seconda, e ρ il raggio di curvatura, ed essendo inoltre $\dot{s} = u = cost.$, $\ddot{s} = 0$ si avrà $\rho = cost.$ Dunque è costante il raggio di curvatura della traiettoria, il che significa che **si tratta di una circonferenza**.
A tale conclusione si può arrivare anche dall'esame delle equazioni cartesiane esplicite del moto. Infatti, indichiamo con θ l'angolo che la velocità $\mathbf{u}$ della particella, ad un istante qualsiasi, forma con l'asse x, le sue componenti cartesiane sugli assi x, y, z del riferimento galileiano $\mathcal{R}$ sono $\mathbf{u} \equiv (u \cos\theta, u \sin\theta, 0)$ e perciò le componenti della forza di Lorentz saranno date da $\mathbf{f} \equiv \left(\frac{q}{c}\, u\, H \sin\theta, -\frac{q}{c}\, u\, H \cos\theta, 0\right)$. Le prime due equazioni cartesiane del moto della particella, per la prima delle (6.112), si possono scrivere così

$$\begin{cases} m_0 \dfrac{d}{dt} \dfrac{u \cos\theta}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \dfrac{q}{c}\, u\, H \sin\theta \\[2ex] m_0 \dfrac{d}{dt} \dfrac{u \sin\theta}{\sqrt{1 - \dfrac{u^2}{c^2}}} = -\dfrac{q}{c}\, u\, H \cos\theta \end{cases} \tag{6.115}$$

mentre la terza equazione si riduce all'identità $0 = 0$.
Tenuto conto che $u = cost. = u_0$, dalla prima delle (6.115) si ottiene

$$\dot{\theta} = -\frac{q\, H}{c\, m_0}\sqrt{1 - \frac{u^2}{c^2}} = cost. \tag{6.116}$$

Introducendo l'ascissa curvilinea s, possiamo scrivere $|\dot{\theta}| = \left|\dfrac{d\theta}{ds}\right||\dot{s}| = \left|\dfrac{d\theta}{ds}\right|u$, cioè

$$\left|\frac{d\theta}{ds}\right| = \frac{q\, H}{c\, m_0 u}\sqrt{1 - \frac{u^2}{c^2}} = cost.$$

Ma $\left|\frac{d\theta}{ds}\right|$ definisce la curvatura $k = \frac{1}{\rho}$ in un punto generico della traiettoria della particella, il cui raggio di curvatura è perciò

$$\rho_{rel} = \frac{c\, m_0}{q\, H} \frac{u}{\sqrt{1 - \frac{u^2}{c^2}}} = cost. \tag{6.117}$$

La traiettoria è dunque una circonferenza di raggio

$$\rho_{rel} = \frac{c\, m_0 u}{q\, H} \gamma(u) = \rho_{class} \gamma(u) = cost., \tag{6.118}$$

dove

$$\rho_{class} = \frac{c\, m_0 u}{q\, H} = cost., \tag{6.119}$$

è il raggio classico della traiettoria: perciò la meccanica relativistica conduce ad un risultato qualitativamente analogo a quello della meccanica classica, ma quantitativamente diverso. In quale semipiano cada il centro C della circonferenza rispetto alla tangente in suo punto generico P, orientata come la velocità $\mathbf{u}$, viene determinato tenendo conto del segno della carica q della particella e dell'espressione della forza di Lorentz.
Il moto ora descritto, come abbiamo già detto, è quello che si realizza in una macchina acceleratrice di particelle.
Per velocità piccole rispetto alla velocità della luce i due risultati differiscono di pochissimo e possono considerarsi coincidenti a meno degli inevitabili errori nell'ambito delle misure sperimentali, perché in tali casi il fattore di Lorentz è pressoché uguale ad 1. Ma se la velocità delle particelle è prossima alla velocità della luce, essendo $\frac{\rho_{rel}}{\rho_{class}} = \gamma(u)$, i raggi delle traiettorie circolari saranno 10, 100, 1000, ecc., volte più grandi nel caso relativistico rispetto al caso classico. Si noti esplicitamente che, data la presenza di c e di u nel numeratore delle formule (6.118) e (6.119), il campo magnetico H, che vi figura invece al denominatore, deve essere molto intenso se si vuole che il raggio della traiettoria circolare abbia un valore ragionevole per una apparecchiatura sperimentale (che non sia cioè dell'ordine di molti km).
Per particelle molto veloci, l'esperienza convalida con estrema precisione e senza possibilità di dubbi le formule relativistiche e non le formule classiche.
L'impiego delle formule relativistiche costituisce di fatto un elemento essenziale ed insostituibile nella stessa progettazione di una macchina acceleratrice come quelle che si trovano, per esempio, al CERN di Ginevra o nei laboratori di fisica nucleare sotto il Gran Sasso nei pressi del capoluogo abruzzese de l'Aquila. **È questa forse la prova più sicura e più decisiva della validità dell'intero corpo della Meccanica Relativistica.**

6.4.4 Cenno alle particelle di massa propria variabile

Finora abbiamo considerato **particelle prive di struttura interna** caratterizzate dall'avere massa propria costante, $m_0 = cost.$. Ora daremo un cenno sulle **particelle con struttura interna** per le quali la massa propria m_0 viene considerata suscettibile di variare durante la storia della particella. Queste ultime vengono inquadrate ancora nello schema **punto materiale** rappresentante macroscopicamente un sistema materiale complesso del quale convenga però, per semplicità, **ignorare l'esatta struttura interna**. Anche ora si ammette che lo schema non ha estensione nello spazio, la sua evoluzione spaziotemporale essendo rappresentata in $\mathcal{M}^4$ da una linea oraria del genere tempo, e la cui struttura viene ancora riassunta in uno scalare assoluto positivo m_0, **massa propria del sistema**. La differenza con il punto materiale privo di struttura interna, sottolineiamolo, sta unicamente nel fatto che la massa propria m_0 può essere variabile.

In altri termini la particella viene considerata suscettibile di diversi "stati interni", ciascuno con una ben determinata energia interna E_0 e una corrispondente massa m_0: al variare dello stato interno, in dipendenza delle azioni esterne agenti sulla particella, varia la massa propria m_0.
Immutata resta la definizione del 4-vettore quantità di moto $(P^i = m_0 U^i)$. Immutata resta l'interpretazione della sua parte spaziale come quantità di moto nel senso ordinario $\left(p^\alpha \equiv P^\alpha = m_0\gamma(u)u^\alpha\right)$ e della sua parte temporale, moltiplicata per c, come energia del sistema $\left(E \equiv cP^0 = m_0c^2\gamma(u)\right)$. Anche la definizione di massa relativistica non cambia $\left(m = m_0\gamma(u)\right)$.
Infine anche la legge di moto permane, non nella seconda forma (6.30) ma nella prima forma (6.29), cioè:

$$\frac{d\mathbf{P}}{d\tau} = \mathbf{F}, \tag{6.120}$$

e la quadriforza $\mathbf{F}$ non è più vincolata, a priori, né alla condizione $\mathbf{F}\cdot\mathbf{U} = \eta_{ik}F^iU^k = 0$ né ad altre condizioni, neppure a quella di essere un vettore del genere spazio.
Tenuto conto dell'espressione di P^i e sviluppando la derivata, la (6.120) può scriversi nella forma

$$\frac{dm_0}{d\tau}U^i + m_0A^i = F^i, \tag{6.121}$$

dove il vettore a primo membro è decomposto in due vettori, uno parallelo, l'altro ortogonale ad U^i. Se effettuiamo una analoga decomposizione per il vettore F^i

$$F^i = F'^{\,i} + F''^{\,i}, \quad \left(F'^{\,i} \parallel U^i, \quad F''^{\,i} \perp U^i\right), \tag{6.122}$$

le $F'^{\,i}$, $F''^{\,i}$, tenuto conto che $\mathbf{U}\cdot\mathbf{U} = \eta_{ik}U^iU^k = c^2$ e che $U_h = \eta_{hk}U^k$, possono essere espresse nella forma

$$F'^i = \frac{1}{c^2}F^hU_hU^i, \quad F''^{\,i} = F^i - \frac{1}{c^2}F^hU_hU^i.$$

La (6.121) si spezza allora nelle due equazioni

$$m_0A^i = F''^{\,i}, \tag{6.123}$$

$$\frac{dm_0}{d\tau}U^i = F'^{\,i} \Longleftrightarrow \frac{dm_0}{d\tau} = \frac{1}{c^2}F^hU_h \tag{6.124}$$

vettoriale la prima, scalare la seconda.
Da esse scaturiscono i due seguenti casi limite:

1) $F'^{\,i} = 0$, $F''^{\,i} \neq 0$, cioè quadriforza $\mathbf{F}$ normale alla linea di universo, in tal caso per la prima delle (6.124) essendo sempre $U^i \neq 0$, si avrà $\dfrac{dm_0}{d\tau} = 0$ e perciò $m_0 = cost.$, dunque: *la massa propria della particella si mantiene allora costante durante il moto.* È quanto accade per una particella priva di struttura interna. Naturalmente la costanza di m_0 può derivare sia dalla struttura irriducibile della particella, incapace di modificare il suo stato interno, sia dalla natura del campo di forza ambiente a cui essa è sottoposta. Una quadriforza ortogonale ad $\mathbf{U}$, capace cioè di accelerare la particella senza modificarne la massa propria si dirà una *quadriforza puramente motrice o puramente meccanica.*

2) $F'^{\,i} \neq 0$, $F''^{\,i} = 0$. Per la (6.123) si avrà $A^i = 0$ e per la (6.121) la quadriforza F^i risultando parallela alla 4-velocità U^i sarà tangente in ogni punto alla linea di universo. *In tal caso la quadriforza non accelera la particella e la lascia pertanto nel suo stato di quiete o di moto rettilineo uniforme. Per contro, essa ha per effetto in base alla* (6.124), *di modificare la massa propria e quindi l'energia propria della particella.*

Una tale forza sta dunque a rappresentare un'azione fisica non strettamente dinamica, nel senso che a tale parola si dà in meccanica classica. per esempio l'azione di una sorgente calorifica che somministra calore ad un corpo è un'azione che in relatività viene rappresentata da una quadriforza **F** parallela ad **U**. Azioni di questo tipo non rientrano nel concetto classico di forza. In relatività il concetto di forza viene dunque ampliato in modo da comprendere anche azioni fisiche del tipo indicato.
Naturalmente in una quadriforza **F** *generica sarà presente sia la componente motrice sia la componente puramente energetica. Nella più estesa accezione che è propria della relatività la quadriforza può considerarsi come la più generale sorgente di quantità di moto e di energia.*
Osservazione:
In meccanica classica la nozione di punto materiale trova, a posteriori, una giustificazione nel teorema del moto del baricentro. Nella dinamica relativistica dei sistemi, che esce peraltro dal nostro programma, si riconosce che un teorema analogo non sussiste in relatività; pertanto in relatività lo schema punto materiale in senso macroscopico va usato con una certa cautela e va limitato a rappresentare quei sistemi le cui dimensioni spaziali siano veramente molto piccole rispetto alle dimensioni dell'ambiente in cui si muovono.

Lo schema **punto materiale** è talora assunto a rappresentare macroscopicamente un sistema materiale complesso del quale convenga, per semplicità, **ignorare l'esatta struttura interna**. Anche ora lo schema non ha estensione nello spazio, la sua evoluzione spaziotemporale è rappresentata in $\mathcal{M}^4$ da una linea oraria del genere tempo, e la sua struttura viene riassunta in uno scalare assoluto positivo m_0, **massa propria del sistema**. La differenza con il punto materiale **privo di struttura interna** sta unicamente nel fatto che la massa propria m_0 viene considerata suscettibile di variare durante la storia della particella.

6.5 APPENDICE

Sulla seconda interpretazione relativistica della 3-quantità di moto di una particella

In questa Appendice, a beneficio del lettore curioso, ci occupiamo brevemente del modo di procedere nel caso in cui si adotti la seconda definizione di quantità di moto introdotta nel paragrafo 6.2.2 con la (6.6), cioè

$$\mathbf{p} \stackrel{\text{def}}{=} m\gamma(u)\mathbf{u}, \tag{6.125}$$

essendo $m \equiv m_0 = cost.$ la massa inerziale della particella (coincidente con la massa propria) definita in ambito newtoniano[18].
Partiamo, anche qui, dalla legge di conservazione della 4-quantità di moto (6.8) - in un processo di collisione fra più particelle - che ora possiamo scrivere così

$$\sum_a{}^* \mathbf{P}_a \equiv \sum_a{}^* (m_a\gamma(u_a)c, \mathbf{p}_a) = 0, \tag{6.126}$$

dove m_a è la massa inerziale newtoniana della generica particella e $\mathbf{p}_a$ la sua quantità di moto espressa nella forma (6.125) e dove l'indice a, ricordiamolo, numera tutte le particelle prima e dopo la collisione. Inoltre il simbolo $\sum_a{}^*$ è una somma che conta i termini prima della collisione positivamente e quelli dopo la collisione negativamente.
Con lo stesso ragionamento usato nel paragrafo 6.2.2, possiamo osservare che il 4-vettore $\sum_a{}^*(m_a\gamma(u_a)c, \mathbf{p}_a)$ quantità di moto totale, essendo nullo, avrà ciascuna delle sue quattro

[18]Per un approfondimento di questo secondo modo di procedere si consiglia il trattato di Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), cap.4, pagg. 112-160, che abbiamo citato più volte.

componenti nulle separatamente, avremo allora:[19]

$$\sum_a{}^* m_a \gamma(u_a) c = 0, \tag{6.127}$$

e

$$\sum_a{}^* \mathbf{p}_a \equiv \sum_a{}^* m_a \gamma(u_a) \mathbf{u}_a = 0 \quad (conserv.\ della\ 3 - quant.\ di\ moto). \tag{6.128}$$

La (6.128) ha una diretta interpretazione fisica in quanto essa è formalmente analoga alla **legge di conservazione della 3-quantità di moto** della meccanica newtoniana, pur di assumere come definizione relativistica di $\mathbf{p}$ la (6.125). Inoltre la (6.128) tende, come è desiderabile, all'analoga legge newtoniana per $\gamma(u_a) \to 1$, cioè per $c \to \infty$, ovvero praticamente per $u_a \ll c$. Essa, inoltre, essendo parte di una legge 4-dimensionale è ovviamente Lorentz-invariante.
Si osservi che la (6.128) dipende dalla definizione della 3-quantità di moto data dalla (6.125) la quale poggia su considerazioni piuttosto formali a partire dalla definizione (6.1) della 4-quantità di moto nello spaziotempo di Minkowski. Abbiamo poi assunto, come principio generale, che nella collisione fra più particelle valga la legge di conservazione della 4-quantità di moto da cui è scaturita la (6.128).
Vogliamo ora stabilire la validità della (6.125), conservando la validità della legge di conservazione della 3-quantità di moto, con le seguenti considerazioni in parte di natura euristica ed in parte fondate su procedimenti d'urto fra particelle attraverso un esperimento ideale.
È del tutto plausibile che la quantità di moto di una particella sia definita, in analogia con la meccanica newtoniana, con una relazione del tipo

$$\mathbf{p} = \mu \mathbf{u}, \tag{6.129}$$

dove però la velocità $\mathbf{u}$ della particella sia moltiplicata per un fattore scalare μ generalmente dipendente dalla sua massa m e dal modulo u della sua velocità tramite il fattore di Lorentz, sia cioè della forma

$$\mu = \mu(m, \gamma(u)). \tag{6.130}$$

Inoltre, è altrettanto plausibile pretendere che nel limite newtoniano, cioè per $c \to \infty$, tale fattore scalare si riduca alla massa inerziale della particella, debba dunque soddisfare alla condizione

$$\lim_{c \to \infty} \mu(m, \gamma(u)) = \mu(m, 1) = m. \tag{6.131}$$

Infine si vuole che, in una collisione fra più particelle, rimanga valida la legge di conservazione della 3-quantità di moto, che per le ipotesi fatte in precedenza sarà ora espressa dalla:

$$\sum_a{}^* \mu_a \mathbf{u}_a = 0 \quad (conserv.\ della\ 3 - quant.\ di\ moto), \tag{6.132}$$

con

$$\mu_a = \mu_a(m_a, \gamma(u_a)). \tag{6.133}$$

Allora necessariamente, per una generica particella, la funzione μ in un qualunque riferimento galileiano $\mathcal{R}$, sarà espressa dalla semplice relazione:

$$\mu(m, \gamma(u)) = m\gamma(u). \tag{6.134}$$

Per dimostrare la (6.134) seguiremo un procedimento analogo a quello usato nel paragrafo 6.2.3.
Precisamente, supponiamo di eseguire un esperimento ideale, considerando l'urto fra due

[19] Per l'interpretazione fisica della (6.127) rimandiamo al libro di Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), cap.4, pagg. 114-117.

Figura 6.9: Urto anelastico fra due particelle aventi la stessa massa inerziale (2^a interpretazione della massa).

particelle identiche $\mathcal{P}_1$, $\mathcal{P}_2$, entrambe di massa inerziale m, mobili, in direzioni opposte ma con velocità di ugual modulo u (Fig.6.9). La prima si muova sul semiasse positivo delle z verso il basso cioè verso l'origine O di un dato sistema galileiano $\mathcal{R}$, mentre la seconda si muova sul semiasse negativo delle z' verso l'alto cioè verso l'origine O' di un altro sistema inerziale $\mathcal{R}'$. I due sistemi $\mathcal{R}$, $\mathcal{R}'$, siano in configurazione x-standard e mobili l'uno rispetto all'altro con velocità relativa v, inoltre supponiamo che prima della collisione fra $\mathcal{P}_1$ e $\mathcal{P}_2$, $\mathcal{R}'$ si trovi a sinistra di $\mathcal{R}$, rispetto ad un osservatore in quiete nell'origine O di $\mathcal{R}$ stesso, in piedi sul piano xy ed orientato dai piedi alla testa come il semiasse positivo delle z, e che abbia lo sguardo rivolto verso il semiasse positivo delle y. Immaginiamo che la collisione sia **anelastica**, cioè che le particelle dopo la collisione rimangano fuse una sull'altra e che ciò avvenga, nell'origine $O \equiv O'$ dei due sistemi di coordinate, nell'istante in cui essi vengono a coincidere. Ammettiamo ancora che, per ragioni di simmetria, la particella composta dopo l'urto rimanga in quiete in $\mathcal{R}$. Utilizzando la legge relativistica di composizione delle velocità nella forma (4.16) o (4.17), cioè

$$\begin{cases} u'_x = \dfrac{u_x - v}{1 - \dfrac{v}{c^2}u_x}, & u_x = \dfrac{u'_x + v}{1 + \dfrac{v}{c^2}u'_x} \\[2ex] u'_y = \dfrac{u_y\gamma^{-1}(v)}{1 - \dfrac{v}{c^2}u_x}, & u_y = \dfrac{u'_y\gamma^{-1}(v)}{1 + \dfrac{v}{c^2}u_x} \\[2ex] u'_z = \dfrac{u_z\gamma^{-1}(v)}{1 - \dfrac{v}{c^2}u_x}, & u_z = \dfrac{u'_z\gamma^{-1}(v)}{1 + \dfrac{v}{c^2}u_x} \end{cases}$$

avremo

PRIMA DELL'URTO

Particella $\mathcal{P}_1$, *in* $\mathcal{R}$: $u_{1x} = 0,\ u_{1y} = 0,\ u_{1z} = -u,$ (componenti assegnate);
Particella $\mathcal{P}_1$, *in* $\mathcal{R}'$: $u'_{1x} = -v,\ u'_{1y} = 0,\ u'_{1z} = -u\gamma^{-1}(v),$ (componenti calcolate);
Particella $\mathcal{P}_2$, *in* $\mathcal{R}'$: $u'_{2x} = 0,\ u'_{2y} = 0,\ u'_{2z} = u,$ (componenti assegnate);
Particella $\mathcal{P}_2$, *in* $\mathcal{R}$: $u_{2x} = v,\ u_{2y} = 0,\ u_{2z} = u\gamma^{-1}(v),$ (componenti calcolate);

DOPO L'URTO (istante dell'urto)

Particelle fuse $\mathcal{P}_1 + \mathcal{P}_2$, *in* $\mathcal{R}$: $U_x = U_y = U_z = 0,$ (componenti assegnate);
Particelle fuse $\mathcal{P}_1 + \mathcal{P}_2$, *in* $\mathcal{R}'$: $U'_x = -v,\ U'_y = U'_z = 0,$ (componenti calcolate);

avendo indicato con lettere minuscole le componenti delle velocità prima dell'urto e con maiuscole le analoghe componenti dopo l'urto.
Applichiamo ora **la legge di conservazione della 3-quantità di moto** in $\mathcal{R}$ lungo l'asse z, in accordo con le (6.132), (6.133) avremo

$$\mu(m, \gamma(u_1))u_{1z} + \mu(m, \gamma(u_2))u_{2z} = \mu(M, \gamma(U))U_z, \tag{6.135}$$

avendo indicato con

$$u_1 = \sqrt{(u_{1x})^2 + (u_{1y})^2 + (u_{1z})^2} = u,$$

e

$$u_2 = \sqrt{(u_{2x})^2 + (u_{2y})^2 + (u_{2z})^2} = \sqrt{v^2 + u^2\gamma^{-2}(v)},$$

rispettivamente i moduli delle velocità delle due particelle $\mathcal{P}_1$ e $\mathcal{P}_2$ in $\mathcal{R}$, prima dell'urto.
Tenuto conto delle componenti delle velocità precedentemente calcolate, la (6.135) diventa

$$\mu(m, \gamma(u))(-u) + \mu(m, \gamma(u_2))u\gamma^{-1}(v) = M(U) \times 0 = 0$$

cioè

$$\mu(m, \gamma(u_2)) = \mu(m, \gamma(u))\gamma(v). \tag{6.136}$$

Se ora facciamo tendere a zero u ($u \to 0$), considerando una sequenza continua di esperimenti con u sempre più piccola, avremo che $u_2 \to v$, $\gamma(u) \to 1$, ciò porta a concludere che la (6.136) diventa

$$\mu(m, \gamma(v)) = \mu(m, 1)\gamma(v),$$

e ricordando la (6.131) otteniamo il risultato:

$$\mu(m, \gamma(v)) = m\gamma(v). \tag{6.137}$$

Nella (6.137), v è la velocità della particella $\mathcal{P}_2$ nel riferimento $\mathcal{R}$; ma quest'ultimo è un generico riferimento inerziale e $\mathcal{P}_2$ è una generica particella che si muove con velocità $|\mathbf{v}| = v$, perciò la (6.137) - a meno dello scambio di v con u - dimostra la nostra tesi.
Perciò possiamo concludere dicendo che: *se in Relatività vogliamo salvare l'aspetto formale* (6.132) *della legge newtoniana della conservazione della quantità di moto dobbiamo accettare che il 3-momento della particella sia definito con la* $\mathbf{p} = m\gamma(u)\mathbf{u}$ ottenuta in precedenza in via del tutto formale.

Capitolo 7

ELETTRODINAMICA RELATIVISTICA NEL VUOTO

Procedendo nello studio di Faraday vidi che la sua concezione dei fenomeni elettromagnetici era anche matematica, sebbene non espressa nelle forme convenzionali dei simboli matematici, e vidi che potevano essere espressi in formule e così comparati con quelli dei matematici di professione.
(James Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 1873.)

Ma chi è stato un Dio a scrivere queste formule?
(Ludwig Boltzmann, riferendosi alle equazioni di Maxwell)

7.1 Introduzione

All'inizio del Cap.2, paragrafo 2.1, abbiamo già messo in luce il fatto che una delle crisi subite dalla fisica classica, nel periodo compreso tra la fine del 1800 e l'inizio del 1900, fu causata da *certe difficoltà presentatesi nell'elettrodinamica di Maxwell* come viene espresso molto bene da Einstein nel suo celebre lavoro del 1905:[1]
"È noto che l'elettrodinamica di Maxwell - come essa attualmente viene concepita - conduce nelle sue applicazioni a corpi in movimento ad **"asimmetrie"** *che paiono non essere aderenti ai fenomeni. Si pensi p.e. alle interazioni elettrodinamiche tra un magnete ed un conduttore. Il fenomeno osservabile dipende qui solo dal moto relativo fra magnete e conduttore, mentre secondo il consueto modo di vedere sono da tener rigorosamente distinti i due casi che l'uno o l'altro di questi corpi sia mosso. Infatti se si muove il magnete e rimane fisso il conduttore, si produce nell'interno del magnete un campo elettrico di certi valori di energia il quale provoca una corrente nei luoghi ove si trovano parti del conduttore. Rimane invece fisso il magnete e si muove il conduttore, non si produce nell'intorno del magnete alcun campo elettrico, ma al contrario [si produce] nel conduttore una forza elettromotrice, alla quale non corrisponde per sè alcuna energia, ma che - supposta l'eguaglianza del moto relativo nei due casi considerati - dà occasione al prodursi di correnti elettriche della stessa grandezza e dello stesso percorso, come nel primo caso [avevano dato] le forze elettriche."*
Ciò che Einstein si propone di fare è stabilire una **simmetria** tra due descrizioni fisicamente equivalenti ma formalmente diverse - fornite dalla fisica classica - dello stesso fenomeno consistente nell'interazione fra il magnete ed il filo conduttore dell'esempio di cui sopra. Ora sappiamo

[1] Albert Einstein, *Zur Elektrodynamik bewegter korper*, Annalen der Physik, **17**, 1905, pp. 891-921; traduzione italiana a cura di Paolo Straneo: *Sull'elettrodinamica dei corpi in moto*, inserito nel volume "Cinquant'anni di Relatività", pp. 480, Vallecchi Editore, Firenze (1955); questo articolo è inserito anche nel volume: A.Einstein, *Opere scelte*, a cura di Enrico Bellone, Bollati Boringhieri, Torino,1988, pp. 148.177.

che queste asimmetrie nascono dall'incompatibilità tra la meccanica newtoniana e l'elettrodinamica di Maxwell, nel senso che queste teorie risultano invarianti rispetto a due gruppi diversi di trasformazioni: le trasformazioni di Galileo per la meccanica e le trasformazioni di Lorentz per l'elettrodinamica. Siccome abbiamo visto, nei precedenti capitoli, dedicati ai principi fondamentali della Relatività Ristretta, che le trasformazioni corrette sono quelle di Lorentz, ne segue che la teoria elettromagnetica resta immutata alla luce di questi principi, mentre le equazioni della meccanica, come abbiamo già visto, devono essere corrette. In altre parole, le leggi dell'elettrodinamica classica[2], sintetizzate nelle equazioni di Maxwell, sono formalmente invarianti per trasformazioni di Lorentz come vedremo fra poco. Storicamente, è opportuno osservare che il problema di trovare una trasformazione che lasci invariate le equazioni di Maxwell condusse Lorentz, nei lavori fondamentali del 1892 e del 1895, alle trasformazioni che ora portano il suo nome.[3] Perciò, ci si potrebbe limitare alla dimostrazione dell'invarianza delle equazioni di Maxwell e terminare così la discussione. Tuttavia questo modo di procedere ci impedirebbe di apprezzare i veri meriti della trattazione relativistica dell'elettromagnetismo. La Relatività ci presenta un nuovo punto di vista che aumenta la nostra comprensione dell'elettromagnetismo, inoltre le tecniche relativistiche per la soluzione dei problemi elettromagnetici sono spesso molto più semplici di quelle classiche. Perciò la Teoria della Relatività costituisce anche un mezzo pratico efficace per la risoluzione dei problemi. A questo proposito è utile citare un passo di Einstein preso da una sua dichiarazione inviata ad un congresso nel 1952 in onore del centenario della nascita di Abraham Michelson: *"L'influenza del cruciale esperimento di Michelson-Morley sulle mie ricerche è stata piuttosto indiretta. Appresi di questo esperimento attraverso lo studio decisivo di H. A. Lorentz sull'elettrodinamica dei corpi in movimento (1895) di cui venni a conoscenza prima di sviluppare la Teoria della Relatività Ristretta....*
Ciò che mi condusse più o meno direttamente alla Teoria della Relatività Ristretta fu la convinzione che la forza elettromotrice agente su un corpo in moto in un campo magnetico non è altro che un campo elettrico."[4]
In questa nostra trattazione dell'elettrodinamica ci limiteremo a mettere in luce i fatti più salienti nella maniera più semplice ed elementare possibile; il lettore interessato ad esposizioni più dettagliate ed approfondite può consultare con profitto i testi classici citati nella nota 2 di questa pagina.[5]

7.2 Densità di carica e densità di corrente

Prima di dimostrare che le equazioni di Maxwell sono compatibili col Principio di Relatività, cioè che esse sono *invarianti* (o come anche si dice *covarianti*) rispetto ad una trasformazione di Lorentz dobbiamo fare una *ipotesi fondamentale* per quanto riguarda la trasformazione della **carica elettrica** q. L'ipotesi che faremo, ben convalidata dall'esperienza, è che **la carica**, *è un*

[2]Per un breve riassunto dell'elettrodinamica classica si rimanda il lettore al paragrafo 1.4 del Cap.1 di questo libro. Per una trattazione più estesa ed approfondita dell'elettrodinamica, sia classica che relativistica, si possono consultare i seguenti volumi: Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino (2004), Cap.8 e Cap.9, oppure L. D. Landau e E. M. Lifschitz, *Teoria dei campi*, Editori Riuniti, Edizione del 2010. Il testo più completo oggi disponibile è, unanimemente considerato, il libro di John David Jackson, *Elettrodinamica Classica*, Zanichelli Editore, 2001.

[3]Tali lavori fondamentali di H. A. Lorentz, sono: *La théorie électromagnétique de Maxwell et son application aux corps mouvants*, Arch. Néerl. Sci. **25**, 363 (1892), e *Versuch einer Theorie der elektrischen und magnetischen Erscheinungen in bewegten Körpern*, Leida (1895).

[4]La dichiarazione di Einstein, riportata sopra, è stata pubblicata da R. S. Shankland in *Michelson-Morley Experiment*, Am. Journ. Physics, **32** (1), 16 (1964).

[5]Semplici trattazioni dell'elettrodinamica sono fornite anche nei seguenti testi molto noti e di larga diffusione: Robert Resnick, *Introduzione alla Relatività Ristretta*, Cap. IV, pag.157, Casa Editrice Ambrosiana, Milano (1979); Wolfgang Rindler, *Introduction to Special Relativity*, Second Edition 1991, Cap.VI, Clarendon Press - Oxford (reprinted in 1996); Wolfgang Rindler, *Relativity, Special, General and Cosmological*, Second Edition, Part I, Cap.7, Oxford University Press, (2006).

invariante, *cioè che una data carica ha lo stesso valore misurato in tutti i sistemi inerziali*. Questa è l'unica ipotesi possibile, compatibile con la covarianza delle equazioni di Maxwell. Infatti è ben noto (e noi lo abbiamo dimostrato nel paragrafo 1.4.3 del Cap.1) che una conseguenza delle equazioni di Maxwell è **l'equazione di continuità** (1.65) che qui riscriviamo

$$\operatorname{div} \mathbf{j} + \frac{\partial \rho}{\partial t} = 0, \tag{7.1}$$

il cui significato fisico è proprio **la legge di conservazione della carica**, *cioè che una data carica conserva lo stesso valore durante tutto il suo moto e quindi a tutte le velocità.* Tale conservazione insieme al Principio di Relatività implica **l'invarianza**: infatti, se il moto di una carica non ne modifica la misura in un dato sistema inerziale $\mathcal{R}$, non c'è nulla che differenzi i vari sistemi rispetto alla carica. *Perciò le equazioni di Maxwell ed il Principio di Relatività non possono essere veri simultaneamente se la carica non è invariante.*
Per renderci conto di come vanno le cose è opportuno richiamare brevemente i concetti fondamentali di: *densità di carica* ρ e di *densità di corrente* $\mathbf{j}$.[6] Consideriamo un certo volume V racchiuso da una superficie S, la carica elettrica macroscopica dq contenuta entro l'elemento infinitesimo di volume $dV \equiv d^3\mathbf{x} \equiv dx\,dy\,dz$ è

$$dq = \rho(\mathbf{x}, t) dV, \tag{7.2}$$

e la corrente dI che fluisce attraverso l'elemento infinitesimo di superficie $d\sigma$ è

$$dI = \mathbf{j}(\mathbf{x}, t) \cdot d\sigma, \tag{7.3}$$

dove $d\sigma = \mathbf{n} d\sigma$, essendo $\mathbf{n}$ il versore della normale esterna alla superficie.
La carica totale Q contenuta all'interno di quel volume in un certo istante t sarà data ovviamente da

$$Q(t) = \int_V \rho(\mathbf{x}, t) dV. \tag{7.4}$$

Invece, la corrente I che fluisce attraverso la superficie S (cioè il flusso di $\mathbf{j}$ attraverso S) all'istante t è dato, in maniera analoga, dall'integrale

$$I(t) = \int_S \mathbf{j}(\mathbf{x}, t) \cdot d\sigma. \tag{7.5}$$

Se facciamo la convenzione di orientare l'elemento di superficie $d\sigma$ verso l'esterno di V (il che significa che il versore $\mathbf{n}$ della normale alla superficie S, in un suo punto generico, è orientato verso l'esterno), *il flusso entrante sarà considerato* **negativo** *mentre il flusso uscente sarà* **positivo**. Perciò se la carica Q all'interno del volume V aumenta, deve esserci un flusso di corrente entrante (quindi **negativo**) attraverso la superficie chiusa S che delimita V. Viceversa, se la carica diminuisce, ci deve essere un flusso di corrente uscente (cioè **positivo**).
Tra ρ e $\mathbf{j}$, come abbiamo osservato più sopra, esiste l'importante relazione espressa dalla legge di continuità (7.1). Il significato fisico di questa equazione diventa chiaro se la riscriviamo in forma integrale. Infatti, integrando primo e secondo membro della (7.1) su di un generico volume V, avremo:

$$\int_V \operatorname{div} \mathbf{j}\, dV + \int_V \frac{\partial \rho}{\partial t} dV = 0. \tag{7.6}$$

Se ora applichiamo al 1° termine del primo membro il teorema della divergenza di Gauss, possiamo trasformare l'integrale di volume in un integrale di superficie, inoltre tenendo conto

[6] Per i concetti fondamentali sull'elettromagnetismo si consiglia il lettore di consultare il libro di *Fisica 2*: Edoardo Amaldi, Romano Bizzarri, Guido Pizzella, *Fisica Generale, Elettromagnetismo, Relatività, Ottica*, Zanichelli Editore, Bologna, 1986, oppure: D. Halliday, R. Resnick, K. S. Krane, *Fisica 2*, Casa Editrice Ambrosiana, Milano, 1994.

che nel secondo integrale l'unica quantità dipendente da t è la densità di carica possiamo permutare l'operazione d'integrazione con quella di derivazione e scrivere la derivata parziale come una derivata totale, avremo perciò:

$$\int_S \mathbf{j}(\mathbf{x}, t) \cdot d\boldsymbol{\sigma} = -\frac{d}{dt} \int_V \rho(\mathbf{x}, t) dV. \tag{7.7}$$

Per le (7.5) e (7.4), la precedente equazione equivale alla

$$I = -\frac{dQ}{dt}, \tag{7.8}$$

dove Q è la carica totale contenuta in V ed I la corrispondente corrente che fluisce attraverso S. La (7.8) è dunque equivalente all'equazione di continuità (7.1) e da essa risulta evidente quanto detto prima, cioè: se Q aumenta ($dQ/dt > 0$), c'è un flusso di corrente entrante, perciò negativo, cioè una certa carica elettrica entra in V; viceversa, se Q diminuisce ($dQ/dt < 0$), c'è un flusso di corrente verso l'esterno, dunque positivo, cioè una certa quantità di carica esce da V.
Se nella (7.7) si fa tendere V all'infinito, cioè si considera la carica Q_{tot} contenuta nell'intero spazio, l'integrale si superficie al secondo membro si annulla, per il semplice fatto che all'infinito non ci può essere un flusso di cariche, si trova perciò che la carica totale dell'universo è costante

$$\frac{dQ_{tot}}{dt} = 0 \Longrightarrow Q_{tot} = cost., \tag{7.9}$$

che costituisce l'aspetto più generale della *legge di conservazione della carica.*

7.3 Covarianza delle equazioni di Maxwell

Ci proponiamo ora il problema di verificare se le equazioni di Maxwell, riferite al vuoto, si trasformano per invarianza oppure no rispetto ad una trasformazione di Lorentz, vale a dire se le leggi fisiche che esse descrivono soddisfano al Principio di Relatività di Einstein.
Si osservi a questo proposito che il fatto, che da tali equazioni si possa dedurre la legge di propagazione dei raggi luminosi (in generale, delle onde elettromagnetiche) con velocità costante rispetto ad un osservatore qualsiasi, non è condizione sufficiente per assicurare che le equazioni di Maxwell sono covarianti rispetto ad una trasformazione di Lorentz. Infatti la legge di propagazione delle onde elettromagnetiche è solo un caso particolare dei fenomeni fisici descritti dalle equazioni di Maxwell, e non è affatto evidente a priori che anche le leggi dei numerosi altri fenomeni derivanti da queste equazioni soddisfino alle richieste condizioni di covarianza.
A suo tempo, dicemmo che Lorentz riuscì a dimostrare, ancor prima della nascita della relatività einsteiniana, che le equazioni di Maxwell sono covarianti rispetto ad una trasformazione speciale di Lorentz.
Ci occuperemo ora di questa dimostrazione, limitandoci alla considerazione delle equazioni di Maxwell nel vuoto per il caso particolare in cui non vi siano né cariche ($\rho = 0$), né correnti ($\mathbf{j} = 0$)[7].
Siano $\mathbf{E} \equiv (E_x, E_y, E_z)$ ed $\mathbf{H} \equiv (H_x, H_y, H_z)$ rispettivamente il *vettore campo elettrico* ed il *vettore campo magnetico* (dello spazio ordinario) che caratterizzano un *campo elettromagnetico* in un certo sistema inerziale $\mathcal{R} \equiv (Oxyz)$. Come abbiamo visto nel Cap.1, paragrafo 1.4.3, le equazioni di

[7]In ciò che segue, per semplicità di calcolo, ci riferiamo al caso più semplice in cui le equazioni di Maxwell sono rappresentate dalle equazioni (7.10), (7.11), ma avvertiamo il lettore che la covarianza per trasformazioni di Lorentz è del tutto generale e vale perciò anche per le equazioni scritte nella forma (1.60), nelle quali anziché i vettori campo elettrico **E** e campo magnetico **H** vengono utilizzati i vettori **induzione elettrica D** ed **induzione magnetica B**.

Maxwell nel vuoto, in assenza di cariche e nel sistema di Gauss, possono essere raggruppate in due coppie:

$$\begin{cases} \operatorname{div} \mathbf{H} = 0, \\ \\ \operatorname{rot} \mathbf{E} + \dfrac{1}{c}\dfrac{\partial \mathbf{H}}{\partial t} = 0, \end{cases} \tag{7.10}$$

$$\begin{cases} \operatorname{div} \mathbf{E} = 0, \\ \\ \operatorname{rot} \mathbf{H} - \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t} = 0, \end{cases} \tag{7.11}$$

che chiameremo, seguendo Landau-Lifschitz, 1^a e 2^a *coppia delle equazioni di Maxwell*. Riscriviamole ora in forma scalare facendo intervenire le componenti dei campi rispetto al sistema galileiano $\mathcal{R}$, avremo per la 1^a coppia:

$$\begin{cases} \dfrac{\partial H_x}{\partial x} + \dfrac{\partial H_y}{\partial y} + \dfrac{\partial H_z}{\partial z} = 0, \\ \\ \dfrac{\partial E_z}{\partial y} - \dfrac{\partial E_y}{\partial z} + \dfrac{1}{c}\dfrac{\partial H_x}{\partial t} = 0, \\ \\ \dfrac{\partial E_x}{\partial z} - \dfrac{\partial E_z}{\partial x} + \dfrac{1}{c}\dfrac{\partial H_y}{\partial t} = 0, \\ \\ \dfrac{\partial E_y}{\partial x} - \dfrac{\partial E_x}{\partial y} + \dfrac{1}{c}\dfrac{\partial H_z}{\partial t} = 0, \end{cases} \tag{7.12}$$

e analogamente per la 2^a coppia:

$$\begin{cases} \dfrac{\partial E_x}{\partial x} + \dfrac{\partial E_y}{\partial y} + \dfrac{\partial E_z}{\partial z} = 0, \\ \\ \dfrac{\partial H_z}{\partial y} - \dfrac{\partial H_y}{\partial z} - \dfrac{1}{c}\dfrac{\partial E_x}{\partial t} = 0, \\ \\ \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} - \dfrac{1}{c}\dfrac{\partial E_y}{\partial t} = 0, \\ \\ \dfrac{\partial H_y}{\partial x} - \dfrac{\partial H_x}{\partial y} - \dfrac{1}{c}\dfrac{\partial E_z}{\partial t} = 0, \end{cases} \tag{7.13}$$

Supponiamo ora di osservare il campo elettromagnetico, di cui stiamo parlando, da un sistema inerziale $\mathcal{R}' \equiv (O'x'y'z')$ in configurazione x-standard rispetto ad $\mathcal{R} \equiv (Oxyz)$, che si muova con velocità di traslazione $\mathbf{v}$ rispetto a quest'ultimo. Si tratta di determinare come si trasformano le componenti del campo elettrico (E_x, E_y, E_z) e del campo magnetico (H_x, H_y, H_z), applicando una trasformazione di Lorentz alle coordinate x, y, z, t affinché conformemente al Principio Einsteiniano di Relatività, le (7.12), (7.13) conservino, nelle componenti trasformate del campo elettrico $(E'_{x'}, E'_{y'}, E'_{z'})$ e del campo magnetico $(H'_{x'}, H'_{y'}, H'_{z'})$, esattamente la stessa forma[8]; vale

[8]Come abbiamo già avvertito nel Cap.4 paragrafo 4.1.10, formula (4.15), riguardo alle componenti cartesiane della velocità, per non appesantire eccessivamente le notazioni, le componenti del campo elettrico e del campo magnetico

a dire affinché siano soddisfatte le equazioni:

$$\begin{cases} \dfrac{\partial H'_x}{\partial x'} + \dfrac{\partial H'_y}{\partial y'} + \dfrac{\partial H'_z}{\partial z'} = 0, \\[2ex] \dfrac{\partial E'_z}{\partial y'} - \dfrac{\partial E'_y}{\partial z'} + \dfrac{1}{c}\dfrac{\partial H'_x}{\partial t'} = 0, \\[2ex] \dfrac{\partial E'_x}{\partial z'} - \dfrac{\partial E'_z}{\partial x'} + \dfrac{1}{c}\dfrac{\partial H'_y}{\partial t'} = 0, \\[2ex] \dfrac{\partial E'_y}{\partial x'} - \dfrac{\partial E'_x}{\partial y'} + \dfrac{1}{c}\dfrac{\partial H'_z}{\partial t'} = 0, \end{cases} \tag{7.14}$$

$$\begin{cases} \dfrac{\partial E'_x}{\partial x'} + \dfrac{\partial E'_y}{\partial y'} + \dfrac{\partial E'_z}{\partial z'} = 0, \\[2ex] \dfrac{\partial H'_z}{\partial y'} - \dfrac{\partial H'_y}{\partial z'} - \dfrac{1}{c}\dfrac{\partial E'_x}{\partial t'} = 0, \\[2ex] \dfrac{\partial H'_x}{\partial z'} - \dfrac{\partial H'_z}{\partial x'} - \dfrac{1}{c}\dfrac{\partial E'_y}{\partial t'} = 0, \\[2ex] \dfrac{\partial H'_y}{\partial x'} - \dfrac{\partial H'_x}{\partial y'} - \dfrac{1}{c}\dfrac{\partial E'_z}{\partial t'} = 0. \end{cases} \tag{7.15}$$

Osserviamo ora che, come conseguenza della trasformazione di Lorentz

$$(L) \qquad \begin{cases} x' = \gamma(v)(x - vt) = \gamma(v)(x - \beta ct), \\ y' = y, \\ z' = z, \\ t' = \gamma(v)\left(t - \dfrac{v}{c^2}x\right) = \gamma(v)\left(t - \dfrac{\beta}{c}x\right), \end{cases}$$

se si calcolano le derivate parziali di quantità che dipendono dalle variabili x, y, z, t legate alle analoghe variabili accentate x', y', z', t' dalla trasformazione di Lorentz, si ottiene

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial}{\partial y'}\frac{\partial y'}{\partial x} + \frac{\partial}{\partial z'}\frac{\partial z'}{\partial x} + \frac{\partial}{\partial t'}\frac{\partial t'}{\partial x} = \gamma\frac{\partial}{\partial x'} - \frac{\gamma v}{c^2}\frac{\partial}{\partial t'}, \text{.....e analoghe.}$$

nel sistema di riferimento $\mathcal{R}'$, d'ora in poi le indicheremo così: (E'_x, E'_y, E'_z), (H'_x, H'_y, H'_z), anziché con le notazioni più precise $(E'_{x'}, E'_{y'}, E'_{z'})$, $(H'_{x'}, H'_{y'}, H'_{z'})$.

A conti fatti avremo:

$$\begin{cases} \dfrac{\partial}{\partial x} = \gamma(v)\Big(\dfrac{\partial}{\partial x'} - \dfrac{\beta}{c}\dfrac{\partial}{\partial t'}\Big) \iff \dfrac{\partial}{\partial x'} = \dfrac{1}{\gamma(v)}\dfrac{\partial}{\partial x} + \dfrac{\beta}{c}\dfrac{\partial}{\partial t'} \\ \\ \dfrac{\partial}{\partial y} = \dfrac{\partial}{\partial y'}, \\ \\ \dfrac{\partial}{\partial z} = \dfrac{\partial}{\partial z'}, \\ \\ \dfrac{\partial}{\partial t} = \gamma(v)\Big(-v\dfrac{\partial}{\partial x'} + \dfrac{\partial}{\partial t'}\Big) \iff \dfrac{\partial}{\partial t'} = \dfrac{1}{\gamma(v)}\dfrac{\partial}{\partial t} + v\dfrac{\partial}{\partial x'} \end{cases} \tag{7.16}$$

Ora è necessario manipolare, in maniera opportuna, le equazioni (7.12) e (7.13) per ottenere il risultato voluto.
Partiamo dalla 1^a delle (7.12) ed utilizziamo le (7.16), otteniamo:

$$\gamma\frac{\partial H_x}{\partial x'} - \frac{\gamma v}{c^2}\frac{\partial H_x}{\partial t'} + \frac{\partial H_y}{\partial y'} + \frac{\partial H_z}{\partial z'} = 0. \tag{7.17}$$

Se invece sostituiamo nella 2^a delle (7.12) le (7.16), avremo:

$$\frac{\partial E_z}{\partial y'} - \frac{\partial E_y}{\partial z'} - \frac{\gamma v}{c}\frac{\partial H_x}{\partial x'} + \frac{\gamma}{c}\frac{\partial H_x}{\partial t'} = 0,$$

isolando in quest'ultima l'ultimo termine avremo:

$$\frac{\gamma}{c}\frac{\partial H_x}{\partial t'} = \frac{\partial E_y}{\partial z'} - \frac{\partial E_z}{\partial y'} + \frac{\gamma v}{c}\frac{\partial H_x}{\partial x'}. \tag{7.18}$$

Sostituiamo nel 2° termine della (7.17) il valore di $\dfrac{\gamma}{c}\dfrac{\partial H_x}{\partial t'}$ dato dalla (7.18), si avrà:

$$\gamma\frac{\partial H_x}{\partial x'} - \frac{v}{c}\frac{\partial E_y}{\partial z'} + \frac{v}{c}\frac{\partial E_z}{\partial y'} - \frac{\gamma v^2}{c^2}\frac{\partial H_x}{\partial x'} + \frac{\partial H_y}{\partial y'} + \frac{\partial H_z}{\partial z'} = 0. \tag{7.19}$$

Semplificando ed ordinando i termini in modo opportuno, in definitiva otteniamo:

$$\frac{\partial H_x}{\partial x'} + \frac{\partial}{\partial y'}\Big[\gamma\Big(H_y + \frac{v}{c}E_z\Big)\Big] + \frac{\partial}{\partial z'}\Big[\gamma\Big(H_z - \frac{v}{c}E_y\Big)\Big] = 0. \tag{7.20}$$

Operando in maniera analoga, le equazioni (7.12) possono, includendovi anche la (7.20), essere riscritte così:

$$\begin{cases} \dfrac{\partial H_x}{\partial x'} + \dfrac{\partial}{\partial y'}\Big[\gamma\Big(H_y + \dfrac{v}{c}E_z\Big)\Big] + \dfrac{\partial}{\partial z'}\Big[\gamma\Big(H_z - \dfrac{v}{c}E_y\Big)\Big] = 0, \\ \\ \dfrac{\partial}{\partial y'}\Big[\gamma\Big(E_z + \dfrac{v}{c}H_y\Big)\Big] - \dfrac{\partial}{\partial z'}\Big[\gamma\Big(E_y - \dfrac{v}{c}H_z\Big)\Big] + \dfrac{1}{c}\dfrac{\partial H_x}{\partial t'} = 0, \\ \\ \dfrac{\partial E_x}{\partial z'} - \dfrac{\partial}{\partial x'}\Big[\gamma\Big(E_z + \dfrac{v}{c}H_y\Big)\Big] + \dfrac{1}{c}\dfrac{\partial}{\partial t'}\Big[\gamma\Big(H_y + \dfrac{v}{c}E_z\Big)\Big] = 0, \\ \\ \dfrac{\partial}{\partial x'}\Big[\gamma\Big(E_y - \dfrac{v}{c}H_z\Big)\Big] - \dfrac{\partial E_x}{\partial y'} + \dfrac{1}{c}\dfrac{\partial}{\partial t'}\Big[\gamma\Big(H_z - \dfrac{v}{c}E_y\Big)\Big] = 0, \end{cases} \tag{7.21}$$

e le (7.13) in quest'altro modo:

$$
\begin{cases}
\dfrac{\partial E_x}{\partial x'} + \dfrac{\partial}{\partial y'}\left[\gamma\left(E_y - \dfrac{v}{c}H_z\right)\right] + \dfrac{\partial}{\partial z'}\left[\gamma\left(E_z + \dfrac{v}{c}H_y\right)\right] = 0, \\[2ex]
\dfrac{\partial}{\partial y'}\left[\gamma\left(H_z - \dfrac{v}{c}E_y\right)\right] - \dfrac{\partial}{\partial z'}\left[\gamma\left(H_y + \dfrac{v}{c}E_z\right)\right] - \dfrac{1}{c}\dfrac{\partial E_x}{\partial t'} = 0, \\[2ex]
\dfrac{\partial H_x}{\partial z'} - \dfrac{\partial}{\partial x'}\left[\gamma\left(H_y + \dfrac{v}{c}E_z\right)\right] - \dfrac{1}{c}\dfrac{\partial}{\partial t'}\left[\gamma\left(E_z + \dfrac{v}{c}H_y\right)\right] = 0, \\[2ex]
\dfrac{\partial}{\partial x'}\left[\gamma\left(H_y + \dfrac{v}{c}E_z\right)\right] - \dfrac{\partial H_x}{\partial y'} - \dfrac{1}{c}\dfrac{\partial}{\partial t'}\left[\gamma\left(E_z + \dfrac{v}{c}H_y\right)\right] = 0.
\end{cases}
\tag{7.22}
$$

Affinché le (7.21) e (7.22) abbiano la forma (7.14) e (7.15) rispettivamente, deve essere, salvo un coefficiente di proporzionalità che evidentemente è uguale ad 1 e ponendo $\dfrac{v}{c} = \beta$:

$$
\begin{cases}
E'_x = E_x, \\[2ex]
E'_y = \gamma(v)\left(E_y - \beta H_z\right) = \dfrac{E_y - \beta H_z}{\sqrt{1-\beta^2}} \\[2ex]
E'_z = \gamma(v)\left(E_z + \beta H_y\right) = \dfrac{E_z + \beta H_y}{\sqrt{1-\beta^2}} \\[2ex]
H'_x = H_x, \\[2ex]
H'_y = \gamma(v)\left(H_y + \beta E_z\right) = \dfrac{H_y + \beta E_z}{\sqrt{1-\beta^2}} \\[2ex]
H'_z = \gamma(v)\left(H_z - \beta E_y\right) = \dfrac{H_z - \beta E_y}{\sqrt{1-\beta^2}}.
\end{cases}
\tag{7.23}
$$

Per il passaggio inverso avremo:

$$\begin{cases} E_x = E'_x, \\ E_y = \gamma(v)\left(E'_y + \beta H'_z\right) = \dfrac{E'_y + \beta H'_z}{\sqrt{1-\beta^2}} \\ E_z = \gamma(v)\left(E'_z - \beta H'_y\right) = \dfrac{E'_z - \beta H'_y}{\sqrt{1-\beta^2}} \\ H_x = H'_x, \\ H_y = \gamma(v)\left(H'_y - \beta E'_z\right) = \dfrac{H'_y - \beta E'_z}{\sqrt{1-\beta^2}} \\ H_z = \gamma(v)\left(H'_z + \beta E'_y\right) = \dfrac{H'_z + \beta E'_y}{\sqrt{1-\beta^2}}. \end{cases} \tag{7.24}$$

Sono queste le formule di trasformazione cercate, le quali permettono di determinare le componenti del campo elettrico e del campo magnetico, o meglio del *campo elettromagnetico* quando si passa da un sistema di riferimento inerziale $\mathcal{R}$ ad un'altro $\mathcal{R}'$ in configurazione x-standard. Come si vede, il campo elettrico ed il campo magnetico sono, come pure la maggior parte delle grandezze fisiche, relativi; le loro proprietà sono differenti in diversi sistemi di riferimento. In particolare un campo elettrico e un campo magnetico possono essere nulli in un sistema di riferimento ed essere nello stesso tempo presenti in un altro sistema.
Le formule di trasformazione (7.24) si semplificano notevolmente nel caso in cui $v \ll c$ e perciò siano trascurabili i termini del 2° ordine in β, cioè i termini contenenti $\beta^2 = \dfrac{v^2}{c^2}$. Con tale approssimazione le precedenti equazioni si riducono alle:

$$\begin{cases} E_x \approx E'_x, \\ E_y \approx E'_y + \beta H'_z, \\ E_z \approx E'_z - \beta H'_y, \\ H_x \approx H'_x, \\ H_y \approx H'_y - \beta E'_z, \\ H_z \approx H'_z + \beta E'_y. \end{cases} \tag{7.25}$$

Queste formule si possono scrivere in forma vettoriale così:

$$\mathbf{E} \approx \mathbf{E}' - \frac{1}{c}\mathbf{v} \wedge \mathbf{H}', \qquad \mathbf{H} \approx \mathbf{H}' + \frac{1}{c}\mathbf{v} \wedge \mathbf{E}'. \tag{7.26}$$

Le formule della trasformazione inversa da $\mathcal{R}'$ ad $\mathcal{R}$ si ricavano dalle (7.25) e (7.26), come al solito, scambiando le lettere accentate con quelle non accentate e cambiando v in $-v$; le (7.26), per esempio, diventano:

$$\mathbf{E}' \approx \mathbf{E} + \frac{1}{c}\mathbf{v} \wedge \mathbf{H}, \qquad \mathbf{H}' \approx \mathbf{H} - \frac{1}{c}\mathbf{v} \wedge \mathbf{E}. \tag{7.27}$$

Osservazione importante:
La legge di trasformazione delle componenti del *campo elettromagnetico* non solo ha il pregio di

mettere in evidenza la covarianza delle equazioni di Maxwell rispetto ad una trasformazione di Lorentz e di permettere l'esecuzione rapida dei calcoli per il passaggio da un sistema di riferimento ad un altro, ma porta altresì ad una nuova e più profonda comprensione del campo elettromagnetico stesso. Questa maggiore comprensione è una conseguenza dell'intima fusione del campo elettrico con quello magnetico, i quali nella Teoria della Relatività costituiscono, a differenza della concezione classica, un unico ente; essa permette inoltre di dare, anche nella approssimazione della fisica classica (vale a dire quando si trascurano termini del 2° ordine in β, cioè contenenti $\beta^2 = \frac{v^2}{c^2}$) un'interpretazione di alcuni fenomeni elettromagnetici diversa e più espressiva di quella ordinaria.
Illustriamo questa circostanza con un semplice esempio. Siano, come al solito, $\mathcal{R}'$ un sistema galileiano che si muove di moto TRU rispetto ad un analogo riferimento $\mathcal{R}$, con velocità $\mathbf{v}$ diretta lungo l'asse $x \equiv x'$, quindi in configurazione x-standard. Si consideri l'azione di un campo elettromagnetico su di una carica q solidale ad $\mathcal{R}'$ e perciò in moto rispetto ad $\mathcal{R}$ con velocità $\mathbf{v}$. Secondo l'ordinaria concezione maxwelliana, in $\mathcal{R}'$ si constaterà che il campo elettrico esercita sulla carica q una forza

$$\mathbf{F}' = q\mathbf{E}', \tag{7.28}$$

mentre il campo magnetico non esercita nessuna forza in quanto la carica è in quiete rispetto ad $\mathcal{R}'$ ($\mathbf{u}' \wedge \mathbf{H}' = 0$ *essendo* $\mathbf{u}' = 0$). Nel sistema $\mathcal{R}$, rispetto al quale la carica si muove con velocità $\mathbf{v}$, la forza $\mathbf{F}$ corrispondente di $\mathbf{F}'$, dal punto di vista relativistico e nell'approssimazione di cui sopra, cioè per basse velocità, si ottiene sostituendo nella (7.28) la prima delle (7.27), si avrà perciò:

$$\mathbf{F} = q(\mathbf{E} + \frac{1}{c}\mathbf{v} \wedge \mathbf{H}), \tag{7.29}$$

che è la ben nota **forza di Lorentz**. Da ciò si vede come, *secondo la concezione relativistica, la validità* **della legge di forza di Lorentz** *non deve essere assunta come un fatto sperimentale nuovo, indipendente dalle equazioni del campo elettromagnetico, come si verifica nell'ordinaria concezione maxwelliana, ma è strettamente legata alle leggi di trasformazione di detto campo.*

7.4 Leggi di trasformazione del campo elettrico E e magnetico H

È istruttivo riottenere le formule (7.23) e (7.24) di trasformazione dei campi elettrici e magnetici in una maniera diversa, utilizzando, come fatto sperimentalmente noto, l'espressione della forza di Lorentz (7.29) e le formule (6.90) di trasformazione delle componenti di una generica forza nel passaggio dal riferimento $\mathcal{R}$ al riferimento $\mathcal{R}'$.
Consideriamo una particella di carica q che si muove con velocità $\mathbf{u}$ nel riferimento $\mathcal{R}$ in un campo elettromagnetico definito dalla coppia di vettori $(\mathbf{E}, \mathbf{H})$; ad un istante generico la particella sarà soggetta alla forza di Lorentz:

$$\mathbf{F} = q(\mathbf{E} + \frac{1}{c}\mathbf{u} \wedge \mathbf{H}). \tag{7.30}$$

Ricordiamo che la forza elettrica $q\mathbf{E}$ non dipende dal moto della particella di carica q, mentre dipende da questo moto la forza magnetica $\frac{q}{c}\mathbf{u} \wedge \mathbf{H}$ che agisce su di essa. Ad ogni modo, il moto di una particella dipende dal riferimento in cui viene descritto, perciò non dovremmo essere sorpresi che anche i campi dipendano dal riferimento in cui sono descritti. Per semplicità, descriveremo le trasformazioni che subiscono i campi considerando alcuni casi particolari; i risultati, tuttavia, hanno validità del tutto generale.
Abbiamo visto nel paragrafo 6.3.9 (formula (6.93)), che le formule di trasformazione per le componenti cartesiane di una forza nel passaggio dal riferimento $\mathcal{R}$ ad un altro, in configurazione

x-standard, in cui la particella, su cui agisce la forza, è istantaneamente in riposo ($\mathbf{u}' = 0$), sono:

$$\begin{cases} F_x &= F'_x, \\ F_y &= \dfrac{F'_y}{\gamma(v)} = F'_y\sqrt{1-\beta_v^2}, \\ F_z &= \dfrac{F'_z}{\gamma(v)} = F'_z\sqrt{1-\beta_v^2}, \end{cases} \tag{7.31}$$

dove $\beta_v = v/c$, ed F_x, F_y, F_z sono le componenti della forza nel riferimento $\mathcal{R}$ ed F'_x, F'_y, F'_z sono le analoghe componenti della forza nel riferimento $\mathcal{R}'$.
Consideriamo dunque una particella di carica q istantaneamente in quiete in $\mathcal{R}'$, dove esiste un campo elettromagnetico di componenti $\mathbf{E}', \mathbf{H}'$. La forza elettromagnetica agente sulla particella sarà $\mathbf{F}' = q\mathbf{E}'$, poiché non si esercita alcuna forza magnetica su una particella carica in riposo, perciò le sue componenti cartesiane sono:

$$F'_x = qE'_x, \qquad F'_y = qE'_y, \qquad F'_z = qE'_z. \tag{7.32}$$

Nel riferimento $\mathcal{R}$ la forza corrispondente è data dalla forza di Lorentz:

$$\mathbf{F} = q(\mathbf{E} + \frac{1}{c}\mathbf{v} \wedge \mathbf{H}), \tag{7.33}$$

in quanto in questo riferimento la carica ha una velocità $\mathbf{u} \equiv \mathbf{v}$ essendo essa solidale con $\mathcal{R}'$. Le componenti cartesiane di $\mathbf{F}$ saranno

$$F_x = q\big[E_x + \frac{1}{c}(\mathbf{v} \wedge \mathbf{H})_x\big], \quad F_y = q\big[E_y + \frac{1}{c}(\mathbf{v} \wedge \mathbf{H})_y\big], \quad F_z = q\big[E_z + \frac{1}{c}(\mathbf{v} \wedge \mathbf{H})_z\big]. \tag{7.34}$$

Poiché le componenti cartesiane di $\mathbf{v}$ sono: $v_x = v$, $v_y = 0, v_z = 0$, avremo:

$$(\mathbf{v} \wedge \mathbf{H})_x = v_y H_z - v_z H_y = 0, \quad (\mathbf{v} \wedge \mathbf{H})_y = v_z H_x - v_x H_z = -vH_z, \quad (\mathbf{v} \wedge \mathbf{H})_z = v_x H_y - v_y H_x = vH_y, \tag{7.35}$$

e le (7.34) diventano:

$$F_x = qE_x, \qquad F_y = q(E_y - vH_z), \qquad F_z = q(E_z + vH_y). \tag{7.36}$$

Se ora, fra le (7.31), (7.32) e (7.36), eliminiamo F_x, F_y, F_z e F'_x, F'_y, F'_z otteniamo il risultato voluto, cioè *le formule di trasformazione per le componenti del campo elettrico*:

$$\begin{cases} E'_x = E_x, \\ E'_y = \gamma(v)\Big(E_y - \beta H_z\Big) \\ E'_z = \gamma(v)\Big(E_z + \beta H_y\Big), \end{cases} \tag{7.37}$$

come si vede esse coincidono con il primo gruppo delle (7.24) già ricavate, per altra via, nel paragrafo precedente.
Per ottenere le trasformazioni inverse, come al solito, si deve cambiare β in $-\beta$ e le quantità accentate con quelle non accentate, si avrà perciò:

$$\begin{cases} E_x = E'_x, \\ E_y = \gamma(v)\Big(E'_y + \beta H'_z\Big) \\ E_z = \gamma(v)\Big(E'_z - \beta H'_y\Big). \end{cases} \tag{7.38}$$

La restrizione che la velocità $\mathbf{v}$ di $\mathcal{R}'$ rispetto ad $\mathcal{R}$ sia diretta lungo l'asse x non è necessaria perché l'orientazione di questo asse è arbitraria. Se, peraltro, introduciamo le componenti dei campi parallela ($\|$) e perpendicolare ($\perp$) alla direzione della velocità relativa, possiamo scrivere le precedenti trasformazioni in modo del tutto generale nella forma:

$$\begin{cases} E'_{\|} = E_{\|}, \\ E'_{\perp} = \gamma(v)\Big[E_{\perp} + \dfrac{1}{c}(\mathbf{v} \wedge \mathbf{H})_{\perp}\Big]. \end{cases} \tag{7.39}$$

Esse affermano che la componente di $\mathbf{E}$ parallela alla velocità relativa dei due riferimenti rimane invariata, mentre le componenti di $\mathbf{E}$ perpendicolari alla stessa velocità si trasformano in campi elettrici e magnetici mescolati.
Occupiamoci ora della trasformazione delle componenti del *campo magnetico.* Consideriamo ancora un caso piuttosto semplice i cui risultati hanno validità generale. Supponiamo ancora che il moto relativo dei due sistemi inerziali $\mathcal{R}$ e $\mathcal{R}'$ avvenga lungo l'asse comune $x \equiv x'$. Si consideri una particella di carica q che si muova nel riferimento $\mathcal{R}'$ nella direzione dell'asse y', con velocità $\mathbf{u}'$. Nel riferimento $\mathcal{R}'$ le componenti della velocità della particella sono perciò $\mathbf{u}' \equiv (u'_x = 0, u'_y = u', u'_z = 0)$, essendo u' il modulo della velocità. In $\mathcal{R}'$ la forza di Lorentz a cui è sottoposta la particella è $\mathbf{F}' = q(\mathbf{E}' + \dfrac{1}{c}\mathbf{u}' \wedge \mathbf{H}')$, ed è facile accertarsi che le sue componenti sono:

$$\begin{cases} F'_x = q(E'_x + \dfrac{u'_y}{c}H'_z), \\ F'_y = qE'_y, \\ F'_z = q(E'_z - \dfrac{u'_y}{c}H'_x). \end{cases} \tag{7.40}$$

Per ottenere le componenti della forza nel riferimento $\mathcal{R}$ dobbiamo sapere qual'è la velocità della particella in questo riferimento. Ciò si ottiene utilizzando le formule relativistiche di trasformazione delle componenti della velocità nel passaggio dal riferimento $\mathcal{R}$ ad $\mathcal{R}'$, cioè le formule (4.17) che, per comodità del lettore, riportiamo qui sotto:

$$\begin{cases} u_x = \dfrac{u'_x + v}{1 + \dfrac{v}{c^2}u'_x} \\[2ex] u_y = \dfrac{u'_y\sqrt{1-\beta^2}}{1 + \dfrac{v}{c^2}u'_x} = \dfrac{u'_y}{\gamma(v)(1 + \dfrac{v}{c^2}u'_x)} \\[2ex] u_z = \dfrac{u'_z\sqrt{1-\beta^2}}{1 + \dfrac{v}{c^2}u'_x} = \dfrac{u'_z}{\gamma(v)(1 + \dfrac{v}{c^2}u'_x)}. \end{cases}$$

Da queste formule si ottiene $u_x = v, u_y = u'_y/\gamma(v) = u'/\gamma(v), u_z = 0$, il che significa che sebbene in $\mathcal{R}'$ la particella abbia solo una componente lungo l'asse y', in $\mathcal{R}$ la velocità della particella ha, diverse da zero, le componenti sia lungo l'asse x sia lungo l'asse y.
Dunque, in $\mathcal{R}$, la forza di Lorentz $\mathbf{F} = q(\mathbf{E} + \dfrac{1}{c}\mathbf{u} \wedge \mathbf{H})$ ha le componenti:

$$\begin{cases} F_x = q(E_x + \dfrac{u_y}{c}H_z), \\ F_y = q(E_y - \beta H_z), \\ F_z = q(E_z + \beta H_y - \dfrac{u_y}{c}H_z). \end{cases} \tag{7.41}$$

Per ottenere le relazioni fra le componenti dei campi è necessario eliminare $F_x, F_y, F_z, F'_x, F'_y, F'_z$ date dalle (7.40), (7.41) con le espressioni (6.89), (6.90) che forniscono le equazioni generali di trasformazione delle componenti di una forza, tenendo conto nello stesso tempo delle (7.37), (7.38). Troveremo la prima e terza componente del campo magnetico

$$H'_x = H_x, \quad e \quad H'_z = \gamma(v)(H_z - \beta E_y).$$

Per trovare come si trasforma la seconda componente del campo magnetico è necessario far muovere la particella di carica q nella direzione z' invece che nella direzione y'. Con un procedimento simile al precedente troveremo $H'_y = \gamma(v)(H_y + \beta E_z)$. Raccogliendo i risultati ottenuti avremo *le formule di trasformazione delle componenti del campo magnetico*:

$$\begin{cases} H'_x = H_x, \\ H'_y = \gamma(v)\Big(H_y + \beta E_z\Big) = \dfrac{H_y + \beta E_z}{\sqrt{1-\beta^2}} \\ H'_z = \gamma(v)\Big(H_z - \beta E_y\Big) = \dfrac{H_z - \beta E_y}{\sqrt{1-\beta^2}}. \end{cases} \tag{7.42}$$

Le formule inverse si trovano, come al solito, scambiando v con $-v$ e le lettere con apici con quelle senza apici:

$$\begin{cases} H_x = H'_x, \\ H_y = \gamma(v)\Big(H'_y - \beta E'_z\Big) = \dfrac{H'_y - \beta E'_z}{\sqrt{1-\beta^2}} \\ H_z = \gamma(v)\Big(H'_z + \beta E'_y\Big) = \dfrac{H'_z + \beta E'_y}{\sqrt{1-\beta^2}}. \end{cases} \tag{7.43}$$

Come si vede le (7.42), (7.43) coincidono perfettamente con le formule (7.23), (7.24) già ottenute, per altra via, nel paragrafo precedente.
Possiamo ancora osservare, che non c'è bisogno di porre la limitazione che la velocità $\mathbf{v}$ di $\mathcal{R}'$ rispetto ad $\mathcal{R}$ sia lungo l'asse x. Prendendo le componenti del campo parallela e perpendicolare alla direzione della velocità $\mathbf{v}$, possiamo scrivere queste trasformazioni in modo del tutto generale, come abbiamo già fatto per le componenti del campo elettrico:

$$\begin{cases} H'_{\parallel} = H_{\parallel}, \\ H'_{\perp} = \gamma(v)\Big[H_{\perp} - \dfrac{1}{c}(\mathbf{v} \wedge \mathbf{E})_{\perp}\Big]. \end{cases} \tag{7.44}$$

Queste relazioni affermano che le componenti di $\mathbf{H}$ parallele alla velocità relativa dei due sistemi di riferimento rimangono invariate, mentre le sue componenti ortogonali si trasformano in campi magnetici ed elettrici mescolati.
Tutto ciò che abbiamo detto finora ci fa capire la interdipendenza fra il campo elettrico $\mathbf{E}$ e il campo magnetico $\mathbf{H}$, poiché le leggi di trasformazione che abbiamo dimostrato legano tutte e sei le componenti di $\mathbf{E}$ ed $\mathbf{H}$, piuttosto che darci due leggi di trasformazione distinte, una per $\mathbf{E}$ ed una per $\mathbf{H}$. *I campi elettrici e magnetici non possono esistere indipendentemente l'uno dall'altro come quantità distinte, ma costituiscono un unico ente:* **il campo elettromagnetico**, come vedremo meglio, dal punto di vista 4-dimensionale, nel prossimo paragrafo allorquando definiremo il *tensore del campo elettromagnetico.*

7.5 Digressione sui 4-tensori doppi dello spazio di Minkowski

Nei paragrafi precedenti abbiamo dimostrato che le equazioni di Maxwell del campo elettromagnetico sono covarianti rispetto ad una trasformazione di Lorentz e ne abbiamo determinato le relative leggi di trasformazione. Rileviamo, ancora una volta, che caratteristica peculiare delle formule di trasformazione (7.37) e loro inverse (7.38) per il campo elettrico e delle (7.42) e loro inverse (7.43) per il campo magnetico è che ciascuno dei vettori $\mathbf{E}'$ ed $\mathbf{H}'$ dipende simultaneamente da $\mathbf{E}$ ed $\mathbf{H}$. In altri termini la separazione di un campo elettromagnetico nelle sue parti elettrica e magnetica dipende dal sistema di riferimento inerziale in cui si opera, per cui un campo che in un sistema può venire considerato puramente elettrico, in un altro sistema, in moto rispetto al precedente, esso viene a possedere anche componenti magnetiche. Ora, per esprimere nel migliore dei modi tutto ciò è necessario dare una formulazione 4-dimensionale, nello spazio $\mathcal{M}^4$ di Minkowski, delle equazioni di Maxwell ed è possibile farlo agevolmente introducendo, in maniera opportuna, un *4-tensore antisimmetrico del 2° ordine* F^{ik}, detto **tensore del campo elettromagnetico.**

Nel Cap.5, paragrafo 5.4.1, abbiamo introdotto il concetto di 4-vettore (o tensore del 1° ordine) e le relative formule di trasformazione per le sue componenti controvarianti, tramite la trasformazione di Lorentz, nel passaggio dal sistema $\mathcal{R}$ al sistema $\mathcal{R}'$, fornite dalle (5.30) e loro inverse. Ora, ci occorre introdurre il concetto di 4-tensore del 2° ordine, espresso per mezzo sia delle componenti controvarianti, sia per mezzo delle componenti covarianti, e stabilire le pertinenti formule di trasformazione.[9]

Si definisce *tensore controvariante del 2° ordine* T^{ik}, nello spazio di Minkowski $\mathcal{M}^4$, un insieme di 16 grandezze dipendenti dalle coordinate spazio- temporali $(x^0, x^1, x^2, x^3) \equiv (x^0, x^\alpha)$:

$$(T^{ik}) = \begin{pmatrix} T^{00} & T^{01} & T^{02} & T^{03} \\ T^{10} & T^{11} & T^{12} & T^{13} \\ T^{20} & T^{21} & T^{22} & T^{23} \\ T^{30} & T^{31} & T^{32} & T^{33} \end{pmatrix} \tag{7.45}$$

i cui valori, nel passaggio da $\mathcal{R}$ ad $\mathcal{R}'$, si trasformano secondo la formula:

$$T'^{jh} = \frac{\partial x'^j}{\partial x^i} \frac{\partial x'^h}{\partial x^k} T^{ik}, \tag{7.46}$$

essendo sottintesa una doppia somma rispetto agli indici muti ripetuti i, k. Inoltre le derivate parziali $L^{i'}_j = \dfrac{\partial x'^j}{\partial x^i}$ devono essere calcolate sulla trasformazione di Lorentz (5.11), che per comodità del lettore, riscriviamo:

$$\begin{cases} x'^0 = \gamma(v)(x^0 - \beta x^1) \\ x'^1 = \gamma(v)(x^1 - \beta x^0) \\ x'^2 = x^2 \\ x'^3 = x^3, \end{cases} \tag{7.47}$$

e perciò coincidono con i coefficienti stessi della trasformazione dati dagli elementi della matrice:

$$(L^{i'}_j) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.48}$$

[9] Ci occuperemo della trattazione generale del Calcolo Tensoriale nel Cap.8.

Si definisce *tensore covariante del 2° ordine* T_{ik}, nello spazio di Minkowski $\mathcal{M}^4$, un insieme di 16 grandezze dipendenti dalle coordinate spazio- temporali $(x^0, x^1, x^2, x^3) \equiv (x^0, x^\alpha)$:

$$(T_{ik}) = \begin{pmatrix} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & T_{11} & T_{12} & T_{13} \\ T_{20} & T_{21} & T_{22} & T_{23} \\ T_{30} & T_{31} & T_{32} & T_{33} \end{pmatrix} \tag{7.49}$$

i cui valori, nel passaggio da $\mathcal{R}$ ad $\mathcal{R}'$, si trasformano secondo la formula:

$$T'_{ik} = \frac{\partial x^j}{\partial x'^i} \frac{\partial x^h}{\partial x'^k} T_{jh}, \tag{7.50}$$

essendo sottintesa una doppia somma rispetto agli indici muti ripetuti j, h. Inoltre le derivate parziali $L^j_{i'} = \dfrac{\partial x^j}{\partial x'^i}$ devono essere calcolate sulla trasformazione inversa della (7.47), cioè:

$$\begin{cases} x^0 = \gamma(v)(x'^0 + \beta x'^1) \\ x^1 = \gamma(v)(x'^1 + \beta x'^0) \\ x^2 = x'^2 \\ x^3 = x'^3, \end{cases} \tag{7.51}$$

e perciò coincidono con i coefficienti stessi della trasformazione dati dagli elementi della matrice inversa della (7.48), cioè della matrice:

$$(L^j_{i'}) = (L^{i'}_j)^{-1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.52}$$

Si può passare facilmente dalla rappresentazione di un 4-tensore controvariante a quella covariante e viceversa utilizzando in maniera opportuna, il tensore metrico dello spazio di Minkowski nella sua rappresentazione covariante η_{ik} o controvariante η^{ik}, effettuando un'importante operazione chiamata *abbassamento o innalzamento degli indici*. Come abbiamo già visto nel paragrafo 5.6.2 del Capitolo 5, per i 4-vettori sussistono le formule:

$$V_i = \eta_{ik} V^k, \quad V_0 = V^0, \quad V_\alpha = -V^\alpha, \quad (\alpha = 1, 2, 3).$$

Per i 4-tensori doppi si procede nella maniera seguente [10]:

$$T_{ik} = \eta_{ij}\eta_{kh} T^{jh},$$

o viceversa

$$T^{ik} = \eta^{ij}\eta^{kh} T_{jh},$$

dove, in entrambe le formule, è sottintesa una doppia sommatoria rispetto agli indici j e h. Utilizzando queste formule è facile dimostrare (esercizio) che:

$$T_{00} = T^{00}, \quad T_{\alpha\,0} = -T^{\alpha\,0}, \quad T_{0\alpha} = -T^{0\alpha}, \quad T_{\alpha\beta} = T^{\alpha\,\beta} \quad (\alpha, \beta = 1, 2, 3). \tag{7.53}$$

Di notevoli proprietà godono due classi particolari di 4-tensori del 2° ordine nello spazio $\mathcal{M}^4$ di Minkowski: quelli *simmetrici* e quelli *antisimmetrici o emisimmetrici*. Un 4-tensore controvariante

[10] L'operazione di abbassamento ed innalzamento degli indici, per tensori qualsiasi, verrà studiata nel paragrafo 8.19 del Cap.8 sul Calcolo Tensoriale, si rimanda perciò il lettore a tale capitolo per maggiori dettagli.

T^{ik} del 2° ordine si dice **simmetrico** se, qualunque siano gli indici $i, k = 0, 1, 2, 3$ si ha $T^{ik} = T^{ki}$; si dice invece **antisimmetrico o emisimmetrico** se, qualunque siano gli indici $i, k = 0, 1, 2, 3$ si ha $T^{ik} = -T^{ki}$. Analoghe definizioni valgono per i 4-tensori covarianti.
Conseguentemente tutte le componenti di un tensore antisimmetrico, controvariante (o covariante), aventi due indici uguali sono nulle $T^{ii} = 0$ (oppure $T_{ii} = 0$) per qualunque valore dell'indice $i = 0, 1, 2, 3$, esse nella matrice rappresentativa occupano la diagonale principale; inoltre, delle restanti 12 componenti, 6 sono uguali in valore assoluto e di segno opposto alle altre 6, precisamente quelle che sono simmetriche rispetto alla diagonale principale, perciò complessivamente si hanno soltanto 6 componenti indipendenti che possono essere identificate come gli elementi che si trovano al disopra della diagonale principale (ovvero quelli che si trovano al disotto della stessa diagonale):

$$T^{01} \; T^{02} \; T^{03}$$

$$T^{12} \; T^{13} \; T^{23}$$

e la matrice rappresentativa delle componenti del tensore si presenterà nella forma:

$$(T^{ik}) = \begin{pmatrix} 0 & T^{01} & T^{02} & T^{03} \\ -T^{01} & 0 & T^{12} & T^{13} \\ -T^{02} & -T^{12} & 0 & T^{23} \\ -T^{03} & -T^{13} & -T^{23} & 0 \end{pmatrix}. \tag{7.54}$$

Per questa ragione i 4-tensori emisimmetrici del 2° ordine sono detti, talvolta, **esavettori**.
Analoghe considerazioni valgono per le componenti di un 4-tensore emisimmetrico covariante.

7.6 Tensore del campo elettromagnetico

Tenuto conto di ciò che abbiamo detto nel paragrafo precedente a proposito dei 4-tensori del 2° ordine nello spazio di Minkowski, siamo ora in grado di scrivere, in maniera esplicita, le formule di trasformazione (7.46) per le componenti di un 4-tensore controvariante. A tal fine, si tenga presente che le derivate che compaiono in tali formule non sono altro che gli elementi della matrice (7.48) che qui riscriviamo in modo dettagliato:

$$(L^{i'}_{j}) = \begin{pmatrix} \dfrac{\partial x'^0}{\partial x^0} = \gamma & \dfrac{\partial x'^0}{\partial x^1} = -\beta\gamma & \dfrac{\partial x'^0}{\partial x^2} = 0 & \dfrac{\partial x'^0}{\partial x^3} = 0 \\ \dfrac{\partial x'^1}{\partial x^0} = -\beta\gamma & \dfrac{\partial x'^1}{\partial x^1} = \gamma & \dfrac{\partial x'^1}{\partial x^2} = 0 & \dfrac{\partial x'^1}{\partial x^3} = 0 \\ \dfrac{\partial x'^2}{\partial x^0} = 0 & \dfrac{\partial x'^2}{\partial x^1} = 0 & \dfrac{\partial x'^2}{\partial x^2} = 1 & \dfrac{\partial x'^2}{\partial x^3} = 0 \\ \dfrac{\partial x'^3}{\partial x^0} = 0 & \dfrac{\partial x'^3}{\partial x^1} = 0 & \dfrac{\partial x'^3}{\partial x^2} = 0 & \dfrac{\partial x'^3}{\partial x^3} = 1 \end{pmatrix}. \tag{7.55}$$

I calcoli da farsi sono di facile esecuzione ma un tantino noiosi, comunque a conti fatti, tutte le 16 componenti di un 4-tensore doppio generico controvariante T^{ik} si trasformano nel modo seguente:

$$\begin{cases} T'^{00} = \gamma^2[T^{00} - \beta(T^{01} + T^{10}) + \beta^2 T^{11}], \\ T'^{01} = \gamma^2[T^{01} - \beta(T^{00} + T^{11}) + \beta^2 T^{10}], \\ T'^{02} = \gamma(T^{02} - \beta T^{12}), \\ T'^{03} = \gamma(T^{03} - \beta T^{13}), \end{cases} \tag{7.56}$$

$$\begin{cases} T'^{10} = \gamma^2[T^{10} - \beta(T^{00} + T^{11}) + \beta^2 T^{01}], \\ T'^{11} = \gamma^2[T^{11} - \beta(T^{01} + T^{10}) + \beta^2 T^{00}], \\ T'^{12} = \gamma(T^{12} - \beta T^{02}), \\ T'^{13} = \gamma(T^{13} - \beta T^{03}), \end{cases} \tag{7.57}$$

$$\begin{cases} T'^{20} = \gamma(T^{20} - \beta T^{21}), \\ T'^{21} = \gamma(T^{21} - \beta T^{20}), \\ T'^{22} = T^{22}, \\ T'^{23} = T^{23}, \end{cases} \tag{7.58}$$

$$\begin{cases} T'^{30} = \gamma(T^{30} - \beta T^{31}), \\ T'^{31} = \gamma(T^{31} - \beta T^{30}), \\ T'^{32} = T^{32}, \\ T'^{33} = T^{33}. \end{cases} \tag{7.59}$$

Nel caso particolare in cui il dato tensore doppio sia **antisimmetrico**, nel qual caso lo indicheremo con F^{ik}, cioè $F^{ik} = -F^{ki}$, gli elementi della diagonale principale rappresentante il tensore sono tutti nulli ($F^{ii} = 0, i = 0, 1, 2, 3$), e le formule precedenti si semplificano notevolmente riducendosi alle:

$$\begin{cases} F'^{00} = F^{00} = 0, \\ F'^{01} = F^{01} = -F^{10}, \\ F'^{02} = \gamma(F^{02} - \beta F^{12}), \\ F'^{03} = \gamma(F^{03} - \beta F^{13}), \end{cases} \tag{7.60}$$

$$\begin{cases} F'^{10} = F^{10} = -F^{01}, \\ F'^{11} = F^{11} = 0, \\ F'^{12} = \gamma(F^{12} - \beta F^{02}), \\ F'^{13} = \gamma(F^{13} - \beta F^{03}), \end{cases} \tag{7.61}$$

$$\begin{cases} F'^{20} = \gamma(F^{20} - \beta F^{21}) = -F'^{02}, \\ F'^{21} = \gamma(F^{21} - \beta F^{20}) = -F'^{12}, \\ F'^{22} = F^{22} = 0, \\ F'^{23} = F^{23}, \end{cases} \tag{7.62}$$

$$\begin{cases} F'^{30} = \gamma(F^{30} - \beta F^{31}) = -F'^{03}, \\ F'^{31} = \gamma(F^{31} - \beta F^{30}) = -F'^{13}, \\ F'^{32} = F^{32} = -F^{23}, \\ F'^{33} = F^{33} = 0. \end{cases} \tag{7.63}$$

Nelle precedenti relazioni soltanto 6 componenti del tensore antissimetrico F^{ik} sono indipendenti, per esempio $F^{01}, F^{02}, F^{03}, F^{23}, F^{31}, F^{12}$. Perciò le formule essenziali di trasformazione delle

componenti del tensore antisimmetrico F^{ik} nel passaggio dal sistema $\mathcal{R}$ al sistema $\mathcal{R}'$ sono:

$$\begin{cases} F'^{01} = F^{01}, \\ F'^{02} = \gamma(F^{02} - \beta F^{12}), \\ F'^{03} = \gamma(F^{03} - \beta F^{13}), \\ \\ F'^{23} = F^{23}, \\ F'^{31} = \gamma(F^{31} - \beta F^{30}), \\ F'^{12} = \gamma(F^{12} - \beta F^{02}). \end{cases} \tag{7.64}$$

Se ora confrontiamo queste formule di trasformazione per le componenti indipendenti del tensore antisimmetrico F^{ik} con le formule di trasformazione per le 6 componenti $E_x, E_y, E_z, H_x, H_y, H_z$ del campo elettrico e del campo magnetico, che qui riscriviamo pe facilitarne il confronto:

$$\begin{cases} E'_x = E_x, \\ E'_y = \gamma(v)\Big(E_y - \beta H_z\Big) \\ E'_z = \gamma(v)\Big(E_z + \beta H_y\Big), \\ H'_x = H_x, \\ H'_y = \gamma(v)\Big(H_y + \beta E_z\Big) \\ H'_z = \gamma(v)\Big(H_z - \beta E_y\Big), \end{cases}$$

si vede immediatamente che tali formule sono le stesse, purché si assuma quali componenti indipendenti, per esempio:

$$F^{10} = E_x, \quad F^{20} = E_y, \quad F^{30} = E_z,$$

$$F^{32} = H_x, \quad F^{13} = H_y, \quad F^{21} = H_z.$$

Si conclude quindi che, *nell'ambito della Teoria della Relatività Ristretta, l'insieme delle componenti del campo elettrico e del campo magnetico costituiscono un 4-tensore antissimmetrico del 2° ordine le cui componenti controvarianti sono gli elementi della matrice*:

$$(F^{ik}) = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -H_z & H_y \\ E_y & H_z & 0 & -H_x \\ E_z & -H_y & H_x & 0 \end{pmatrix}. \tag{7.65}$$

Il tensore antissimetrico F^{ik} è detto **tensore del campo elettromagnetico**: *esso fonde in un unico ente assoluto il campo elettrico* **E** *e il campo magnetico* **H** *che separatamente hanno invece carattere relativo.* Il tensore campo elettromagnetico può essere espresso anche per mezzo delle sue componenti covarianti F_{ik} che si possono ricavare dalle componenti controvarianti abbassando gli indici col tensore metrico, cioè $F_{ik} = \eta_{ij}\eta_{kh}F^{jh}$, ovvero utilizzando le formule (7.53), avremo perciò:[11]

$$(F_{ik}) = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -H_z & H_y \\ -E_y & H_z & 0 & -H_x \\ -E_z & -H_y & H_x & 0 \end{pmatrix}. \tag{7.66}$$

[11] Avvertiamo il lettore che non tutti gli autori scrivono le due matrici (7.65), (7.66), che rappresentano il campo elettromagnetico, esattamente nella stessa forma utilizzata da noi (seguendo Landau-Lifschitz, *Teoria dei campi*, Editori Riuniti, 1976, pag.89). Ciò dipende dalla scelta *arbitraria* che si fa delle 6 componenti indipendenti fra le 12, diverse da zero, che compongono le matrici.

7.7 Forma tensoriale delle equazioni di Maxwell

Utilizzando, in maniera opportuna, il tensore campo elettromagnetico, sia nella sua forma controvariante F^{ik} sia nella sua forma covariante F_{ik} siamo ora in grado di scrivere in forma tensoriale le equazioni di Maxwell. Lo faremo, dapprima, per uno spazio vuoto senza sorgenti, in cui tali equazioni, in forma vettoriale ordinaria, sono le (7.10), (7.11). In un secondo tempo ci occuperemo anche della formulazione tensoriale delle equazioni di Maxwell nel caso che ci si trovi in presenza di sorgenti.
Si può constatare facilmente che la 1ª coppia delle equazioni di Maxwell, scritte in forma scalare, cioè (7.12), è espressa dall'uguaglianza tensoriale:

$$\partial_i F_{kl} + \partial_k F_{li} + \partial_l F_{ik} = 0, \qquad (i,k,l = 0,1,2,3), \tag{7.67}$$

dove i simboli di derivata parziale, come ∂_i, significano derivazione rispetto alla corrispondente coordinata spaziotemporale x^i, cioè $\partial_i \equiv \dfrac{\partial}{\partial x^i}$ e dove gli indici i,k,l, nei tre termini della sommatoria, si susseguono in ordine ciclico.
Infatti, se scegliamo $(i,k,l) \equiv (1,2,3)$, l'equazione (7.67) diventa:

$$\partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12} = 0,$$

o più esplicitamente:

$$\frac{\partial F_{23}}{\partial x^1} + \frac{\partial F_{31}}{\partial x^2} + \frac{\partial F_{12}}{\partial x^3} = 0. \tag{7.68}$$

Ricordando che $x^1 = x, x^2 = y, x^3 = z$ e tenendo conto dell'espressione (7.66) del tensore campo elettromagnetico sotto forma covariante, si constata che $F_{23} = -H_x, F_{31} = -H_y, F_{12} = -H_z$, e perciò la (7.68) equivale alla prima delle equazioni (7.12), cioè alla:

$$\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} + \frac{\partial H_z}{\partial z} = 0.$$

che esprime il fatto che la divergenza del vettore campo magnetico **H** è nulla.
In maniera analoga si procede nel caso che gli indici (i,k,l) che figurano nella (7.67) siano $(i,k,l) \equiv (2,3,0)$, $(i,k,l) \equiv (0,1,3)$, $(i,k,l) \equiv (0,1,2)$, con $x^0 = ct$; si otterranno rispettivamente la 2^a, la 3^a e la 4ª delle (7.12), cioè l'equazione tensoriale (7.67) è perfettamente equivalente alla 1ª coppia delle equazioni di Maxwell, c.d.d.
La (7.67) si può scrivere in una forma più compatta se si introduce il cosiddetto **indicatore o simbolo di Levi-Civita** in forma controvariante ϵ^{ijhk} (densità tensoriale di peso $W = +1$) così definito:[12] *le sue componenti sono nulle se almeno due degli indici ijhk sono uguali, le altre valgono* $+1$ *o* -1 *a seconda che ijhk sia rispettivamente una permutazione di classe pari o dispari della permutazione fondamentale* 0123.
Per esempio:

$$\epsilon^{0023} = 0, \quad \epsilon^{0113} = 0 \quad \epsilon^{0123} = +1, \quad \epsilon^{1032} = +1, \quad \epsilon^{0321} = -1, \quad \epsilon^{1230} = -1.$$

Il numero delle componenti non nulle di ϵ^{ijhk} è ovviamente uguale al numero di permutazioni di 4 oggetti distinti, cioè $4! = 24$. Allo stesso modo si può definire *l'indicatore o simbolo di Levi-Civita* in forma covariante ϵ_{ijhk} (densità di peso $W = -1$), così:

$$\epsilon_{ijhk} \stackrel{\text{def}}{=} -\epsilon^{ijhk}. \tag{7.69}$$

[12] Per la definizione di **densità tensoriale** rimandiamo al Cap.11, paragrafo 11.15.4 di questo libro e per uno studio più approfondito **dell'indicatore di Levi-Civita** si può consultare con profitto Vincenzo Barone, *Relatività*, Bollati Boringhieri, Torino, 2004, pagg. 186-188; oppure L. D. Landau e E. M. Lifchitz, *Teoria dei Campi*, Editori Riuniti, Edizione del 2010, pagg. 34-36; o ancora Moshe Carmeli, *Classical Fields: General Relativity and Gauge Theory*, John Wiley and Sons, New York, 1982, pagg. 37-38.

Con il simbolo di Levi-Civita, ora introdotto, è facile constatare che la 1^a coppia delle equazioni di Maxwell, cioè l'equazione tensoriale (7.67) si scrive nella forma compatta:

$$\epsilon^{ijhk}\partial_j F_{hk} = 0. \tag{7.70}$$

Passiamo ora a considerare la 2^a coppia delle equazioni di Maxwell, cioè le (7.13). Anche per queste ultime è facile verificare che possono essere rappresentate in forma tensoriale con l'equazione:

$$\partial_k F^{ik} = 0, \qquad (i, k = 0, 1, 2, 3), \tag{7.71}$$

dove l'indice k, ripetuto in alto e in basso, significa sommatoria rispetto a quell'indice, perciò tale equazione equivale a:

$$\partial_0 F^{i0} + \partial_1 F^{i1} + \partial_2 F^{i2} + \partial_3 F^{i3} = 0, \qquad (i = 0, 1, 2, 3), \tag{7.72}$$

cioè alle quattro equazioni:

$$\begin{cases} \partial_0 F^{00} + \partial_1 F^{01} + \partial_2 F^{02} + \partial_3 F^{03} = 0, \\ \partial_0 F^{10} + \partial_1 F^{11} + \partial_2 F^{12} + \partial_3 F^{13} = 0, \\ \partial_0 F^{20} + \partial_1 F^{21} + \partial_2 F^{22} + \partial_3 F^{23} = 0, \\ \partial_0 F^{30} + \partial_1 F^{31} + \partial_2 F^{32} + \partial_3 F^{33} = 0. \end{cases} \tag{7.73}$$

Con ragionamenti del tutto analoghi a quelli fatti in precedenza, si verifica facilmente che queste equazioni rappresentano, in forma cartesiana, la 2^a coppia delle equazioni di Maxwell (7.13).
Concludiamo questo paragrafo occupandoci, più in generale, della formulazione tensoriale delle equazioni di Maxwell nel caso in cui il campo elettromagnetico (nel vuoto) sia determinato da una distribuzione di sorgenti, costituite da cariche elettriche di densità $\rho(\mathbf{x}, t)$ e da correnti di densità $\mathbf{j}(\mathbf{x}, t)$. Nel caso in esame, nel sistema di unità di misura di Gauss, la 1^a coppia delle equazioni di Maxwell consiste nelle due equazioni:

$$\begin{cases} \operatorname{div} \mathbf{H} = 0, \\ \\ \operatorname{rot} \mathbf{E} + \dfrac{1}{c}\dfrac{\partial \mathbf{H}}{\partial t} = 0, \end{cases} \tag{7.74}$$

mentre la 2^a coppia è data dalle:

$$\begin{cases} \operatorname{div} \mathbf{E} = 4\pi\rho, \\ \\ \operatorname{rot} \mathbf{H} - \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{4\pi}{c}\mathbf{j}. \end{cases} \tag{7.75}$$

Particolarmente interessante per le applicazioni è il caso in cui le correnti che determinano il campo sono dovute al moto di cariche libere che si muovono con velocità $\mathbf{u}$, per cui si ha:

$$\mathbf{j} = \rho\mathbf{u}. \tag{7.76}$$

Con ciò le equazioni (7.74), (7.75) diventano le note equazioni di Maxwell-Lorentz che stanno alla base della teoria degli elettroni.
Si tratta ora di determinare le leggi con cui si trasformano le varie grandezze che compaiono in queste equazioni, allo scopo di assicurarsi della loro covarianza relativistica. Le considerazioni fatte per il caso di un campo senza sorgenti ci hanno mostrato che i vettori ordinari campo elettrico $\mathbf{E}$ e campo magnetico $\mathbf{H}$ si trasformano come un complesso unico avente le proprietà

di un 4- tensore doppio antisimmetrico F^{ik}. Si tratta ora di trovare le leggi di varianza della densità di corrente $\mathbf{j}$ e della densità di carica ρ. Dette grandezze compaiono nelle equazioni di Maxwell-Lorentz ai secondi membri:

$$\begin{cases} \operatorname{div} \mathbf{E} = 4\pi\rho, \\ \\ \operatorname{rot} \mathbf{H} - \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{4\pi}{c}\rho\mathbf{u}. \end{cases} \tag{7.77}$$

Ricordando che, come abbiamo visto all'inizio del paragrafo, i primi membri di queste equazioni, dal punto di vista tensoriale non sono altro che la divergenza del tensore F^{ik}, vale a dire $\partial_k F^{ik}$, che si trasforma come un 4-vettore, ne consegue che $\mathbf{j}$ e ρ dovranno pure costituire un unico ente avente le proprietà di varianza di un 4-vettore. Ciò può essere verificato con le considerazioni dirette che ora faremo.
L'equazione (7.76) si può ottenere con le semplici considerazioni che seguono.
Ricordiamoci che in RR, la massa di un corpo in moto con velocità $\mathbf{u}$ è maggiore della sua massa a riposo di un fattore $\gamma(u)$ (essendo $\gamma(u) > 1$) (almeno nell'interpretazione da noi accettata), cioè:

$$m = m_0\gamma(u).$$

Si verifica facilmente che un'analoga relazione vale per la densità di carica ρ. Infatti, se consideriamo un elemento di volume cubico 3-dimensionale che a riposo, in un dato sistema di riferimento $\mathcal{R}_0$, abbia il lato uguale ad L_0 e perciò di volume $V_0 = L_0^3$, quando esso è in moto (nel riferimento $\mathcal{R}$) nella direzione di un suo spigolo, con velocità $\mathbf{u}$, questo subisce la contrazione di Lorentz, assumendo la lunghezza $\dfrac{L_0}{\gamma(u)}$ (rispetto all'osservatore fisso in $\mathcal{R}_0$), mentre gli altri due lati, muovendosi ortogonalmente alla direzione di moto, non subiscono alcuna contrazione. Perciò il volume del cubo in moto, dal punto di vista del riferimento $\mathcal{R}$, sarà

$$V = \frac{L_0}{\gamma(u)} \cdot L_0 \cdot L_0 = \frac{L_0^3}{\gamma(u)} = \frac{V_0}{\gamma(u)}. \tag{7.78}$$

Per un'osservatore *co-mobile* col flusso di cariche, cioè mobile assieme ad esse, *la densità propria di carica* è $\rho_0 = \dfrac{q_0}{V_0}$, essendo q_0 la carica contenuta in V_0; mentre *la densità di carica* misurata da un'osservatore fisso, tenuto conto *dell'invarianza della carica* $(q_0 = q)$ nel passaggio da un riferimento inerziale ad un altro, sarà:

$$\rho = \frac{q}{V} = \frac{q_0}{V_0}\gamma(u),$$

cioè:

$$\rho = \rho_0\gamma(u) = \frac{\rho_0}{\sqrt{1-\beta_u^2}}. \tag{7.79}$$

Se quindi un flusso di cariche di densità ρ_0, solidali col sistema $\mathcal{R}_0$, scorre con velocità $\mathbf{u}$ di fronte ad un osservatore fisso in $\mathcal{R}$, quest'ultimo misurerà una densità di carica alterata di un fattore $\gamma(u)$ come dalla (7.79).
Utilizzando le considerazioni precedenti è facile introdurre in maniera naturale un 4-vettore intimamente legato alla densità di corrente ρ ed al suo modo di trasformarsi compendiato nella (7.79).
Infatti, siano x^i, $(i = 0, 1, 2, 3)$ le coordinate di una carica in moto in $\mathcal{R}$: moltiplicando la quadri-velocità $dx^i/d\tau$ per lo scalare ρ_0, si costruisce un nuovo quadri-vettore, detto **quadri-vettore densità di corrente** direttamente proporzionale alla 4-velocità della carica:

$$J^i = \rho_0\frac{dx^i}{d\tau} = \rho_0 U^i. \tag{7.80}$$

Quindi, per le (7.79), (7.80) e ricordando la (5.56), abbiamo:

$$J^i = \rho_0 \gamma(u)(c, \mathbf{u}) = (\rho c, \rho \mathbf{u}) = (\rho c, \mathbf{j}), \tag{7.81}$$

dove

$$\mathbf{j} = \rho \mathbf{u},$$

denota l'usuale tri-vettore **densità di corrente**, cioè la (7.76).
Osserviamo che utilizzando il 4-vettore J^i appena introdotto, l'equazione di conservazione della carica, cioè l'equazione di continuità (7.1), si può scrivere in forma tensoriale come:

$$\partial_i J^i = 0.$$

Ovviamente le equazioni di Maxwell-Lorentz omogenee (7.74) (1ª coppia delle equazioni di Maxwell) per un campo generato da cariche in moto, si scriveranno in forma tensoriale nello stesso modo di prima:

$$\partial_i F_{kl} + \partial_k F_{li} + \partial_l F_{ik} = 0, \qquad (i, k, l = 0, 1, 2, 3), \tag{7.82}$$

ovvero nella forma compatta

$$\epsilon^{ijhk} \partial_j F_{hk} = 0. \tag{7.83}$$

da cui è facile constatarne la coincidenza con le equazioni scritte in forma vettoriale ordinaria, cioè con le (7.74).
Per quanto concerne le equazioni non omogenee (7.75) (2^a coppia delle equazioni di Maxwell) esse sono date dall'equazione tensoriale:

$$\partial_i F^{ik} = \frac{4\pi}{c} J^k, \qquad (i, k = 0, 1, 2, 3). \tag{7.84}$$

Infatti è facile verificare che, per $k = 0$, si riottiene *la legge di Gauss* in quanto il primo membro della (7.84) ponendovi $k = 0$ diventa:

$$\partial_\alpha F^{\alpha 0} = \partial_\alpha E^\alpha = \mathrm{div}\mathbf{E}, \qquad (\alpha = 1, 2, 3), \tag{7.85}$$

avendo indicato con (E^α) $(\alpha = 1, 2, 3)$ la terna delle componenti cartesiane (E_x, E_y, E_z) del campo elettrico. Il secondo membro della (7.84) è $(4\pi/c)J^0 = 4\pi\rho$, cosiché si ricava effettivamente

$$\mathrm{div}\mathbf{E} = 4\pi\rho.$$

Per $k = \alpha = 1, 2, 3$, al 1° membro della (7.84) si trova:

$$\partial_i F^{i\alpha} = \partial_0 F^{0\alpha} + \partial_\beta F^{\beta\alpha} = -\partial_0 E^\alpha + \partial_1 F^{1\alpha} + \partial_2 F^{2\alpha} + \partial_3 F^{3\alpha} = -\frac{1}{c}\frac{\partial E^\alpha}{\partial t} + (\mathrm{rot}\mathbf{H})^\alpha, \tag{7.86}$$

e quindi

$$-\frac{1}{c}\frac{\partial E^\alpha}{\partial t} + (\mathrm{rot}\mathbf{H})^\alpha = \frac{4\pi}{c} J^\alpha,$$

cioè

$$\mathrm{rot}\mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j}, \tag{7.87}$$

in cui si riconosce *la legge di Ampère-Maxwell.*

7.8 Formulazione potenziale delle equazioni di Maxwell

A conclusione di questo capitolo vogliamo accennare brevemente a due concetti fondamentali. Spesso, piuttosto che lavorare con i campi **E** ed **H**, è più conveniente lavorare con i cosiddetti **potenziali elettromagnetici**: Il **potenziale scalare** φ e il **potenziale vettore** **A**, l'uso dei quali permette di scrivere le equazioni di Maxwell in maniera notevolmente più semplice.
A tal fine riprendiamo in considerazione la 1^a coppia delle equazioni di Maxwell, cioè delle due equazioni omogenee:

$$\begin{cases} \operatorname{div} \mathbf{H} = 0, \\ \\ \operatorname{rot} \mathbf{E} + \dfrac{1}{c}\dfrac{\partial \mathbf{H}}{\partial t} = 0, \end{cases} \tag{7.88}$$

e la 2^a coppia di tali equazioni, cioè quelle non omogenee:

$$\begin{cases} \operatorname{div} \mathbf{E} = 4\pi\rho, \\ \\ \operatorname{rot} \mathbf{H} - \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{4\pi}{c}\mathbf{j}. \end{cases} \tag{7.89}$$

La possibilità di introdurre i potenziali elettromagnetici, è fondata sull'uso di due noti teoremi di calcolo vettoriale. Se $\varphi(\mathbf{x})$ è un generico campo scalare e $\mathbf{\Phi}(\mathbf{x})$ un generico campo vettoriale, posto:

$$\mathbf{F} = \operatorname{grad} \varphi, \tag{7.90}$$

$$\mathbf{G} = \operatorname{rot} \mathbf{\Phi}, \tag{7.91}$$

si hanno le identità:

$$\operatorname{rot} \mathbf{F} = 0, \tag{7.92}$$

$$\operatorname{div} \mathbf{G} = 0. \tag{7.93}$$

Inversamente si dimostra che:

1) Se **F** è un campo vettoriale soddisfacente la (7.92) (in una regione semplicemente connessa), esiste un campo scalare φ per cui è soddisfatta la (7.90);

2) Se **G** è un campo vettoriale soddisfacente la (7.93), esiste un secondo campo vettoriale $\mathbf{\Phi}$ per cui è soddisfatta la (7.91).

Per la $(7.88)_1$, in virtù del secondo teorema esiste un vettore **A** tale che

$$\mathbf{H} = \operatorname{rot} \mathbf{A}. \tag{7.94}$$

Sostituendo la (7.94) nella $(7.88)_2$ si trova allora:

$$\operatorname{rot}\left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) = 0, \tag{7.95}$$

ed in virtù del primo teorema deve esistere uno scalare φ tale che

$$\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = -\operatorname{grad} \varphi. \tag{7.96}$$

I campi **E** e **H** si possono dunque sempre esprimere nella forma:

$$\begin{cases} \mathbf{E} = -\text{grad}\,\varphi - \dfrac{1}{c}\dfrac{\partial \mathbf{A}}{\partial t}, \\ \mathbf{H} = \text{rot}\,\mathbf{A}. \end{cases} \tag{7.97}$$

Le grandezze φ e **A** prendono rispettivamente il nome di **potenziale scalare** e di **potenziale vettore**.
Bisogna osservare che che per assegnati valori del campo elettrico **E** e del campo magnetico **H** i potenziali φ e **A** non sono univocamente determinati dalle (7.97). Infatti se indichiamo con φ, **A** e con φ', $\mathbf{A}'$ due distinte soluzioni delle (7.97) per dati **E** e **H**, dalla $(7.88)_2$ abbiamo

$$\text{rot}\,(\mathbf{A}' - \mathbf{A}) = 0;$$

deve quindi esistere uno scalare Λ tale che

$$\mathbf{A}' - \mathbf{A} = \text{grad}\,\Lambda. \tag{7.98}$$

Risolvendo la (7.98) rispetto ad **A** e sostituendo nella $(7.97)_1$, possiamo allora scrivere:

$$\mathbf{E} = -\text{grad}\left(\varphi - \frac{1}{c}\frac{\partial \Lambda}{\partial t}\right) - \frac{1}{c}\frac{\partial \mathbf{A}'}{\partial t}. \tag{7.99}$$

A meno di una inessenziale costante additiva, φ' deve perciò coincidere con $\varphi - \dfrac{1}{c}\dfrac{\partial \Lambda}{\partial t}$ e si ha che le due possibili scelte dei potenziali devono essere legate da una relazione del tipo

$$\begin{cases} \varphi' = \varphi - \dfrac{1}{c}\dfrac{\partial \Lambda}{\partial t}, \\ \mathbf{A}' = \mathbf{A} + \text{grad}\,\Lambda. \end{cases} \tag{7.100}$$

Inversamente se φ, **A** soddisfano le (7.97) anche le grandezze φ', $\mathbf{A}'$ definite dalla (7.100) le soddisfano qualunque sia Λ per gli stessi **E**, **H**.
La trasformazione (7.100) prende il nome di **trasformazione di gauge** (lett. *trasformazione di calibrazione*) ed a Λ ci si riferisce come alla **funzione di gauge**.
Osserviamo che per φ, **A** arbitrari le (7.97) forniscono in sostanza la più generale soluzione della 1^a coppia delle equazioni di Maxwell (7.88). È necessario successivamente determinare φ e **A** in modo che sia soddisfatta anche la 2^a coppia delle equazioni di Maxwell (7.89).
Se ora sostituiamo le (7.97) nella 2^a coppia delle equazioni di Maxwell (7.75) e teniamo conto delle identità vettoriali:

$$\text{div}\,\text{grad}\varphi \equiv \Delta_2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2},$$

$$\text{rot}\,\text{rot}\mathbf{A} = \text{grad}\,\text{div}\mathbf{A} - \Delta_2\mathbf{A},$$

intendendo in quest'ultima che sia $\Delta_2\mathbf{A} = (\Delta_2 A_x, \Delta_2 A_y, \Delta_2 A_z)$, avremo:

$$\begin{cases} \Delta_2\varphi - \dfrac{1}{c^2}\dfrac{\partial^2\varphi}{\partial t^2} + \dfrac{1}{c}\dfrac{\partial}{\partial t}\left(\text{div}\mathbf{A} + \dfrac{1}{c}\dfrac{\partial\varphi}{\partial t}\right) = -4\pi\rho \\ \Delta_2\mathbf{A} - \dfrac{1}{c^2}\dfrac{\partial^2\mathbf{A}}{\partial t^2} - \text{grad}\left(\text{div}\mathbf{A} + \dfrac{1}{c}\dfrac{\partial\varphi}{\partial t}\right) = -\dfrac{4\pi}{c}\mathbf{j}. \end{cases} \tag{7.101}$$

Le (7.101) possono essere semplificate se si usa l'arbitrarietà nella scelta di φ e **A**, infatti basta fare una tale scelta in modo tale che sia verificata la condizione

$$\text{div}\mathbf{A} + \frac{1}{c}\frac{\partial\varphi}{\partial t} = 0. \qquad \text{(gauge di Lorentz)} \tag{7.102}$$

Ammessa la (7.102), le (7.101) diventano *equazioni di propagazione ondosa per i potenziali*:

$$\begin{cases} \Delta_2 \varphi - \dfrac{1}{c^2}\dfrac{\partial^2 \varphi}{\partial t^2} = -4\pi\rho \\ \\ \Delta_2 \mathbf{A} - \dfrac{1}{c^2}\dfrac{\partial^2 \mathbf{A}}{\partial t^2} = -\dfrac{4\pi}{c}\mathbf{j}, \end{cases} \tag{7.103}$$

ovvero in maniera più compatta:

$$\begin{cases} \Box\varphi = -4\pi\rho \\ \Box\mathbf{A} = -\dfrac{4\pi}{c}\mathbf{j}, \end{cases} \tag{7.104}$$

utilizzando il cosiddetto **operatore di d'Alembert o d'Alembertiano**.
Il problema della risoluzione delle equazioni di Maxwell (7.88), (7.89) è così ricondotto a quello della ricerca delle soluzioni delle equazioni ondulatorie (7.104), per i potenziali, che soddisfano la (7.102). Il pregio principale delle (7.104) rispetto alle (7.101) è quello di essere *completamente disaccoppiate*, cioè abbiamo un'equazione distinta per ciascuna delle quattro grandezze indipendenti φ, A_x, A_y, A_z. Alla restrizione introdotta nella scelta di φ e $\mathbf{A}$ dalla (7.102) si fa riferimento dicendo che questa definisce il **gauge di Lorentz**.
A questo punto possiamo introdurre il concetto di **4-potenziale** e riattaccarci alla trattazione tensoriale data nei paragrafi precedenti. Diciamo **quadripotenziale** o semplicemente **4-potenziale** il 4-vettore le cui componenti controvarianti sono:

$$A^i = (A^0, A^\alpha) \equiv (\varphi, A^\alpha), \quad (i = 0,1,2,3; \alpha = 1,2,3). \tag{7.105}$$

La prima componente (la componente temporale) coincide col potenziale scalare φ, e le componenti spaziali $A^\alpha \equiv \mathbf{A}$ coincidono con le componenti ordinarie del potenziale vettore. Le corrispondenti componenti covarianti, definite dalla $A_i = \eta_{ik}A^k$ sono:

$$A_i = (A_0, A_\alpha) = (A_0, -A^\alpha) = (\varphi, -A^\alpha). \tag{7.106}$$

Si trova allora che le (7.97) equivalgono ad esprimere le componenti del tensore campo elettromagnetico in funzione delle componenti del 4-potenziale (esercizio):

$$F_{ik} = \frac{\partial A_k}{\partial x^i} - \frac{\partial A_i}{\partial x^k} = \partial_i A_k - \partial_k A_i. \tag{7.107}$$

Abbiamo osservato in ciò che precede, che il potenziale scalare ed il potenziale vettore non sono definiti in modo univoco ma a meno di una trasformazione di gauge, analogamente il 4-potenziale è definito a meno di una trasformazione di gauge:

$$A'_i = A_i + \frac{\partial \psi}{\partial x^i} = A_i + \partial_i \psi, \tag{7.108}$$

dove ψ è un campo scalare arbitrario. Anche se una trasformazione di gauge modifica i potenziali, lascia invariato il tensore elettromagnetico F_{ik}, e quindi $\mathbf{E}$ e $\mathbf{H}$ che sono le quantità misurabili. Per semplificare la discussione sulla radiazione elettromagnetica di solito si utilizza il *gauge di Lorentz* (7.102), facilmente esprimibile in termini del 4-potenziale:

$$\eta^{ik}\partial_k A_i = 0. \qquad \text{(gauge di Lorentz)} \tag{7.109}$$

Applicando questo vincolo alla (7.108), troviamo che il campo scalare φ non è più arbitrario ma deve essere soluzione **dell'equazione d'onda**:

$$\Box\psi \equiv \eta^{ik}\frac{\partial^2 \psi}{\partial x^i \partial x^k} = \eta^{ik}\partial_{ik}\psi = 0, \tag{7.110}$$

dove $\Box$ è l'operatore di d'Alembert:

$$\Box\psi \equiv \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \Big(\frac{\partial^2\psi}{\partial(x^0)^2} + \frac{\partial^2\psi}{\partial(x^1)^2} + \frac{\partial^2\psi}{\partial(x^2)^2} + \frac{\partial^2\psi}{\partial(x^3)^2}\Big). \tag{7.111}$$

La definizione (7.106) rende la 1^a coppia delle equazioni tensoriali di Maxwell (7.67) automaticamente soddisfatte, cioè esse diventano delle identità (esercizio). La 2^a coppia delle equazioni tensoriali di Maxwell (7.84) diventano, in funzione del 4-potenziale:

$$\partial_k[\eta^{il}\eta^{km}(\partial_m A_l - \partial_l A_m)] = \frac{4\pi}{c}J^i. \tag{7.112}$$

Nel gauge di Lorentz (7.109) questa si riduce a:

$$\Box A^i = \frac{4\pi}{c}J^i. \tag{7.113}$$

In regioni prive di sorgenti essendo $J^i = 0$ la precedente equazione diventa:

$$\Box A^i = 0, \tag{7.114}$$

da cui segue che A^i e F^{ik}, e quindi **E** e **H**, soddisfano tutti l'equazione d'onda.
La nostra trattazione dell'elettrodinamica termina qui ritenendo di aver detto sull'argomento le cose più importanti e le più semplici per un principiante.

Capitolo 8

ALGEBRA TENSORIALE

I mezzi matematici necessari per la Teoria della Relatività Generale erano già pronti nel "Calcolo Differenziale Assoluto", il quale si basa sulle ricerche di Gauss, Riemann e Christoffel sulle varietà non euclidee, ed è stato eretto a sistema da Ricci e Levi-Civita e da essi applicato a problemi della Fisica teorica.
(A. Einstein, *I fondamenti della Teoria della Relatività Generale*, **Opere scelte**, a cura di Enrico Bellone, Bollati Boringhieri, 1988, pag. 282.)

8.1 Introduzione

L'algoritmo più espressivo e più agile per la impostazione e la discussione matematica di molti problemi della Meccanica e di altre teorie fisiche è fornito, com'è noto, dalla **Teoria dei vettori**. Per questo motivo nella meccanica classica si fa largo uso del calcolo vettoriale per esprimere in maniera sintetica alcune leggi fisiche in quanto un'equazione vettoriale compendia tre equazioni scalari ed anche perché il linguaggio dei vettori permette di risolvere più rapidamente alcuni problemi e ne aiuta notevolmente la comprensione.
Esattamente nello stesso modo, nella Teoria della Relatività sia Ristretta sia Generale - ma anche in altre importanti branche della Fisica matematica, come la Teoria dell'elasticità, la Meccanica dei continui, la Fluidodinamica, ecc. e nella Geometria Differenziale, in particolare nella Geometria Riemanniana - si ha bisogno del *linguaggio dei tensori, del* **Calcolo tensoriale**, *articolato in due capitoli fondamentali:* **l'Algebra Tensoriale e l'Analisi tensoriale**, *quest'ultima fondata sul concetto di* **derivata covariante**.
Svilupperemo gli elementi di base della Teoria Algebrica dei tensori, dell'Analisi tensoriale e della Geometria Riemanniana essenziali per comprendere la Relatività Generale.
Il concetto di **tensore** ha avuto la sua origine negli sviluppi della Geometria Differenziale per opera di Karl Friedrich Gauss, Georg Friedrich Bernhard Riemann ed Elwin Bruno Christoffel.
La nascita del **Calcolo Tensoriale**, originariamente chiamato **Calcolo differenziale assoluto**, come un ramo sistematico della Matematica è stata stimolata non solo da ricerche di natura geometrica ma anche da concreti problemi di Fisica matematica, inizialmente per opera di Gregorio Ricci-Curbastro, nel decennio 1887-1896, e successivamente sviluppato dal suo allievo Tullio Levi-Civita, all'epoca entrambi professori all'università di Padova. In collaborazione essi pubblicarono la prima memoria fondamentale su questo soggetto, intitolata: *Méthodes de calcul différentiel absolu et leurs applications* (Mathematische Annalen, vol.54, 1901).
Lo scopo principale del Calcolo Tensoriale è la ricerca di **relazioni invarianti** *quando si passa da un sistema di coordinate ad un altro.* Le leggi della fisica non possono dipendere dal sistema di riferimento che i fisici scelgono per scopo di descrizione. Perciò utilizzare il calcolo tensoriale come mezzo matematico per formulare le leggi della fisica è desiderabile da un punto di vista

estetico e spesso anche molto conveniente. In particolare Einstein lo trovò un mezzo eccellente per la formulazione della sua Teoria della Relatività Generale (RG).
Nella Relatività Speciale (RS) o Ristretta (RR) le trasformazioni di coordinate che si considerano sono esclusivamente lineari, mentre la RG considera trasformazioni del tutto generali.
L'impostazione che adotteremo è quella più seguita dalla stragrande maggioranza degli autori di testi sulla Relatività Generale.[1] Essa consiste soprattutto nel concentrarsi sulle tecniche di manipolazione dei tensori senza prendere in considerazione il significato geometrico più profondo che sta dietro la teoria. *Ci si interessa cioè più a quello che si può fare con i tensori piuttosto che al loro significato geometrico intrinseco*.
Noi condividiamo pienamente ciò che scrive Ray d'Inverno nel suo libro, come introduzione al capitolo di Algebra Tensoriale:
"Esistono due diversi approcci all'insegnamento dei tensori: l'approccio astratto, o senza indici (senza coordinate) e l'approccio tradizionale, basato sugli indici. Da alcuni anni si sta facendo strada in alcuni ambienti la tendenza ad introdurre i tensori fin dall'inizio con l'approccio astratto, più moderno che ha il principale vantaggio di offrire un'idea più chiara del significato geometrico (anche se alcuni hanno poi cambiato idea e sono ritornati all'approccio tradizionale). Soffre tuttavia di due svantaggi: prima di tutto richiede molta più preparazione matematica, che richiede tempo per essere sviluppata; l'altro svantaggio è che, malgrado la sua eleganza, quando poi si deve affrontare un calcolo reale con i tensori, come spesso accade, allora bisogna fare ricorso agli indici. Noi adotteremo l'approccio più tradizionale con gli indici, perché si rivelerà più veloce e più pratico. Consigliamo comunque a chi voglia approfondire ulteriormente lo studio dell'argomento di considerare l'approccio senza indici come la prima possibilità".

8.2 Varietà differenziabili e sistemi di coordinate

È difficile immaginare un problema di fisica che non involga qualche sorta di **spazio continuo** in cui si svolgono i fenomeni oggetto del nostro studio. Potrebbe essere *lo spazio fisico 3-dimensionale, lo spazio-tempo 4-dimensionale della Relatività Ristretta (spazio di Minkowski) o della Relatività Generale*, *lo spazio delle configurazioni* di un sistema olonomo in Meccanica Lagrangiana, *lo spazio delle fasi* per un problema di Meccanica Hamiltoniana Classica o Quantistica, *lo spazio di tutti gli stati termodinamici di equilibrio*, oppure qualche altro spazio più astratto.
Tutti questi spazi hanno differenti proprietà geometriche, ma tutti condividono qualche proprietà in comune, qualche cosa che ha a che fare col fatto che essi sono tutti spazi continui, piuttosto che reticoli di punti discreti.
La chiave dell'importanza della **Geometria Differenziale** nella fisica moderna è che essa studia precisamente quelle proprietà che sono comuni a tutti tali spazi. La più importante di queste proprietà riguarda la definizione di **varietà differenziabile**, che è l'oggetto matematicamente preciso che sostituisce la parola **"spazio".**
Lo studio approfondito delle varietà si fa nell'ambito dei corsi di Geometria, qui, per i nostri

[1] La nostra trattazione dell'Algebra Tensoriale è ispirata, almeno nell'approccio iniziale e nelle prime definizioni, all'impostazione del Cap. 5 del bel libro del relativista inglese dell'università di Southampton, Ray d'Inverno *Introducing Einstein's Relativity*, Oxford University Press, Oxford (1992), tradotto in italiano col titolo *Introduzione alla Relatività di Einstein*, CLUEB (2001).
Nel seguito, la nostra esposizione, è fortemente ispirata al classico e bellissimo libro di uno dei fondatori del Calcolo Tensoriale nonché uno dei più grandi matematici del secolo XX°, l'italiano: Tullio Levi-Civita, *Lezioni di Calcolo Differenziale Assoluto*, Alberto Stock Editore, Roma (1925).
Un altro libro classico, che abbiamo tenuto d'occhio, notevole per la chiarezza dell'esposizione e per la sua completezza, è: J. L. Synge and A. Schild, *Tensor Calculus*, Dover Publications, Inc., New York, ristampa del 1978 fatta sull'edizione originale del 1948.
Infine, lo studente che ha molta fretta, può consultare il Cap. 2 dell'interessante libro di Moshe Carmeli, *Classical Fields: General Relativity and Gauge Theory*, John Wiley and Sons, Inc., New York, 1982, in cui vi si trova una concisa ma esauriente esposizione dei tensori e della Geometria Riemanniana (con esercizi svolti) necessaria per capire la RG.

scopi, dobbiamo accontentarci di un semplice cenno in quanto una trattazione rigorosa richiederebbe una digressione troppo estesa e molto tempo[2].
Ci limitiamo a dire che **una varietà ad n dimensioni** è un ente geometrico che localmente assomiglia ad una porzione di spazio euclideo n-dimensionale $\mathbb{R}^n$. Se, per esempio, confrontiamo una 2-sfera S^2 (cioè la **superficie di una sfera** della geometria euclidea elementare) con il piano euclideo $\mathbb{R}^2$ ci rendiamo conto che essi sono palesemente differenti, ma una piccola porzione di sfera somiglia molto ad una piccola porzione di piano (trascurando naturalmente le proprietà metriche in quanto, com'è noto, la sfera non è sviluppabile su di un piano) (vedi Fig.8.1).

Figura 8.1: Una 2-sfera S^2 come varietà.

Il fatto che S^2 è *compatta*, cioè in un certo senso finita, mentre $\mathbb{R}^2$ *si estende all'infinito* è una proprietà globale e non locale. Le proprietà globali delle varietà sono oggetto della **topologia** il cui studio esula dalle nostre considerazioni.
Semplificando molto, diremo che una **varietà n-dimensionale**, che indicheremo con V^n oppure con M^n - M essendo l'iniziale della parola inglese **manifold** che significa appunto **varietà** - è un insieme di punti tali che ciascuno sia individuato da un insieme ordinato di n numeri reali $(x^1, x^2, ..., x^n)$, o sinteticamente (x^i), $i = 1, 2, ..., n$, dette **coordinate**, dove ogni coordinata assume valori in un sottoinsieme dei numeri reali che può, in particolare, estendersi da $-\infty$ a $+\infty$. Per fissare le idee, possiamo pensare a queste coordinate come corrispondenti a distanze o angoli dello spazio euclideo. Il motivo per cui le coordinate sono indicate con indici in alto (apici) e non con indici in basso (pedici), come si usa di consueto, risulterà chiaro in seguito.
Il concetto di **varietà** deriva dalla **cartografia** cioè da quella scienza che notoriamente studia la rappresentazione di porzioni di superficie terrestre su delle carte piane (carte geografiche, topografiche, nautiche, ecc.). Le basi matematiche della cartografia sono dovute a Gauss. L'eccezionale generalità del concetto di varietà è dimostrato dalle sue applicazioni su una vasta classe di figure geometriche molto complicate.
Nella pratica della cartografia le persone incaricate dei rilevamenti si dividono in gruppi in modo che:

a) ogni porzione del territorio da rilevare venga assegnata ad un gruppo, indicato per es. con l'indice p;

b) ogni volta che due porzioni di territorio assegnato a gruppi diversi (per esempio con indici p e q) presentino punti in comune, i gruppi devono marcare tutte le linee di corrispondenza sulle loro rispettive carte. In pratica ciò significa indicare un certo numero di località o di luoghi per poter individuare immediatamente le corrispondenze.

[2]Le precedenti considerazioni di natura intuitiva sono desunte dall'interessante libro di Bernard Schutz, *Geometrical methods of mathematical physics*, Cambridge University Press, 1980, Chap. 2, da pag. 23 in poi. Invece le considerazioni del testo successive alla presente nota sono una nostra rielaborazione ispirata al bel libro di B. A. Dubrovin, S. P Novikov, A. T Fomenko, *Geometria Contemporanea, vol.2°, Geometria e Topologia delle varietà*, da pag.11 in poi.

Sulla carta viene, di regola, indicato anche il modo per calcolare le distanze reali, ma ciò è un aspetto sul quale non ci soffermiamo - almeno per ora - in quanto non ha nulla a che fare col concetto generale di varietà, ci ritorneremo in seguito quando parleremo delle **varietà riemanniane** che sarà il principale obbiettivo del nostro studio.
Noi sappiamo che il supporto materiale di ogni **carta** è un **foglio quadrettato**, cioè dotato di **coordinate** (comunemente dette *latitudine e longitudine*). L'insieme dei fogli o carte geografiche, com'è noto, si chiama **atlante**. Questo linguaggio verrà mantenuto anche nello studio astratto delle varietà.
Una osservazione importante a proposito di una varietà è che **può non essere possibile coprirla con un'unica carta** (cioè con un unico sistema di coordinate **non degenere**), cioè un sistema che assegni **un'unica** n-pla di coordinate ad ogni punto. A volte, per certi scopi, può essere conveniente usare sistemi di coordinate con **punti degeneri**. Per esempio le coordinate polari (r, θ) definite su di un piano presentano una degenerazione nell'origine perché in tale punto $r = 0$ e θ è indeterminato. In questo caso è però possibile evitare tale degenerazione nell'origine usando coordinate cartesiane, ma in altre circostanze non sempre è possibile evitare i punti degeneri. Si dimostra, per esempio, che non esiste un solo sistema di coordinate che copra l'intera sfera S^2 senza degenerazioni o punti singolari che dir si voglia! **Per ricoprire l'intera sfera senza degenerazioni occorrono due distinti sistemi di coordinate**. Lavoreremo quindi con sistemi di coordinate che coprono solo una porzione della varietà e che sono chiamate (proprio come in cartografia) **carte coordinate o sistemi di coordinate**. Un insieme di carte coordinate che copra l'intera varietà è detto ovviamente (come nella geografia) **atlante**. La teoria delle varietà ci dice come passare da una carta ad un'altra con una trasformazione di coordinate nelle regioni di sovrapposizione. *Il comportamento delle grandezze geometriche, analitiche o fisiche, sotto l'azione delle trasformazioni di coordinate è alla base del calcolo tensoriale*.
Per chi non si accontenta di una descrizione intuitiva delle varietà come quella data più sopra, riportiamo qui di seguito una breve trattazione matematicamente più soddisfacente[3].

Definizione 1:
Uno **spazio topologico** è una coppia $(M, \mathcal{U})$ costituita da un "insieme astratto" M di "elementi qualsiasi", che diremo "punti" di M e da una famiglia $\mathcal{U}$ di sottoinsiemi di M, soddisfacente le seguenti condizioni:

i) $\emptyset \in \mathcal{U}$, $M \in \mathcal{U}$;

ii) L'intersezione di un numero finito di sottoinsiemi di $\mathcal{U}$, appartiene ad $\mathcal{U}$;
cioè se $\mathcal{U}_1, \mathcal{U}_2, \ldots\ldots, \mathcal{U}_n \in \mathcal{U}$, risulta $\bigcap_{i=1}^{n} \mathcal{U}_i \in \mathcal{U}$;

iii) Unioni arbitrarie di elementi di $\mathcal{U}$ stanno in $\mathcal{U}$.

Gli elementi di $\mathcal{U}$ vengono chiamati **insiemi aperti** o semplicemente gli **aperti** di M. La collezione $\mathcal{U}$ si chiama **topologia** in M. Talvolta, per semplificare, lo spazio topologico si indica semplicemente con M anziché con $(M, \mathcal{U})$.

Definizione 2:
Dicesi **varietà topologica** ad n dimensioni ogni **spazio topologico** M che sia:

a) **di Hausdorff** (cioè tale che ogni coppia di punti distinti ammetta intorni disgiunti);

[3]Per una trattazione rigorosa delle varietà differenziabili consigliamo al lettore di consultare oltre ai libri, già citati nella nota precedente, cioè il libro di Bernard Schutz e quelli di Dubrovin, Novikov, Fomenko, il seguente volume:
William M. Boothby, *An introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, Inc., 1986, Chap.I and Chap. III.

b) tale che ogni punto P possiede un intorno aperto U_P **omeomorfo** (cioè in corrispondenza biunivoca e continua) ad un **aperto** dello spazio euclideo n-dimensionale $\mathbb{R}^n$ ($M = \bigcup U_P$) (vedi Fig.8.2).

L'omeomorfismo $\varphi : U_P \to \mathbb{R}^n$ si dice **applicazione coordinata relativa a** U_P. Chiameremo **carta locale o sistema di coordinate** di M ogni coppia del tipo (U, φ); U si dirà **dominio della carta**.
Con le locuzioni introdotte la condizione b) si esprime anche dicendo che **i domini di tutte le carte ricoprono** M (cioè ($M = \bigcup U_P$).
Usando le cosiddette **funzioni di proiezione** $\pi^k : \mathbb{R}^n \to \mathbb{R}$, definite da:

$$\pi^1(x^1, x^2, ..., x^n) = x^1, \quad \pi^2(x^1, x^2, ..., x^n) = x^2, ..., \pi^n(x^1, x^2, ..., x^n) = x^n,$$

una carta definisce nel suo dominio U un insieme di **funzioni coordinate**:

$$x^i = \pi^i \circ \varphi : U \to \mathbb{R}, \qquad (i = 1, 2, ..., n).$$

Per ogni $P \in U$, x^i si dice **coordinata i-esima** e gli n numeri reali

$$\varphi(P) = (x^1(P), x^2(P), ..., x^n(P))$$

si chiamano **coordinate** di P nella carta (U, φ). L'intero n si dice **dimensione della varietà** $dim\, M = n$.

Figura 8.2: Carta locale (U, φ) relativa ad un generico punto P di una varietà M.

Se P giace anche in un secondo intorno V è possibile definire una seconda carta locale (V, ψ) e perciò altre coordinate per il punto P: $\psi(P) = (x'^1(P), ..., x'^n(P))$. Poiché φ e ψ sono omeomorfismi essi definiscono per composizione un nuovo omeomorfismo (vedi Fig.8.3)[4]:

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V) \qquad (\textbf{funzione di transizione}), \tag{8.1}$$

il cui dominio $\varphi(U \cap V)$ ed il cui codominio $\psi(U \cap V)$ sono due insiemi aperti di $\mathbb{R}^n$ corrispondenti ai punti di $U \cap V$ mediante le due applicazioni coordinate φ, ψ rispettivamente. In "coordinate" l'applicazione $\psi \circ \varphi^{-1}$ è data dalle *funzioni continue*:

$$x'^i = h^i(x^1, x^2, ..., x^n) \equiv x'^i(x^1, x^2, ..., x^n), \quad (i = 1, 2, ..., n) \tag{8.2}$$

le quali danno le coordinate x' del punto $P \in U \cap V$ in funzione delle coordinate x.
Analogamente l'applicazione

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V) \quad (\textbf{funzione di transiz. inversa della } \psi \circ \varphi^{-1}), \tag{8.3}$$

[4]Si ricordi che l'applicazione composta $\psi \circ \varphi^{-1}$ deve intendersi che si applica prima φ^{-1} e poi ψ, cioè in ordine inverso da come sono scritte.

Figura 8.3: Carte locali (U, φ), (V, ψ) su di una varietà M e funzione di transizione.

dà la funzione inversa che esprime la coordinata x di P in funzione delle x'

$$x^j = g^j(x'^1, x'^2, ..., x'^n) \equiv x^j(x'^1, x'^2, ..., x'^n); \quad (j = 1, 2, ..., n) \tag{8.4}$$

le (8.4) ci danno il passaggio di coordinate inverso di (8.2).
Il fatto che le applicazioni $\varphi \circ \psi^{-1}$ e $\psi \circ \varphi^{-1}$ sono **omeomorfismi** *e sono* **inverse una dell'altra** *è equivalente ad affermare la continuità delle funzioni $h^i(x)$ e $g^j(x)$ $i, j = 1, 2, ...n$ insieme con la validità delle identità*

$$h^i\big(g^1(x'), ..., g^n(x')\big) \equiv x'^i \quad (i = 1, 2, ..., n),$$
$$g^j\big(h^1(x), ..., h^n(x)\big) \equiv x^j \quad (j = 1, 2, ..., n).$$

Allora ogni punto di una varietà topologica M giace in una grande **collezione di carte** detta, in analogia con il linguaggio geografico, **atlante**, ma allorquando due intorni di un punto hanno intersezione non vuota, noi possediamo le formule (8.2) e (8.4) per il passaggio di coordinate da un sistema all'altro.
L'idea di base che conduce alle **varietà differenziabili** è quella di selezionare quelle carte che godono della proprietà di essere definite sempre da **funzioni differenziabili** e non semplicemente continue.

Definizione 3 (carte compatibili):
Diremo che due carte (U, φ) e (V, ψ) sono $C^\infty -$ **compatibili** se $U \cap V \neq \emptyset$ implica che le funzioni $h^i(x)$ e $g^j(x')$, che forniscono il cambio di coordinate, siano C^∞; ciò equivale a richiedere che le applicazioni $\varphi \circ \psi^{-1}$ e $\psi \circ \varphi^{-1}$ siano **diffeomorfismi** dei due sottoinsiemi $\varphi(U \cap V)$ e $\psi(U \cap V)$ di $\mathbb{R}^n$ (vedi Fig.8.3).

Definizione 4 (struttura differenziabile):
Una **struttura differenziabile** (o C^∞ o **liscia**) su di una varietà topologica M è una famiglia $\mathcal{U} = \{(U_\alpha, \varphi_\alpha)\}$ di **carte locali** tali che:

1) gli insiemi U_α **ricoprono** M: $M = \bigcup_{\alpha \in I} U_\alpha$, con I un dato insieme di indici;

2) per ogni coppia di indici α, β le carte $(U_\alpha, \varphi_\alpha)$ e (U_β, φ_β) sono $C^\infty -$ compatibili;

3) ogni carta (V, ψ) compatibile con ogni $(U_\alpha, \varphi_\alpha) \in \mathcal{U}$ è essa stessa in $\mathcal{U}$.

Una varietà C^∞ è una varietà topologica insieme con una struttura differenziabile C^∞. Essa è una **varietà differenziabile**. Possiamo perciò porre la fondamentale definizione:

Definizione 5 (varietà differenziabile):
Una **varietà differenziabile** *è una* **varietà topologica** *munita di una* **struttura differenziabile**.
L'esempio più elementare di varietà differenziabile è lo spazio euclideo stesso di dimensione $n = 1, 2, 3, ..., n$ o un suo dominio qualunque: le curve e le superfici regolari degli spazi euclidei sono anch'esse delle varietà differenziabili.

8.3 Esempi di varietà differenziabili

L'utilità del concetto di varietà risiede veramente nella sua generalità, infatti esso abbraccia insiemi di **"oggetti"** di natura qualsiasi (chiamati usualmente **"punti"**), insiemi che ordinariamente non verrebbero riguardati come **"spazi"**.
Per definizione, *ogni insieme M che può essere* **"parametrizzato"** *in modo continuo è una* **varietà** *la cui dimensione è il numero dei parametri indipendenti.*
Diamo ora alcuni esempi di varietà.

Figura 8.4: Semplici esempi di varietà 2-dimensionali: il piano, il cilindro, il cono e la sfera.

Sono esempi semplici ma di notevole importanza nelle applicazioni **le curve** di $\mathbb{R}^2$ ed $\mathbb{R}^3$ e **le superfici** di $\mathbb{R}^3$.
Il concetto di **curva o linea** trae la sua origine dall'immagine di un filo flessibile ed inestensibile (per es. un cavetto sottile di acciaio come quelli che si usano sulle barche a vela), oppure dall'immagine della traiettoria descritta da un punto materiale in moto rispetto ad un prefissato sistema di riferimento.
Il concetto di **superficie** trae invece la sua origine da un foglio di materiale *flessibile ed inestendibile*, su cui si possano disegnare figure, si possa deformarlo, curvandolo e piegandolo in infiniti modi, ma *non mai lacerarlo né dilatarlo.*
Consideriamo, in modo, più dettagliato alcuni esempi di varietà.

1) La più semplice di tutte le superfici è ovviamente il piano $\mathbb{R}^2$.
   Altre superfici semplici ma di notevole importanza per le applicazioni sono: il cilindro, il

cono e la sfera della geometria elementare (Fig.8.4). Col termine **"sfera"**, o meglio 2- sfera $\mathbb{S}^2$, intenderemo sempre la **"superficie della sfera"** e non il suo interno e cioè il *"solido sfera o sfera piena"*. La 2-sfera $\mathbb{S}^2$ può essere definita come l'insieme dei punti di $\mathbb{R}^3$ per i quali sia

$$(x^1)^2 + (x^2)^2 + (x^3)^2 = cost. > 0$$

Ogni punto di $\mathbb{S}^2$ possiede un intorno sufficientemente piccolo in corrispondenza biunivoca con un disco di $\mathbb{R}^2$.
Come specifico esempio, consideriamo le usuali coordinate sferiche, con $\theta = x^1$ e $\varphi = x^2$. Allora la 2-sfera $\mathbb{S}^2$ appare essere applicata al rettangolo $0 \leq x^1 \leq \pi, 0 \leq x^2 \leq 2\pi$.
La sfera è l'esempio più semplice che illustra l'importanza di permettere l'esistenza di più carte.

2) L'insieme delle **rotazioni** di un corpo rigido in 3 dimensioni è una **varietà** 3-dimensionale poiché esso può essere parametrizzato, per esempio con i 3 *angoli di Eulero*, ovvero con altri 3 parametri indipendenti qualsiasi, che potrebbero essere 3 coseni direttori indipendenti di una terna solidale al corpo rigido, oppure i cosiddetti *parametri di Andoyer,* o ancora i *parametri di Cayley-Klein*, ecc.

3) L'insieme di tutte le trasformazioni di Lorentz (gruppo di Lorentz) è una **varietà a 3 dimensioni**: i parametri sono le 3 componenti della velocità di un riferimento inerziale rispetto ad un altro;

4) Lo **spazio** $\mathbb{Q}^n$ **delle configurazioni** di un sistema olonomo ad n gradi di libertà è definito come l'insieme delle n-ple $q^i, i = 1, 2, ..., n$ di parametri lagrangiani, ciascuno dei quali caratterizza una configurazione del sistema. Esso è una varietà differenziabile di dimensione n;

5) Lo **spazio** Φ^{2n} **delle fasi** di un sistema olonomo è definito come l'insieme delle **"fasi" o "stati"** caratterizzati dalle 2 n-ple $(q^i, p^i), i = 1, 2, ..., n$, le q^i rappresentanti la configurazione e le p^i i momenti coniugati (ovvero le velocità) ad un determinato istante.
Se il sistema materiale è formato da N particelle libere, lo spazio delle fasi è una varietà differenziabile a $6N$ dimensioni perché la posizione di una particella richiede 3 parametri e la sua velocità altri 3.

6) Una varietà molto comune è uno **spazio vettoriale** V (per es. uno spazio vettoriale di dimensione finita sopra i numeri reali). Per vedere che tale spazio è una varietà, consideriamo un'applicazione da esso a $\mathbb{R}^n$ per qualche n. Supponiamo che lo spazio vettoriale V sia n-dimensionale, e scegliamo una base $\{\mathbf{e}_i\}$. Ogni vettore $\mathbf{v}$ è allora rappresentabile in modo univoco come una combinazione lineare dei vettori della data base:

$$\mathbf{v} = v^1\mathbf{e}_1 + v^2\mathbf{e}_2 + ... + v^n\mathbf{e}_n = \sum_{i=1}^{n} v^i\mathbf{e}_i.$$

Ma $\mathbf{v}$ è un punto di V, perciò questo stabilisce un'applicazione da V ad $\mathbb{R}^n$, $\mathbf{v} \mapsto (v^1, v^2, ..., v^n)$. Infatti ogni punto di $\mathbb{R}^n$ corrisponde ad un unico vettore di V sotto questa applicazione, perciò non solo V è coperto interamente da un singolo sistema di coordinate, che abbiamo appena costruito, ma V è identico, come varietà, ad $\mathbb{R}^n$. Nel linguaggio della teoria dei gruppi, V ed $\mathbb{R}^n$ sono *isomorfi*. Questo è un risultato importante. Ciò significa che ogni spazio vettoriale può essere pensato, quando ciò sia conveniente, semplicemente come $\mathbb{R}^n$.

7) Data un'equazione (algebrica o differenziale) per una *variabile dipendente* y in termini di una *variabile indipendente* x, si può definire l'insieme di tutte le coppie (y, x) come una varietà. Questo concetto si può estendere facilmente ad un numero arbitrario di variabili *dipendenti ed indipendenti.*

8) L'esempio 2) del gruppo delle rotazioni di un corpo rigido, dato più sopra, è un esempio di **gruppo di Lie**, che viene definito nel seguente modo:
Dicesi gruppo di Lie G, un gruppo che è anche una varietà C^∞, con la restrizione che l'operazione di gruppo induce un'applicazione C^∞ della varietà in sé stessa.
Ciò significa che: se fissiamo un elemento arbitrario a del gruppo, questo induce una mappa di G in sé stesso, mandando ogni elemento $b \in G$ in ba, $b \mapsto ba$. Questa mappa deve essere C^∞, ecc., ecc.

8.4 Curve e superfici di una varietà differenziabile

Data una varietà differenziabile di dimensione finita, fisseremo l'attenzione sui punti che la compongono e sui sottoinsiemi di punti che definiscono curve e superfici di dimensioni differenti. Spesso definiremo queste curve e queste superfici in modo **parametrico**, perciò (esattamente nello stesso modo in cui si procede nello spazio euclideo a due o tre dimensioni), poiché una curva ha un solo grado di libertà, i suoi punti dipendono da un solo parametro, e perciò definiamo una **curva** con le seguenti **equazioni parametriche**:

$$x^i = x^i(u), \qquad (i = 1, 2, ..., n) \tag{8.5}$$

dove u è il parametro e $x^1(u), x^2(u), ..., x^n(u)$ denotano n funzioni di u. Analogamente, un **sottospazio** o **superficie** ad m dimensioni $(m < n)$ ha equazioni parametriche dipendenti da m parametri:

$$x^i = x^i(u^1, u^2, ..., u^m), \qquad (i = 1, 2, ..., n). \tag{8.6}$$

Se, in particolare, $m = n - 1$, il sottospazio è chiamato **ipersuperficie** le cui equazioni parametriche sono:

$$x^i = x^i(u^1, u^2, ..., u^{n-1}), \qquad (i = 1, 2, ..., n) \tag{8.7}$$

e gli $n - 1$ parametri possono essere eliminati da queste n equazioni ottenendo una sola equazione che lega le n coordinate $x^1, x^2, ..., x^n$, cioè

$$f(x^1, x^2, ..., x^n) = 0, \tag{8.8}$$

questa è l'equazione cartesiana della ipersuperficie.
Possiamo considerare le cose da un punto di vista diverso ma equivalente dicendo che un punto generico di una varietà n-dimensionale ha n gradi di libertà. Se lo si costringe a giacere su una ipersuperficie, che è uno spazio $(n-1)$-dimensionale, allora le sue coordinate devono soddisfare ad $n - (n-1) = 1$ vincoli, cioè ad **un solo vincolo**, espresso dall'equazione (8.8). Analogamente, punti di un sottospazio m-dimensionale $(m < n)$ devono soddisfare ad $\mathbf{n - m}$ **vincoli**

$$\begin{cases} f^1(x^1, x^2, ..., x^n) = 0 \\ f^2(x^1, x^2, ..., x^n) = 0 \\ ..... \\ f^{n-m}(x^1, x^2, ..., x^n) = 0 \end{cases} \tag{8.9}$$

che è una definizione alternativa alla rappresentazione parametrica (8.6).

8.5 Trasformazioni di coordinate su di una varietà differenziabile

Come abbiamo visto, un punto di una varietà può essere coperto da diverse carte coordinate ma *lo scopo essenziale del calcolo tensoriale consiste nel fatto che quando facciamo un'affermazione che riguarda i tensori non vogliamo che valga per un solo sistema di coordinate, ma pretendiamo che essa sia valida in ogni sistema di coordinate*. Sicché dobbiamo scoprire come si comportano le varie grandezze e quantità, di natura analitica, geometrica o fisica, che prenderemo in considerazione, nel passaggio da un sistema di coordinate ad un altro. Consideriamo perciò una generica **trasformazione di coordinate** $x^i \to x'^{\,i}$ definita dalle equazioni:

$$x'^{\,i} = f^i(x^1, x^2, ..., x^n) \quad (i = 1, 2, ..., n), \tag{8.10}$$

dove le f^i sono funzioni continue e differenziabili, in un certo campo di variabilità delle coordinate $x^1, x^2, ..., x^n$. Per il momento consideriamo una trasformazione di coordinate **passivamente**, cioè come l'assegnazione ad un punto della varietà delle **nuove coordinate** accentate $(x'^{\,1}, x'^{\,2}, ..., x'^{\,n})$, le cui **vecchie coordinate** erano $(x^1, x^2, ..., x^n)$. Possiamo scrivere le (8.10) più sinteticamente come

$$x'^{\,i} = f^i(x)$$

dove, da adesso in poi, si sottintende che gli indici minuscoli latini varino da 1 ad n, essendo n la dimensione della varietà, e le f^i sono tutte funzioni delle vecchie coordinate non accentate. La stessa equazione può essere scritta ancora più sinteticamente come

$$x'^{\,i} = x'^{\,i}(x), \tag{8.11}$$

dove le $x'^{\,i}(x)$ indicano le n funzioni $f^i(x)$. Alcuni autori utilizzano un altro modo di scrivere le formule (8.11) di trasformazione delle coordinate, ponendo l'accento anziché sulla lettera x, sull'indice i, scrivendo $x^{i'} = f^{i'}(x)$ o più semplicemente $x^{i'} = x^{i'}(x)$. In questo caso si deve intendere che l'indice i' assuma gli stessi valori dell'indice i, cioè $i' = 1, 2, ..., n$, perciò quando un indice non è scritto genericamente con una lettera ma assume i valori numerici concreti $1, 2, 3, ...$ si deve intendere che $(x^{1'}, x^{2'}, ..., x^{n'})$ e $(x'^{\,1}, x'^{\,2}, ..., x'^{\,n})$ siano due scritture del tutto equivalenti. Questa notazione si rivelerà molto utile soprattutto nell'economizzare i molti indici che è necessario usare nel calcolo tensoriale e di ciò ci si renderà conto in seguito.
Calcoliamo ora le derivate parziali, rispetto alle variabili non accentate x^j, dei secondi membri delle formule di trasformazione (8.11) e con esse formiamo la seguente matrice

$$\left(\frac{\partial x'^{\,i}}{\partial x^j}\right) = \begin{pmatrix} \frac{\partial x'^{\,1}}{\partial x^1} & \frac{\partial x'^{\,1}}{\partial x^2} & \cdots \frac{\partial x'^{\,1}}{\partial x^n} \\ \frac{\partial x'^{\,2}}{\partial x^1} & \frac{\partial x'^{\,2}}{\partial x^2} & \cdots \frac{\partial x'^{\,2}}{\partial x^n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x'^{\,n}}{\partial x^1} & \frac{\partial x'^{\,n}}{\partial x^2} & \cdots \frac{\partial x'^{\,n}}{\partial x^n} \end{pmatrix} \tag{8.12}$$

Il determinante J' di questa matrice si chiama lo **jacobiano** della trasformazione:

$$J' = det\left(\frac{\partial x'^{\,i}}{\partial x^j}\right). \tag{8.13}$$

Supporremo che questo sia non nullo ($J' \neq 0$) in un certo campo di variabilità delle x^j. Per il teorema di Dini sulle funzioni implicite, le equazioni (8.11) si possono risolvere (almeno in linea di principio) rispetto alle vecchie coordinate x^i ed ottenere le equazioni della **trasformazione inversa**

$$x^i = x^i(x'). \tag{8.14}$$

Lo jacobiano della trasformazione inversa è

$$J = det\left(\frac{\partial x^i}{\partial x'^j}\right),$$

e dalla regola sul prodotto dei determinanti ne segue che $J = 1/J' \neq 0$.
Nello spazio euclideo tridimensionale $\mathbb{R}^3$ l'equazione cartesiana esplicita di una superficie (varietà 2-dimensionale) nelle coordinate (x, y, z), è $z = f(x, y)$, quindi il differenziale totale di z è così definito

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

Se $f(x^1, ..., x^n)$ è una funzione di più variabili il suo differenziale totale è dato da:

$$df = \frac{\partial f}{\partial x^1}dx^1 + \frac{\partial f}{\partial x^2}dx^2 + ... + \frac{\partial f}{\partial x^n}dx^n.$$

In maniera analoga, dalle (8.11) possiamo definire il differenziale totale delle x'^i come

$$dx'^i = \frac{\partial x'^i}{\partial x^1}dx^1 + \frac{\partial x'^i}{\partial x^2}dx^2 + ... + \frac{\partial x'^i}{\partial x^n}dx^n,$$

per ogni i da 1 a n. Il precedente differenziale può essere scritto più sinteticamente facendo uso del simbolo di sommatoria:

$$dx'^i = \sum_{j=1}^{n} \frac{\partial x'^i}{\partial x^j}dx^j. \tag{8.15}$$

La formula precedente si può scrivere in maniera più sintetica introducendo la **convenzione di sommazione di Einstein**: *ogni espressione monomia, cioè formata da un certo numero di fattori, in cui certi indici letterali figurano ciascuno due volte, rispettivamente in posizione superiore in un fattore e in posizione inferiore in un altro, significa la somma di tutti i monomi analoghi in cui ciascun indice prende gli n valori* $1, 2, ..., n$.
Per esempio

$$A_iB^i = \sum_{i=1}^{n} A_iB^i = A_1B^1 + A_2B^2 + A_3B^3 + ... + A_nB^n$$

$$A_{ijk}B^{ij} = \sum_{i,j=1}^{n} A_{ijk}B^{ij} =$$

$$\begin{cases} = A_{11k}B^{11} + A_{12k}B^{12} + \cdots + A_{1nk}B^{1n} \\ +A_{21k}B^{21} + A_{22k}B^{22} + \cdots + A_{2nk}B^{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ +A_{n1k}B^{n1} + A_{n2k}B^{n2} + \cdots + A_{nnk}B^{nn}. \end{cases}$$

Con una piccola estensione sottintenderemo la somma anche in espressioni del tipo

$$A^i_i, \quad \frac{\partial u^i}{\partial x^i}, \quad \frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial t}, \ ecc.$$

Possiamo quindi scrivere la (8.15) molto più sinteticamente così

$$dx'^i = \frac{\partial x'^i}{\partial x^j}dx^j, \tag{8.16}$$

oppure, tenendo conto di ciò che abbiamo detto in precedenza, anche più semplicemente

$$dx^{i'} = \frac{\partial x^{i'}}{\partial x^i} dx^i, \tag{8.17}$$

col risparmio di un indice.
L'indice non ripetuto i che compare in entrambi i membri della (8.16), ovvero l'indice i' che compare in entrambi i membri della (8.17), si dice **libero** e può assumere qualsiasi valore da 1 ad n. L'indice j nel membro destro della (8.16), ovvero l'indice i nel membro destro della (8.17) sono indici ripetuti e perciò indicano una sommatoria da 1 a n. Un indice ripetuto si chiama **saturato** o **muto** perché può essere sostituito da qualsiasi altro non ancora usato. Per esempio si può scrivere

$$dx'^{\,i} = \frac{\partial x'^{\,i}}{\partial x^j} dx^j = \frac{\partial x'^{\,i}}{\partial x^h} dx^h = \frac{\partial x'^{\,i}}{\partial x^k} dx^k = ... \tag{8.18}$$

Un simbolo molto usato in matematica, e nel calcolo tensoriale in particolare, è la cosiddetta **delta di Kronecker** così definita

$$\delta^i_j = \delta_{ij} = \delta^{ij} = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases} \tag{8.19}$$

Per esempio è facile constatare che

$$\frac{\partial x'^{\,i}}{\partial x'^{\,j}} = \frac{\partial x^i}{\partial x^j} = \delta^i_j, \tag{8.20}$$

poiché sia le coordinate x^i, sia le $x'^{\,i}$, sono indipendenti.
Un'altra proprietà ovvia della delta di Kronecker è che $\delta^k_j A^j = A^k$, poiché nella parte sinistra di questa equazione, quando si fa la sommatoria, il solo termine che sopravvive è quello per cui $j = k$.

Esercizio: Dimostrare che

$$\delta^i_j \delta^j_k = \delta^i_k; \quad \delta^i_i = n; \quad \frac{\partial x'^{\,i}}{\partial x^j} \frac{\partial x^j}{\partial x'^{\,k}} = \frac{\partial x^i}{\partial x'^{\,j}} \frac{\partial x'^{\,j}}{\partial x^k} = \delta^i_k.$$

8.6 Digressione sulla convenzione di sommazione di Einstein

Nel paragrafo precedente abbiamo visto che passare da un sistema di coordinate, definito su di una varietà differenziabile, ad un altro comporta necessariamente l'uso di notazioni diverse per indicare i due sistemi di coordinate: noi abbiamo scelto le notazioni x^i ed $x'^{\,i}$ indicando il secondo sistema di coordinate con un apice sovrapposto alla stessa lettera x che indica il primo sistema di coordinate.[5] Ciò induce ad usare per la matrice jacobiana della trasformazione di coordinate $x^i \to x^{i'}$ una notazione in cui gli elementi della matrice $A' = \left(\frac{\partial x^{i'}}{\partial x^j}\right) \equiv (A^{i'}_j)$ sono contrassegnati da un indice di riga posto in alto e accentato i' ed un indice di colonna posto in basso e non accentato j. In maniera analoga si indica la matrice jacobiana della trasformazione inversa $x^{i'} \to x^i$ e cioè $(A')^{-1} = \left(\frac{\partial x^i}{\partial x^{j'}}\right) \equiv (A^i_{j'})$. L'uso degli indici accentati si rivelerà molto conveniente nello sviluppo del calcolo tensoriale ove si tenga presente che le notazioni sono scelte in modo tale da soddisfare al seguente requisito: **scrivere le formule nel modo più chiaro e compatto possibile**. Ciò viene da noi realizzato rispettando rigorosamente le seguenti convenzioni:

[5] Alcuni autori, usano per esempio le lettere x e $\bar{x}$ per indicare i due diversi sistemi di coordinate, o altre notazioni.

1) gli elementi di una matrice saranno indicati con una lettera munita di due indici: uno in alto (apice) ed uno in basso (pedice) di cui uno accentato, l'indice in alto è sempre di riga e quello in basso sempre di colonna;

2) Il prodotto di matrici si effettuerà sempre righe per colonne;

3) la convenzione di sommazione di Einstein verrà applicata esclusivamente per indici muti ripetuti una volta in alto ed una volta in basso; per es. $\sum_{i=1}^{n} u^i v_i = u^i v_i$. Nelle somme formate da prodotti di termini con indici ripetuti solo in alto o solo in basso conserveremo il simbolo di sommatoria: $\sum_{i=1}^{n} u^i v^i$ oppure $\sum_{k=1}^{n} u_k v_k$.[6] Se si vuole scrivere un monomio del tipo $u^i v^i$ con i fissato, oppure $u_j v_j$ con j fissato, perciò senza intendere alcuna sommazione, scriveremo $u^i v^i (NS)$ ovvero $u_j v_j (NS)$, dove NS sta per non sommare.

L'espressione monomia di cui si parla nella convenzione di somma di Einstein vista nel paragrafo precedente è chiamata da qualche autore un **falso monomio** in quanto si tratta in realtà di un polinomio somma di n termini se comporta un solo indice muto, di n^2 termini se ne comporta due, ecc.
Osserviamo anche che: *ogni uguaglianza ove figurano alcuni indici liberi, alla stessa "altezza" nei due membri, s'intenderà valida per tutti i valori* $1, 2, ..., n$ *degli indici: essa rappresenterà dunque in realtà, se comporta* p *indici liberi, un sistema di* n^p *uguaglianze*. Per es. l'uguaglianza $a^i_h x^h = b^i$ con a^i_h, b^i coefficienti assegnati e le x^i incognite rappresenta un sistema di n equazioni lineari in n incognite

$$\begin{cases} a^1_1 x^1 + a^1_2 x^2 + \cdots + a^1_n x^n = b^1 \\ a^2_1 x^1 + a^2_2 x^2 + \cdots + a^2_n x^n = b^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a^n_1 x^1 + a^n_2 x^2 + \cdots + a^n_n x^n = b^n \end{cases} \tag{8.21}$$

Alcune osservazioni.

a) L'impiego degli indici in alto può creare rischio di confusione con la scrittura delle potenze, ma generalmente il contesto chiarisce sempre di che cosa si tratti; si può comunque fare uso di parentesi quando è necessario. Per esempio se devo elevare a potenza p-esima la coordinata x^i posso scrivere $(x^i)^p$.

b) Se si vuole designare un falso monomio con una sola lettera, si dovrà munirla degli stessi indici liberi del falso monomio, purché la notazione introdotta si estenda a tutti i valori di questi indici. Per esempio:

$$u^i v_i = w, \quad a^i_j x^j = p^i, \quad a^i_j x^j x^k = q^{ik}, \ ecc.$$

Indichiamo infine due precauzioni che rendono indispensabile l'impiego sistematico di tutte le convenzioni descritte in precedenza:

i) per dire che "$a^i b_j = 1$ se $i = j$", si deve scrivere effettivamente così, e non $a^i b_i = 1$ che significa $a^1 b_1 + a^2 b_2 + ... + a^n b_n = 1$, cioè tutt'altra cosa.

ii) rispettare nelle scritture l'imperativo di **non usare indici già utilizzati per una coppia di indici muti**: la scrittura $A^i_j B_i X^i$, per esempio, non ha alcun senso, mentre ha senso ciascuna delle scritture $A^h_j B_h X^i$ e $A^i_j B_h X^h$.

[6] Per l'esattezza storica dobbiamo osservare che Einstein introdusse la convenzione di somma su indici ripetuti, senza fare distinzione se essi sono tutti e due in basso o tutti e due in alto, oppure uno in alto ed uno in basso; alcuni autori infatti applicano la convenzione così come è stata enunciata da Einstein, ciò può essere, in alcuni casi, effettivamente utile come vedremo alla fine del paragrafo 8.10. Tuttavia noi useremo la convenzione così come l'abbiamo enunciata per evitare ambiguità.

8.7 Spazio vettoriale tangente $T_P(M)$ ad una varietà differenziabile

Data una varietà differenziabile M^n di dimensione n e di classe C^2 almeno, facciamo vedere come, per ciascun suo punto P si possa definire uno **spazio vettoriale tangente** $T_P(M)$ i cui elementi sono i "**vettori tangenti**" alla varietà in P definiti nella maniera che segue.
Sia $\overline{P}$ un generico punto della varietà M^n, $U_{\overline{P}}$ un suo intorno munito di coordinate x^i (*carta locale*), $\overline{x}^i$ le coordinate di $\overline{P}$. Si pensi poi ad una dipendenza delle x^i da una variabile numerica t

$$x^i = x^i(t) \tag{8.22}$$

che per comodità interpreteremo come un tempo e che chiameremo *"ascissa temporale"*, definita in un intervallo comprendente lo zero ($-\varepsilon \leq t \leq \varepsilon, \varepsilon > 0$). Diremo che le (8.22) sono le equazioni parametriche del *"moto"* di un punto sulla varietà. Ci limitiamo a considerare quei particolari "moti" che hanno inizio da $\overline{P}$, cioè tali che $x^i(0) = \overline{x}^i$ e che sono sufficientemente regolari perché abbia senso uno sviluppo del tipo

$$x^i = \overline{x}^i + \left(\frac{dx^i}{dt}\right)_{t=0} t + \psi^i(t), \tag{8.23}$$

essendo $\psi^i(t)$ un resto infinitesimo d'ordine superiore a t per $t \to 0$.

Definizione 1: *Due moti del tipo* (8.22) *si diranno tra loro* **tangenti** *in* $\overline{P}$ *se i corrispondenti sviluppi associati* (8.23) *hanno la stessa parte lineare in* t*, o, ciò che è lo stesso, ammettono gli stessi coefficienti* $a^i = \left(\frac{dx^i}{dt}\right)_{t=0}$.
Proprietà: *È facile riconoscere che la tangenza di due moti è un* **carattere intrinseco**, *cioè indipendente dalla scelta delle coordinate in* $U_{\overline{P}}$.
Infatti in un cambiamento di coordinate $x^i = x^i(x^{i'})$, ovvero $x^{i'} = x^{i'}(x^i)$, si ha, pensando le $x^{i'}$ funzioni di t per il tramite delle x^i,

$$\frac{dx^{i'}}{dt} = \sum_i \frac{\partial x^{i'}}{\partial x^i}\frac{dx^i}{dt} = \theta^{i'}_i \frac{dx^i}{dt},$$

avendo posto $\theta^{i'}_i = \frac{\partial x^{i'}}{\partial x^i}$. Si vede bene di qui che se due moti hanno le stesse $\left(\frac{dx^i}{dt}\right)_{t=0}$ essi vengono ad ammettere anche le stesse $\left(\frac{dx^{i'}}{dt}\right)_{t=0}$, in quanto le $\theta^{i'}_i$ sono calcolate in entrambi i casi, in $\overline{P}$.
È ovvio che la relazione di tangenza è una relazione di equivalenza.
Definizione 2: *Chiameremo* **elemento lineare** *della varietà in* $\overline{P}$ *l'ente astratto che ha i caratteri*

Figura 8.5: Spazio tangente $T_P(M)$ ad una varietà in un suo punto P.

comuni a tutti i "moti" tra loro tangenti passanti per $\overline{P}$.
È chiaro che un elemento lineare in $\overline{P}$ è caratterizzato, in un sistema di coordinate, dalle

quantità numeriche $a^i = \left(\frac{dx^i}{dt}\right)_{t=0}$. Il complesso degli elementi lineari in $\overline{P}$ definisce uno spazio vettoriale. Infatti, siano $(a^1, a^2, ..., a^n)$ e $(b^1, b^2, ..., b^n)$ le n-ple corrispondenti a due elementi lineari in $\overline{P}$. Si può chiamare **somma** dei due elementi lineari l'elemento lineare individuato dalla n-pla $(a^1 + b^1, a^2 + b^2, ..., a^n + b^n)$; **prodotto** di un elemento lineare $(a^1, a^2, ..., a^n)$ per un numero reale α, l'elemento lineare $(\alpha a^1, \alpha a^2, ..., \alpha a^n)$. Si riconosce agevolmente che tali operazioni si conservano in ogni cambiamento di coordinate, hanno cioè un significato intrinseco; e che inoltre esse soddisfano a tutti gli assiomi di uno spazio vettoriale. Dunque un elemento lineare in $\overline{P}$ può essere concepito come un "**vettore tangente**" alla varietà. L'insieme degli **elementi lineari tangenti (vettori tangenti)** in $\overline{P}$ alla varietà costituisce dunque uno spazio vettoriale che chiameremo **spazio vettoriale tangente** o semplicemente **spazio tangente** alla varietà nel punto $\overline{P}$ e lo indicheremo con $T_{\overline{P}}(M)$ (Fig.8.5).

8.8 Dimensione dello spazio tangente. Base naturale. Cenno allo spazio cotangente $T^*_P(M)$

Un elemento lineare (vettore tangente) può essere rappresentato da uno qualsiasi dei moti della classe che lo definisce.
Consideriamo allora il particolare moto definito dalle seguenti equazioni

$$x^1 = \overline{x}^1 + t, \;\; x^2 = \overline{x}^2, ..., x^n = \overline{x}^n, \tag{8.24}$$

moto che evidentemente si svolge lungo la linea coordinata $x^1 = var$. L'elemento lineare associato è, nel sistema di coordinate considerato, individuato dalla n-pla $(1, 0, ..., 0)$. Viene spesso indicato con $\frac{\partial}{\partial x^1}$, altre notazioni sono $\frac{\partial P}{\partial x^1} \equiv \partial_1 P$ o più sinteticamente $\mathbf{e}_1$. In modo analogo consideriamo l'elemento lineare $\partial_2 P \equiv (0, 1, 0, ..., 0)$ corrispondente al moto

$$x^1 = \overline{x}^1, \;\; x^2 = \overline{x}^2 + t, ..., x^n = \overline{x}^n, \tag{8.25}$$

lungo la linea $x^2 = var$. Così procedendo si ottengono n vettori linearmente indipendenti $\mathbf{e}_i \equiv \partial_i P$, per mezzo dei quali è possibile esprimere, mediante combinazione lineare, il generico vettore dello spazio tangente:

$$(a^1, ..., a^n) = a^1(1, 0, ..., 0) + ... + a^n(0, 0, ..., 0, 1). \tag{8.26}$$

resta così messo in evidenza che lo spazio tangente ha in ogni caso una **dimensione pari a quella della varietà**.
Fa comodo indicare con dP il generico vettore dello spazio tangente in $\overline{P}$ (notazione che ci consente di immaginare il vettore tangente come quello che va da $\overline{P}$ ad un punto P infinitamente vicino) e indicare in conformità con dx^i le sue componenti nella base $\partial_i P$:

$$dP = dx^i \partial_i P = dx^i \mathbf{e}_i. \tag{8.27}$$

Definizione 3: *La base $\{\mathbf{e}_i\}$che risulta costituita dai vettori tangenti alle linee coordinate uscenti da P, si suole chiamare la* **base naturale** *associata al sistema di coordinate considerato; e dx^i le* **componenti naturali** *del vettore dP.*
Un sistema di coordinate definito da una carta associa quindi una base naturale agli spazi tangenti nei singoli punti del dominio della carta.

Cenno allo spazio cotangente.
A partire dallo spazio vettoriale tangente $T_P(M)$ in un punto generico P di una data varietà

differenziabile è facile costruire **il duale** definito come l'insieme di tutte le applicazioni lineari da $T_P(M)$ in $\mathbb{R}$. Lo spazio vettoriale duale di un dato spazio vettoriale V si suole indicare con V^*, perciò il duale di $T_P(M)$ sarà $T_P^*(M)$: questo spazio si chiama **spazio cotangente** alla varietà in P ed ha la stessa dimensione del corrispondente spazio tangente da cui proviene. L'unione di tutti gli spazi tangenti ai vari punti della varietà si chiama il **fibrato tangente**, mentre l'unione di tutti gli spazi cotangenti si chiama **il fibrato cotangente**, essi costituiscono ancora delle varietà differenziabili ma di dimensione doppia $2n$ rispetto alle corrispondenti varietà.
Gli spazi tangente e cotangente in un punto di una data varietà sono gli ingredienti geometrici fondamentali per poter definire in maniera del tutto generale il concetto di **tensore**. Non possiamo percorrere un così lungo cammino, che ci porterebbe via molto tempo, per definire i tensori e studiare sia l'Algebra che l'Analisi tensoriale su di una varietà differenziabile; d'altra parte ciò si fa, in maniera approfondita, nei corsi di Geometria. Nei paragrafi che seguono noi ci accontenteremo di un approccio meno rigoroso ma più pratico che conduce ai concetti utili per formulare la Teoria della Relatività Generale che è il nostro obbiettivo principale.

8.9 Tensori controvarianti di rango 1

In maniera un po' vaga e concisa possiamo dire che un **tensore** *è un "oggetto" definito su una* **varietà differenziabile** *in termini delle sue proprietà di trasformazione per un cambiamento di coordinate.*
Cominceremo, per semplicità, con un **prototipo** e poi enunceremo la definizione generale. Consideriamo due punti P e Q infinitamente vicini su di una data varietà differenziabile M^n le cui coordinate siano rispettivamente x^i e $x^i + dx^i$ $(i = 1, 2, ..., n)$ in una assegnata carta locale. I due punti definiscono uno **spostamento infinitesimo** o **vettore infinitesimo** $\overrightarrow{PQ} \equiv dP$ appartenente allo spazio tangente $T_P(M^n)$ in P alla varietà. Il vettore non va considerato libero, ma legato al punto P (Fig.8.6).

Figura 8.6: Spostamento infinitesimo dP di un punto P su di una varietà.

Le componenti di questo vettore, rispetto alla base naturale $\mathbf{e}_i = \dfrac{\partial P}{\partial x^i} \equiv \partial_i P$ dello spazio tangente $T_P(M^n)$, associata al sistema di coordinate x^i, sono dx^i, cioè

$$dP = dx^i \mathbf{e}_i. \tag{8.28}$$

Le componenti dello stesso vettore in un'altra base $\mathbf{e}_{i'} = \dfrac{\partial P}{\partial x^{i'}} \equiv \partial_{i'} P$ ovvero in un altro sistema di coordinate, poniamo $x^{i'}$, sono $dx^{i'}$ che sono legate alle dx^i dalle (8.17), cioè

$$dx^{i'} = \frac{\partial x^{i'}}{\partial x^i} dx^i. \tag{8.29}$$

La matrice di trasformazione che compare in questa equazione deve essere calcolata nel punto P, e a rigore si dovrebbe scrivere

$$dx^{i'} = \left(\frac{\partial x^{i'}}{\partial x^i} \right)_P dx^i, \tag{8.30}$$

ma per non appesantire troppo le formule adotteremo la notazione più semplice (8.29). La matrice che figura nella (8.30) è perciò una matrice $n \times n$ di numeri reali, ovvero se consideriamo P come un punto generico della nostra varietà, essa può essere concepita come una matrice formata da n^2 funzioni delle variabili x^i. La formula (8.29) costituisce quindi una trasformazione lineare omogenea fra le due n-ple di differenziali (dx^i) e $(dx^{i'})$ cioè ci fornisce la legge con cui si trasformano i differenziali delle coordinate per un generico cambiamento di variabili. Diremo che la n-pla (dx^i) - con un leggero abuso di linguaggio - è un ente (o oggetto) chiamato **tensore controvariante di rango (o ordine) 1** o anche **vettore controvariante**: questo è il nostro **prototipo**. In generale possiamo dunque dare la seguente:

Def.: Un **tensore controvariante di rango 1 (o del 1° ordine)** detto anche **vettore controvariante** è un ente (o oggetto) geometrico caratterizzato, nel sistema di coordinate x^i, da un insieme di n funzioni delle x^i, *dette componenti del vettore*, che indichiamo con $V^i(x)$ associate al punto P della varietà, che si trasformano sotto un cambiamento di coordinate $x \to x'$ allo stesso modo dei differenziali (dx^i), cioè

$$V^{i'} = \left(\frac{\partial x^{i'}}{\partial x^i}\right)_P V^i. \tag{8.31}$$

dove la matrice di trasformazione è calcolata nel punto P.

La (8.30) è un caso particolare della (8.31) dove le componenti $V^i \equiv dx^i$ sono infinitesime. Un esempio di vettore con componenti finite è dato dal vettore tangente dx^i/du alla curva di equazioni parametriche $x^i = x^i(u)$, infatti per un cambiamento di coordinate si ha

$$\frac{dx^{i'}}{du} = \frac{\partial x^{i'}}{\partial x^i}\frac{dx^i}{du},$$

che è proprio una legge di trasformazione del tipo (8.31).

È quasi superfluo avvertire che **è necessario distinguere fra effettivi oggetti geometrici come il vettore tangente, rappresentato graficamente da una freccia, e la sua rappresentazione in un determinato sistema di coordinate**, come gli n numeri $(dx^i/du)_P$ nel sistema di coordinate x^i o i numeri, in generale diversi, $(dx^{i'}/du)_P$ nel sistema di coordinate $x^{i'}$.

Nota: *Osserviamo che, per alleggerire il linguaggio, spesso diremo semplicemente* **vettore controvariante** V^i *anziché dire: le* V^i *sono le componenti di un vettore controvariante. Insomma identificheremo l'oggetto geometrico* **vettore controvariante** *con la* $n-pla$ *delle sue componenti in un dato sistema di coordinate. Lo stesso faremo per i* **vettori covarianti** *e per i* **tensori** *di qualsiasi tipo.*

Per l'esatezza, un vettore o un tensore sono degli "oggetti intrinseci" per i quali non ha senso applicare i concetti di controvarianza o covarianza che invece sono riferiti esclusivamente al modo di variare delle loro componenti in un cambiamento di base ovvero in una trasformazione di coordinate. Ma la terminologia di cui sopra è ormai entrata nell'uso e sebbene sia imprecisa viene usata correntemente: l'importante è che se ne sia consapevoli e non ce se ne dimentichi.

Se ora moltiplichiamo membro a membro le (8.31) per $\dfrac{\partial x^k}{\partial x^{i'}}$ e sommiamo rispetto all'indice i', da 1 a n, otteniamo

$$\frac{\partial x^k}{\partial x^{i'}} V^{i'} = \frac{\partial x^k}{\partial x^{i'}}\frac{\partial x^{i'}}{\partial x^i} V^i = \frac{\partial x^k}{\partial x^i} V^i = \delta^k_i V^i = V^k. \tag{8.32}$$

Perciò risolvere l'equazione tensoriale (8.31) rispetto alle V^i conduce alla trasformazione inversa che fa passare dal sistema delle coordinate $x^{i'}$ a quello delle x^i, cioè

$$V^i = \frac{\partial x^i}{\partial x^{i'}} V^{i'}, \tag{8.33}$$

che altro non è che la (8.32) scritta cambiando l'indice k in i. *A questo proposito è bene avvertire che nella pratica del calcolo tensoriale è spesso conveniente ed utile cambiare frequentemente il nome degli*

indici rispettando ovviamente tutte le regole che abbiamo illustrato in precedenza sugli indici liberi e su quelli muti, soprattutto un indice muto non deve essere mai ripetuto più di due volte, una volta come apice ed una volta come pedice.
Inoltre osserveremo la convenzione che un vettore controvariante venga sempre indicato con una lettera avente un indice in alto, a meno che non si stabilisca esplicitamente il contrario. È comunque ovvio che non tutte le quantità indicate con un indice in alto sono dei vettori controvarianti, per esserlo devono soddisfare alla legge di trasformazione (8.31) o (8.33). Per esempio la n-pla delle x^i per una generica trasformazione di coordinate, non costituisce affatto un vettore controvariante. Questo significa, in particolare, che se noi scegliamo $V^i = x^i$, allora le nuove componenti $V^{i'}$ rispetto al sistema di coordinate $x^{i'}$ non soddisfano, in generale, le equazioni $V^{i'} = x^{i'}$.
Le coordinate x^i si comportano come le componenti di un vettore controvariante solo rispetto alle trasformazioni lineari omogenee del tipo

$$x^{i'} = A^{i'}_j x^j, \tag{8.34}$$

dove le $A^{i'}_j$ costituiscono un insieme di n^2 costanti, le quali non necessariamente sono le componenti di un tensore di rango 2 la cui definizione sarà data nei paragrafi seguenti. Nel caso di trasformazioni lineari si ha $A^{i'}_j = \dfrac{\partial x^{i'}}{\partial x^j}$, e la trasformazione può essere scritta nella forma

$$x^{i'} = \frac{\partial x^{i'}}{\partial x^j} x^j,$$

che dimostra che la n-pla delle coordinate x^i si comporta in modo tensoriale soddisfacendo ad una legge di trasformazione del tipo (8.31).

8.10 Tensori covarianti di rango 1

Analogamente a ciò che si è fatto nel paragrafo precedente cominciamo considerando, come si trasforma sotto un cambiamento di coordinate, **il gradiente di una funzione** che fungerà da **prototipo** per la definizione generale di **tensore covariante di rango 1 o del 1° ordine (o vettore covariante)**. Sia

$$\phi = \phi(x^i) \tag{8.35}$$

una funzione a valori reali definita sulla varietà, cioè tale che $\phi(P)$ associa un numero reale ad ogni punto P della varietà (*campo di scalari*). Supponiamo anche che ϕ sia continua e differenziabile, e che perciò esistano continue le derivate parziali $\partial\phi/\partial x^i$ che costituiscono per definizione il **gradiente di** ϕ: $grad\phi \stackrel{def}{=} (\partial\phi/\partial x^i)$. Effettuiamo ora un cambio di variabili $x^i \to x^{i'}$, per le (8.14) le x^i sono funzioni delle $x^{i'}$, perciò la (8.35) può essere riscritta così

$$\phi = \phi(x^i(x')).$$

Derivando ambo i membri della relazione precedente rispetto ad $x^{j'}$, per la regola di derivazione delle funzioni composte si ottiene

$$\frac{\partial\phi}{\partial x^{j'}} = \frac{\partial\phi}{\partial x^i}\frac{\partial x^i}{\partial x^{j'}}$$

e scambiando, nel 2° membro, l'ordine dei fattori otteniamo l'espressione equivalente

$$\frac{\partial\phi}{\partial x^{j'}} = \frac{\partial x^i}{\partial x^{j'}}\frac{\partial\phi}{\partial x^i}. \tag{8.36}$$

Questa equazione ci dice come si trasforma il gradiente di una funzione di più variabili ed è la legge di trasformazione **prototipo** che ci serve: osserviamo che essa contiene la matrice jacobiana inversa di quella che compare nella (8.31). Se, per semplificare la scrittura, poniamo $\partial\phi/\partial x^i = \phi_i$, la (8.36) si scriverà più semplicemente così

$$\phi_{j'} = \frac{\partial x^i}{\partial x^{j'}}\phi_i. \tag{8.37}$$

Diremo che **il gradiente** di una funzione di più variabili è un **tensore covariante di rango 1 o vettore covariante (detto anche covettore)**. Possiamo dunque dare la seguente definizione generale:
Def.: Dicesi **tensore covariante di rango 1 (o del 1°ordine)** detto anche **vettore covariante (o covettore)** un insieme di n quantità, che indichiamo con V_i nel sistema di coordinate x^i, associato al punto P della varietà, che si trasforma sotto un cambiamento di coordinate allo stesso modo del gradiente ϕ_i di una funzione, cioè

$$V_{i'} = \left(\frac{\partial x^j}{\partial x^{i'}}\right)_P V_j, \tag{8.38}$$

dove la matrice di trasformazione è calcolata nel punto P.
Se si vuole la formula inversa della (8.38) basta moltiplicare, primo e secondo membro, per la matrice inversa della $\frac{\partial x^j}{\partial x^{i'}}$ cioè per $\frac{\partial x^{i'}}{\partial x^k}$ e sommare rispetto all'indice i' da 1 ad n, otteniamo

$$\frac{\partial x^{i'}}{\partial x^k}V_{i'} = \frac{\partial x^j}{\partial x^{i'}}\frac{\partial x^{i'}}{\partial x^k}V_j = \frac{\partial x^j}{\partial x^k}V_j = \delta^j_k V_j = V_k,$$

che volendo, può essere riscritta cambiando l'indice k in i ed i' in j' così

$$V_i = \frac{\partial x^{j'}}{\partial x^i}V_{j'}. \tag{8.39}$$

Analogamente a quanto abbiamo fatto nel paragrafo precedente per i vettori controvarianti, anche qui osserveremo la convenzione che un vettore covariante venga sempre indicato con una lettera munita di un indice in basso, a meno che non si stabilisca esplicitamente il contrario. È comunque ovvio che non tutte le quantità indicate con un indice in basso sono dei vettori covarianti, per esserlo devono soddisfare alla legge di trasformazione (8.38) o (8.39).
Osservazione: Dimostriamo ora che non c'è distinzione tra vettori covarianti e controvarianti quando ci si limiti a trasformazioni lineari *ortogonali*

$$x^{i'} = a^{i'}_h x^h + b^{i'}, \tag{8.40}$$

dove $b^{i'}$ sono n costanti che non costituiscono necessariamente le componenti di un vettore controvariante ed $a^{i'}_h$ sono costanti, non formanti necessariamente un tensore misto del 2° ordine (che definiremo nel paragrafo 8.12) ma soddisfacenti alla *condizione di ortogonalità della matrice* $a^{i'}_h$

$$\sum_{i'} a^{i'}_r a^{i'}_s = \delta_{rs}. \tag{8.41}$$

Siccome le leggi di controvarianza (8.31) e di covarianza (8.38) fanno intervenire due matrici l'una l'inversa dell'altra si tratta di dimostrare che esse sono uguali, cioè: $\frac{\partial x^{i'}}{\partial x^j} = \frac{\partial x^j}{\partial x^{i'}}$. A tal fine facciamo le seguenti considerazioni. Se gli elementi della matrice inversa (detti *i reciproci* degli $a^{i'}_h$) si indicano con la notazione $a^h_{i'}$, avremo com'è noto

$$a^h_{i'} = \frac{cof(a^{i'}_h)^t}{det(a^{i'}_h)}.$$

Sicché sono verificate le seguenti uguaglianze

$$a^{i'}_h a^h_{j'} = \delta^{i'}_{j'}, \quad a^{i'}_h a^k_{i'} = \delta^k_h.$$

Moltiplichiamo ora membro a membro la (8.40) per $a^k_{i'}$ e sommiamo sugli indici i', avremo

$$a^k_{i'} x^{i'} = a^k_{i'} a^{i'}_h x^h + a^k_{i'} b^{i'} = \delta^k_h x^h + a^k_{i'} b^{i'} = x^k + a^k_{i'} b^{i'},$$

dalla quale si ricava

$$x^k = a^k_{i'} x^{i'} - a^k_{i'} b^{i'}, \tag{8.42}$$

che è la trasformazione inversa della (8.40). Se ora deriviamo la (8.40) rispetto ad x^k avremo

$$\frac{\partial x^{i'}}{\partial x^k} = a^{i'}_h \frac{\partial x^h}{\partial x^k} = a^{i'}_h \delta^h_k = a^{i'}_k, \tag{8.43}$$

mentre se deriviamo la (8.42) rispetto ad $x^{j'}$ otteniamo

$$\frac{\partial x^k}{\partial x^{j'}} = a^k_{i'} \frac{\partial x^{i'}}{\partial x^{j'}} = a^k_{i'} \delta^{i'}_{j'} = a^k_{j'}, \tag{8.44}$$

che sono ovviamente due matrici inverse l'una dell'altra.
Infine se moltiplichiamo membro a membro la condizione di ortogonalità (8.41) per $a^r_{j'}$ e sommiamo, non solo rispetto all'indice i', ma anche rispetto all'indice r avremo

$$\sum_{i'} a^r_{j'} a^{i'}_r a^{i'}_s = a^r_{j'} \delta_{rs},$$

cioè

$$\sum_{i'} \delta^{i'}_{j'} a^{i'}_s = a^s_{j'},$$

ovvero

$$a^{j'}_s = a^s_{j'}.$$

Quest'ultima uguaglianza, in base alle (8.43), (8.44) equivale alla

$$\frac{\partial x^{j'}}{\partial x^s} = \frac{\partial x^s}{\partial x^{j'}}, \tag{8.45}$$

il che dimostra che le equazioni (8.31) e (8.38) definiscono lo stesso ente.

8.11 Invarianti

Una funzione $\phi(x)$ delle n coordinate x^i è detta un **invariante** o uno **scalare** o **un tensore di rango zero** se rispetto alla trasformazione di coordinate $x^i \to x^{i'}$ essa non cangia valore e si ha

$$\phi(x) = \phi'(x'),$$

dove $\phi'(x')$ è l'espressione che assume $\phi(x)$ nel nuovo sistema di coordinate $x^{i'}$.
Per esempio, con le componenti A^i e B_i di un vettore controvariante e, rispettivamente di un vettore covariante possiamo formare la somma $A^i B_i$. Quando si passa alle nuove coordinate $x^{i'}$, questa somma si trasforma in $A^{i'} B_{i'}$. Ma

$$A^{i'} B_{i'} = \frac{\partial x^{i'}}{\partial x^j} A^j \frac{\partial x^k}{\partial x^{i'}} B_k = \frac{\partial x^{i'}}{\partial x^j} \frac{\partial x^k}{\partial x^{i'}} A^j B_k = \delta^k_j A^j B_k = A^k B_k,$$

cioè, cambiando nell'ultima sommatoria k in i:

$$A^{i'}B_{i'} = A^i B_i,$$

perciò $A^i B_i$ è un invariante per trasformazioni arbitrarie di coordinate. Un altro invariante è, per esempio, in un dato sistema di coordinate x^i, la delta di Kronecker δ^i_k nel caso che si pongano uguali l'indice superiore con quello inferiore $k = i$, o come anche si dice se si effettua **una contrazione** sugli indici i, k :

$$\delta^i_i = \delta^1_1 + \delta^2_2 + ... + \delta^n_n = n;$$

in un qualsiasi altro sistema di coordinate $x^{i'}$ si avrà analogamente

$$\delta^{i'}_{i'} = \delta^1_1 + \delta^2_2 + ... + \delta^n_n = n,$$

e perciò δ^i_i è un invariante non cambiando il suo valore per una trasformazione arbitraria di coordinate.

8.12 Tensori del secondo ordine e di ordine superiore

Generalizziamo ora le definizioni date nei numeri precedenti per ottenere tensori controvarianti o covarianti di rango o ordine 2 o, più in generale, di rango maggiore di 2.
Formiamo a tal fine le n^2 quantità $W^{ij} = U^i V^j$, dove U^i e V^j sono le componenti di due vettori controvarianti. Segue dalla (8.31) che le W^{ij} si trasformano così

$$W^{i'j'} = U^{i'}V^{j'} = \frac{\partial x^{i'}}{\partial x^h} U^h \frac{\partial x^{j'}}{\partial x^k} V^k = \frac{\partial x^{i'}}{\partial x^h}\frac{\partial x^{j'}}{\partial x^k} U^h V^k = \frac{\partial x^{i'}}{\partial x^h}\frac{\partial x^{j'}}{\partial x^k} W^{hk},$$

cioè

$$W^{i'j'} = \frac{\partial x^{i'}}{\partial x^h}\frac{\partial x^{j'}}{\partial x^k} W^{hk}. \tag{8.46}$$

Diremo che le n^2 quantità W^{ij}, che in un generico cambio di coordinate si trasformano secondo le (8.46), sono **le componenti di un tensore controvariante di rango (o ordine) 2**, o, per semplicità di linguaggio, che costituiscono esse stesse **un tensore controvariante di rango (o ordine) 2.**
Generalizzando ciò che si è detto più sopra, possiamo dunque dare la seguente:
Def.: Un **tensore controvariante di rango (o ordine) 2** è un insieme di n^2 quantità, che indichiamo, per esempio, con T^{ij} nel sistema di coordinate x^i, associate al punto P della varietà, che si trasformano, sotto un cambiamento di coordinate, allo stesso modo del prodotto $W^{ij} = U^i V^j$ di due vettori controvarianti, come nella formula (8.46), cioè nel modo seguente

$$T^{i'j'} = \frac{\partial x^{i'}}{\partial x^h}\frac{\partial x^{j'}}{\partial x^k} T^{hk}, \tag{8.47}$$

dove le $T^{i'j'}$ sono le componenti del tensore nel sistema di coordinate $x^{i'}$ e dove le matrici di trasformazione sono calcolate entrambe nel punto P.
Dovrebbe ormai essere chiaro come è possibile passare dalla formula di trasformazione (8.47) alla sua inversa

$$T^{ij} = \frac{\partial x^i}{\partial x^{h'}}\frac{\partial x^j}{\partial x^{k'}} T^{h'k'}, \tag{8.48}$$

nella quale intervengono le matrici inverse di quelle che compaiono nella (8.47).
In modo analogo possiamo dare la seguente:

Def.: Dicesi **tensore covariante di rango (o ordine) 2** un insieme di n^2 quantità, che indichiamo con T_{ij} nel sistema di coordinate x^i, associate al punto P della varietà, che si trasformano, sotto un cambiamento di coordinate, al modo seguente

$$T_{i'j'} = \frac{\partial x^h}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{j'}} T_{hk}, \tag{8.49}$$

dove le matrici di trasformazione sono calcolate entrambe nel punto P.
Oltre ai tensori controvarianti e covarianti si definiscono anche i tensori misti così:
Def.: Dicesi **tensore misto di rango 2 o del secondo ordine** un insieme di n^2 quantità, che indichiamo con T^i_j nel sistema di coordinate x^i, associate al punto P della varietà, che si trasformano, sotto un cambiamento di coordinate, al modo seguente

$$T^{i'}_{j'} = \frac{\partial x^{i'}}{\partial x^h} \frac{\partial x^k}{\partial x^{j'}} T^h_k, \tag{8.50}$$

dove le matrici di trasformazione sono, come al solito, calcolate entrambe nel punto P.
Osserviamo che, nel calcolo tensoriale, è universalmente adottata la convenzione che gli indici sui tensori sono soprascritti (*apici*) quando indicano controvarianza e sono sottoscritti (*pedici*) quando indicano covarianza. In particolare un tensore misto di rango 2, T^i_j si trasforma come un vettore controvariante rispetto all'indice i e come un vettore covariante rispetto all'indice j.
Tutto ciò che abbiamo detto in precedenza per i tensori covarianti, controvarianti e misti di rango 2 può essere facilmente generalizzato in modo ovvio a tensori di rango superiore. Per esempio un tensore misto di rango 3, una volta controvariante e due volte covariante, soddisfa la seguente equazione

$$T^{i'}_{j'k'} = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^j}{\partial x^{j'}} \frac{\partial x^k}{\partial x^{k'}} T^i_{jk}. \tag{8.51}$$

Se un tensore misto ha rango di controvarianza p e rango di covarianza q, si dice che è un tensore di **tipo (o valenza)** (p, q), ovvero che è p volte controvariante e q volte covariante.
Siamo ora in grado di dare una:
Prima defnizione generale di tensore (Legge generale di trasformazione di un tensore)
Dicesi **tensore (o campo tensoriale)** *di tipo* (p, q) *e rango (o ordine)* $r = p + q$ *un ente* **T** *definito in ogni sistema di coordinate* (x^i) *da un insieme di* n^r *numeri (o funzioni)*[7] $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}$, *dette componenti di* **T** *nel sistema* (x), *Che per un qualsiasi cambiamento delle coordinate* $x \to x'$ *si trasformano secondo la legge:*

$$T^{i'_1 \ldots\ldots i'_p}_{j'_1 \ldots\ldots j'_q} = \frac{\partial x^{i'_1}}{\partial x^{i_1}} \ldots\ldots\ldots \frac{\partial x^{i'_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial x^{j'_1}} \ldots\ldots\ldots \frac{\partial x^{j_q}}{\partial x^{j'_q}} T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}. \tag{8.52}$$

Uguaglianza di due tensori dello stesso tipo:
Due tensori $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}$ *e* $Q^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}$ *dello stesso tipo* (p, q) *diconsi* **uguali** *se, per un qualunque insieme fissato di indici* $(i_1 \ldots\ldots i_p, j_1 \ldots\ldots j_q)$ *le corrispondenti componenti di questi tensori coincidono in ogni sistema di coordinate:* $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q} = Q^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}$.
È sufficiente, fra l'altro, verificare questo fatto in un sistema di coordinate qualsiasi, poiché la coincidenza delle componenti in un altro sistema deriva dalla legge di trasformazione tensoriale.
Si osservi anche che *l'ordine in cui sono scritti gli indici è essenziale*; in generale, per esempio si ha $T^k_{ij} \neq T^k_{ji} \neq T^j_{ik}$.

[7] $n^r = n^{p+q}$ è uguale al numero delle disposizioni con ripetizione di n oggetti ad r ad r, cioè $D'_{n,r} = n^r$.

Osservazione importantissima:
Siamo giunti al punto in cui è possibile spiegare il motivo per cui i tensori hanno un'importanza fondamentale in Meccanica, in Relatività, in Geometria ecc., ed, in generale, in tutti i rami della Fisica Matematica!
Il motivo risiede nel fatto che sussistono le seguenti due proprietà:

1ª Proprietà: *Un'equazione tensoriale che vale in un sistema di coordinate, vale necessariamente in* **tutti** *i sistemi di coordinate.*
Infatti, limitandoci al caso di tensori del 2° ordine, supponiamo di trovare che in un prefissato sistema di coordinate due tensori, per esempio covarianti di rango 2, diciamo T_{ik} ed S_{ik}, siano uguali, cioè soddisfino all'equazione tensoriale

$$T_{ik} = S_{ik}, \tag{8.53}$$

la quale rappresenti, per esempio, una qualsiasi legge fisica o geometrica.
Moltiplichiamo entrambi i membri della precedente equazione per le matrici $\dfrac{\partial x^i}{\partial x^{j'}}$ e $\dfrac{\partial x^k}{\partial x^{h'}}$, e sommiamo rispetto agli indici ripetuti, otterremo

$$\frac{\partial x^i}{\partial x^{j'}}\frac{\partial x^k}{\partial x^{h'}}T_{ik} = \frac{\partial x^i}{\partial x^{j'}}\frac{\partial x^k}{\partial x^{h'}}S_{ik}. \tag{8.54}$$

Poiché T_{ik} ed S_{ik} sono entrambi tensori covarianti di rango 2, ne segue per la definizione di tensorialità vista in precedenza, che la precedente uguaglianza equivale all'altra

$$T_{j'h'} = S_{j'h'}, \tag{8.55}$$

ciò significa che l'equazione (8.53) valida nel sistema di coordinate x^i, vale in qualsiasi altro sistema di coordinate $x^{i'}$, c.d.d.
Se ora poniamo $R_{ik} = T_{ik} - S_{ik}$ la (8.53) equivale alla $R_{ik} = 0$ e la (8.55) diventa $R_{i'k'} = 0$, perciò il risultato precedente può essere così rienunciato:

2ª Proprietà: *Se un tensore è nullo in un dato sistema di coordinate, esso è necessariamente nullo in tutti i sistemi di coordinate.*
Quindi anche se introduciamo, per comodità, particolari sistemi di coordinate, nel trattare problemi specifici, se lavoriamo con equazioni tensoriali queste valgono poi in tutti i sistemi di coordinate. **Le equazioni tensoriali sono insomma indipendenti dalle coordinate e perciò particolarmente adatte a rappresentare leggi fisiche**. Ciò viene spesso espresso con una locuzione entrata ormai nell'uso, dicendo che *le leggi fisiche scritte in forma tensoriale sono* **covarianti** *nel passaggio da un sistema di coordinate ad un altro anche se sarebbe più corretto dire che esse sono* **invarianti**.
La spiegazione della validità dei due precedenti risultati fondamentali è molto semplice: essa risiede nel fatto che **le leggi di trasformazione dei tensori, viste in precedenza, sono trasformazioni lineari ed omogenee.**
Esercizio:
Facciamo ora vedere, come utile esercizio, che il delta di Kronecker se viene scritto con un indice in alto ed uno in basso, cioè δ^i_k, ha carattere tensoriale (rispetto ad un generico cambiamento di coordinate) e precisamente si tratta di un tensore misto del secondo ordine, infatti, tenuto conto che in un cambiamento arbitrario di coordinate $x^i \to x^{i'}$ le due matrici $\dfrac{\partial x^{i'}}{\partial x^h}$, $\dfrac{\partial x^k}{\partial x^{j'}}$ sono l'una l'inversa dell'altra, avremo

$$\delta^{i'}_{k'} = \frac{\partial x^{i'}}{\partial x^j}\frac{\partial x^j}{\partial x^{k'}} = \frac{\partial x^{i'}}{\partial x^j}\frac{\partial x^j}{\partial x^h}\frac{\partial x^h}{\partial x^{k'}} = \frac{\partial x^{i'}}{\partial x^j}\delta^j_h\frac{\partial x^h}{\partial x^{k'}} = \frac{\partial x^{i'}}{\partial x^j}\frac{\partial x^h}{\partial x^{k'}}\delta^j_h, \tag{8.56}$$

che dimostra il carattere tensoriale di δ^i_k.
Lasciamo al lettore l'utile e facile esercizio di dimostrare che δ_{ik} e δ^{ik} non sono tensori generali, cioè non si trasformano con legge tensoriale per un generico cambiamento di coordinate; essi possono avere carattere tensoriale, cioè comportarsi come tensori soltanto per particolari trasformazioni di coordinate, per es. nel caso di trasformazioni lineari o in particolare per trasformazioni ortogonali di coordinate (si può vedere che δ_{ik} è il tensore metrico di uno spazio euclideo).

8.13 Tensori simmetrici e tensori antisimmetrici

Un'importante proprietà di un tensore è la sua simmetria od antisimmetria rispetto allo scambio di alcuni suoi indici.

A) **Tensori di ordine 2**
Un tensore doppio, con indici entrambi di covarianza (o di controvarianza) si dice **simmetrico** se risulta

$$T_{ik} = T_{ki} \tag{8.57}$$

in una base, e quindi in tutte com'è immediato verificare:

$$T_{i'k'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{k'}} T_{ik} = \frac{\partial x^k}{\partial x^{k'}} \frac{\partial x^i}{\partial x^{i'}} T_{ki} = T_{k'i'};$$

si dice invece **antisimmetrico (o emisimmetrico)** se in una base, e quindi in tutte, si ha

$$T_{ik} = -T_{ki}. \tag{8.58}$$

È immediato constatare che un tensore simmetrico ha soltanto $n(n+1)/2$ componenti indipendenti (anziché n^2), mentre uno antisimmetrico ne ha $n(n-1)/2$. Infatti, per la simmetria, le componenti T_{ik} indipendenti sono ovviamente tante quante le combinazioni con ripetizione di n oggetti a 2 a 2, cioè $C'_{n,2} = \binom{n+2-1}{2} = \binom{n+1}{2} = n(n+1)/2$. Nel caso dell'antisimmetria, possiamo fare lo stesso ragionamento di prima, ma dal numero ottenuto dobbiamo togliere n che è il numero degli elementi nulli della diagonale principale della matrice quadrata formata dalle componenti $T_{ik} = -T_{ki}$ del tensore antisimmetrico: avremo perciò $\dfrac{n(n+1)}{2} - n = \dfrac{n(n-1)}{2}$.
Ora, poiché

$$\frac{n(n+1)}{2} + \frac{n(n-1)}{2} = n^2,$$

segue che ogni tensore doppio T_{ik} è univocamente decomponibile nella somma di un tensore simmetrico S_{ik} e di un tensore antisimmetrico A_{ik}:

$$T_{ik} = S_{ik} + A_{ik}. \tag{8.59}$$

La (8.59), tenuto conto della simmetria ed antisimmetria, possiamo riscriverla così:

$$T_{ik} = S_{ki} - A_{ki}, \tag{8.60}$$

sommando e sottraendo membro a membro la (8.59) con la (8.60) otteniamo

$$S_{ik} = \frac{1}{2}(T_{ik} + T_{ki}), \tag{8.61}$$

$$A_{ik} = \frac{1}{2}(T_{ik} - T_{ki}). \tag{8.62}$$

Il tensore simmetrico S_{ik} ed il tensore antisimmetrico A_{ik} sono usualmente indicati così:

$$S_{ik} = T_{(ik)}, \qquad A_{ik} = T_{[ik]},$$

e vengono chiamati rispettivamente *il simmetrizzato* e *l'antisimmetrizzato* del tensore T_{ik}, avremo dunque per definizione:

$$T_{(ik)} = \frac{1}{2}(T_{ik} + T_{ki}), \tag{8.63}$$

$$T_{[ik]} = \frac{1}{2}(T_{ik} - T_{ki}). \tag{8.64}$$

Analoghe definizioni possiamo darle anche per i tensori contravarianti del 2° ordine.

B) **Tensori di ordine maggiore di 2.**
Se lo scambio di due indici *qualunque* non modifica il valore delle componenti, il tensore si dice **simmetrico**; se tale scambio ne altera soltanto il segno, il tensore si dice **antisimmetrico**. Se la simmetria o l'antisimmetria riguarda solo due indici particolari, allora il tensore si dirà simmetrico (rispettivamente antisimmetrico) rispetto a quei due indici.
Nel caso dei tensori simmetrici e antisimmetrici di ordine maggiore di 2, per esempio del 3° ordine, T_{ijk}, non vale più una decomposizione nella somma e nella differenza di due termini analoga a quella rappresentata dalle (8.61), (8.62) perché il numero delle componenti indipendenti di un tensore simmetrico del 3° ordine è $C'_{n,3} = \binom{n+3-1}{3} = \binom{n+2}{3} = \dfrac{n(n+1)(n+2)}{6}$, mentre quello di un tensore antisimmetrico è $C_{n,3} = \binom{n}{3} = \dfrac{n(n-2)(n-1)}{6}$ e la loro somma non è n^3 ma $\dfrac{n(n^2+2)}{3} < n^3$.
Tuttavia possiamo effettuare su di un qualunque tensore un'operazione di simmetrizzazione ed una di antisimmetrizzazione analoghe alle (8.63), (8.64), che definiremo fra poco.
Osservazione: Si noti che il rango di un tensore antisimmetrico non nullo non può superare n. Per riconoscerlo basta osservare che, in un qualunque sistema di coordinate, ciascuna componente del tensore, se esso fosse di rango superiore ad n, dovrebbe contenere necessariamente qualche indice ripetuto, e perciò sarebbe nulla.

C) **Simmetrizzazione di tensori di ordine maggiore di 2**.
L'operazione si simmetrizzazione di un tensore del 2° ordine vista nel punto **A)** si può generalizzare a tensori di ordine maggiore. Per esempio, per tensori del 3° ordine T_{ijk} si può definire il suo **simmetrizzato** così:

$$T_{(ijk)} = \frac{1}{3!}\sum_{\sigma(ijk)} T_{ijk} = \frac{1}{3!}(T_{ijk} + T_{jki} + T_{kij} + T_{jik} + T_{ikj} + T_{kji}) \tag{8.65}$$

dove $\sum_{\sigma(ijk)}$ indica che la somma è estesa a tutte le permutazioni della fondamentale ijk. Se i tensori di partenza fossero già simmetrici, la simmetrizzazione rigenererebbe ovviamente i tensori medesimi (è a tale scopo che è stato premesso il fattore numerico).
La formula (8.65) si può generalizzare in modo ovvio al caso che il tensore sia di rango r qualsiasi $T_{i_1 i_2 \ldots\ldots i_r}$.

D) **Antisimmetrizzazione di tensori di ordine maggiore di 2**
Anche l'operazione di antisimmetrizzazione può essere generalizzata a tensori di ordine maggiore di 2. Per esempio, per tensori del 3° ordine T_{ijk} si può definire il suo

antisimmetrizzato così:

$$T_{[ijk]} = \frac{1}{3!} \sum_{\sigma(ijk)} (-1)^p T_{ijk} = \frac{1}{3!}(T_{ijk} + T_{jki} + T_{kij} - T_{jik} - T_{ikj} - T_{kji}), \tag{8.66}$$

dove la somma s'intende estesa a tutte le permutazioni della fondamentale ijk, e p indica la classe di ciascuna di esse.

8.14 Operazioni algebriche sui tensori

Oltre alle due operazioni di simmetrizzazione ed antisimmetrizzazione studiate nel paragrafo precedente, possiamo operare sui tensori con altre operazioni di carattere algebrico che ora vedremo.

1) **Addizione, Sottrazione e Moltiplicazione per uno scalare**
È facile stabilire che la **somma o la differenza** fra due tensori dello stesso tipo, è ancora un tensore dello stesso tipo degli addendi. Per esempio dai due tensori A^i_{jk} e B^i_{jk} di tipo $(1,3)$ possiamo constatare che anche la loro somma $S^i_{jk} = A^i_{jk} + B^i_{jk}$ e la loro differenza $D^i_{jk} = A^i_{jk} - B^i_{jk}$ sono tensori dello stesso tipo. Infatti applicando la legge tensoriale di trasformazione nel passaggio da un sistema di coordinate x ad un'altro generico x' si ha:

$$S^i_{jk} = A^i_{jk} + B^i_{jk} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial x^{k'}}{\partial x^k} A^{i'}_{j'k'} + \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial x^{k'}}{\partial x^k} B^{i'}_{j'k'} = \left(\frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial x^{k'}}{\partial x^k}\right)\left(A^{i'}_{j'k'} + B^{i'}_{j'k'}\right)$$

cioè

$$S^i_{jk} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial x^{k'}}{\partial x^k} S^{i'}_{j'k'},$$

S^i_{jk} è dunque un tensore. In maniera analoga si procede per la differenza di due tensori.
È facile verificare che moltiplicando un tensore $T^{i\ldots}_{j\ldots}$ di qualunque tipo per uno scalare α si ottiene ancora un tensore $\alpha T^{i\ldots}_{j\ldots}$ dello stesso tipo.
Ne segue che ogni combinazione lineare a coefficienti costanti di tensori di un certo tipo, è ancora un tensore dello stesso tipo. Per esempio dai due tensori A^i_{jk}, B^i_{jk} prima considerati possiamo formare il tensore $\lambda A^i_{jk} \pm \mu B^i_{jk}$, purché λ e μ siano invarianti.

2) **Moltiplicazione fra due tensori**
Dati due tensori misti qualsiasi, per esempio un tensore di tipo (p,q) ed uno di tipo (r,s), si può vedere facilmente che i prodotti delle loro componenti formano un tensore misto di tipo $(p+r, q+s)$, questo tensore è detto **prodotto esterno** dei due tensori dati.
Facciamo un semplice esempio: consideriamo due covettori, T_i e Q_i, e formiamo l'oggetto $P_{ik} = T_i Q_k$ costruito facendo il loro prodotto; risulta:

$$P_{i'k'} = T_{i'} Q_{k'} = \frac{\partial x^i}{\partial x^{i'}} T_i \frac{\partial x^k}{\partial x^{k'}} Q_k = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{k'}} T_i Q_k = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{k'}} P_{ik},$$

cioè P_{ik} è un tensore covariante di rango 2.
Non è difficile constatare che *il prodotto di tensori non è commutativo*; rifacendoci all'esempio precedente, poniamo $P_{ik} = T_i Q_k$ ed $S_{ik} = Q_i T_k$. Nel caso semplice in cui gli indici assumono i valori $i, k = 1, 2$, si avrà:

$$P_{11} = T_1 Q_1,\ P_{12} = T_1 Q_2,\ P_{21} = T_2 Q_1,\ P_{22} = T_2 Q_2,$$

$$S_{11} = Q_1 T_1,\ S_{12} = Q_1 T_2,\ S_{21} = Q_2 T_1,\ S_{22} = Q_2 T_2,$$

da cui si vede che

$$P_{11} = S_{11},\ P_{12} \neq S_{12},\ P_{21} \neq S_{21},\ P_{22} = S_{22},$$

da cui risulta manifesto che $P_{ik} \neq S_{ik}$, cioè $T_i Q_k \neq Q_i T_k$.

3) **Contrazione o saturazione di due indici**
Riferiamoci per esempio ad un tensore misto di rango 2, S_i^j; se si pone $i = j$ e si somma rispetto a questo indice si ottiene uno scalare (invariante). Infatti si ha

$$S_{i'}^{i'} = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{i'}} S_j^i = \delta_i^j S_j^i = S_i^i.$$

Si può analogamente dimostrare che la stessa operazione, se viene applicata a due indici di valenza opposta (cioè uno di covarianza ed uno di controvarianza) di un generico tensore di rango $n > 2$, dà luogo ad un tensore di rango $n - 2$. Per esempio T_{ijh}^h è un tensore doppio covariante.

Spesso accade di dover moltiplicare due o più tensori e poi di saturare due indici appartenenti a due diversi fattori. Si dice allora di aver effettuato una *moltiplicazione contratta*.

8.15 Criteri di tensorialità e seconda definizione di tensore

Capiterà di imbatterci talvolta in **"quantità"** (denominate talvolta **"oggetti" o "enti"**), tipicamente **insiemi di funzioni di più variabili**, caratterizzate, in ciascun sistema di coordinate, da certe componenti, senza sapere a priori se si tratta di tensori o no. Si potrebbe riconoscerlo, applicando la **prima definizione di tensore** che abbiamo dato in precedenza, verificando cioè come tali enti si trasformano in un cambiamento di coordinate. Ciò però, in generale, può risultare difficile o macchinoso. Perciò è conveniente e molto più sbrigativo servirsi dei cosiddetti **criteri di tensorialità** che risolvono il problema senza effettuare una trasformazione di coordinate e forniscono in definitiva una **seconda definizione di tensore** in quanto si tratta di **condizioni necessarie e sufficienti per la tensorialità**. In pratica, per verificare se un dato ente è un tensore o meno si applica quasi sempre un opportuno criterio di tensorialità **usato come condizione sufficiente**.
Cominciamo col considerare il caso più semplice:

Teorema 8.1 (Primo criterio di tensorialità) *C.N.E.S. affinché una n-pla di quantità ad un solo indice in basso $T_i (i = 1, 2, ..., n)$ siano le componenti di un tensore covariante del 1° ordine è che* **comunque si scelga un vettore controvariante** X^i *il prodotto contratto $T_i X^i$ risulta invariante per trasformazioni arbitrarie di coordinate $x \to x'$, cioè*

$$T_i X^i = T_{i'} X^{i'} = \textit{Invariante}. \tag{8.67}$$

Dimostrazione.
La condizione è necessaria (C.N.):
Cioè se T_i è un dato tensore covariante ed X^i è un vettore arbitrario allora necessariamente si avrà

$$T_i X^i = \text{Invariante}.$$

Infatti se operiamo una arbitraria trasformazione di coordinate $x \to x'$, tenuto conto che T_i ed X^i sono tensori si trasformano con legge tensoriale, perciò:

$$T_i X^i = \frac{\partial x^{j'}}{\partial x^i} T_{j'} \frac{\partial x^i}{\partial x^{h'}} X^{h'} = \frac{\partial x^{j'}}{\partial x^i}\frac{\partial x^i}{\partial x^{h'}} T_{j'} X^{h'} = \delta_{h'}^{j'} T_{j'} X^{h'} = T_{h'} X^{h'} = \text{Invariante}.$$

c.d.d.
La condizione è sufficiente (C.S.):
Invero, dall'ipotesi che $T_i X^i$ è invariante per trasformazioni arbitrarie di coordinate $x \to x'$, si ha:

$$T_i X^i = T_{i'} X^{i'}$$

con X^i vettore arbitrario. Essendo quest'ultimo un vettore controvariante, per la legge tensoriale, la precedente relazione può riscriversi così:

$$T_i X^i = T_{i'} \frac{\partial x^{i'}}{\partial x^i} X^i$$

e portando tutti i termini al 1° membro avremo:

$$\left(T_i - T_{i'} \frac{\partial x^{i'}}{\partial x^i} \right) X^i = 0.$$

Quest'ultima deve essere soddisfatta per una scelta arbitraria del vettore X^i, ciò comporta l'annullarsi dell'espressione in parentesi. Infatti se poniamo $T_i - T_{i'} \dfrac{\partial x^{i'}}{\partial x^i} = R^i$, avremo:

$$R_i X^i = 0, \Longleftrightarrow R_1 X^1 + R_2 X^2 + ... + R_n X^n = 0,$$

con le X^i arbitrarie, ma allora se scegliamo:

$$X^1 \neq 0, X^2 = 0, X^2 = 0, ..., X^n = 0,$$

si ottiene $R_1 X^1 = 0$ con $X^1 \neq 0$, il che implica $R^1 = 0$.
Se scegliamo:

$$X^1 = 0, X^2 \neq 0, X^2 = 0, ..., X^n = 0,$$

si ottiene $R_2 X^2 = 0$ con $X^2 \neq 0$, il che implica $R_2 = 0$.
Continuando con analoghe scelte rimane dimostrato che:

$$R_i = 0, \quad (i = 1, 2, ..., n)$$

il che equivale a scrivere:

$$T_i = \frac{\partial x^{i'}}{\partial x^i} T_{i'}.$$

Ciò dimostra che la n-pla delle T_i si trasforma con legge tensoriale, si tratta perciò delle componenti di un tensore covariante del 1° ordine. c.d.d.
Ossevazione: Analogo procedimento si segue per dimostrare che una n-pla di quantità con un indice in alto T^i è un tensore controvariante di rango 1 se essendo X_i un arbitrario covettore, il prodotto contratto $T^i X_i$ è invariante per trasformazioni arbitrarie di coordinate.
Dimostriamo ora un secondo criterio di tensorialità per enti a due indici. Per semplificarne la dimostrazione premettiamo il seguente lemma algebrico, tanto semplice quanto importante, che generalizza il ragionamento che abbiamo svolto nella dimostrazione della condizione sufficiente del teorema precedente.

Lemma 8.2 *Se qualunque sia l'ente a* 2 *indici* R_{ik} *risulta*

$$R_{ik} X^i Y^k = 0, \quad \forall X^i, Y^k, \quad i, k = 1, 2, \ldots, n,$$

allora $R_{ik} = 0$.

Dimostrazione. Deve chiaramente risultare

$$(R_{ik}X^i)Y^k = 0, \quad \forall Y^k,$$

ovvero

$$(R_{i1}X^i)Y^1 + (R_{i2}X^i)Y^2 + \ldots + (R_{in}X^i)Y^n = 0, \quad \forall Y^k, \quad k = 1, 2, ...n.$$

Essendo le Y arbitrarie, possiamo scegliere dapprima:

$$Y^1 \neq 0, Y^2 = Y^3 = ... = Y^n = 0,$$

ciò implica

$$(R_{i1}X^i)Y^1 = 0, \quad \text{con} \quad Y^1 \neq 0$$

perciò sarà

$$R_{i1}X^i = 0.$$

Con ragionamento analogo, scegliendo

$$Y^1 = 0, Y^2 \neq 0, Y^3 = ... = Y^n = 0,$$

si otterrà

$$R_{i2}X^i = 0,$$

e così via, fino ad ottenere

$$R_{in}X^i = 0.$$

Ciascuna delle precedenti relazioni deve essere soddisfatta per valori arbitrari delle X^i, perciò con un ragionamento analogo al precedente si avrà:

$$R_{i1} = 0, \quad R_{i2} = 0, \quad \ldots, \quad R_{in} = 0, \quad i = 1, 2, \ldots, n,$$

ovvero

$$R_{ik} = 0, \quad i, k = 1, 2, \ldots, n$$

cioè la tesi.

Teorema 8.3 (Secondo criterio di tensorialità) *C.N.E.S. affinché le n^2 quantità a due indici in basso T_{ij} siano le componenti di un tensore doppio covariante è che* **comunque si scelgano** *due vettori controvarianti X^i ed Y^j il prodotto completamente contratto $T_{ij}X^iY^j$ sia un invariante per trasformazioni arbitrarie di coordinate $x \to x'$, cioè*

$$T_{ij}X^iY^j = T_{i'j'}X^{i'}Y^{j'} = \textit{Invariante}. \tag{8.68}$$

Dimostrazione.

La condizione è necessaria (C.N.):
Cioè se T_{ij} è un dato tensore doppio covariante ed X^i, Y^j sono due vettori arbitrari allora necessariamente si avrà:

$$T_{ij}X^iY^j = T_{i'j'}X^{i'}Y^{j'} = \text{Invariante}.$$

Infatti se operiamo una arbitraria trasformazione di coordinate $x \to x'$, tenuto conto che T_{ij}, X^i, Y^j sono tensori si trasformano con legge tensoriale, perciò:

$$T_{ij}X^iY^j = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{j'}}{\partial x^j}T_{i'j'}\frac{\partial x^i}{\partial x^{h'}}X^{h'}\frac{\partial x^j}{\partial x^{k'}}Y^{k'} = \delta^{i'}_{h'}\delta^{j'}_{k'}T_{i'j'}X^{h'}Y^{k'} = T_{i'j'}X^{i'}Y^{j'} = \text{Invariante}.$$

c.d.d.
La condizione è sufficiente (C.S.):
Invero, dall'ipotesi che $T_{ij}X^iY^j$ è invariante per trasformazioni arbitrarie di coordinate $x \to x'$, si ha

$$T_{ij}X^iY^j = T_{i'j'}X^{i'}Y^{j'},$$

con X^i, Y^j vettori controvarianti arbitrari. Per la legge tensoriale possiamo scrivere

$$T_{ij}X^iY^j = T_{i'j'}\frac{\partial x^{i'}}{\partial x^i}X^i\frac{\partial x^{j'}}{\partial x^j}Y^j = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{j'}}{\partial x^j}T_{i'j'}X^iY^j \tag{8.69}$$

e portando tutti i termini dell'ultima espressione al 1° membro avremo:

$$\Big(T_{ij} - \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{j'}}{\partial x^j}T_{i'j'}\Big)X^iY^j = 0. \tag{8.70}$$

Indichiamo con R_{ij} l'espressione che figura dentro parentesi nella (8.70) avremo allora

$$R_{ij}X^iY^j = 0$$

con X^i e Y^j arbitrari. Possiamo dunque applicare il Lemma 8.2 dimostrato prima ottenendo

$$T_{ij} = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{j'}}{\partial x^j}T_{i'j'}. \tag{8.71}$$

La (8.71) dimostra che T_{ij} è un tensore covariante di rango 2 come si voleva dimostrare.
Osservazione: Qualora per ottenere un invariante fosse necessario saturare i due indici delle componenti dell'ente assegnato con quelle di due arbitrari vettori covarianti, si tratterebbe di un tensore doppio controvariante; mentre se i due indici saturati fossero l'uno covariante, l'altro controvariante, l'ente in questione sarebbe un tensore misto.
Il precedente criterio si semplifica quando l'ente a due indici T_{ik} di cui si vuol riconoscere il carattere tensoriale è simmetrico, ossia quando per le sue componenti, comunque si scelga il sistema di coordinate, sussistono le uguaglianze $T_{ik} = T_{ki}$. Prima di vederlo, dimostriamo ancora un altro lemma.

Lemma 8.4 *Se per l'ente* **simmetrico** R_{ik} *($R_{ik} = R_{ki}$) risulta*

$$R_{ik}X^iX^k = 0, \quad \forall X^i, \quad i,k = 1,2,\dots,n,$$

allora $R_{ik} = 0$.

Dimostrazione.
In questo caso non si può concludere immediatamente, ripetendo il ragionamento del Lemma precedente, che data l'arbitrarietà delle X^i, si ha $R_{ik} = 0$ in quanto la doppia sommatoria sugli indici i,k **non involge quantità indipendenti**, perché le X^i e le X^k rappresentano la stessa n-pla di variabili e ò non possono essere indipendenti.
Scomponiamo la doppia sommatoria nel modo seguente:

$$R_{ik}X^iX^k = \sum_i R_{ii}(X^i)^2 + \sum_{i\neq k} R_{ik}X^iX^k = \sum_i R_{ii}(X^i)^2 + \sum_{i<k}(R_{ik}+R_{ki})X^iX^k,$$

dove nell'ultima sommatoria si è tenuto conto della simmetria di R_{ik}.
Possiamo quindi lavorare sulla relazione

$$\sum_i R_{ii}(X^i)^2 + \sum_{i<k}(R_{ik}+R_{ki})X^iX^k = 0, \quad \forall X^i.$$

Poiché le X^i sono del tutto arbitrarie, possiamo scegliere le n-uple in cui una componente è diversa da zero e tutte le altre nulle:

$$X^j \neq 0, \quad X^k = 0 \text{ per } k \neq j,$$

per $j = 1, 2, \ldots, n$, che, sostituite una alla volta nella relazione precedente, permettono di concludere che deve essere

$$R_{ii} = 0, \quad i = 1, 2, \ldots, n.$$

La relazione di partenza si riduce quindi alla

$$\sum_{i<k} (R_{ik} + R_{ki}) X^i X^k = 0, \quad \forall X^i.$$

Come prima, le X^i sono del tutto arbitrarie, per cui possiamo scegliere le n-uple in cui due componenti sono diverse da zero e tutte le altre nulle:

$$X^j \neq 0, \quad X^h \neq 0, \quad X^k = 0 \text{ per } k \neq j \text{ e } k \neq h,$$

per $j, h = 1, 2, \ldots, n$, che, sostituite una alla volta nella relazione precedente, permettono di concludere che deve essere

$$R_{ik} + R_{ki} = 0, \quad i \neq k, \quad i, k = 1, 2, \ldots, n.$$

Poiché il caso degli R_{ii} rientra in questo caso, come caso particolare, possiamo concludere che deve essere

$$R_{ik} = -R_{ki}, \quad i, k = 1, 2, \ldots, n.$$

Ma per ipotesi le R_{ik} sono simmetriche, quindi l'unica possibilità è che risulti

$$R_{ik} = 0, \quad i, k = 1, 2, \ldots, n,$$

ovvero la tesi.
Siamo ora in grado di dimostrare il seguente:

Teorema 8.5 (Criterio di tensorialità per sistemi doppi simmetrici)
Un sistema di n^2 quantità **simmetriche** *T_{ik} è un tensore doppio covariante se, comunque si scelga un vettore controvariante X^i, il prodotto completamente contratto $T_{ik} X^i X^k$ è uno scalare intrinseco.*

Dimostrazione.
Se la quantità $T_{ik} X^i X^k$ è uno scalare intrinseco, allora, considerato un qualsiasi cambiamento di coordinate $x \to x'$, si avrà

$$T_{ik} X^i X^k = T_{i'k'} X^{i'} X^{k'}.$$

Essendo X^i un vettore controvariante esso si trasformerà conformemente alla $X^{i'} = \dfrac{\partial x^{i'}}{\partial x^i} X^i$, la precedente relazione si trasforma nella

$$T_{ik} X^i X^k = T_{i'k'} \frac{\partial x^{i'}}{\partial x^i} X^i \frac{\partial x^{k'}}{\partial x^k} X^k,$$

cioè

$$\left(T_{ik} - T_{i'k'} \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{k'}}{\partial x^k} \right) X^i X^k, \quad \forall X^i, \quad i, k = 1, 2, \ldots, n,$$

Se ora poniamo $R_{ik} = T_{ik} - T_{i'k'} \dfrac{\partial x^{i'}}{\partial x^i} \dfrac{\partial x^{k'}}{\partial x^k}$ la precedente relazione diventa

$$R_{ik} X^i X^k = 0, \quad \forall X^i, \quad i, k = 1, 2, \ldots, n,$$

ed applicando il lemma precedente avremo la tesi

$$T_{ik} = T_{i'k'} \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{k'}}{\partial x^k}.$$

Osservazione:
Facciamo notare che il precedente Criterio di tensorialità per enti doppi simmetrici lo abbiamo enunciato come una **condizione sufficiente** perché è sotto questa forma che si usa in pratica, ma è ovvio che esso si può enunciare come una C.N.E.S.
È evidente come il criterio di tensorialità 8.5 possa estendersi a tensori doppi controvarianti e misti; basta infatti considerare moltiplicazioni completamente contratte, nel primo caso con due covettori, e nel secondo caso con un vettore ed un covettore.
È immediata la generalizzazione dei Criteri precedenti al caso di tensori arbitrari. Sussiste infatti il seguente:

Teorema 8.6 *C.N.E.S. affinché un sistema di n^{p+q} componenti a p indici superiori e q indici inferiori sia un tensore di tipo (p, q) e rango $p + q$ è che il suo prodotto completamente contratto per p covettori arbitrari e q vettori arbitrari sia uno scalare intrinseco.*

Un'ulteriore generalizzazione porta ai seguenti altri **Criteri sufficienti di tensorialità** le cui dimostrazioni lasciamo come esercizio allo studente:

Criterio speciale di tensorialità:
Supponiamo che $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}$ sia un ente a più indici che possa essere rappresentato in diversi sistemi di coordinate e sia X^{j_q} un vettore controvariante arbitrario, se il prodotto contratto

$$S^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_{q-1}} = T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q} X^{j_q}$$

è un **tensore**, *allora anche $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}$ è un* **tensore**.

Criterio generale di tensorialità:
Supponiamo che $T^{i_1 \ldots\ldots i_{k-1}\ i_k \ldots\ldots i_p}_{j_1 \ldots\ldots j_{l-1}\ j_l \ldots\ldots j_q}$ sia un ente a più indici, e $A^{j_l \ldots\ldots j_q}_{i_k \ldots\ldots i_p}$ un tensore arbitrario, se il prodotto contratto

$$S^{i_1 \ldots\ldots i_{k-1}}_{j_1 \ldots\ldots j_{l-1}} = T^{i_1 \ldots\ldots i_{k-1}\ i_k \ldots\ldots i_p}_{j_1 \ldots\ldots j_{l-1}\ j_l \ldots\ldots j_q} A^{j_l \ldots\ldots j_q}_{i_k \ldots\ldots i_p}$$

è un **tensore**, *allora anche $T^{i_1 \ldots\ldots i_{k-1}\ i_k \ldots\ldots i_p}_{j_1 \ldots\ldots j_{l-1}\ j_l \ldots\ldots j_q}$ è un* **tensore**.
Possiamo concludere questa sezione sintetizzando ciò che abbiamo detto con le due definizioni generali seguenti, la prima delle quali già enunciata nel paragrafo 8.12 viene qui ripetuta per comodità del lettore:

1ª Definizione di tensore (per mezzo di un cambiamento di coordinate)
In ogni punto di una varietà differenziabile M^n (o in una sua regione) sia definito un campo di n^{p+q} funzioni $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}(x)$, espresse in un dato sistema di coordinate (x^i).
Queste funzioni si dice che sono le **componenti** *di un* **tensore misto** *(o meglio di un* **campo tensoriale***), p* **volte controvariante e** *q* **volte covariante**, *ovvero di tipo (p, q) e di* **rango o ordine** *$r = p + q$, se in un generico cambiamento invertibile di coordinate $x \to x'$ sulla varietà, esse si trasformano secondo la legge*:

$$T^{i'_1 \ldots\ldots i'_p}_{j'_1 \ldots\ldots j'_q}(x') = \frac{\partial x^{i'_1}}{\partial x^{i_1}} \ldots\ldots \frac{\partial x^{i'_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial x^{j'_1}} \ldots\ldots\ldots \frac{\partial x^{j_q}}{\partial x^{j'_q}} T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}(x). \tag{8.72}$$

2ª Definizione di tensore (per mezzo dell'invarianza di una forma multilineare)
c.n.s affinché un sistema di n^{p+q} funzioni $T^{i_1 \ldots\ldots i_p}_{j_1 \ldots\ldots j_q}(x)$, a p indici superiori e q indici inferiori, definite su

di una varietà differenziabile M^n, sia un **tensore misto** p **volte controvariante e** q **volte covariante**, *ovvero di tipo* (p, q) *e di* **rango o ordine** $r = p + q$, *è che* **comunque si scelgano** p **vettori covarianti** $Y_{i_1}, Y_{i_2}, ..., Y_{i_p}$ *e* q **vettori controvarianti** $X^{j_1}, X^{j_2}, ..., X^{j_q}$ *il prodotto completamente contratto*

$$T^{i_1.....i_p}_{j_1.....j_q} Y_{i_1} Y_{i_2} ..... Y_{i_p} X^{j_1} X^{j_2} ..... X^{j_q} = Invariante \tag{8.73}$$

sia uno **scalare intrinseco**, *sia cioè un* **invariante** *per trasformazioni arbitrarie di coordinate.*

8.16 Varietà riemanniane

Il concetto di varietà riemanniana è fondamentale nell'ambito della geometria differenziale e nelle sue applicazioni soprattutto alla meccanica ed alla fisica matematica. Esempi notevoli di varietà riemanniane sono: le superfici dello spazio ordinario, lo spazio delle configurazioni in meccanica lagrangiana, l'insieme delle rotazioni di un corpo rigido con un punto fisso, lo spaziotempo di Minkowski in Relatività Ristretta. Soprattutto Einstein trovò le varietà riemanniane un mezzo eccellente per formulare la sua teoria della Relatività Generale, come vedremo in seguito dettagliatamente nel Cap.11.
Consideriamo una varietà differenziabile M^n di dimensione finita n. Una volta associato a ciascun punto di M^n lo spazio vettoriale tangente $T_P(M^n)$, è chiaro che si potranno considerare **campi vettoriali definiti sulla varietà**. Per averne uno basterà associare a ciascun punto di M^n un vettore dello spazio tangente. In modo analogo si potranno considerare **campi tensoriali definiti su** M^n, perché ha senso parlare di prodotto tensoriale dello spazio tangente per sé stesso e per il suo duale $T^*_P(M^n)$ (spazio cotangente), ecc. In particolare, si può pensare assegnato, in ogni punto della varietà, un **tensore simmetrico del 2° ordine** $g_{ik}(x)$, atto a definire, in ciascuno spazio tangente, il prodotto scalare fra due vettori, la norma di un vettore, ecc. È chiaro che il tensore $g_{ik}(x)$ definisce un campo tensoriale che è funzione del punto $P \in M^n$.

Definizione: *Una varietà differenziabile che sia munita di un tale tensore* $g_{ik}(x)$ detto — **tensore metrico** o **tensore fondamentale** — *non degenere (ovvero tale che* $g = \det(g_{ik}) \neq 0$*) prende il nome di* **varietà riemanniana**. La forma quadratica avente per coefficienti le $g_{ik}(x)$

$$\mathbf{ds}^2 = g_{ik}(x) dx^i dx^k, \tag{8.74}$$

viene chiamata **prima forma quadratica fondamentale** o anche, frequentemente, **metrica della varietà**.
Essa definisce la norma $\mathbf{ds}^2$ del generico vettore tangente $\overrightarrow{dP}$, cioè $\|\overrightarrow{dP}\| \equiv \overrightarrow{dP} \cdot \overrightarrow{dP} = \mathbf{ds}^2$, o come suol dirsi, *"il quadrato della distanza"* fra due punti infinitamente vicini della varietà. Tale locuzione, presa in prestito dalla geometria ordinaria, è giustificata nel caso che la forma quadratica (8.74) sia definita positiva. Invece nel caso, che pur spesso si presenta, specialmente in Relatività Ristretta (RR) e Relatività Generale (RG), che la (8.74) non sia definita di segno, la locuzione è impropria, essa è comunque nell'uso. Sono di uso corrente anche le seguenti locuzioni: se la forma quadratica $\mathbf{ds}^2$ è definita positiva, la varietà si dice **propriamente riemanniana o riemanniana** tout court, se invece essa è indefinita, la varietà si dice **pseudoriemanniana**.
Osservazione importante:
La forma quadratica (8.74) espressa nelle coordinate x^i deve considerarsi **invariante** per trasformazioni generali di coordinate $x^i \to x^{i'}$, cioè deve essere verificata l'uguaglianza

$$\mathbf{ds}^2 = g_{ik}(x) dx^i dx^k = g_{i'k'}(x') dx^{i'} dx^{k'} = Invariante. \tag{8.75}$$

È proprio in virtù dell'invarianza di tale forma quadratica, tenuto conto che dx^i è un vettore infinitesimo arbitrario, che discende (in virtù del criterio di tensorialità per sistemi doppi

simmetrici) il carattere tensoriale di g_{ik}.
A titolo di esempio, data una superficie σ dello spazio ordinario, siano

$$y^i = y^i(x^1, x^2) \qquad (i = 1, 2, 3) \tag{8.76}$$

equazioni parametriche di una sua porzione regolare, dove le y^i $(i = 1, 2, 3)$ sono coordinate cartesiane, nello spazio 3-dimensionale dove è immersa la superficie, e le x^α $(\alpha = 1, 2)$ coordinate curvilinee sulla superficie. Consideriamo, sulla superficie, due punti P e P', infinitamente vicini, individuati rispettivamente dai valori x^α e $x^\alpha + dx^\alpha$ $(\alpha = 1, 2)$ dei parametri x^1, x^2. Il quadrato della loro distanza è dato, in forma cartesiana, da $\mathbf{ds}^2 = \sum_i (dy^i)^2$, ovvero, essendo $dy^i = \dfrac{\partial y^i}{\partial x^\alpha} dx^\alpha$, da

$$\mathbf{ds}^2 = \sum_{i=1}^{3} \sum_{\alpha,\beta=1}^{2} \frac{\partial y^i}{\partial x^\alpha} \frac{\partial y^i}{\partial x^\beta} dx^\alpha \, dx^\beta. \tag{8.77}$$

Quest'ultima relazione è equivalente alla

$$\mathbf{ds}^2 = g_{\alpha\beta} dx^\alpha \, dx^\beta, \tag{8.78}$$

con

$$g_{\alpha\beta} = \sum_{i=1}^{3} \frac{\partial y^i}{\partial x^\alpha} \frac{\partial y^i}{\partial x^\beta}.$$

Si tratta evidentemente di una forma quadratica definita positiva.
Quanto precede giustifica la seguente definizione di *lunghezza di un arco di curva* γ tracciato sulla varietà M^n. Nell'ipotesi che lungo di essa il $\mathbf{ds}^2$ sia sempre $\neq 0$. Precisamente siano

$$x^i = x^i(\lambda), \qquad \lambda_1 \leq \lambda \leq \lambda_2, \tag{8.79}$$

equazioni parametriche di tale curva (a $\mathbf{ds}^2$ di segno costante, ad esempio positivo). Generalizzando un concetto familiare nello spazio ordinario, chiamiamo *lunghezza di* γ l'integrale

$$\ell = \int_{s_1}^{s_2} ds = \int_{\lambda_1}^{\lambda_2} \frac{ds}{d\lambda} d\lambda,$$

cioè:

$$\ell = \int_{\lambda_1}^{\lambda_2} \sqrt{g_{ik} \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda}} d\lambda. \tag{8.80}$$

8.17 Varietà euclidee e varietà globalmente euclidee

Definizione 1: *Una varietà riemanniana ad n dimensioni M^n si dice* **euclidea** *quando nell'intorno di ogni suo punto essa ammette coordinate che rendono costanti le componenti g_{ik} del tensore metrico.*
Tali coordinate speciali si dicono *cartesiane*.
In una varietà euclidea esistono evidentemente infiniti sistemi di coordinate cartesiane, e si può passare dall'uno all'altro mediante una trasformazione lineare. Particolarizzando opportunamente tale trasformazione è poi sempre possibile trovare coordinate cartesiane - *dette coordinate cartesiane ortogonali* - che diano al $\mathbf{ds}^2$ forma pitagorica (o pseudopitagorica), cioè:

$$\mathbf{ds}^2 = \sum_{i=1}^{n} \epsilon_i (dx^i)^2, \tag{8.81}$$

con $\epsilon_i = \pm 1$.

Definizione 2: *Una varietà euclidea M^n si dice* **globalmente euclidea** *se può essere posta in corrispondenza biunivoca e bicontinua con uno spazio euclideo E^n.*
Essa allora non si distingue da uno spazio euclideo nel senso or ora richiamato.
Per mettere in evidenza la differenza tra una varietà euclidea e una varietà globalmente euclidea (spazio euclideo) si pensi ad un cilindro circolare indefinito nello spazio ordinario. Il fatto che esso, tagliato lungo una generatrice, si può sviluppare su di un piano, assicura il suo carattere di varietà euclidea (a 2 dimensioni). Non si tratta però evidentemente di un 2-spazio euclideo (cioè di un piano euclideo) non potendosi i suoi punti porre in corrispondenza biunivoca e bicontinua con quelli di un piano, a causa della sua diversa struttura topologica.

8.18 Il tensore coniugato del tensore metrico

Dal *tensore metrico* di una varietà riemanniana M^n, dato in forma **covariante** g_{ik}, possiamo ottenere un altro tensore, ancora del 2° ordine, ma assegnato in forma **controvariante**.
A tal fine, consideriamo il determinante

$$g = det(g_{ik}(x)) = det \begin{pmatrix} g_{11} & g_{12} & \cdots\cdots & g_{1n} \\ g_{21} & g_{22} & \cdots\cdots & g_{2n} \\ & \cdots\cdots\cdots & & \\ g_{n1} & g_{n2} & \cdots\cdots & g_{nn} \end{pmatrix}, \tag{8.82}$$

e supponiamo, qui e nel seguito, che sia $g \neq 0$.
Indichiamo con $\Delta^{ik} = cofg_{ik}$, *il cofattore o complemento algebrico* di g_{ik} in questo determinante; applicando il 1° ed il 2° *teorema di Laplace sui determinanti*, avremo:

$$g_{ih}\Delta^{ik} = g_{hi}\Delta^{ki} = \delta^k_h g, \tag{8.83}$$

dove $g_{ih}\Delta^{ik}$ è lo sviluppo per colonne e $g_{hi}\Delta^{ki}$ è lo sviluppo per righe del determinante.
Definiamo ora il cosiddetto **elemento reciproco** di g_{ik} nel determinante $g = det(g_{ik})$, che indicheremo con g^{ik}, ponendo quindi, per definizione,

$$g^{ik} \stackrel{\text{def}}{=} \frac{\Delta^{ik}}{g} = \frac{cofg_{ik}}{g}. \tag{8.84}$$

Utilizzando la (8.84), ne segue, dalla (8.83), in cui si consideri l'uguaglianza fra il 1° ed il 3° membro, che:

$$g_{ih}g^{ik} = \delta^k_h, \tag{8.85}$$

e similmente, dall'uguaglianza fra il 2° ed il 3° membro della (8.83), tenuto conto della (8.84), si avrà:

$$g_{hi}g^{ki} = \delta^k_h. \tag{8.86}$$

Poiché g_{ik} è un *tensore simmetrico*, anche g^{ik}, per definizione, è un sistema simmetrico; che si tratti di un tensore si può accertare col seguente ragionamento. L'uguaglianza (8.85) afferma che il prodotto contratto del tensore del 2° ordine g_{ih} per il sistema a due indici g^{ik} (di cui non si conosce, per il momento, il carattere tensoriale), è uguale al tensore misto di Kronecker δ^k_h, dunque si tratta, per un noto criterio di tensorialità, di un tensore anch'esso simmetrico: lo chiameremo *tensore metrico in forma controvariante.*
Osserviamo che sia la (8.85) sia la (8.86) possono riguardarsi come definizione del tensore controvariante g^{ik} poiché entrambi questi due sistemi di equazioni lineari determinano le g^{ik} univocamente. Infatti, se moltiplichiamo ambo i membri della (8.85) per Δ^{lh} e sommiamo rispetto ad h avremo

$$g_{ih}g^{ik}\Delta^{lh} = \delta^k_h\Delta^{lh},$$

che, tenuto conto della seconda uguaglianza che figura nella (8.83), equivale alla

$$\delta_i^l g g^{ik} = \Delta^{kl}.$$

Per la simmetria sugli indici l, k si ottiene infine:

$$g^{lh} = \frac{\Delta^{lh}}{g}, \tag{8.87}$$

che è la stessa della (8.84), e ciò prova la nostra asserzione.
Il tensore controvariante g^{ik}, talvolta è chiamato **il tensore coniugato** del tensore fondamentale g_{ik}.

8.19 Abbassamento ed innalzamento degli indici mediante il tensore metrico

Avendo ora a disposizione il tensore covariante fondamentale g_{ik} ed il suo coniugato controvariante g^{ik}, siamo in grado di introdurre il cosiddetto procedimento di **abbassamento ed innalzamento** degli indici di un tensore.
Una stessa grandezza fisica vettoriale può essere rappresentata sia in componenti controvarianti sia in componenti covarianti utilizzando proprio il tensore metrico nell'una o nell'altra forma. Tale legame, per un vettore, è dato dalle formule

$$A_i \stackrel{\text{def}}{=} g_{ik} A^k, \qquad A^i \stackrel{\text{def}}{=} g^{ik} A_k, \tag{8.88}$$

dove nella prima è stato abbassato l'indice in alto di A^k (che è diventato i), sommando su k ed usando il tensore metrico g_{ik} in forma covariante, mentre nella seconda è stato innalzato l'indice k di A_k usando il tensore metrico in forma controvariante.
Ciò che abbiamo detto per i vettori vale anche per tensori di rango arbitrario. Il passaggio tra le diverse forme di un tensore, rappresentante una grandezza fisica, viene effettuato mediante l'uso del tensore metrico, da utilizzarsi una o più volte a seconda del numero di indici da abbassare od alzare, in base a formule del tipo

$$A^i{}_k \stackrel{\text{def}}{=} g^{il} A_{lk}, \qquad A^{ik} \stackrel{\text{def}}{=} g^{il} g^{km} A_{lm}, \tag{8.89}$$

e analogamente per tensori di ogni rango.
Nello stesso ordine di idee, possiamo scrivere il prodotto scalare di due vettori A, B, definito come prodotto contratto, usando sia le componenti covarianti sia quelle controvarianti dei dati vettori, cioè

$$A_i B^i = A_i g^{ik} B_k = A^k B_k. \tag{8.90}$$

Infine, è utile ricordare che innalzando entrambi gli indici del tensore metrico covariante g_{ik} si ottiene il tensore metrico in forma controvariante

$$g^{ih} g^{jk} g_{hk} = g^{ij}, \tag{8.91}$$

analogamente abbassando entrambi gli indici di g^{ij} si riottiene g_{hk}

$$g_{hi} g_{kj} g^{ij} = g_{hk}. \tag{8.92}$$

Capitolo 9

PARALLELISMO DI LEVI-CIVITA

A distanza di anni il concetto di parallelismo di Levi-Civita introdotto durante le sue ricerche sulla relatività e sulle varietà riemanniane appare sempre di più come una di quelle idee centrali, ricche di generalizzazioni e di utili conseguenze, che hanno aperto nuovi fruttuosi indirizzi di ricerca sia in fisica sia in matematica. .... Nella sua forma originale il trasporto parallelo fu introdotto per unificare e semplificare **il calcolo differenziale assoluto**, *con il risultato addizionale di dare ad esso una interpretazione geometrica intuitiva di grande immediatezza.*
(Dalla Postfazione, pag.183, al libro di Tullio Levi-Civita, *Fondamenti di Meccanica Relativistica*, Zanichelli Editore, Bologna, 1982, scritta da Tullio Regge in occasione della ristampa di questo classico.)

Si racconta che Einstein in visita in Italia, interrogato su quali fossero le cose che gli piacessero di più del nostro paese, abbia risposto: gli spaghetti e Levi-Civita.
(Aneddoto)

9.1 Parallelismo di Levi-Civita

In questo paragrafo e nel successivo introdurremo, prima su di una superficie dello spazio euclideo 3-dimensionale ordinario e poi su di una varietà riemanniana, un concetto fondamentale, quello di **parallelismo**, dovuto al grande matematico italiano Tullio Levi-Civita[1].
Tale concetto non solo è di per sé notevole dal punto di vista geometrico, ma è importantissimo anche dal punto di vista analitico perché è intimamente connesso col fondamentale concetto di **derivata covariante.**
A questo proposito riportiamo un brano illuminante di Einstein:[2]
"L'apporto importantissimo di Levi-Civita mi sembra consista nella seguente scoperta teorica: il risultato essenziale della teoria della relatività generale, e precisamente il superamento dello spazio rigido,

[1]Il concetto di parallelismo fu introdotto da Levi-Civita nel lavoro fondamentale: *Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura Riemanniana*, Rendiconti del Circolo Matematico di Palermo, fascicolo XLII, 1917, pagg. 173-215, riprodotto in Tullio Levi-Civita, *Opere Matematiche, Memorie e Note*, vol. 4, Nota I, pagg. 1-39, Nicola Zanichelli Editore, Bologna 1960.
In ciò che segue ci siamo attenuti alla bellissima trattazione del parallelismo superficiale che Levi-Civita ha svolto nel suo notevole libro: *Lezioni di Calcolo Differenziale Assoluto*, Alberto Stock Editore, Roma (1925), Cap. V. A nostro modesto avviso, riteniamo, che questa trattazione, per la limpida e chiara visione geometrica che la ispira, sia, ancora oggi, insuperata.

[2]Dalla Prefazione di Albert Einstein al libro *Cinquant'anni di Relatività* pubblicato da Mario Pantaleo, Editrice Universitaria, Firenze, 1955, pagg. XVIII-XIX.

ossia del sistema inerziale, è legato solo indirettamente con l'introduzione di una metrica riemanniana. L'immediato essenziale elemento concettuale è il **campo degli spostamenti** Γ^l_{ik} *il quale permette di esprimere lo spostamento infinitesimale di un vettore. Questo precisamente sostituisce il parallelismo spaziale di vettori qualsiasi separati (cioè l'eguaglianza delle corrispondenti componenti) per mezzo di una operazione infinitesimale. Con ciò viene resa possibile* **la derivazione tensoriale [derivazione covariante]**, *e così resa superflua l'introduzione dello spazio "rigido" (del sistema inerziale). Di fronte a questo risultato appare in un certo senso di importanza secondaria che un particolare campo* Γ *risulti deducibile dall'esistenza di una metrica riemanniana.* [3].
Come prerequisito ad una buona comprensione di ciò che segue, si richiede, al lettore, la conoscenza dei fondamenti della *Teoria delle superfici* che viene studiata nei corsi di geometria.[4]

9.1.1 Osservazione fondamentale di Gauss circa la geometria intrinseca di una superficie

Lo scopo delle considerazioni che seguono è quello di fare *un'osservazione fondamentale* che metterà in luce, nello studio di una superficie (Fig.9.1), tutta l'importanza che ha la forma quadratica 2-dimensionale

$$\mathbf{ds}^2 = g_{ik}(x)dx^i dx^k, \qquad g_{ik} = g_{ki}, \quad (i,k = 1,2) \tag{9.1}$$

dove le $g_{ik}(x) \equiv g_{ik}(x^1, x^2)$ devono considerarsi come funzioni assegnate delle coordinate gaussiane x^1, x^2. Tale forma possiamo scriverla più esplicitamente così

$$\mathbf{ds}^2 = g_{11}(x^1, x^2)(dx^1)^2 + 2g_{12}(x^1, x^2)dx^1 dx^2 + g_{22}(x^1, x^2)(dx^2)^2. \tag{9.2}$$

Figura 9.1: Nozione intuitiva di superficie.

Per formulare la suddetta *osservazione fondamentale* ci baseremo su alcune considerazioni intuitive dovute a Karl Friedrich Gauss,che poggiano sul concetto di **geometria intrinseca di una superficie.**

[3]Quello che Einstein chiama il campo Γ (ovvero l'insieme delle funzioni Γ^l_{ik} (campo degli spostamenti), oggi è universalmente noto come **la connessione di Levi-Civita** della quale ci occuperemo dettagliatamente nel Cap. 10, paragrafo 10.7.2

[4]Un'ottima trattazione della teoria delle superfici, fatta nello spirito del calcolo tensoriale, si può trovare nel libro di Erwin Kreyszig, *Differential Geometry,* Dover Publications, Inc., New York, 1991. Un'altra buona trattazione si trova anche nel libro di James J. Stoker, *Differential Geometry,* Wiley Interscience, Inc., New York, 1989. Infine, un'ampio trattato, in due volumi, di circa 1000 pagine, moderno, ottimo sotto tutti i punti di vista, con un'amplissima raccolta di bellissime figure a colori, ottenute con Mathematica, ed una bibliografia pressoché completa è quello di Renzo Caddeo e Alfred Gray, *Lezioni di Geometria Differenziale su curve e superfici,* C.U.E.C. (*CooperativaUniversitaria Editrice Cagliaritana*), vol. I, 2001, vol. II, 2002.

A tal fine si materializzi il concetto di superficie, pensando ad un foglio di materia flessibile ed inestensibile, su cui si possano disegnare delle figure: si possa deformarlo, curvandolo e piegandolo in infiniti modi, *ma non mai lacerarlo, né dilatarlo.* Le figure disegnate su tale superficie verranno ad assumere, in conseguenza delle deformazioni di quella, diverse configurazioni spaziali: tuttavia alcune loro proprietà rimarranno invariate. Per esempio, se due linee s'intersecano, esse conservano questa proprietà comunque si deformi il foglio; la lunghezza di una porzione di linea resta la stessa, e quindi la distanza fra due punti, *contata sulla superficie* (cioè lungo la più breve linea superficiale che li congiunge) non varia, così pure non varia l'angolo di due linee, che si incrocino in un punto, ecc. Brevemente possiamo enunciare la seguente
Osservazione fondamentale di Gauss:
Tutte quelle proprietà nelle quali non intervengono enti estranei alla superficie (o, come suol dirsi, quelle che si possono studiare senza uscire da questa) sono indipendenti dalle deformazioni della superficie, e costituiscono la sua **geometria intrinseca**. *Essa dipende esclusivamente dai coefficienti $g_{ik}(x)$ della cosiddetta* 1^a **forma quadratica fondamentale ds**$^2 = g_{ik}(x)dx^i dx^k$ *della superficie.*
Anche nella geometria elementare abbiamo esempi di tal genere: la geometria piana si costruisce tutta senza fare uso di punti esterni al piano, ed è quindi *intrinseca* di tal superficie: essa vale (almeno in regioni convenientemente limitate) anche se il piano si piega, o si avvolge su un cono o su un cilindro.
Ora si osservi che gli elementi fondamentali per lo studio delle proprietà metriche di una figura sono: la distanza di due punti infinitamente vicini e l'angolo di due direzioni. Infatti, la lunghezza di una linea qualunque si ottiene per integrazione da quella dei suoi elementi infinitesimi, l'area di una figura si può calcolare decomponendola in parallelogrammi elementari, ecc.

9.1.2 Superfici sviluppabili

Definizione 1:
Si chiamano **sviluppabili** *quelle superfici che, supposte flessibili ed inestensibili, possono essere portate a coincidere con una regione di piano, senza lacerazioni né duplicature.* Esempi di superfici sviluppabili sono il cilindro circolare retto, il cono circolare, e qualunque superficie formata da più pezzi di piano. Nel caso del cilindro e del cono lo sviluppo si ottiene immaginando di tagliare queste superfici lungo una loro generatrice g e poi distenderle su di un piano senza lacerazioni né duplicature (Fig.9.2).

La geometria intrinseca di tali superfici è identica a quella del piano e l'elemento lineare di esse può assumere le stesse forme di quello del piano, per esempio può assegnarsi un sistema di coordinate superficiali x^1, x^2, tale che $ds^2 = (dx^1)^2 + (dx^2)^2$.
Si consideri una semplice infinità (∞^1) di piani di $\mathbb{R}^3$ (che possono pensarsi rappresentati da un'equazione lineare nelle coordinate cartesiane x, y, z i cui coefficienti siano funzioni continue di un parametro u):

$$a(u)x + b(u)y + c(u)z + d(u) = 0. \tag{9.3}$$

Possiamo perciò dire che:
*L'***inviluppo** *di questa famiglia di piani è una superficie sviluppabile, di cui i piani rappresentati dall'equazione* (9.3) *sono i piani tangenti.*
Questa proposizione può rendersi intuitiva con le seguenti considerazioni infinitesimali. Siano $\omega_1, \omega_2, \omega_3, \ldots$ alcuni piani della famiglia, corrispondenti a successivi incrementi infinitesimi del parametro u, diciamo g_1 l'intersezione di ω_1 con ω_2, g_2 l'intersezione di ω_2 con ω_3 e così via (Fig. 9.3).
Definizione 2:
Il luogo geometrico di tutte queste rette è la **superficie inviluppo***; le rette $g_1, g_2, \ldots$ vengono chiamate le sue* **caratteristiche o generatrici**. *Ognuno dei piani ω contiene due generatrici, racchiudenti tra loro*

Figura 9.2: Esempi elementari di superfici sviluppabili: cono e cilindro.

Figura 9.3: Superfici sviluppabili come inviluppo di una famiglia di piani.

un angolo infinitesimo; l'inviluppo può pensarsi costituito da queste infinite regioni piane infinitesime (Fig.9.3).

È allora chiaro che lo sviluppo su di un piano di una superficie sviluppabile può essere realizzato mediante successive rotazioni infinitesime intorno alle sue generatrici $g_1, g_2, \ldots$**.**

Consideriamo ora l'inviluppo di una particolare famiglia (ad un parametro) di piani, e precisamente di tutti i piani tangenti ad una superficie qualunque S, lungo una sua curva prefissata γ.

Definizione 3:
L'inviluppo di questi piani è una superficie sviluppabile S_γ*, che viene chiamata la* **sviluppabile circoscritta ad** S **lungo** γ (Fig.9.4).
Poiché i piani $T_P(S)$ tangenti ad S lungo γ sono anche tangenti a S_γ ($T_P(S) \equiv T_P(S_\gamma)$) ne risulta

Figura 9.4: Sviluppabile circoscritta.

che **la sviluppabile circoscritta è tangente ad S lungo la curva γ.**
Esempi:

1) se la superficie S è una sfera e la curva γ è un suo circolo massimo, la sviluppabile circoscritta S_γ è il cilindro circolare retto indefinito, avente lo stesso diametro della sfera e tangente ad essa lungo γ, le cui generatrici sono le rette tangenti ad S lungo γ e perpendicolari al piano contenente il circolo γ (vedi Fig.9.5);

2) se la superficie S è ancora una sfera e la curva γ è invece un suo circolo minore, la sviluppabile circoscritta S_γ è il cono circolare retto tangente alla sfera S lungo γ (vedi Fig.9.5).

Figura 9.5: Sviluppabili circoscritte ad una sfera lungo un circolo massimo o lungo un circolo minore.

9.1.3 Parallelismo superficiale: Definizione geometrica

Nella geometria piana euclidea è ben noto come, fissati due punti P e P_1, ad ogni direzione spiccata da P ne corrisponda una, ed una sola, spiccata da P_1, **parallela** alla prima (per il V° postulato di Euclide)(vedi Fig. 9.6).

Ci proponiamo ora di estendere questa nozione dalla geometria intrinseca del piano a quella di una superficie qualunque S. A tal fine consideriamo un punto P di S, il relativo piano tangente $T_P(S) \equiv \omega$ e una direzione generica, spiccata da P tangenzialmente ad S, e quindi contenuta in ω: penseremo individuata la direzione mediante il corrispondente versore (vettore

Figura 9.6: Rette e vettori paralleli in senso euclideo.

unitario) $\mathbf{u}$, e in conformità diremo semplicemente direzione $\mathbf{u}$, anziché direzione il cui versore è $\mathbf{u}$. Siano ancora P_1 un altro punto qualsiasi di S, e $T_{P_1}(S) \equiv \omega_1$ il piano tangente in P_1 (Fig. 9.7).
Distingueremo il caso di una superficie sviluppabile da quello di una superficie non sviluppabile.

Definizione di parallelismo per superfici sviluppabili:
Se la superficie S è **sviluppabile**, *si può ovviamente stabilire fra le direzioni uscenti tangenzialmente da P e le analoghe uscenti da P_1 una corrispondenza, che diremo* **per parallelismo**, *chiamandosi* **parallela ad u in senso superficiale** *quella* $\mathbf{u}_1$ *che diviene parallela ad* $\mathbf{u}$ *in senso ordinario, quando si fa lo sviluppo di S sopra un piano.* (Fig. 9.7).

Figura 9.7: Parallelismo superficiale per superfici sviluppabili.

È chiaro che, in generale, le due direzioni, viste nello spazio ordinario in cui è immersa la superficie S, non risulteranno parallele in senso ordinario.
Quando si fa lo sviluppo di S sopra un piano, questo lo chiameremo **piano di sviluppo** e lo indicheremo con π^*, i punti P e P_1 di S si trasformano rispettivamente nei punti P^* e P_1^* di π^* ed analogamente i vettori $\mathbf{u}$ e $\mathbf{u}_1$ vengono trasformati nei vettori $\mathbf{u}^*$ e $\mathbf{u}_1^*$ paralleli in senso ordinario.
Un tale criterio non è più applicabile nel caso di una superficie S **non sviluppabile.** (anche del tipo più elementare come, ad es. la sfera), ed è naturale di cercarne una generalizzazione adeguata. Vi si è condotti nel modo più spontaneo purché agli elementi posizionali già presi in considerazione (i quali sono sufficienti senza ulteriore specificazione per le superfici sviluppabili) si aggiunga una legge, *a priori arbitraria*, di concatenamento, raffigurandosi P_1 come proveniente da P attraverso una determinata curva γ di S, detta **curva di trasporto** (vedi Fig. 9.8).

Con riferimento a questa curva γ, si può allora definire il trasporto per parallelismo da P a P_1 nel modo seguente. Si consideri **la sviluppabile circoscritta** ad S lungo γ: questa superficie,

Figura 9.8: Parallelismo superficiale per superfici non sviluppabili.

che indicheremo con S_γ, è, come sappiamo tangente ad S lungo γ, e, in particolare, in P e P_1, dove le direzioni tangenziali ad S, spiccate da questi due punti, sono anche tangenziali ad S_γ.Criterio speciale di tensorialità **Definizione di parallelismo per superfici non sviluppabili:** *Chiameremo parallelismo superficiale su S lungo una sua linea γ, quello subordinato dalla* **sviluppabile circoscritta** *S_γ, convenendo di chiamare parallela in P_1 ad una direzione generica superficiale* $\mathbf{u}$ *in P lungo la linea γ, la direzione superficiale* $\mathbf{u}_1$ *che, sulla sviluppabile S_γ, è parallela a* $\mathbf{u}$ *nel senso prima definito.*[5]

9.1.4 Prime conseguenze. Equipollenza superficiale di vettori

Conseguenza necessaria della precedente definizione è che, a differenza di quanto accade per le sviluppabili, la direzione $\mathbf{u}_1$, superficialmente parallela ad $\mathbf{u}$ in P, *non è univocamente determinata dai soli dati P,* $\mathbf{u}$, *P_1, ma, in generale, dipende anche dalla curva di trasporto*[6].
Dimostriamo ora alcune notevoli proprietà del trasporto per parallelismo.

- **Proprietà 1:** *Nel trasporto per parallelismo si conservano gli angoli.*
  Cioè se $\mathbf{a}$, $\mathbf{b}$ sono due generiche direzioni superficiali in P formanti un angolo θ, le loro parallele superficiali in P_1, $\mathbf{a}_1$, $\mathbf{b}_1$, formano il medesimo angolo θ. Ciò è del tutto evidente se il trasporto per parallelismo avviene su di una superficie sviluppabile. Infatti, il vettore di partenza $\mathbf{a}$, ed il vettore trasportato per parallelismo su S, $\mathbf{a}_1$ diventano, dopo lo sviluppo, paralleli in senso ordinario e lo stesso succede per $\mathbf{b}$ e $\mathbf{b}_1$, inoltre lo sviluppo non altera gli angoli (Fig. 9.9).

  Nel caso di una superficie non sviluppabile il trasporto per parallelismo viene a dipendere dalla curva γ di trasporto. In tal caso è necessario considerare la sviluppabile circoscritta S_γ che è tangente ad S in ogni punto di γ e perciò sia i vettori di partenza $\mathbf{a}$, $\mathbf{b}$ che quelli trasportati $\mathbf{a}_1$, $\mathbf{b}_1$, oltre che essere tangenti ad S sono altresì tangenti ad S_γ. Perciò è ancora valido il ragionamento precedente facendo ora lo sviluppo su di un piano, non di S, ma della sua sviluppabile circoscritta S_γ lungo γ.

Nelle considerazioni che abbiamo fatto finora ci siamo riferiti unicamente alle direzioni, considerando i corrispondenti versori. È chiaro che la stessa costruzione con cui si passa da $\mathbf{u}$ ad $\mathbf{u}_1$ può essere applicata ad un vettore tangenziale $\mathbf{v}$ di lunghezza qualsiasi v (non unitaria). Se $\mathbf{u}$

[5]Un tracciamento. per così dire automatico, delle direzioni parallele si può conseguire assai semplicemente mediante rotolamento della superficie S sopra un piano. Cfr. Enrico Persico, *Realizzazione cinematica del parallelismo superficiale*, Rend. della R. Acc. dei Lincei, vol. XXX (2° semestre 1921, pagg. 127-128.

[6]Osserviamo che *la nozione geometrica di parallelismo* si può ravvicinare a *quella fisica di lavoro*, in cui interviene l'integrale di un'espressione della forma $X_1dx^1 + X_2dx^2$, (dove x^1, x^2 sono coordinate curvilinee generali su S), integrale che dipende, in generale dalla linea γ di integrazione: solo nel caso particolare che $X_1dx^1 + X_2dx^2$ sia un differenziale esatto, questa dipendenza non ha più luogo.

Figura 9.9: Nel trasporto per parallelismo si conservano gli angoli.

è il corrispondente versore, si ha $\mathbf{v} = v\mathbf{u}$, e nel trasporto per parallelismo si ottiene un vettore $\mathbf{v}_1 = v\mathbf{u}_1$, cioè un vettore applicato in P_1, avente la stessa lunghezza di $\mathbf{v}$ e per direzione quella della parallela superficiale $\mathbf{u}_1$. Diremo, com'è naturale, che $\mathbf{v}$ e $\mathbf{v}_1$ sono *vettori superficialmente equipollenti*, con riferimento al cammino γ. Perciò definiamo **equipollenti due vettori tangenziali quando sono paralleli ed hanno la stessa lunghezza.** Dunque, per definizione, valgono le seguenti proprietà:

- **Proprietà 2:** *Nel trasporto per parallelismo si conservano le lunghezze.*
- **Proprietà 3:** *Nel trasporto per parallelismo si conservano i prodotti scalari fra vettori.*

Un cenno particolare, nei riguardi del parallelismo, merita il caso in cui **la curva di trasporto sia una geodetica di** S, che indicheremo con $\mathcal{G}$.[7]

- **Proprietà 4:** *Una geodetica di S è geodetica anche della sviluppabile circoscritta $S_{\mathcal{G}}$ e viceversa.* Per convicersene, basta pensare che S ed $S_{\mathcal{G}}$ hanno, per definizione, gli stessi piani tangenti lungo $\mathcal{G}$, e quindi le rette perpendicolari ad una delle superfici, nei punti di $\mathcal{G}$, sono pure perpendicolari all'altra. Ne segue che, se le normali principali nei punti di $\mathcal{G}$ (considerata come curva spaziale) coincidono con le normali ad S, cioè se, per definizione, la $\mathcal{G}$ è geodetica di S allora la stessa proprietà vale per $S_{\mathcal{G}}$, dunque tali curve $\mathcal{G}$ sono anche geodetiche della sviluppabile circoscritta, e viceversa.

Figura 9.10: Parallelismo e geodetiche.

Nello sviluppo su di un piano, la geodetica $\mathcal{G}$ diviene una retta (come risulta immediatamente dalla sua proprietà caratteristica di segnare il minimo cammino fra due suoi punti

[7]Ricordiamo che *chiamasi* **geodetica** *ogni linea di S che goda della proprietà che in ogni suo punto* **la normale principale** *n alla linea coincida con* **la normale** *ν* **alla superficie**, *o, equivalentemente, che abbia, in ogni suo punto, il* **piano osculatore** *perpendicolare al* **piano tangente** *ad S. Le linee che segnano il più breve percorso (sulla superficie) fra due punti dati, godono sempre di tale proprietà*: d'altra parte, anche la proposizione reciproca - sotto certe restrizioni - è vera. Perciò, per definire le geodetiche si adotta talvolta l'uno, talvolta l'altro criterio.

qualunque) e le direzioni $\mathbf{u}$ e $\mathbf{u}_1$, che, per effetto dello sviluppo, divengono parallele nel piano, formeranno con tale retta angoli uguali (Fig. 9.10). Poiché lo sviluppo non altera gli angoli, possiamo enunciare anche la seguente:

- **Proprietà 5:** *Direzioni parallele su S lungo una geodetica $\mathcal{G}$, formano con questa geodetica angoli uguali.*

In particolare, se la $\mathbf{u}$ coincide con la direzione di $\mathcal{G}$ in P, anche la $\mathbf{u}_1$ coinciderà con la direzione di $\mathcal{G}$ in P_1, cioè le direzioni di una geodetica nei vari suoi punti, vale a dire le direzioni delle rette tangenti alla geodetica nei suoi punti, sono tutte parallele lungo la geodetica stessa, cioè vale la:

- **Proprietà 6:** *Le direzioni di una geodetica nei vari suoi punti, sono tutte parallele, cioè* **le geodetiche sono curve autoparallele.** (Fig. 9.11).

Figura 9.11: Le geodetiche sono curve autoparallele.

Da queste considerazioni risulta che **l'autoparallelismo** è proprietà **caratteristica** delle geodetiche, e può servire a definirle.[8]

9.1.5 Trasporto di un vettore lungo una linea chiusa. Angolo di parallelismo

Nel paragrafo precedente abbiamo visto che, in generale, il trasporto per parallelismo lungo una linea γ tracciata sulla data superficie S dipende essenzialmente da questa linea, ciò significa che: se un vettore si trasporta da un punto A ad un punto B di una superficie (non sviluppabile) lungo due linee diverse, si arriva in generale a due vettori distinti in B che formano fra loro un certo angolo θ detto *angolo di parallelismo*. **Quest'angolo è nullo se la superficie è sviluppabile.** Si dimostra che, in generale, l'angolo di parallelismo è legato alla curvatura della superficie.[9] Possiamo verificare direttamente sopra un esempio quanto abbiamo appena affermato, utilizzando la proprietà rilevata, nel paragrafo precedente, che le geodetiche sono curve autoparallele. A tal fine, consideriamo sopra una sfera un triangolo geodetico formato da un arco AB di cerchio massimo (equatore) e da due archi AC e BC, anch'essi di cerchio massimo (meridiani) perpendicolari al primo e formanti fra di loro un angolo θ (Fig. 9.12).

Consideriamo in A un vettore $\mathbf{u}$, tangente all'arco AB, e trasportiamolo parallelamente in C lungo il percorso ABC: esso arriverà in B mantenendosi sempre tangente all'arco AB (perché il trasporto avviene lungo una geodetica), mentre, nel trasporto da B a C si manterrà perpendicolare all'arco BC (per la stessa ragione). Esso arriverà dunque in C mantenendosi perpendicolare

[8]Si riconosce in questo enunciato un'ovvia estensione a superfici di natura qualsiasi dell'intuizione primordiale della retta, espressa da Euclide nell'assioma: *linea retta è quella che giace ugualmente rispetto ai suoi punti.*

[9]Precisamente si può dimostrare la seguente interessante relazione tra parallelismo e curvatura della superficie: *il rapporto tra l'angolo di parallelismo e l'area racchiusa entro un circuito infinitesimo di forma qualunque, uguaglia la curvatura della superficie.* Per la dimostrazione di questa proprietà si può vedere: Bruno Finzi e Maria Pastori, *Calcolo Tensoriale e Applicazioni*, pag. 228, Zanichelli Editore Bologna, 1971.

Figura 9.12: Trasporto di un vettore lungo una linea chiusa. Angolo di parallelismo.

all'arco BC.
Se ora trasportiamo lo stesso vettore $\mathbf{u}$ da A in C lungo l'arco AC esso arriverà in C con la direzione perpendicolare all'arco AC. Perciò le due nuove posizioni del vettore $\mathbf{u}$, ottenute per trasporto parallelo lungo il cammino ABC o lungo il cammino AC essendo il primo perpendicolare a BC ed il secondo perpendicolare ad AC, formeranno dunque un angolo uguale a θ.
Se invertiamo il tratto di percorso AC, così da trasportare il vettore $\mathbf{u}$ lungo il circuito chiuso $ABCA$, esso ritorna in A formando l'angolo θ rispetto alla posizione iniziale, perciò quest'angolo è l'angolo di parallelismo.
È evidente che, se si ripete la stessa operazione sopra una superficie sviluppabile, prendendo come percorso un triangolo geodetico, si verifica che l'angolo di parallelismo è nullo; basta infatti distendere la superficie sopra un piano, così che il triangolo geodetico diventi un ordinario triangolo rettilineo, e rifare su di esso l'operazione precedente.

9.1.6 Trasporto infinitesimo. Forma differenziale della legge di parallelismo

Supponiamo ora che il punto P_1 sia infinitamente vicino a P, riducendo il cammino γ all'arco elementare PP_1 (che a meno di infinitesimi d'ordine superiore al primo, rimane univocamente determinato dagli estremi). Siano inoltre $\omega \equiv T_P(S)$ e $\omega_1 \equiv T_{P_1}(S)$ i due corrispondenti piani tangenti alla superficie S (vedi Fig. 9.13 in cui tutto è stato ingrandito). Per avere lo sviluppo, basta in tal caso far subire al piano ω_1 una rotazione infinitesima attorno alla retta r, secondo cui esso interseca ω.

La direzione di tale retta dicesi *coniugata* della direzione PP_1, (in P o in P_1, ciò che è lo stesso, a meno d'infinitesimi). Indicheremo con $-\boldsymbol{\psi}$ il vettore infinitesimo, parallelo ad r, che rappresenta, in grandezza, direzione e verso, la rotazione elementare mediante la quale si adagia ω_1 su $\omega \equiv \omega^*$ (piano di sviluppo). Sarà allora $\boldsymbol{\psi}$ la rotazione elementare che fa ripassare ω_1 dal piano di sviluppo ω^* alla sua effettiva posizione di piano tangente ad S in P_1. Se $\mathbf{v} = PQ$ è un generico vettore tangenziale spiccato da P, per avere l'equipollente $\mathbf{v}_1 = P_1Q_1$ in P_1, si spiccherà da P_1, quando questo si trova nel piano di sviluppo (e perciò coincide con P_1^*), il vettore $\mathbf{v}_1^* = P_1^*Q_1^*$ equipollente (nel senso ordinario) a $\mathbf{v} = PQ$. Poi si ricondurrà il piano ω_1, adagiato su ω, alla sua posizione primitiva, trasportando con esso il vettore $\mathbf{v}_1^* = P_1^*Q_1^*$, talché il vettore $\mathbf{v}_1 = P_1Q_1$ proviene da $\mathbf{v} = PQ$, cui si è fatta subire, oltre a una traslazione, da $\mathbf{v} = PQ$ a $\mathbf{v}_1^* = P_1^*Q_1^*$ (che non ci interessa se consideriamo il vettore indipendentemente dal suo punto di applicazione), anche la rotazione $\boldsymbol{\psi}$. Gli spostamenti presi in considerazione sono manifestamente spostamenti rigidi, perciò ricordando la formula fondamentale della cinematica dei sistemi rigidi, riferita a spostamenti infinitesimi, avremo:

$$dQ = dP + \boldsymbol{\psi} \wedge (Q - P), \tag{9.4}$$

che possiamo scrivere anche così:

$$d(Q - P) = \boldsymbol{\psi} \wedge (Q - P), \tag{9.5}$$

Figura 9.13: Trasporto infinitesimo. Forma differenziale della legge di parallelismo.

dove $d(Q-P) = d\mathbf{v}$ rappresenta il divario dei due vettori $\mathbf{v}$ e $\mathbf{v}_1$, cioè l'incremento vettoriale $d\mathbf{v} = \mathbf{v}_1 - \mathbf{v}$ che subisce il vettore $\mathbf{v}$ nel trasporto per parallelismo da P a P_1. Tale divario sarà allora espresso dal prodotto vettoriale:

$$d\mathbf{v} = \boldsymbol{\psi} \wedge \mathbf{v}. \tag{9.6}$$

Siccome tanto $\boldsymbol{\psi}$ che $\mathbf{v}$ sono vettori appartenenti al piano ω l'incremento $d\mathbf{v}$ risulta perpendicolare a questo piano (o, in particolare, nullo).[10] Perciò $d\mathbf{v}$ sarà perpendicolare a qualunque vettore del piano $\omega \equiv T_P(S)$, in particolare a $\mathbf{v}$ medesimo, cioè:

$$\mathbf{v} \cdot d\mathbf{v} = 0. \tag{9.7}$$

Ora faremo vedere che la condizione (9.6), unita all'altra che $\mathbf{v}_1$, sia un vettore tangenziale, cioè appartenga ad ω_1, individua completamente il vettore $\mathbf{v}_1$, sicché, **come definizione differenziale del parallelismo superficiale, possiamo assumere l'insieme delle due relazioni geometriche seguenti**, in cui con n abbiamo indicato la normale ad ω:

$$\begin{cases} d\mathbf{v} \parallel n \\ \mathbf{v}_1 \parallel \omega_1 \end{cases} \tag{9.8}$$

o equivalentemente:

$$\begin{cases} d\mathbf{v} \perp T_P(S) \\ \mathbf{v}_1 \in T_{P_1}(S). \end{cases} \tag{9.9}$$

Si pensi infatti che deve risultare (Fig.9.14)

$$\mathbf{v} = \mathbf{v}_1 - d\mathbf{v}, \tag{9.10}$$

[10]Quest'ultima circostanza si persenta quando $\mathbf{v}$ abbia la direzione di $\boldsymbol{\psi}$, cioè la coniugata di PP_1; in questo caso, e in questo solo, la parallela superficiale $\mathbf{v}_1$, coincide con la parallela euclidea.

Figura 9.14: Decomposizione di un vettore in un componente parallelo ad un piano e ad una direzione non complanare.

cioè il vettore $\mathbf{v}$ deve essere decomponibile in un componente $\mathbf{v}_1$ parallelo ad un piano dato, e in un altro $-d\mathbf{v}$, parallelo ad una direzione data (non contenuta in quel piano): tale decomposizione, come è noto, può effettuarsi in un solo modo.
Osservazione:
Osserviamo che quando ci si sposta da un punto generico $P \in S$ ad un punto infinitamente vicino, l'incremento $d\mathbf{v}$ che subisce un vettore superficiale $\mathbf{v}$, supposto applicato in P, in generale, non appartiene al piano tangente ad S in P, cioè $d\mathbf{v} \notin T_P(P)$ (vedi Fig. 9.15).

Figura 9.15: Decomposizione dell'incremento $d\mathbf{v}$ che subisce un vettore superficiale nel trasporto parallelo, in uno spostamento tangenziale ed in uno spostamento normale (dove si è posto $(d\mathbf{v})_\perp \equiv (d\mathbf{v})_1$.)

Perciò $d\mathbf{v}$ può essere decomposto in un vettore tangenziale ed in un vettore normale alla superficie S, cioè:

$$d\mathbf{v} = (d\mathbf{v})_{tang.} + (d\mathbf{v})_{norm.} = (d\mathbf{v})_{\parallel} + (d\mathbf{v})_{\perp}. \tag{9.11}$$

Il componente tangenziale viene generalmente indicato con $D\mathbf{v}$, cioè:

$$(d\mathbf{v})_{\parallel} = D\mathbf{v}, \tag{9.12}$$

e prende il nome di **differenziale superficiale o differenziale assoluto di v**. Quest'ultimo nome sarà giustificato in seguito essendo legato al fondamentale concetto di **derivata assoluta o derivata covariante**.
Utilizzando questo concetto è evidente che la condizione geometrica (9.9) di parallelismo potrà essere espressa dalla semplice equazione

$$D\mathbf{v} = 0. \tag{9.13}$$

9.1.7 Equazione simbolica ed equazioni intrinseche di parallelismo

Alla definizione differenziale del parallelismo superficiale, riassunta nelle relazioni geometriche (9.9), si può dare una forma più espressiva, osservando che essa equivale a dire che *il vettore* $d\mathbf{v}$ *è perpendicolare a qualunque direzione tangenziale ad* S *in* P, ovvero, se pensiamo tale direzione individuata da uno spostamento infinitesimo del punto P lungo la superficie, che $d\mathbf{v}$ è perpendicolare a tutti questi spostamenti. In simboli, se con δP indichiamo il vettore infinitesimo rappresentante lo spostamento, avremo:

$$d\mathbf{v} \cdot \delta P = 0, \qquad \forall \delta P \in T_P(S) \tag{9.14}$$

equazione che ricorda formalmente *il Principio dei lavori virtuali* [11], e che chiameremo **equazione simbolica del parallelismo.**
L'equazione simbolica (9.14)*, facendo intervenire elementi geometrici estranei alla superficie, non mette in evidenza* **il carattere intrinseco della nozione di parallelismo***: però si presta assai bene a ricavarne delle equazioni, che abbiano questa notevole particolarità.*
A tale scopo, introduciamo la base locale $\left\{\mathbf{e}_j = \dfrac{\partial P}{\partial x^j}\right\}$ $(j = 1, 2)$ definita sul piano tangente $T_P(S)$, espressa in funzione delle coordinate curvilinee (gaussiane) x^i definite sulla superficie, e determiniamo le componenti controvarianti dei vettori $d\mathbf{v}$ e δP che figurano nell'equazione simbolica. Avremo allora, essendo $\mathbf{v}$ un vettore superficiale:

$$\mathbf{v} = v^j \mathbf{e}_j, \qquad (j = 1, 2) \tag{9.15}$$

e perciò

$$d\mathbf{v} = d(v^j \mathbf{e}_j) = dv^j \mathbf{e}_j + v^j d\mathbf{e}_j, \qquad (j = 1, 2) \tag{9.16}$$

con

$$d\mathbf{e}_j \stackrel{\text{def}}{=} \frac{\partial \mathbf{e}_j}{\partial x^h} dx^h \equiv \partial_h \mathbf{e}_j dx^h, \qquad (j, h = 1, 2). \tag{9.17}$$

D'altra parte, essendo anche δP un vettore superficiale, si avrà:

$$\delta P = \delta x^k \mathbf{e}_k, \qquad (k = 1, 2) \tag{9.18}$$

con le δx^k *arbitrarie*.
L'equazione simbolica di parallelismo (9.14) si può perciò esprimere nella forma

$$d\mathbf{v} \cdot \delta P = 0 \Longleftrightarrow (dv^j \mathbf{e}_j + v^j d\mathbf{e}_j) \cdot \delta x^k \mathbf{e}_k = 0, \tag{9.19}$$

ovvero

$$(dv^j \mathbf{e}_j \cdot \mathbf{e}_k + v^j d\mathbf{e}_j \cdot \mathbf{e}_k) \delta x^k = 0. \tag{9.20}$$

Ricordando la (9.17) e ponendo, com'è d'uso nella teoria delle superfici, $\mathbf{e}_j \cdot \mathbf{e}_k \stackrel{\text{def}}{=} g_{jk}$, la (9.20) si scriverà così:

$$(dv^j g_{jk} + v^j \partial_h \mathbf{e}_j \cdot \mathbf{e}_k dx^h) \delta x^k = 0. \tag{9.21}$$

Se ora introduciamo i simboli a 3 indici, simmetrici rispetto ai primi due[12]:

$$\Gamma_{hj,k} \stackrel{\text{def}}{=} \partial_h \mathbf{e}_j \cdot \mathbf{e}_k \qquad \text{con} \quad (\Gamma_{hj,k} = \Gamma_{jh,k}). \tag{9.22}$$

[11] Ricordiamo al lettore l'enunciato del **Principio dei lavori virtuali per una sola particella**: *La reazione vincolare* $\mathbf{R}$*, dovuta ad un vincolo liscio, è tale che il lavoro* $\delta\Lambda$ *da essa effettuato, per ogni spostamento virtuale reversibile* δP *è:* $\delta\Lambda = \mathbf{R} \cdot \delta P = 0$.

[12] I simboli a tre indici $\Gamma_{hj,k}$ sono simmetrici rispetto ai primi due indici h, j perché $\partial_h \mathbf{e}_j = \dfrac{\partial}{\partial x^h}\dfrac{\partial P}{\partial x^j} = \dfrac{\partial^2 P}{\partial x^h \partial x^j} = \dfrac{\partial^2 P}{\partial x^j \partial x^h} = \partial_j \mathbf{e}_h$. Si osservi inoltre che, mentre i vettori $\mathbf{e}_j$ sono superficiali, cioè tangenti alla superficie S in P, i vettori derivati $\partial_h \mathbf{e}_j$ generalmente escono fuori dalla superficie.

detti **simboli di Christoffel di** 1^a **specie** (componenti dei vettori derivati $\partial_h \mathbf{e}_j$ della base locale secondo i vettori della base stessa)[13], la (9.21) potrà scriversi più semplicemente così:

$$(dv^j g_{jk} + v^j \Gamma_{hj,k} dx^h)\delta x^k = 0. \tag{9.23}$$

Ponendo, per brevità

$$\tau_k = dv^j g_{jk} + v^j \Gamma_{hj,k} dx^h \qquad (j, h, k = 1, 2) \tag{9.24}$$

la (9.23) e perciò la (9.14) può essere ulteriormente scritta

$$\tau_k \delta x^k = 0, \qquad (k = 1, 2) \tag{9.25}$$

e *l'equazione simbolica del parallelismo*, per la completa arbitrarietà delle δx^k equivale alle due seguenti equazioni:

$$\tau_k = 0, \qquad (k = 1, 2). \tag{9.26}$$

Sono queste le due equazioni che definiscono gli incrementi dv^1, dv^2 da attribuirsi alle componenti controvarianti di un generico vettore $\mathbf{v}$ quando lo si trasporta per parallelismo lungo il cammino elementare definito da (dx^1, dx^2), che possiamo sinteticamente indicare come uno spostamento infinitesimo $dP \equiv (dx^i)$. Il loro carattere intrinseco sarà dimostrato più avanti.
Ora è importante rilevare il carattere tensoriale (covariante) delle τ_k, che segue dall'invarianza (per trasformazione di coordinate) della forma lineare $\tau_k \delta x^k$ attestata dall'identità

$$d\mathbf{v} \cdot \delta P \equiv \tau_k \delta x^k$$

in quanto si tratta di un prodotto scalare e si tiene conto della completa arbitrarietà delle δx^k (1° criterio di tensorialità).
Il sistema reciproco delle τ_k così definito:

$$\tau^i \stackrel{\text{def}}{=} g^{ik} \tau_k, \qquad (k = 1, 2) \tag{9.27}$$

risulta, in conformità, *controvariante*. In base alle (9.24) esso si esplicita come segue

$$\tau^i = dv^j g^{ik} g_{jk} + v^j g^{ik} \Gamma_{hj,k} dx^h = dv^i + v^j g^{ik} \Gamma_{hj,k} dx^h, \tag{9.28}$$

ovvero, introducendo un altro insieme di simboli a tre indici:

$$\Gamma^i_{hj} \stackrel{\text{def}}{=} g^{ik} \Gamma_{hj,k}, \qquad (\Gamma^i_{hj} = \Gamma^i_{jh}), \tag{9.29}$$

detti **simboli di Christoffel di** 2^a **specie**[14], sotto la forma:

$$\tau^i = dv^i + \Gamma^i_{jh} v^j dx^h, \tag{9.30}$$

tenuto conto che $\Gamma^i_{hj} = \Gamma^i_{jh}$.
Si osservi che i simboli di Christoffel di 2ª specie (9.29), a 3 indici, si sono ottenuti a partire

[13]Riguardo al simbolismo, si deve avvertire lo studioso del fatto che essendo il prodotto scalare fra vettori commutativo si ha $\partial_h \mathbf{e}_j \cdot \mathbf{e}_k = \mathbf{e}_k \cdot \partial_h \mathbf{e}_j$, perciò è altrettanto lecito indicare i **simboli di Christoffel di** 1^a **specie**, anziché con $\Gamma_{hj,k}$, come abbiamo fatto noi, con $\Gamma_{k,hj}$, notazione quest'ultima molto usata dai fisici, per es. da Landau-Lifschitz. Storicamente, val la pena notare, che l'originaria notazione usata da Ricci, Levi-Civita e da tutti gli autori dell'epoca, risalente probabilmente a Christoffel, era $\left[{hj \atop k}\right]$, che però è stata abbandonata essendo in contrasto con la comodissima convenzione di sommazione di Einstein, oggi universalmente adottata. Alcuni autori, per conformarsi a questa convenzione, hanno adottato il simbolo $[hj, k]$ o ancora (hj, k).

[14]Anche qui val la pena osservare che l'originaria notazione per i **simboli di Christoffel di** 2^a **specie** era $\left\{{hj \atop i}\right\}$, sostituita successivamente dalla $\left\{{i \atop hj}\right\}$ per conformarsi alla convenzione di sommazione di Einstein, e usata ancora oggi da alcuni autori.

dai simboli di Christoffel di 1ª specie innalzando il terzo indice, col tensore metrico in forma controvariante, come se si trattasse di un tensore. Dimostreremo però nel paragrafo 9.1.12 che sia i simboli di 1ª specie, sia quelli di 2ª specie, in generale non sono tensori.
Le equazioni di parallelismo scritte nella forma covariante (9.26), possono ovviamente scriversi anche nella forma controvariante

$$\tau^i \equiv dv^i + \Gamma^i_{jh} v^j dx^h = 0, \qquad (i, j, h = 1, 2). \tag{9.31}$$

Nell'una e nell'altra forma, esse hanno, come è a priori prevedibile per il loro significato geometrico, carattere invariantivo di fronte a qualsiasi cambiamento di sistema di coordinate curvilinee (x^1, x^2). Risolvendo le (9.31) rispetto ai differenziali dv^i avremo il **risultato fondamentale**:

$$dv^i = -\Gamma^i_{jh} v^j dx^h, \qquad (i, j, h = 1, 2). \tag{9.32}$$

Questa è la forma finale delle equazioni differenziali del parallelismo superficiale: *esse forniscono l'incremento subito dalle componenti controvarianti di un vettore superficiale nel trasporto per parallelismo lungo il cammino elementare dx^i, espresso mediante le dx^i stesse, le componenti del vettore, e certe funzioni del posto (da considerarsi assegnate: i cosiddetti* **coefficienti di connessione**, *più precisamente* **connessione di Levi-Civita***) dipendenti solo dai coefficienti del* $\mathbf{ds}^2$, *cioè dal tensore metrico $g_{ik}(x)$, e quindi dalla* **natura intrinseca della superficie.**
Per completezza ci rimane da dimostrare quest'ultima affermazione: lo faremo fra poco.

9.1.8 Proprietà dei simboli di Christoffel. Identità di Ricci. 1^a e 2^a formula di Christoffel

Si tratta ora di far vedere che i simboli di Christoffel sia di 1^a che di 2^a specie *sono esprimibili con elementi intrinseci*, e, precisamente, esprimibili con combinazioni lineari delle derivate del tensore metrico.
A tal fine, deriviamo, rispetto ad x^i, l'identità

$$g_{jk} = \mathbf{e}_j \cdot \mathbf{e}_k,$$

avremo

$$\partial_i g_{jk} = \partial_i \mathbf{e}_j \cdot \mathbf{e}_k + \mathbf{e}_j \cdot \partial_i \mathbf{e}_k,$$

e per la (9.22) otteniamo la cosiddetta **identità di Ricci**:

$$\partial_i g_{jk} = \Gamma_{ij,k} + \Gamma_{ik,j}, \tag{9.33}$$

che è di uso corrente.
Osserviamo, per mnemotecnica, che sul 1° termine del 2° membro gli indici i, j, k si susseguono nello stesso ordine del 1° membro, mentre nel 2° termine del 2° membro i rimane al suo posto mentre si scambiano j con k.
Se ora, nella (9.33) operiamo nel primo membro una permutazione circolare sugli indici i, j, k, otteniamo:

$$\partial_j g_{ki} = \Gamma_{jk,i} + \Gamma_{ji,k}, \tag{9.34}$$

$$\partial_k g_{ij} = \Gamma_{ki,j} + \Gamma_{kj,i}. \tag{9.35}$$

Sommiamo ora membro a membro la (9.33) con la (9.34) e sottraiamo da tale somma la (9.35), otteniamo l'importantissima 1^a **formula di Christoffel**:

$$\Gamma_{ij,k} = \frac{1}{2}\Big(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}\Big), \tag{9.36}$$

nella quale si deve intendere $\partial_i \equiv \dfrac{\partial}{\partial x^i}$ e analoghe.
Abbiamo così espresso i simboli di Christoffel di 1ª specie in termini delle derivate parziali prime delle componenti del tensore metrico della superficie S, perciò essi rappresentano certe funzioni delle coordinate gaussiane x^1, x^2 **dipendenti solo dalla 1ª forma quadratica fondamentale** *della data superficie aventi dunque* **carattere intrinseco.**
Naturalmente anche i simboli di Christoffel di 2ª specie, in virtù delle (9.29) e (9.36), si esprimeranno in funzione del tensore metrico mediante l'altrettanto importante 2^a **formula di Christoffel**:

$$\Gamma^i_{hj} = g^{ik}\Gamma_{hj,k} = \frac{1}{2}g^{ik}\Big(\partial_h g_{jk} + \partial_j g_{hk} - \partial_k g_{hj}\Big). \tag{9.37}$$

Avvertiamo che sia i simboli di Christoffel di 1^a specie, sia quelli di 2^a specie, in generale, non sono tensori. Si comportano come tensori soltanto per trasformazioni lineari di coordinate. Vedremo in seguito come si trasformano in un cambiamento di coordinate.
Osserviamo che ogni indice assume i valori 1 o 2, avremo allora che in totale i simboli di Christoffel di 1^a specie sono tanti quante le disposizioni con ripetizione di 2 oggetti a 3 a 3, cioè $D'_{2,3} = 2^3 = 8$, ma data la simmetria rispetto ai primi due indici, essi si riducono a 6 simboli distinti.
È facile esprimere i simboli di Christoffel di 1^a specie in funzione di quelli di 2^a specie, basta moltiplicare 1° e 2° membro della (9.37) per g_{il} e sommare rispetto all'indice i, si ottiene

$$g_{il}\Gamma^i_{hj} = g_{il}g^{ik}\Gamma_{hj,k} = \delta^k_l\Gamma_{hj,k} = \Gamma_{hj,l},$$

cioè:

$$\Gamma_{hj,l} = g_{il}\Gamma^i_{hj}. \tag{9.38}$$

A questo punto possiamo fare una semplice osservazione sui simboli di Christoffel. Se sulla superficie S (per es. su di un piano) è possibile introdurre coordinate tali da rendere il tensore metrico a coefficienti costanti (coordinate cartesiane), è ovvio che tutti i simboli di Christoffel saranno nulli, $\Gamma_{ij,k} \equiv 0$, $\Gamma^k_{ij} \equiv 0$, perché sono nulle tutte le derivate parziali che figurano nelle (9.36) e (9.37), e quindi l'equazione di parallelismo (9.32) fornisce l'ordinario concetto euclideo. In tal caso, infatti, avremo

$$dv^i = 0,$$

relazione che implica $v^i = \text{cost}$.
Osservazione:
Le definizioni che abbiamo dato per i simboli di Christoffel di 1^a e di 2ª specie sono immediatamente estensibili anche al caso di forme quadratiche in n indeterminate e perciò riferibili a varietà riemanniane di dimensione n. Ma su ciò torneremo in seguito.

9.1.9 Equazioni di parallelismo nelle componenti covarianti

È facile trovare delle equazioni, analoghe alle (9.32), **per i differenziali delle componenti covarianti** di un vettore **v** che venga trasportato per parallelismo su di una superficie S. Queste componenti, infatti, si ottengono da quelle controvarianti, per mezzo della ben nota relazione

$$v_i = g_{ij}v^j,$$

da cui, differenziando e cambiando l'indice j in k nell'ultima sommatoria, avremo:

$$dv_i = \partial_l g_{ij}dx^l v^j + g_{ik}dv^k.$$

Ora si sostituisca dv^k con la sua espressione, data dalla (9.32), cioè, previo opportuno aggiustamento degli indici, $dv^k = -\Gamma^k_{jl}v^j dx^l$, si avrà allora:

$$dv_i = \partial_l g_{ij}dx^l v^j - g_{ik}\Gamma^k_{jl}v^j dx^l.$$

Applicando al 1 termine del 2° membro, l'identità di Ricci, e tenendo conto che $g_{ik}\Gamma^k_{jl} = \Gamma_{jl,i}$, si ha:

$$dv_i = (\Gamma_{li,j} + \Gamma_{lj,i})dx^l v^j - \Gamma_{jl,i} v^j dx^l$$

semplificando otteniamo:

$$dv_i = \Gamma_{li,j} v^j dx^l = \Gamma_{li,j} g^{jk} v_k dx^l = \Gamma^k_{li} v_k dx^l.$$

Infine, per la simmetria dei simboli di Christoffel di 2^a specie rispetto ai due indici in basso si ha $\Gamma^k_{li} = \Gamma^k_{il}$ e scambiando, nell'ultimo termine l'indice k in j ed l in h, per mettere meglio in evidenza l'analogia con le (9.32), otteniamo[15]

$$dv_i = \Gamma^j_{ih} v_j dx^h \qquad (i, j, h = 1, 2), \tag{9.39}$$

sono queste anch'esse **equazioni differenziali di parallelismo** equivalenti alle (9.32); è ovvio che reciprocamente da esse, con calcoli analoghi, si possono ricavare le (9.32).

Osservazione:

Si può giungere a stabilire le equazioni differenziali di parallelismo (9.39) nelle componenti covarianti del vettore trasportato utilizzando la proprietà, già stabilita, che nel trasporto per parallelismo il prodotto scalare di due vettori superficiali si conserva.

Siano $\mathbf{u}$ e $\mathbf{v}$ due vettori superficiali applicati in un generico punto P di S e quindi appartenenti al piano tangente $T_P(S)$, supponiamo di trasportarli per parallelismo in un punto P_1 infinitamente vicino a P. Chiediamoci ora come si altera il prodotto scalare $\mathbf{u} \cdot \mathbf{v}$ nel trasporto infinitesimo. È del tutto ovvio che, per esso, vale la seguente identità:

$$d(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot d\mathbf{v} + \mathbf{v} \cdot d\mathbf{u}. \tag{9.40}$$

Ora, in virtù dell'equazione simbolica del parallelismo superficiale (9.14) i due vettori $\mathbf{u}$ e $d\mathbf{v}$ ed i due vettori $\mathbf{v}$ e $d\mathbf{u}$ sono fra di loro ortogonali, perciò i loro prodotti scalari sono entrambi nulli. Ma allora è

$$d(\mathbf{u} \cdot \mathbf{v}) = 0, \tag{9.41}$$

[15] A proposito della (9.39) val la pena far osservare al lettore che alcuni autori, una volta stabilita, in qualche modo, la validità dell'equazione (9.32) di parallelismo per le componenti controvarianti di un vettore superficiale $\mathbf{u}$, per dedurre la validità dell'analoga equazione espressa nelle componenti covarianti, cioè della (9.39), fanno il seguente ragionamento.

"L'equazione $du^i = -\Gamma^i_{jl} u^j dx^l$ definisce il trasporto parallelo di un vettore superficiale $\mathbf{u}$, da un punto P ad un'altro P_1 infinitamente vicino di una data superficie S, espresso per mezzo delle sue componenti controvarianti u^i. Il trasporto parallelo del vettore superficiale $\mathbf{v}$, applicato allo stesso punto P, espresso invece per mezzo delle sue componenti covarianti v_i è determinato dalla condizione *che il prodotto scalare* $\mathbf{u} \cdot \mathbf{v} = u^i v_i$ *dei due vettori* **essendo, per sua definizione, uno scalare (invariante)** *si mantiene costante durante il trasporto parallelo.* Perciò risulterà $d(u^i v_i) = u^i dv_i + v_i du^i = 0$, dunque, in virtù dell'equazione di parallelismo valida per u^i, si avrà $u^i dv_i - \Gamma^i_{jl} v_i u^j dx^l = 0$, ovvero (scambiando al 2° membro i con j) $u^i(dv_i - \Gamma^j_{il} v_j dx^l) = 0$. Per l'arbitrarietà di u^i se ne deduce che deve essere $dv_i = \Gamma^j_{il} v_j dx^l$."

Questo ragionamento però non è corretto perché il prodotto scalare di due vettori applicati in uno stesso punto P di una superficie **non è, per sua definizione, uno scalare invariante nel passare da un punto** P **ad un altro** P_1 **anche se infinitamente vicini**, esso è un invariante per trasformazioni di coordinate qualora si rimanga nello stesso punto. Nel caso delle superfici il prodotto scalare fra due vettori (superficiali) applicati nello stesso punto, è dato dall'espressione $\mathbf{u} \cdot \mathbf{v} = g_{ik}(x^j) u^i v^k$ che, in generale, *è una funzione del posto*, cioè delle coordinate x^j del punto ove sono applicati i due vettori, a meno che tutti gli elementi g_{ik} del tensore metrico siano delle costanti, come succede negli spazi euclidei o anche nello spazio di Minkowski in coordinate cartesiane. Perciò nel trasporto per parallelismo su di una superficie (o su una varietà qualsiasi) non si può dire, a priori, che il prodotto scalare fra due vettori rimanga costante. Insomma l'invarianza del prodotto scalare **è una conseguenza del trasporto per parallelismo** e non una sua proprietà intrinseca!

e ciò significa che *nel trasporto per parallelismo il prodotto scalare di due vettori superficiali si conserva*, cioè, rimane inalterato.
La (9.41) riscritta in termini di componenti equivale alla

$$d(u^i v_i) = u^i dv_i + v_i du^i = 0. \tag{9.42}$$

Tenendo conto della (9.32), la precedente equazione possiamo riscriverla

$$u^i dv_i - \Gamma^i_{jl} v_i u^j dx^l = 0,$$

scambiando, nel 2° termine, fra di loro, gli indici i e j e mettendo in evidenza u^i, avremo

$$u^i (dv_i - \Gamma^j_{il} v_j dx^l) = 0. \tag{9.43}$$

Per l'arbitrarietà di u^i se ne deduce che

$$dv_i = \Gamma^j_{il} v_j dx^l, \tag{9.44}$$

che è l'equazione voluta, cioè la (9.39), c.d.d.

9.1.10 Alcune conferme analitiche

Siamo ora in grado di dimostrare per via analitica alcune proprietà del parallelismo, già dedotte come conseguenze immediate della definizione geometrica.
Consideriamo intanto il trasporto per parallelismo di un vettore **v** lungo un tratto finito γ di curva, da P in P_1. Questa curva sia definita dalle equazioni parametriche

$$x^i = x^i(\lambda) \qquad (i = 1, 2), \tag{9.45}$$

dove λ rappresenta un parametro qualunque (in particolare essa può rappresentare *l'ascissa curvilinea*, cioè la lunghezza dell'arco di curva a partire da un'origine arbitraria P_0).
Le v^i sono da considerarsi funzioni di λ di cui sono assegnati arbitrariamente i valori in P. Le equazioni differenziali di parallelismo (9.32) possono allora riscriversi così:

$$\frac{dv^i}{d\lambda} + \Gamma^i_{jh} v^j \frac{dx^h}{d\lambda} = 0, \qquad (i, j, h = 1, 2), \tag{9.46}$$

dove le $\dfrac{dx^h}{d\lambda}$ sono, naturalmente, ricavate derivando, rispetto a λ, le (9.45), e quindi da considerarsi funzioni assegnate della variabile λ.
Le (9.46) *costituiscono un sistema di due equazioni differenziali ordinarie, lineari del primo ordine, in forma normale rispetto alle derivate delle due funzioni incognite* $v^1(\lambda), v^2(\lambda)$, e pertanto, come si sa dal Calcolo, **queste funzioni restano da esse univocamente determinate, una volta assegnati ad arbitrio i valori iniziali.** Si ha così la conferma del fatto, geometricamente evidente, della **possibilità del trasporto per parallelismo di un vettore superficiale arbitrariamente assegnato, e della unicità del risultato per ogni curva γ assegnata sulla superficie S e congiungente i punti P e P_1.**
Verifichiamo ora (sempre giovandosi delle equazioni differenziali trovate) l'invarianza nel trasporto per parallelismo di certe grandezze, che già rilevammo, nel paragrafo 9.1.4, come un fatto ovvio dal punto di vista geometrico, e cioè:

Proprietà:
Il trasporto per parallelismo conserva il prodotto scalare fra due vettori e perciò conserva la lunghezza di un vettore e l'angolo di due vettori.

Tali verifiche si possono fare contemporaneamente nel modo che segue.
Che il prodotto scalare di due vettori superficiali si conserva nel parallelismo lo abbiamo già stabilito ampiamente nel paragrafo precedente ma qui viene ridimostrato utilizzando le equazioni differenziali del parallelismo.
Siano $\mathbf{u}$, $\mathbf{v}$ due vettori superficiali applicati nello stesso punto P: trasportiamoli per parallelismo lungo un cammino infinitesimo, e *calcoliamo l'alterazione subita, in questo trasporto, dal loro prodotto scalare.* Avremo:

$$d(\mathbf{u} \cdot \mathbf{v}) = d(u^i v_i) = u^i dv_i + v_i du^i,$$

e ponendo per du^i e dv_i le espressioni che le competono in virtù delle equazioni di parallelismo (9.32) e (9.39), si ha

$$d(\mathbf{u} \cdot \mathbf{v}) = u^i \Gamma^j_{il} v_j dx^l - v_i \Gamma^i_{jl} u^j dx^l.$$

Scambiando in una delle due sommatorie, gli indici i e j, si riconosce che esse sono uguali, e quindi

$$d(\mathbf{u} \cdot \mathbf{v}) = 0,$$

cioè il trasporto infinitesimo (e quindi anche il trasporto finito) non altera il prodotto scalare. Ora basta supporre $\mathbf{u}$ coincidente con $\mathbf{v}$ (nel qual caso $\mathbf{u} \cdot \mathbf{v} = u^2 = v^2$) per concludere che il trasporto per parallelismo conserva la lunghezza di un vettore. Se poi si tiene presente questo risultato, e che $\mathbf{u} \cdot \mathbf{v} = uvcos\theta$, si riconosce che, restando invariato il prodotto scalare di due vettori, e le loro rispettive lunghezze, deve, se si tratta di due vettori entrambi diversi da zero, restare invariato anche il loro angolo.

9.1.11 Trasporto parallelo su di una varietà riemanniana

Siamo ora in grado di generalizzare tutto ciò che abbiamo detto sul parallelismo superficiale, relativo cioè a varietà 2-dimensionali $M^2 \equiv S$ (superfici), al caso di varietà riemanniane arbitrarie M^n di dimensione n. In tal caso, com'è ovvio, per vettori tangenti alla varietà in un dato punto P si devono intendere i vettori appartenenti allo spazio tangente $T_P(M^n)$ alla varietà in P, che, com'è noto, è uno spazio vettoriale di dimensione n. Un vettore tangente $\mathbf{v}(P)$ in un punto generico P della varietà possiamo rappresentarlo in una base locale sia in componenti controvarianti v^i, sia in componenti covarianti $v_i = g_{ik} v^k$ utilizzando il tensore metrico g_{ik} della varietà.
Per estendere alle varietà riemanniane M^n la nozione di parallelismo non è più possibile svolgere le semplici considerazioni geometriche a partire dal concetto di sviluppabile circoscritta poiché quest'ultima, per varietà generiche, non esiste. È invece adattabile immediatamente anche alle M^n riemanniane la legge differenziale del parallelismo espressa dall'equazione simbolica (9.14): il ragionamento sarebbe però un tantino lungo, risulta invece agevole operare l'estensione dal punto di vista analitico *utilizzando le equazioni differenziali del parallelismo* (9.32), (9.39) che abbiamo stabilito nel caso delle superficie facendo ora variare gli indici da 1 ad n.
Porremo la seguente

Definizione di parallelismo su una varietà riemanniana:
Diciamo che il vettore $\mathbf{v}$, *tangente ad* M^n, *di componenti controvarianti* v^i *e covarianti* v_i, *è* **trasportato per parallelismo**, *su* M^n, *dal punto* P *di coordinate* x^i, $(i = 1, 2, ..., n)$ *al punto infinitamente vicino* P_1 *di coordinate* $x^i + dx^i$, $(i = 1, 2, ..., n)$ *se gli incrementi* dv^i *delle componenti controvarianti, ovvero gli incrementi* dv_i *delle componenti covarianti, soddisfano rispettivamente ai seguenti sistemi differenziali lineari del 1° ordine:*

$$dv^i = -\Gamma^i_{jh} v^j dx^h, \qquad (i, j, h = 1, 2, ..., n), \tag{9.47}$$

$$dv_i = \Gamma^j_{ih} v_j dx^h, \qquad (i, j, h = 1, 2, ..., n), \tag{9.48}$$

dove i simboli a 3 indici Γ^i_{jh}, *detti* **simboli di Christoffel di** 2^a **specie o coefficienti di connessione**, *devono considerarsi come funzioni ben determinate delle coordinate* $x^i, (i = 1, 2, ..., n)$ *esprimibili in funzione del tensore metrico* g_{ik} *mediante* **la** 2^a **formula di Christoffel**:

$$\Gamma_{hj}{}^i \stackrel{\text{def}}{=} \frac{1}{2} g^{ik} \Big(\partial_h g_{jk} + \partial_j g_{hk} - \partial_k g_{hj} \Big) \qquad (i, h, j, k = 1, 2, ...n). \tag{9.49}$$

Naturalmente è possibile definire, in conformità della (9.49), anche i **simboli di Christoffel di** 1^a **specie** nel modo seguente,

$$\Gamma_{hj,l} = g_{il} \Gamma_{hj}{}^i \qquad (i, j, h, l = 1, 2, ..., n), \tag{9.50}$$

e viceversa si possono definire quelli di 2^a specie in funzione di quelli di di 1^a specie:

$$\Gamma_{hj}{}^k = g^{lk} \Gamma_{hj,l} \qquad (h, k, j, l = 1, 2, ..., n). \tag{9.51}$$

La 1^a **formula di Christoffel**, analoga a quella data per le superfici, è ovviamente:

$$\Gamma_{ij,k} = \frac{1}{2} \Big(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} \Big), \qquad (i, j, k = 1, 2, ..., n). \tag{9.52}$$

Avvertiamo, che anche nel caso delle varietà, sia i simboli di Christoffel di 1^a *specie, sia quelli di* 2^a *specie, in generale, non sono tensori. Si comportano come tensori soltanto per trasformazioni lineari di coordinate.* Vedremo, nel prossimo paragrafo, quale sia la legge di trasformazione per un arbitrario cambiamento di coordinate.
Le formule (9.47) e (9.48) definiscono **il trasporto infinitesimo per parallelismo** di un vettore $\mathbf{v}(P)$ sulla varietà M^n dal punto P_1 di coordinate $x^i_1, (i = 1, 2, ..., n)$ al punto infinitamente vicino P_2 di coordinate $x^i_2 = x^i_1 + dx^i, (i = 1, 2, ..., n)$. Se il trasporto per parallelismo avviene lungo un tratto finito da, P_1 a P_2, di una data curva γ della varietà di equazioni parametriche

$$x^i = x^i(\lambda) \qquad \lambda_1 \leq \lambda \leq \lambda_2 \qquad (i = 1, 2, ..., n), \tag{9.53}$$

dove λ rappresenta un parametro qualunque (in particolare può essere *l'ascissa curvilinea* s, lunghezza dell'arco di curva a partire da un'origine arbitraria P_0).
Le v^i e le v_i sono allora da considerarsi funzioni di λ di cui sono assegnati arbitrariamente i valori in P_1. Le equazioni differenziali di parallelismo (9.47) e (9.48) possono allora riscriversi così:

$$\frac{dv^i}{d\lambda} + \Gamma^i_{jh} v^j \frac{dx^h}{d\lambda} = 0, \qquad (i, j, h = 1, 2, ..., n), \tag{9.54}$$

$$\frac{dv_i}{d\lambda} - \Gamma^j_{ih} v_j \frac{dx^h}{d\lambda} = 0, \qquad (i, j, h = 1, 2, ..., n), \tag{9.55}$$

rispettivamente. Le $\dfrac{dx^h}{d\lambda}$ sono, naturalmente, ricavate derivando, rispetto a λ, le (9.53), e quindi da considerarsi funzioni assegnate della variabile λ.
Le (9.54) *e* (9.55) *costituiscono ciascuna un sistema di n equazioni differenziali ordinarie, lineari del primo ordine, di forma normale rispetto alle derivate delle n funzioni incognite* $(v^1(\lambda), v^2(\lambda), ..., v^n(\lambda))$, ovvero $(v_1(\lambda), v_2(\lambda), ..., v_n(\lambda))$, e pertanto, come si sa dal Calcolo, **queste funzioni restano, da tali equazioni, univocamente determinate, una volta assegnati ad arbitrio i valori iniziali,** da cui discende l'unicità del trasporto per parallelismo anche su di una varietà riemanniana. *Naturalmente il parallelismo sulla varietà riemanniana* M^n *dipende dalla curva* γ *di trasporto come nel caso di una superficie.* In particolare, trasportando un vettore lungo un cammino chiuso non si

riottiene generalmente il vettore di partenza; ciò, come si dimostra, è in relazione stretta con la curvatura della varietà. Soltanto quando la varietà è euclidea (curvatura nulla) il trasporto per parallelismo non dipende dal cammino scelto (*parallelismo assoluto*).
Osservazione:
Le espressioni al primo membro delle (9.54) e (9.55) si chiamano **le derivate assolute**[16], rispetto al parametro λ, dei campi v^i e v_i rispettivamente e si sogliono indicare con le notazioni

$$\frac{Dv^i}{d\lambda} \equiv \frac{dv^i}{d\lambda} + \Gamma^i_{jh} v^j \frac{dx^h}{d\lambda}, \qquad (i, j, h = 1, 2, ..., n), \tag{9.56}$$

$$\frac{Dv_i}{d\lambda} \equiv \frac{dv_i}{d\lambda} - \Gamma^j_{ih} v_j \frac{dx^h}{d\lambda}, \qquad (i, j, h = 1, 2, ..., n), \tag{9.57}$$

rispettivamente. Perciò le equazioni di parallelismo (9.54) e (9.55) possono scriversi, in forma molto sintetica, così:

$$\frac{Dv^i}{d\lambda} = 0 \qquad (i = 1, 2, ..., n), \tag{9.58}$$

$$\frac{Dv_i}{d\lambda} = 0 \qquad (i = 1, 2, ..., n), \tag{9.59}$$

o, ancora:

$$Dv^i = 0 \qquad (i = 1, 2, ..., n), \tag{9.60}$$

$$Dv_i = 0 \qquad (i = 1, 2, ..., n). \tag{9.61}$$

È facile verificare che l'operazione di trasporto parallelo su di una varietà riemanniana, come nel caso delle superfici, non altera il prodotto scalare di due vettori, né l'angolo da essi formato, né la norma di un vettore.[17] Torneremo fra breve sul concetto di **derivata assoluta** che scaturisce spontaneamente dal concetto fondamentale di **derivata covariante** di un campo di vettori o di tensori, definiti su di una varietà riemanniana, *concetto quest'ultimo legato intimamente, come vedremo, al trasporto per parallelismo più sopra definito.*

9.1.12 Ancora su alcune proprietà dei simboli di Christoffel e loro formule di trasformazione

Come abbiamo già osservato i sistemi a 3 indici introdotti, per una varietà riemanniana, nel paragrafo precedente e che abbiamo chiamato simboli di Christoffel di 1^a e di 2^a specie, definiti dalle (9.52) e (9.49), pur rappresentando degli **enti geometrici**[18] ben definiti, non sono tuttavia tensori. I simboli di Christoffel si comportano come tensori solo nella classe delle trasformazioni lineari. Questo giustifica per essi la denominazione di **tensori affini**. In tale accezione più ristretta nei simboli di 1^a specie tutti gli indici hanno carattere di covarianza, in quelli di 2^a specie due hanno carattere covariante e uno controvariante: di qui la posizione attribuita agli indici anche nel caso in cui questi non hanno carattere tensoriale.

[16]Le derivate assolute sono talvolta chiamate da alcuni autori (per es. Synge e Schild) **derivate intrinseche** ed indicate con $\frac{\delta v^i}{\delta\lambda}$ e $\frac{\delta v_i}{\delta\lambda}$.

[17]Lasciamo al lettore, come esercizio, la verifica di queste proprietà seguendo ciò che abbiamo fatto, alla fine del paragrafo precedente, nel caso delle superfici.

[18]Si dice che un sistema S ad uno o più indici (definito su una varietà M^n) è un **ente geometrico** quando sono verificate le seguenti condizioni:
1) Note le componenti di S in un riferimento prefissato si possono determinare, con una certa legge di trasformazione, quelle relative ad un qualunque altro sistema di coordinate, non appena si conoscano le formule di passaggio dal primo riferimento al secondo;
2) La legge di trasformazione costituisce un **gruppo**.

I simboli di Christoffel si chiamano anche **coefficienti della connessione riemanniana** appunto perché permettono di **connettere** tensori definiti in un punto con quelli definiti in un punto infinitamente vicino. Essi hanno, come vedremo, un ruolo essenziale in tutta la geometria riemanniana.
Osserviamo, ancora una volta, che essi godono delle **proprietà di simmetria**

$$\Gamma_{ij,k} = \Gamma_{ji,k}, \qquad \Gamma^i_{jk} = \Gamma^i_{kj},$$

che riducono da n^3 a $\dfrac{n^2(n+1)}{2}$ il numero dei simboli di Christoffel di 1^a o di 2^a specie distinti.[19]
D'altra parte anche le derivate $\partial_i g_{jk}$ distinte sono in pari numero e si capisce che esse si potranno esprimere mediante i suddetti simboli, allo stesso modo che per le superfici e con identica dimostrazione, per mezzo dell'importante **identità di Ricci**:

$$\partial_i g_{jk} = \Gamma_{ij,k} + \Gamma_{ik,j}, \qquad (i, j, k = 1, 2, ..., n). \tag{9.62}$$

Se la varietà M^n si riduce ad uno spazio euclideo e si adottano in esso coordinate cartesiane ortogonali si avrà

$$ds^2 = \sum_{i=1}^{n} \pm (dx^i)^2,$$

e la costanza delle g_{ik} fa sì che si annullino tutti i simboli di Christoffel. La stessa circostanza si presenta ovviamente se nello spazio euclideo si adottano coordinate cartesiane oblique; ma è chiaro che in coordinate curviliee generali i simboli di Christoffel non saranno tutti nulli, ciò in accordo col fatto che essi non sono dei tensori. Infatti l'annullarsi dei simboli di Christoffel in un particolare sistema di coordinate non implica l'annullarsi in ogni altro sistema di coordinate.
Ora ci rimane da dimostrare che i simboli di Christoffel sia di 1ª che di 2ª specie, definiti dalle (9.52) e (9.49), in generale, cioè per trasformazioni arbitrarie di coordinate, **non sono tensori** in quanto non si trasformano con legge tensoriale.
A tal fine, partiamo dal modo di trasformarsi del tensore fondamentale g_{ik}, che essendo covariante, per un generico cambiamento $x \to x'$ di coordinate si trasforma nel modo seguente:

$$g_{i'k'} = \theta^i_{i'} \theta^k_{k'} g_{ik}, \tag{9.63}$$

dove, per semplicità di scrittura, abbiamo posto $\theta^i_{i'} \equiv \dfrac{\partial x^i}{\partial x^{i'}}$ e porremo anche $\theta^k_{r'k'} \equiv \dfrac{\partial^2 x^k}{\partial x^{r'} x^{k'}}$ e analoghe. Derivando rispetto ad $x^{r'}$, 1° e 2° membro della (9.63), si ha:

$$\partial_{r'} g_{i'k'} = \theta^i_{i'} \theta^k_{k'} \theta^r_{r'} \partial_r g_{ik} + \theta^i_{i'r'} \theta^k_{k'} g_{ik} + \theta^i_{i'} \theta^k_{r'k'} g_{ik}, \tag{9.64}$$

facciamo ora due successive permutazioni circolari sugli indici $r, i, k; r', i', k'$,, avremo:

$$\partial_{i'} g_{k'r'} = \theta^k_{k'} \theta^r_{r'} \theta^i_{i'} \partial_i g_{kr} + \theta^k_{k'i'} \theta^r_{r'} g_{kr} + \theta^k_{k'} \theta^r_{i'r'} g_{kr}, \tag{9.65}$$

$$\partial_{k'} g_{r'i'} = \theta^r_{r'} \theta^i_{i'} \theta^k_{k'} \partial_k g_{ri} + \theta^r_{r'k'} \theta^i_{i'} g_{ri} + \theta^r_{r'} \theta^i_{k'i'} g_{ri}. \tag{9.66}$$

[19] Se i simboli di Christoffel, sia di 1^a che di 2^a specie, non godessero della proprietà di simmetria, il loro numero sarebbe uguale al *numero delle disposizioni con ripetizione di n oggetti a 3 a 3*, cioè $D'_{n,3} = n^3$, ma in virtù della simmetria questo numero si riduce. Precisamente se fissiamo, per esempio, l'attenzione su $\Gamma_{ij,k}$, e pensiamo che k sia fissato il numero delle $\Gamma_{ij,k}$ sarebbe uguale *al numero delle combinazioni con ripetizione di n oggetti a 2 a 2*, cioè $C'_{n,2} = \binom{n+2-1}{2} = \binom{n+1}{2} = \frac{n(n+1)}{2}$. Siccome però k può assumere n valori distinti il precedente numero va moltiplicato per n, dunque il numero dei simboli di Christoffel di 1ª specie, $\Gamma_{ij,k}$ è $\frac{n^2(n+1)}{2}$. Lo stesso ragionamento si fa per calcolare il numero dei simboli di Christoffel di 2ª specie che sono ovviamente in numero pari a quelli di 1^a specie.

Addizioniamo ora membro a membro le (9.64) e (9.65) e sottraiamo la (9.66), moltiplicando inoltre 1° e 2° membro dell'espressione ottenuta per $\frac{1}{2}$, otterremo:

$$\frac{1}{2}\left(\partial_{r'}g_{i'k'}+\partial_{i'}g_{k'r'}-\partial_{k'}g_{r'i'}\right)=\theta^r_{r'}\theta^i_{i'}\theta^k_{k'}\left[\frac{1}{2}\left(\partial_r g_{ik}+\partial_i g_{kr}-\partial_k g_{ri}\right)\right]+$$

$$+\frac{1}{2}\left[\theta^i_{i'r'}\theta^k_{k'}g_{ik}+\theta^i_{i'}\theta^k_{r'k'}g_{ik}+\theta^k_{k'i'}\theta^r_{r'}g_{kr}+\theta^k_{k'}\theta^r_{i'r'}g_{kr}-\theta^r_{r'k'}\theta^i_{i'}g_{ri}-\theta^r_{r'}\theta^i_{k'i'}g_{ri}\right],$$

e semplificando si avrà:

$$\Gamma_{r'i',k'}=\theta^r_{r'}\theta^i_{i'}\theta^k_{k'}\Gamma_{ri,k}+\theta^k_{k'}\theta^i_{i'r'}g_{ik}. \tag{9.67}$$

Tornando alle vecchie notazioni, la (9.67) si può scrivere così:

$$\Gamma_{r'i',k'}=\frac{\partial x^r}{\partial x^{r'}}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^k}{\partial x^{k'}}\Gamma_{ri,k}+\frac{\partial x^k}{\partial x^{k'}}\frac{\partial^2 x^i}{\partial x^{i'}\partial x^{r'}}g_{ik}. \tag{9.68}$$

Naturalmente vale anche la trasformazione inversa:

$$\Gamma_{ri,k}=\frac{\partial x^{r'}}{\partial x^r}\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{k'}}{\partial x^k}\Gamma_{r'i',k'}+\frac{\partial x^{k'}}{\partial x^k}\frac{\partial^2 x^{i'}}{\partial x^i\partial x^r}g_{i'k'}. \tag{9.69}$$

Ricaviamo ora le leggi di trasformazione per i simboli di Christoffel di 2^a specie.
Partiamo dalla legge di trasformazione del tensore metrico sotto forma controvariante:

$$g^{k'h'}=\theta^{k'}_l\theta^{h'}_m g^{lm}, \tag{9.70}$$

e moltiplichiamo ambo i membri della (9.67) per i corrispondenti membri della (9.70), otteniamo dopo facili calcoli:

$$g^{k'h'}\Gamma_{r'i',k'}=\theta^r_{r'}\theta^i_{i'}\theta^k_{k'}\theta^{k'}_l\theta^{h'}_m g^{lm}\Gamma_{ri,k}+\theta^k_{k'}\theta^i_{i'r'}\theta^{k'}_l\theta^{h'}_m g^{lm}g_{ik}, \tag{9.71}$$

cioè

$$\Gamma^{h'}_{i'r'}=\theta^r_{r'}\theta^i_{i'}\theta^{h'}_m\Gamma^m_{ri}+\theta^{h'}_i\theta^i_{i'r'}, \tag{9.72}$$

cambiando, nel 1° termine del 2° membro, l'indice m in h, e nel nel 2° termine del 2° membro, l'indice i in h e tornando alle vecchie notazioni, otteniamo infine:

$$\Gamma^{h'}_{i'r'}=\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^r}{\partial x^{r'}}\frac{\partial x^{h'}}{\partial x^h}\Gamma^h_{ir}+\frac{\partial x^{h'}}{\partial x^h}\frac{\partial^2 x^h}{\partial x^{i'}\partial x^{r'}}, \tag{9.73}$$

è questa la formula cercata e vale naturalmente anche la sua inversa:

$$\Gamma^h_{ir}=\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{r'}}{\partial x^r}\frac{\partial x^h}{\partial x^{h'}}\Gamma^{h'}_{i'r'}+\frac{\partial x^h}{\partial x^{h'}}\frac{\partial^2 x^{h'}}{\partial x^i\partial x^r}. \tag{9.74}$$

Le (9.68) e (9.69) sono le formule di trasformazione dei simboli di Christoffel di 1ª specie, mentre le (9.73) e (9.74) sono le formule di trasformazione dei simboli di Christoffel di 2ª specie. Come si vede, tali simboli, pur rappresentando degli **enti geometrici** ben definiti, non sono tuttavia dei tensori: ciò è chiaramente dovuto alla presenza, nelle formule di trasformazione, del termine contenente le derivate parziali seconde delle coordinate. Soltanto se la trasformazione di coordinate fosse lineare, le derivate seconde si annullerebbero, ed i simboli di Christoffel, sia di 1ª che di 2ª specie diventerebbero tensori: questo giustifica per essi la denominazione di *tensori affini* che talvolta viene usata.
Esercizio:
Esprimere le derivate parziali seconde delle coordinate in funzione dei simboli di Christoffel di

2^a specie (queste formule sono talvolta molto utili).
Risultato:

$$\frac{\partial^2 x^l}{\partial x^{i'} \partial x^{r'}} = \frac{\partial x^l}{\partial x^{l'}} \Gamma^{l'}_{i'r'} - \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^r}{\partial x^{r'}} \Gamma^l_{ir}, \tag{9.75}$$

$$\frac{\partial^2 x^{l'}}{\partial x^i \partial x^r} = \frac{\partial x^{l'}}{\partial x^l} \Gamma^l_{ir} - \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{r'}}{\partial x^r} \Gamma^{l'}_{i'r'}. \tag{9.76}$$

Capitolo 10

GEOMETRIA RIEMANNIANA E DERIVAZIONE COVARIANTE

La grandezza di Riemann come matematico risiede nel possente carattere di generalità e nella portata illimitata dei metodi e dei nuovi punti di vista da lui rivelati tanto in matematica pura che in matematica applicata.
(Eric T. Bell, *I grandi matematici*, Sansoni Editore, Firenze, 1966, pag.510)

LA DERIVAZIONE COVARIANTE

La derivazione covariante è una delle operazioni fondamentali della moderna **Analisi tensoriale sulle varietà** la cui denominazione originale, dovuta ai suoi fondatori, i matematici italiani Gregorio Ricci Curbastro ed il suo allievo Tullio Levi-Civita, era: **Calcolo differenziale assoluto**. Essa è assolutamente necessaria, come ebbe a dire Einstein, per la formulazione della Teoria della Relatività Generale.
Attualmente ci sono diversi approcci all'introduzione del concetto di **derivata covariante**, noi seguiremo quello originale di Levi-Civita che lega in maniera intima e circostanziata questo concetto con quello di parallelismo che non tutte le trattazioni attuali mettono nel giusto rilievo. Nell'APPENDICE 1 a questo capitolo presenteremo comunque una trattazione assiomatica moderna che può essere utile per uno studioso che desideri fare un confronto fra le due impostazioni.

10.1 Derivazione ordinaria di un campo tensoriale

Finora abbiamo incontrato (nell'ambito dell'Algebra tensoriale) operazioni algebriche sui tensori che sono tensoriali, cioè che convertono tensori in tensori. Tali operazioni sono: l'addizione, la sottrazione, la moltiplicazione, la contrazione, la moltiplicazione contratta, ma anche l'innalzamento ed abbassamento di indici mediante il tensore metrico.
Il problema che ora ci si pone è il seguente: *Quali operazioni differenziali esistono che siano tensoriali?* La risposta è piuttosto complessa e verrà impostata in questo paragrafo ed esaminata dettagliatamente nei prossimi paragrafi.
La necessità di differenziare un campo tensoriale si presenta frequentemente in problemi applicativi. Prima di tutto si presenta la necessità di calcolare le derivate parziali ordinarie delle componenti di un tensore in coordinate curvilinee generali.
Cominciamo col fare alcune considerazioni di natura generale. È noto che in uno spazio vettoriale E_n, di dimensione n, si può parlare di uguaglianza di vettori, di tensori, ecc. Ha senso

altresì assegnare in un medesimo E_n un vettore funzione di un parametro λ e svolgere l'ordinaria analisi vettoriale (differenziale di un vettore, derivata rispetto al parametro λ e così via). Così su di una varietà M^n *ha senso il confronto tra enti tensoriali definiti in uno stesso punto P di* M^n, cioè in uno stesso spazio vettoriale tangente $T_P(M^n)$.
Consideriamo invece due punti distinti di M^n, siano P e Q, e gli spazi tangenti ad essi associati $T_P(M^n)$ e $T_Q(M^n)$. Fissati in P e Q, cioè nei rispettivi spazi tangenti, due vettori **v** e **w**, vediamo se sia possibile un immediato confronto fra essi (Fig.10.1).

Figura 10.1: Confronto fra vettori tangenti in due punti distinti P e Q di una varietà.

A priori si potrebbe dire che i due vettori sono uguali se, in un particolare sistema di coordinate, hanno componenti uguali

$$v_j = w_j. \tag{10.1}$$

Ma è chiaro che tale uguaglianza non ha *significato intrinseco* perché non si conserva in un generico cambiamento di coordinate. Se infatti si opera una trasformazione di coordinate $x \to x'$, per le nuove componenti sarà

$$v_{j'}(P) = \left(\frac{\partial x^j}{\partial x^{j'}}\right)_P v_j(P); \qquad v_{j'}(Q) = \left(\frac{\partial x^j}{\partial x^{j'}}\right)_Q v_j(Q). \tag{10.2}$$

Si vede bene che, dall'uguaglianza $v_j = w_j$ non segue affatto l'altra $v_{j'} = w_{j'}$, in quanto i coefficienti $\left(\frac{\partial x^j}{\partial x^{j'}}\right)_P$ e $\left(\frac{\partial x^j}{\partial x^{j'}}\right)_Q$ non sono calcolati nello stesso punto, ma in P e Q rispettivamente. In altri termini la somma $v_j + w_j$ o la differenza $v_j - w_j$ non rappresenta un vettore.
La prima cosa che possiamo constatare è che *la derivazione parziale dei tensori non è una operazione tensoriale nel senso che a partire da tensori non produce tensori.*
Infatti, se consideriamo su di una varietà differenziabile (in particolare riemanniana) M^n, per esempio un campo di vettori covarianti $v_j(P)$, sappiamo che in un cambiamento di variabili $x \to x'$ tali vettori si trasformano con la legge seguente:

$$v_j = \frac{\partial x^{j'}}{\partial x^j} v_{j'}. \tag{10.3}$$

Se ora deriviamo 1° e 2° membro della (10.3) rispetto alla variabile x^i avremo:

$$\partial_i v_j = \frac{\partial}{\partial x^i}\left(\frac{\partial x^{j'}}{\partial x^j} v_{j'}\right) = \frac{\partial x^{j'}}{\partial x^j}\frac{\partial v^{j'}}{\partial x^{i'}}\frac{\partial x^{i'}}{\partial x^i} + \frac{\partial^2 x^{j'}}{\partial x^i \partial x^j} v_{j'},$$

cioè

$$\partial_i v_j = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^{j'}}{\partial x^j}\partial_{i'} v^{j'} + \frac{\partial^2 x^{j'}}{\partial x^i \partial x^j} v_{j'}. \tag{10.4}$$

Se nel 2° membro della (10.4) ci fosse solo il 1° termine allora questa sarebbe la solita legge tensoriale per un tensore doppio covariante. La presenza del 2° termine, contenente le derivate

seconde delle coordinate, distrugge però, in generale, la tensorialità a meno che la trasformazione di coordinate non sia lineare.
C'è un motivo essenziale per cui accade questo. Per definizione, il processo di derivazione riguarda il confronto tra una quantità calcolata in punti vicini della varietà M^n, poniamo P e Q, fatto calcolando la differenza tra tali valori, dividendo poi per qualche parametro che rappresenta la distanza fra P e Q e considerando infine il limite per questo parametro che tende a zero. Nel caso di un vettore covariante v_j questo significherebbe calcolare il

$$\lim_{\delta u \to 0} \frac{v_j(P) - v_j(Q)}{\delta u},$$

per un appropriato parametro δu.
La legge di trasformazione di un vettore covariante è espressa dalle (10.2) già viste, dalle quali è chiaro che $v_j(P) - v_j(Q)$ non è un vettore (cioè non è un tensore del 1° ordine).

Considerazioni analoghe valgono per la derivazione dei tensori in generale.
Infatti, consideriamo, per semplicità, un campo tensoriale covariante di rango k, $T_{i_1 \dots i_k}$. Fissiamo un sistema di coordinate curvilinee $(x) = (x^1, \dots x^n)$ e costruiamo l'insieme di funzioni $P_{i_1 \dots i_k, \alpha} = \partial_\alpha T_{i_1 \dots i_k}$, ovvero le derivate parziali (in senso ordinario) delle componenti del dato campo tensoriale. Resta ora da verificare se questo insieme forma o meno un campo tensoriale P; in altre parole ci si pone la domanda: **le derivate parziali di un tensore si trasformano, per un arbitrario cambiamento di coordinate, come un tensore?** Verifichiamo questo fatto con un calcolo diretto.

Il cambiamento di coordinate $(x) \to (x')$ conduce alla trasformazione

$$\begin{aligned} \partial_{\alpha'} T_{i'_1 \dots i'_k} &= \frac{\partial}{\partial x^{\alpha'}} \left(\frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_k}}{\partial x^{i'_k}} T_{i_1 \dots i_k} \right) = \\ &= \frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_k}}{\partial x^{i'_k}} \frac{\partial T_{i_1 \dots i_k}}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial x^{\alpha'}} + T_{i_1 \dots i_k} \frac{\partial}{\partial x^{\alpha'}} \left(\frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_k}}{\partial x^{i'_k}} \right). \end{aligned}$$

Allora

$$P_{i'_1 \dots i'_k, \alpha'} = P_{i_1 \dots i_k, \alpha} \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_k}}{\partial x^{i'_k}} + S_{i'_1 \dots i'_k, \alpha'}(T, x, x'), \tag{10.5}$$

avendo posto

$$S_{\alpha'; i'_1 \dots i'_k}(T, x, x') = T_{i_1 \dots i_k} \frac{\partial}{\partial x^{\alpha'}} \left(\frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_k}}{\partial x^{i'_k}} \right).$$

Per un'arbitraria sostituzione $(x) \to (x')$ la quantità a $S_{i'_1 \dots i'_k, \alpha'}$ è, in generale, diversa da zero, cioè $P_{i_1 \dots i_k, \alpha'}$ non si trasforma con legge tensoriale. Perciò la derivazione parziale ordinaria ∂_α è un'operazione *non* tensoriale.
L'espressione di $S_{i'_1 \dots i'_k, \alpha'}$ mostra quali ipotesi sono necessarie affinché l'operazione ∂_α sia tensoriale, ovvero *soltanto per trasformazioni lineari di coordinate* $(x) \to (x')$; in questo caso $S_{i'_1 \dots i'_k, \alpha'} \equiv 0$.

Un altro esempio eloquente è il seguente. Consideriamo un campo vettoriale controvariante T^i riferito ad un sistema di coordinate curvilinee (x) e calcoliamo la divergenza di questo campo mediante la formula

$$\mathrm{div}\mathbf{T} = \sum_i \frac{\partial T^i}{\partial x^i}, \tag{10.6}$$

che è l'analoga della definizione di divergenza in un sistema di coordinate cartesiane. Vediamo se questa espressione è uno scalare invariante per una arbitraria trasformazione di coordinate

$(x) \to (x')$, ovvero se $\mathrm{div}\mathbf{T}_{(x')} = \mathrm{div}\mathbf{T}_{(x)}$.
Abbiamo

$$\begin{aligned}\mathrm{div}\mathbf{T}_{(x')} &= \sum_{i'} \frac{\partial T^{i'}}{\partial x^{i'}} = \frac{\partial}{\partial x^{i'}}\left(\frac{\partial x^{i'}}{\partial x^i} T^i\right) = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial T^i}{\partial x^k}\frac{\partial x^k}{\partial x^{i'}} + T^i \frac{\partial x^k}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^k \partial x^i} = \\ &= \frac{\partial T^i}{\partial x^k}\delta_i^k + T^i \frac{\partial x^k}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^k \partial x^i} = \sum_i \frac{\partial T^i}{\partial x^i} + T^i \frac{\partial x^k}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^k \partial x^i}.\end{aligned}$$

Ma allora

$$\mathrm{div}\mathbf{T}_{(x')} = \mathrm{div}\mathbf{T}_{(x)} + R(T, x, x'), \tag{10.7}$$

avendo posto

$$R(T, x, x') = T^i \frac{\partial x^k}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^k \partial x^i}.$$

Per una arbitraria trasformazione di coordinate la quantità R non è nulla perché in generale non sono nulle le derivate seconde delle (x') rispetto alle (x), cioè la data definizione di divergenza non è invariante. Il termine R è nullo per una trasformazione lineare di coordinate. Perciò ribadiamo che, per trasformazioni non lineari, ∂_i non è una operazione tensoriale. Il nostro compito è dunque quello di definire un'opportuna operazione di derivazione che sia tensoriale nel senso che trasformi tensori in tensori. È ciò che faremo nei prossimi paragrafi.

10.2 Derivazione covariante di un campo vettoriale

Iniziamo la nostra trattazione da uno dei casi più semplici: consideriamo un campo di tensori covarianti di rango 1 (vettori covarianti) $v_i(P)$ definito sulla varietà riemanniana M^n.
Fissata l'attenzione su un determinato punto P della varietà (prescindendo cioè dal fatto che le $v_i(P)$ sono definite come funzioni del posto) ricordando (per il 1° Criterio di tensorialità visto nel Cap.8, paragrafo 8.15) che *la legge di trasformazione delle v_i in un cambiamento di coordinate $x \to x'$, è definita dall'invarianza della forma lineare*

$$F = v_i \xi^i = Invariante, \tag{10.8}$$

in cui le ξ^i costituiscono un *generico vettore controvariante*, o, se si vuole, sono le componenti controvarianti di un *generico* vettore ¸.
Ora, poiché ad ogni punto di M^n è coordinato un sistema di valori delle v_i, potremo in ogni punto scegliere *un arbitrario vettore controvariante* ξ^i e comporre con esso e le v_i *una forma invariante*.
Immaginiamo fatta questa scelta in un ben determinato, ma qualunque punto P, e prendiamo in considerazione *un generico punto* $P_1 \in M^n$ infinitamente vicino a P, *convenendo di assumere in* P_1, *come* ξ^i, *il vettore* **"parallelo"** *a quello che abbiamo scelto in P (trattandosi di trasporto infinitesimo non interviene la curva di trasporto).*
Indichiamo genericamente **con l'operatore** δ **l'incremento di una quantità** nel passaggio da P a P_1, e proponiamoci di calcolare il δF. Avremo differenziando la (10.8) col simbolo δ:

$$\delta F = \delta(v_i \xi^i) = \delta v_i \xi^i + v_i \delta \xi^i. \tag{10.9}$$

Ora, per la convenzione fatta poc'anzi sul campo vettoriale ξ^i, le $\delta\xi^i$ si debbono calcolare con **le formule di parallelismo** (9.47), cioè:

$$\delta \xi^i = -\Gamma^i_{jl} \xi^j \delta x^l, \tag{10.10}$$

mentre le v_i, essendo per ipotesi funzioni del posto, il δv_i è dato dalla consueta regola di differenziazione cioè:

$$\delta v_i = \frac{\partial v_i}{\partial x^l}\delta x^l. \tag{10.11}$$

Tenendo conto delle (10.10) e (10.11), si ha:

$$\delta F = \frac{\partial v_i}{\partial x^l}\delta x^l \xi^i - v_i \Gamma^i_{jl}\xi^j \delta x^l. \tag{10.12}$$

Possiamo mettere in evidenza, in ciascuna sommatoria, $\xi^i \delta x^l$: basta a tal fine scambiare i con j nella 2^a sommatoria, avremo:

$$\delta F = \left(\frac{\partial v_i}{\partial x^l} - v_j \Gamma^j_{il}\right)\xi^i \delta x^l. \tag{10.13}$$

Poiché il 1° membro, per il suo significato, è *invariante*[1], mentre ξ^i e δx^l sono *arbitrari*, **i coefficienti di questa forma (cioè le espressioni dentro parentesi nella (10.13)) costituiscono, per definizione, un tensore covariante rispetto agli indici i ed l, che indicheremo con**:

$$\nabla_l v_i \stackrel{\text{def}}{=} \partial_l v_i - v_j \Gamma^j_{il}. \tag{10.14}$$

Questo tensore covariante di tipo $(0,2)$ e rango 2, dicesi **il derivato covariante (o la derivata covariante) del vettore covariante v_i.**
È evidente che in coordinate cartesiane (le quali esistono quando la varietà in considerazione è euclidea) **la derivata covariante si riduce alla derivata parziale ordinaria** essendo nulli tutti i simboli di Christoffel di 2^a specie: $\Gamma^j_{il} = 0$.
Se il campo vettoriale è dato in **forma controvariante** v^i si costruisce, in modo analogo al precedente, la forma lineare $F = v^i \xi_i$ invariante per trasformazioni di coordinate in cui le ξ_i costituiscono un *generico vettore covariante*. Si segue poi un ragionamento del tutto simile al precedente utilizzando, questa volta, le equazioni di parallelismo nella forma (9.48): $\delta \xi_i = \Gamma^j_{il}\xi_j \delta x^l$, si perviene dunque al seguente tensore

$$\nabla_l v^i \stackrel{\text{def}}{=} \partial_l v^i + v^j \Gamma^i_{jl}. \tag{10.15}$$

Questo tensore misto di tipo $(1,1)$ e di rango 2, dicesi **il derivato covariante (o la derivata covariante) del vettore controvariante v^i.** Il termine derivata covariante, sia nel 1° caso che nel 2°, dipende dal fatto che il tensore ottenuto ha un indice di covarianza in più rispetto al tensore di partenza[2].
La derivazione covariante è una estensione della derivata parziale ordinaria e si riduce a questa se la varietà M^n è uno spazio euclideo riferito a coordinate cartesiane (perché allora si annullano tutti i simboli di Chrisoffel di 2^a specie: $\Gamma^i_{jl} = 0$). Naturalmente anche in uno spazio euclideo la coincidenza fra derivata parziale e derivata covariante viene a mancare se si adoperano coordinate curvilinee generali.
A differenza della derivazione parziale ordinaria ∂_i, l'operazione ∇_i ha carattere assoluto: agendo su enti assoluti (vettori), essa dà luogo ad enti assoluti (tensori).
Si noti anche che, mentre la derivata parziale $\partial_i v_j$ opera solo su una componente del vettore, la derivata covariante $\nabla_i v_j$ fa intervenire in blocco tutte le componenti di $\mathbf{v}$.
Le due formule precedenti (10.14) e (10.15) sono fondamentali e costituiscono il prototipo delle

[1]Si potrebbe essere tentati di affermare che essendo la forma lineare F un invariante, sia $\delta F = 0$. Questa conclusione è però falsa, perché si deve osservare che la F, e perciò anche δF, **sono invarianti per trasformazioni di coordinate**, ma essendo funzioni del posto variano da un punto ad un altro della varietà M^n.

[2]Per quanto riguarda le notazioni correntemente usate per indicare le derivate covarianti è opportuno avvertire il lettore che la notazione ∇_i da noi introdotta è quella più usata dai matematici. I fisici teorici, per distinguere una derivata parziale ordinaria, quella che noi abbiamo indicato con $\partial_i \equiv \frac{\partial}{\partial x^i}$, usano spesso la *notazione virgola*, $v^i_{,j} \equiv \frac{\partial v^i}{\partial x^j}$ e per la derivata covariante la cosiddetta *notazione punto e virgola*. In tal caso le due formule fondamentali (10.14) e (10.15) si scrivono così: $v_{i;l} = v_{i,l} - v_j \Gamma^j_{il}$, $v^i_{;l} = v^i_{,l} + v^j \Gamma^i_{jl}$ rispettivamente. Alcuni autori, seguendo Levi-Civita, per la derivata covariante, usano le notazioni $v_{i|l}$ e $v^i_{|l}$.

derivate covarianti dei campi tensoriali definiti su una varietà riemanniana. Abbiamo così risolto completamente, almeno nei casi più semplici dei campi vettoriali, il problema che ci siamo posti all'inizio del paragrafo precedente: **ottenere con un'appropriata operazione di derivazione, che produca tensori a partire da tensori dati.**
Alcune utili osservazioni:
Si vede facilmente che il tensore doppio $\nabla_l v_i$ definito dalla (10.14) in generale *non è simmetrico*, è invece tale il tensore che si ottiene da quello per differenza, dopo aver scambiato fra di loro gli indici i ed l, ed anzi si ha la notevole relazione

$$\nabla_l v_i - \nabla_i v_l = \partial_l v_i - \partial_i v_l, \tag{10.16}$$

in cui la differenza delle derivate covarianti coincide con la differenza delle derivate parziali ordinarie.
Ł'annullarsi della derivata covariante del vettore v_i

$$\nabla_l v_i = 0,$$

ha un significato geometrico semplice: *esprime il fatto che il vettore di componenti v_i si trasporta per parallelismo lungo la linea coordinata l definita da $x^l = variabile, x^h = 0, con\ h \neq l$.*
In tal caso, moltiplicando 1° e 2° membro della (10.14) (nella quale si è posto $\nabla_l v_i = 0$), per dx^l (con l fissato), si ottiene:

$$\frac{\partial v_i}{\partial x^l} dx^l = \Gamma^j_{il} v_j dx^l \ \text{ (non c'è somma su } l \text{ in quanto fissato)},$$

nella quale il primo membro si identifica con dv_i, cioè:

$$dv_i = \Gamma^j_{il} v_j dx^l, \ \text{ (non c'è somma su } l\text{)}.$$

Se confrontiamo quest'ultima relazione con l'equazione (9.48) di parallelismo:

$$dv_i = \Gamma^j_{ih} v_j dx^h,$$

in cui si suppongono nulle tutte le dx^h tranne la $l - esima$, ci si rende conto che il vettore di componenti covarianti v_i è trasportato per parallelismo lungo la linea l. c.d.d.
Consideriamo infine un tensore di ordine zero, cioè uno scalare $f(x) \equiv f(x^1, x^2, ...x^n)$, la derivazione covariante coincide, per definizione, con la ordinaria derivazione parziale

$$\nabla_l f \stackrel{\text{def}}{=} \partial_l f \equiv f_l \equiv \frac{\partial f}{\partial x^l} \equiv \text{grad}\, f. \tag{10.17}$$

Se da queste, formiamo il sistema delle derivate seconde applicando la (10.14),avremo

$$\nabla_k f_l = \frac{\partial^2 f}{\partial x^l \partial x^k} - \Gamma^j_{lk} \frac{\partial f}{\partial x^j}, \tag{10.18}$$

queste non coincidono con le derivate seconde ordinarie, però sono, come quelle, simmetriche rispetto agli indici l, k.

10.3 Derivazione covariante di un tensore misto del 2° ordine

Consideriamo ora il caso, leggermente più generale dei due visti nel paragrafo precedente, e precisamente consideriamo il tensore misto del 2° ordine A^h_i, con un solo indice di controvarianza h ed un solo indice di covarianza i.

Fissata l'attenzione su un ben determinato punto P della varietà M^n (prescindendo cioè dal fatto che le A_i^h sono definite come funzioni del posto), ricordiamo (per il 2° Criterio di tensorialità visto nel Cap. 8, paragrafo 8.15) che *la legge di trasformazione delle A_i^h in un cambiamento di coordinate è definita dall'invarianza della forma bilineare*:

$$F = A_i^h \xi^i u_h = Invariante, \tag{10.19}$$

in cui le ξ^i sono le componenti controvarianti di un *generico* vettore ¸ ed analogamente le u_h si possono considerare componenti covarianti di un *generico* vettore $\mathbf{u}$.
Ora, poiché ad ogni punto della varietà M^n è coordinato un sistema di valori delle A_i^h, potremo in ogni punto scegliere due arbitrari vettori uno controvariante ξ^i e l'altro covariante u_h, e comporre con essi e le A_i^h una forma invariante.
Immaginiamo fatta questa scelta in un ben determinato, ma qualunque punto P, e prendiamo in considerazione un *generico* punto $P_1 \in M^n$ infinitamente vicino a P, *convenendo di assumere, in P_1, come ξ^i e u_h i vettori "* **paralleli***" a quelli che ci siamo scelti in P (trattandosi di trasporto infinitesimo non interviene la curva di trasporto).*
Come abbiamo fatto nel paragrafo precedente, indichiamo genericamente **con l'operatore δ l'incremento di una quantità** nel passaggio da P a P_1, e proponiamoci di calcolare il δF. Avremo differenziando la (10.19) col simbolo δ:

$$\delta F = \delta(A_i^h \xi^i u_h) = (\delta A_i^h)\xi^i u_h + A_i^h \delta\xi^i u_h + A_i^h \xi^i \delta u_h. \tag{10.20}$$

Ora, per la convenzione fatta poc'anzi sui campi vettoriali ξ^i e u_h, le $\delta\xi^i$ e le δu_h si debbono calcolare con **le formule di parallelismo** (9.47) e (9.48), cioè:

$$\delta\xi^i = -\Gamma_{jl}^i \xi^j \delta x^l, \qquad \delta u_h = \Gamma_{hl}^j u_j \delta x^l, \tag{10.21}$$

mentre essendo le A_i^h, per ipotesi, funzioni del posto, il δA_i^h va calcolato con la consueta regola di differenziazione cioè:

$$\delta A_i^h = \frac{\partial A_i^h}{\partial x^l}\delta x^l. \tag{10.22}$$

Tenendo conto delle (10.21) e (10.22), si ha:

$$\delta F = \frac{\partial A_i^h}{\partial x^l}\delta x^l \xi^i u_h - A_i^h \Gamma_{jl}^i \xi^j \delta x^l u_h + A_i^h \Gamma_{hl}^j \xi^i \delta x^l u_j. \tag{10.23}$$

Possiamo mettere in evidenza, in ciascuna sommatoria, $\xi^i u_h \delta x^l$: basta a tal fine scambiare i con j nella 2^a sommatoria e h con j nell'ultima, e raccogliendo in un'unica somma avremo:

$$\delta F = \left(\frac{\partial A_i^h}{\partial x^l} - A_j^h \Gamma_{il}^j + A_i^j \Gamma_{jl}^h\right)\xi^i u_h \delta x^l. \tag{10.24}$$

Poiché il 1° membro, per il suo significato, è *invariante*, mentre ξ^i, u_h, δx^l sono *arbitrari*, i coefficienti di questa forma (cioè le espressioni dentro parentesi nella (10.24)) **costituiscono, per definizione, un tensore covariante rispetto agli indici i ed l, controvariante rispetto ad h** che indicheremo con:

$$\nabla_l A_i^h \stackrel{\text{def}}{=} \partial_l A_i^h - A_j^h \Gamma_{il}^j + A_i^j \Gamma_{jl}^h. \tag{10.25}$$

Questo tensore misto di tipo $(1,2)$ e di rango 3, dicesi **il derivato covariante (o la derivata covariante) del tensore dato A_i^h.** Il termine derivata covariante, come abbiamo già detto nel paragrafo precedente, dipende dal fatto che il tensore ottenuto ha un indice di covarianza in più rispetto al tensore di partenza.
Come già detto, sottolineiamo che *la derivazione covariante è una estensione della derivata parziale*

ordinaria e si riduce a questa se la varietà M^n è uno spazio euclideo riferito a coordinate cartesiane (perché allora si annullano tutti i simboli di Chrisoffel di 2^a specie: $\Gamma^j_{il} = 0$. Naturalmente anche in uno spazio euclideo tale coincidenza viene a mancare se si adoperano coordinate generali.
A differenza della derivazione parziale ordinaria ∂_l, l'operazione ∇_l ha carattere assoluto: agendo su enti assoluti (tensori), essa dà luogo ad enti assoluti (tensori).

10.4 Derivazione covariante di un tensore qualunque

Considerazioni del tutto analoghe a quelle svolte nei due paragrafi precedenti portano ad estendere l'operazione di derivazione covariante al caso di un tensore qualunque $T^{h_1 \ldots h_\mu}_{i_1 \ldots i_m}$ di tipo (μ, m) e rango $\mu + m$.
Il procedimento seguito nei paragrafi precedenti si può ripetere in modo analogo, generalizzandolo, per un tensore misto generico: si otterrà sempre per δF (come risulta immediatamente dal calcolo materiale) una forma plurilineare invariante i cui coefficienti costituiscono la **derivata covariante del dato tensore misto**. La formula generale cui si perviene è la seguente:

$$\nabla_l T^{h_1 \ldots h_\mu}_{i_1 \ldots i_m} \stackrel{\text{def}}{=} \partial_l T^{h_1 \ldots h_\mu}_{i_1 \ldots i_m} - \sum_{r=1}^{m} T^{h_1 \ldots h_\mu}_{i_1 \ldots i_{r-1}\, j i_{r+1} \ldots i_m} \Gamma^j_{i_r l} + \sum_{\rho=1}^{\mu} T^{h_1 \ldots h_{\rho-1}\, j\, h_{\rho+1} \ldots h_\mu}_{i_1 \ldots i_m} \Gamma^{h_\rho}_{jl}, \tag{10.26}$$

essa è formata da un primo termine, che è la derivata ordinaria, cui seguono tante sommatorie col segno meno $(-)$, quanti sono gli indici di covarianza del dato tensore, e tante sommatorie col segno più $(+)$, quanti sono gli indici di controvarianza.
Utilizzando la notazione con *multi-indici*, cioè denotando con (i) il complesso degli indici $(i_1, \ldots i_m)$, e con (h) il complesso degli indici $(h_1, \ldots h_\mu)$, la formula precedente si può scrivere , più sinteticamente, così:

$$\nabla_l T^{(h)}_{(i)} \stackrel{\text{def}}{=} \frac{\partial T^{(h)}_{(i)}}{\partial x^l} - \sum_{r=1}^{m} T^{(h)}_{i_1 \ldots i_{r-1}\, j\, i_{r+1} \ldots i_m} \Gamma^j_{i_r l} + \sum_{\rho=1}^{\mu} T^{h_1 \ldots h_{\rho-1}\, j\, h_{\rho+1} \ldots h_\mu}_{(i)} \Gamma^{h_\rho}_{jl}. \tag{10.27}$$

Osserviamo, per mnemotecnica, che nella prima sommatoria l'indice di covarianza i_r viene spostato dal tensore ai simboli di Christoffel $\Gamma^j_{i_r l}$, dove compare come primo indice in basso, al suo posto, nel tensore, ci va un indice muto j che viene ovviamente ripetuto in alto nei simboli di Christoffel perché si deve sommare rispetto a questo indice. Il secondo indice in basso l che figura in $\Gamma^j_{i_r l}$ è l'indice relativo alla variabile x^l rispetto alla quale si effettua la derivazione covariante e che può essere chiamato *indice di derivazione*. In maniera analoga si procede per gli indici di controvarianza nei termini della seconda sommatoria.

10.5 Proprietà formali della derivazione covariante e teorema di Ricci

Vediamo ora quali sono le proprietà più importanti della derivazione covariante.

1) La derivazione covariante si riduce alla ordinaria derivazione parziale se le g_{ik} si riducono a costanti (per il che occorre che lo spazio in cui si opera sia euclideo e che lo si riferisca a coordinate cartesiane, rettangolari od oblique).

2) La derivata covariante applicata ad un sistema tensoriale le cui componenti sono costanti in M^n non dà generalmente il tensore nullo. Ciò è ben naturale se si pensa che la costanza delle componenti di un tensore non ha carattere invariantivo rispetto ad un cambiamento di coordinate.

3) La derivata della somma algebrica di più tensori è la somma delle derivate.

4) La derivata di un prodotto si effettua con la regola ordinaria di Leibniz. Ad esempio

$$\nabla_i(v_j T^k_l) = \nabla_i v_j T^k_l + v_j \nabla_i T^k_l.$$

5) La derivazione covariante e la contrazione sono operazioni permutabili. Cioè dato un tensore misto $T^{j\ldots}_{\ kl\ldots}$ si può calcolarne la derivata $\nabla_i T^{j\ldots}_{\ kl\ldots}$ e poi contrarre due indici di tipo diverso (non i), ad esempio j ed l; al medesimo risultato si perviene se si effettua prima la contrazione e poi la derivazione covariante. **6)** La formula (10.17) applicata allo scalare $f = v_j w^j$, per il teorema di Leibniz, ha per conseguenza l'identità

$$v_j \nabla_i w^j + w^j \nabla_i v_j = v_j\, \partial_i w^j + w^j \partial_i v_j.$$

Si noti però che mentre al 1° membro ciascuno dei due termini è un tensore, i termini a 2° membro non hanno singolarmente carattere tensoriale, pur avendolo la loro somma.

Sussiste inoltre il seguente importante:
Teorema di Ricci:
La derivata covariante del tensore metrico, nelle sue tre forme: covariante g_{ij}, controvariante g^{ij} e mista $g^i_j \equiv \delta^i_j$ è identicamente nulla, cioè [3]:

$$\nabla_k g_{ij} = 0, \quad \nabla_k g^{ij} = 0, \quad \nabla_k g^i_j = 0. \tag{10.28}$$

Dimostrazione:
La 3^a delle (10.28) si dimostra facilmente applicando la definizione di derivata covariante, avremo:

$$\nabla_k g^i_j \equiv \nabla_k \delta^i_j = \partial_k \delta^i_j - \delta^i_n \Gamma^n_{jk} + \delta^m_j \Gamma^i_{mk} = 0 - \Gamma^i_{jk} + \Gamma^i_{jk} = 0. \tag{10.29}$$

Per la 1^a delle (10.28) procediamo ancora con calcolo diretto ed introducendo i simboli di Christoffel di 1^a specie:

$$\nabla_k g_{ij} = \partial_k g_{ij} - g_{mj}\Gamma^m_{ik} - g_{in}\Gamma^n_{jk} = \partial_k g_{ij} - \left(\Gamma_{ik,j} + \Gamma_{jk,i}\right),$$

applicando l'identità di Ricci (9.62) ai termini che stanno in parentesi, avremo

$$\nabla_k g_{ij} = \partial_k g_{ij} - \partial_k g_{ij} = 0.$$

Infine per la 2ª delle (10.28) procediamo così:

$$\nabla_k\left(g_{ij} g^{jl}\right) = \nabla_k\left(\delta^l_i\right) = 0,$$

in virtù della 3ª delle (10.28). Ma, applicando la regola di Leibniz di derivazione covariante del prodotto di due tensori, otteniamo:

$$\nabla_k\left(g_{ij} g^{jl}\right) = g_{ij} \nabla_k g^{jl} + g^{jl} \nabla_k g_{ij}.$$

[3] Ricordiamo che il tensore metrico in forma covariante g_{ij} serve per abbassare un indice in alto di un dato tensore, mentre nella forma controvariante g^{ij} serve per innalzare qualche indice in basso, perciò avremo $g_{ij} g^{ik} = g^k_j$ in cui g_{ij} viene usato come operatore di abbassamento dell'indice i che figura in g^{ik}, quest'ultimo essendo il tensore del quale si vuole abbassare l'indice i che diventa j dopo essere stato abbassato in virtù della contrazione su i. D'altra parte g_{ij} e g^{ik} sono due tensori l'uno l'inverso dell'altro e dunque $g_{ij} g^{ik} = \delta^k_j$ per definizione. Dalle precedenti due relazioni ne segue che $g^k_j \equiv \delta^k_j$, si tratta dunque del simbolo di Kronecker in forma mista del quale, già sappiamo, che è un tensore misto di rango 2.

Dalla precedente relazione e dalla 1ª delle (10.28) si deduce che

$$g_{ij}\nabla_k g^{jl} = 0,$$

moltiplicando 1° e 2° membro di quest'ultima relazione per g^{im} e sommando su i, avremo

$$g^{im}g_{ij}\nabla_k g^{jl} = \delta^m_j \nabla_k g^{jl} = \nabla_k g^{ml} = 0,$$

che è ciò che volevamo dimostrare.
Possiamo concludere che: *di fronte alla derivazione covariante, il tensore metrico, qualunque sia la sua rappresentazione (covariante, controvariante o mista), si comporta dunque come una costante. Quindi, se in un prodotto compare il tensore fondamentale, con qualche indice saturato o no, esso si può portare fuori dal segno di derivazione covariante.*
Ad esempio:

$$\nabla_i\left(g_{jk}T^{lm}\right) = g_{jk}\nabla_i T^{lm},$$

oppure contraendo j con l:

$$\nabla_i\left(g_{jk}T^{jm}\right) = g_{jk}\nabla_i T^{jm}.$$

10.6 Derivazione controvariante

Il carattere tensoriale dell'indice di derivazione covariante giustifica la seguente definizione di **derivata controvariante** di un campo di vettori $v_j(P)$:

$$\nabla^i v_j = g^{ir}\nabla_r v_j. \tag{10.30}$$

Si tratta di una operazione che può essere ovviamente estesa ad un campo di tensori qualunque. Essa gode di tutte le proprietà della derivata covariante, in particolare risulta:

$$\nabla^i g_{jk} = 0, \qquad \nabla^i g^k_j = 0, \qquad \nabla^i g^{jk} = 0, \tag{10.31}$$

e quindi:

$$\nabla^i v_j = g_{jr}\nabla^i v^r. \tag{10.32}$$

Dalla (10.30) segue naturalmente:

$$\nabla^i v_j = g_{ir}\nabla^r v_j. \tag{10.33}$$

Insomma i quattro sistemi doppi $\nabla_i v_j$, $\nabla_i v^j$, $\nabla^i v_j$, $\nabla^i v^j$, in corrispondenza biunivoca fra di loro, costituiscono solo una diversa rappresentazione di un medesimo ente tensoriale; ciò rivela il profondo contenuto intrinseco delle operazioni differenziali introdotte.
Osservazione:
Se operiamo in uno spazio di Minkowski alla derivazione controvariante ∇^i le si fa corrispondere la derivazione $\partial^i = \eta^{ir}\partial_r$, dove $\eta_{ik} = \eta^{ik}$ è il tensore metrico dello spazio di Minkowski: questa derivazione ∂^i talvolta può risultare comoda al fine di scrivere in modo compatto certe formule.

10.7 Differenziale assoluto e derivata assoluta

È noto dall'analisi ordinaria cosa si intenda per **differenziale totale** di un campo scalare $\varphi(x)$ [calcolato in un punto P e relativamente ad un determinato vettore infinitesimo $dP \equiv (dx^i)$] lo scalare:

$$d\varphi = \partial_i\varphi dx^i.$$

Analogamente, in uno spazio euclideo riferito a coordinate cartesiane, suole chiamarsi **differenziale totale** di un campo vettoriale v^j il vettore:

$$dv^j = \left(\partial_i v^j\right)_P dx^i.$$

Poiché queste definizioni fanno riferimento alla derivazione parziale ordinaria, si presenta spontanea un'analoga definizione in una varietà riemanniana M^n in cui sia definita la derivazione covariante.

Definizione 1:

Si chiama **differenziale assoluto** di un campo vettoriale controvariante v^j [calcolato in un punto P e relativamente ad un determinato vettore tangente $dP \equiv (dx^i)$], il campo:

$$Dv^j = \nabla_i v^j dx^i. \tag{10.34}$$

Analogamente, per un campo covariante v_j si avrà:

$$Dv_j = \nabla_i v_j dx^i. \tag{10.35}$$

Più in generale, per un campo tensoriale $T^{i_1 \dots i_p}_{j_1 \dots j_q}$, sarà:

$$DT^{i_1 \dots i_p}_{j_1 \dots j_q} = \nabla_\alpha T^{i_1 \dots i_p}_{j_1 \dots j_q} dx^\alpha, \tag{10.36}$$

dove l'indice muto di sommazione si è indicato con α per maggiore chiarezza di scrittura e per distinguerlo dagli altri indici liberi.

La differenziazione assoluta è evidentemente un'estensione della differenziazione ordinaria e si riduce a questa se la varietà M^n è una varietà euclidea, riferita a coordinate cartesiane.

A differenza dell'ordinaria differenziazione d, l'operazione D ha carattere assoluto come la derivazione covariante; in altri termini, agendo su enti assoluti (tensori), essa dà luogo ancora a dei tensori. Si noti anche che l'operazione di differenziazione assoluta non altera il rango del tensore di partenza e si riduce alla differenziazione ordinaria se agisce su un tensore di rango zero. Dal concetto di differenziale assoluto si passa spontaneamente a quello di **derivata assoluta** rispetto ad un parametro λ.

Pensiamo assegnato su M^n un campo di vettori $v^j(P)$ e consideriamo una curva γ sulla varietà di equazioni parametriche $x^i = x^i(\lambda)$, uscente da un determinato punto P di M^n (il parametro λ potrà talvolta coincidere con l'ascissa curvilinea s). Indicando con $dP \equiv (dx^i)$ un vettore tangente in P alla curva γ, ha senso allora considerare i differenziali

$$dx^i = \frac{dx^i}{d\lambda} d\lambda$$

e quindi il differenziale assoluto

$$Dv^j = \nabla_i v^j dx^i = \nabla_i v^j \frac{dx^i}{d\lambda} d\lambda.$$

Definizione 2:

Si chiama **derivata assoluta** in P, rispetto al parametro λ, del campo vettoriale v^j il vettore

$$\frac{Dv^j}{d\lambda} = \nabla_i v^j \frac{dx^i}{d\lambda} = \frac{dv^j}{d\lambda} + \Gamma^j_{ri} v^r \frac{dx^i}{d\lambda}. \tag{10.37}$$

Ovviamente questo concetto estende la nozione di derivata ordinaria.

Più in generale, per un campo tensoriale $T^{i_1 \dots i_p}_{j_1 \dots j_q}$ qualunque, si chiamerà derivata assoluta l'ente tensoriale

$$\frac{DT^{i_1 \dots i_p}_{j_1 \dots j_q}}{d\lambda} = \nabla_\alpha T^{i_1 \dots i_p}_{j_1 \dots j_q} \frac{dx^\alpha}{d\lambda} = \frac{dT^{i_1 \dots i_p}_{j_1 \dots j_q}}{d\lambda} - \Gamma^r_{j_1 \alpha} T^{i_1 \dots i_p}_{r j_2 \dots j_q} \frac{dx^\alpha}{d\lambda} - \dots + \Gamma^{i_1}_{r\alpha} T^{r i_2 \dots i_p}_{j_1 j_2 \dots j_q} \frac{dx^\alpha}{d\lambda} + \dots. \tag{10.38}$$

Si osservi che la definizione (10.38) ha significato anche se i vettori v^j sono definiti soltanto nei punti della linea γ e non fuori di essa.

10.7.1 Cenno sulle varietà a connessione affine

In questo paragrafo desideriamo dare un cenno sulle cosiddette *varietà a connessione affine* il cui studio, pur non essendo strettamente necessario per i nostri scopi (inerenti la formulazione della Teoria della Relatività Generale) è di per sé interessante e può contribuire ad una migliore comprensione della geometria riemanniana.
Def.1:
Dicesi **varietà a connessione affine** *una varietà differenziabile M^n sulla quale, punto per punto è definito un insieme di n^3 funzioni $\Gamma^k_{ij}(x), (k, i, j = 1, 2, \dots n)$ del posto, chiamate* **coefficienti della connessione**, *generalmente non simmetrici rispetto agli indici in basso (i, j). Tale* **connessione** *si indica generalmente con $\nabla \equiv \Gamma^k_{ij}(x)$.*
In questo breve cenno sulle varietà a connessione affine considereremo, per semplicità, soltanto *connessioni simmetriche per le quali si ha $\Gamma^k_{ij}(x) = \Gamma^k_{ji}(x)$.*
Va immediatamente notato che la nozione di **connessione** e la nozione di **metrica riemanniana** non sono, in generale, in alcun modo legate fra loro: sono due strutture indipendenti introdotte sulla stessa varietà differenziabile M^n. Il legame ci può essere soltanto se è soddisfatta una particolare condizione di cui ora ci occuperemo.
Sia M^n una varietà riemanniana con **metrica** assegnata $g_{ij}(x)$. Consideriamo, su M^n, una **connessione simmetrica** ∇ definita dai coefficienti $\Gamma^k_{ij}(x)$.[4]
Def.2:
Si dice che *la connessione Γ^k_{ij} è* **associata** *alla metrica g_{ij}, o* **compatibile** *con essa*, se la derivata covariante del tensore metrico g_{ij} è identicamente nulla, cioè se è soddisfatto il **teorema di Ricci**:

$$\nabla_k g_{ij} \equiv 0. \tag{10.39}$$

Tenuto conto che applicando la definizione di derivata covariante al tensore di Kronecker δ^i_j si ottiene:

$$\nabla_k \delta^i_j = \partial_k \delta^i_j + \delta^h_j \Gamma^i_{hk} - \delta^i_h \Gamma^h_{jk} = \Gamma^i_{jk} - \Gamma^i_{jk} \equiv 0$$

possiamo enunciare il seguente lemma il cui contenuto è comunque incluso nel teorema di Ricci.

Lemma 10.1 *Le condizioni $\nabla_k g_{ij} \equiv 0$ e $\nabla_k g^{ij} \equiv 0$ sono equivalenti.*

Dimostrazione. Partiamo dalla relazione:

$$0 \equiv \nabla_k \delta^i_j = \nabla_k (g^{ih} g_{hj}) = g^{ih} \nabla_k g_{hj} + g_{hj} \nabla_k g^{ih}.$$

Se è vera la prima condizione, allora risulta

$$g_{hj} \nabla_k g^{ih} \equiv 0,$$

da cui, facendo una moltiplicazione contratta in ambo i membri per g^{lj} si avrà:

$$g^{lj} g_{hj} \nabla_k g^{ih} = \delta^l_h \nabla_k g^{ih} = \nabla_k g^{il} \equiv 0,$$

cioè è vera la seconda condizione.
Se invece è vera la seconda, allora risulta

$$g^{ih} \nabla_k g_{hj} \equiv 0,$$

[4] Avvertiamo lo studioso, che è d'uso corrente utilizzare la stessa notazione Γ^k_{ij}, sia per i simboli di connessione sia per i simboli di Christoffel di 2^a specie che sono definiti, come sappiamo, sulla varietà riemanniana utilizzando la metrica g_{ij}, i due concetti a priori non devono essere confusi. Il legame tra il concetto di connessione e quello di metrica lo esamineremo fra un momento.

da cui, facendo una moltiplicazione contratta in ambo i membri per g_{il} si ottiene:

$$g_{il}g^{ih}\nabla_k g_{hj} = \delta^l_h \nabla_k g_{hj} = \nabla_k g_{lj} \equiv 0,$$

cioè è vera anche la prima condizione. Il lemma è così dimostrato.
Se la connessione data è associata alla metrica il tensore metrico, qualunque sia la sua rappresentazione (covariante, controvariante o mista), *si comporta come una costante di fronte alla derivazione covariante*. In questo caso, quindi, *l'operazione di derivazione covariante commuta con quella di innalzamento o di abbassamento degli indici*, visto che il tensore metrico può essere portato fuori dal segno di derivazione covariante.

Data una connessione ∇ mediante i suoi coefficienti Γ^k_{ij} che diremo **simboli di connessione di 2ª specie**, si possono definire altri simboli che diremo **simboli di connessione di 1ª specie** nel modo seguente:

$$\Gamma_{ij,k} \stackrel{\text{def}}{=} \Gamma^h_{ij} g_{hk}. \tag{10.40}$$

Se la connessione è simmetrica, allora evidentemente i simboli di connessione di prima specie sono simmetrici rispetto ai primi due indici, ovvero agli indici prima della virgola.
Ora, facendo una moltiplicazione contratta su ambo i membri per g^{kl}:

$$\Gamma_{ij,k}g^{kl} = \Gamma^h_{ij} g_{hk} g^{kl} = \Gamma^h_{ij}\delta^l_h,$$

si ottiene l'espressione dei simboli di connessione di seconda specie in funzione di quelli di prima specie:

$$\Gamma^l_{ij} = \Gamma_{ij,k}g^{kl}. \tag{10.41}$$

Per le connessioni associate alla metrica vale anche il seguente:

Lemma 10.2 *La connessione Γ^k_{ij} è associata alla metrica g_{ij} se e solo se vale la seguente* **identità di Ricci***:*

$$\partial_k g_{ij} = \Gamma_{ki,j} + \Gamma_{kj,i}. \tag{10.42}$$

Dimostrazione. Se la connessione Γ^k_{ij} è associata alla metrica g_{ij} si ha, per definizione:

$$\nabla_k g_{ij} \equiv 0. \tag{10.43}$$

Ricordando la definizione di derivata covariante e la definizione dei simboli di connessione di prima specie, avremo:

$$\nabla_k g_{ij} = \partial_k g_{ij} - g_{hj}\Gamma^h_{ik} - g_{ih}\Gamma^h_{jk} = \partial_k g_{ij} - \Gamma_{ik,j} - \Gamma_{jk,i} = 0,$$

da cui, tenendo conto della simmetria dei simboli di connessione di 1^a specie rispetto ai primi due indici otteniamo:

$$\partial_k g_{ij} = \Gamma_{ki,j} + \Gamma_{kj,i}.$$

Il viceversa è immediato. c.d.d.

10.7.2 Connessione di Levi-Civita o riemanniana

Cominciamo col dimostrare subito il seguente:

Teorema 10.3 (Teorema fondamentale della geometria riemanniana) *Data su una varietà riemanniana M^n una metrica g_{ij} non degenere (cioè $g = det(g_{ij}) \neq 0$),* **esiste una ed una sola connessione simmetrica***, associata alla metrica g_{ij}. Questa connessione è definita, in qualsiasi sistema di coordinate $(x^1, ..., x^n)$, mediante la* **formula di Christoffel***:*

$$\Gamma^k_{ij} = g^{hk}\Gamma_{ij,h} = \frac{1}{2}g^{hk}(\partial_i g_{jh} + \partial_j g_{hi} - \partial_h g_{ij}).$$

Dimostrazione. Supponiamo che una connessione simmetrica associata alla metrica g_{ij} esista; siano $\Gamma^k_{ij} = \Gamma^k_{ji}$ i coefficienti che la definiscono. Per il Lemma 10.2 del paragrafo precedente, vale l'identità di Ricci (10.42). Su questa relazione, ora, eseguiamo una permutazione circolare sugli indici a primo membro, otteniamo così le tre relazioni:

$$\partial_k g_{ij} = \Gamma_{ki,j} + \Gamma_{kj,i}, \qquad \partial_i g_{jk} = \Gamma_{ij,k} + \Gamma_{ik,j}, \qquad \partial_j g_{ki} = \Gamma_{jk,i} + \Gamma_{ji,k}.$$

Sottraendo la prima dalla somma delle altre due, tenendo conto che la connessione è simmetrica, avremo:

$$\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} = \Gamma_{ij,k} + \Gamma_{ik,j} + \Gamma_{jk,i} + \Gamma_{ji,k} - \Gamma_{ki,j} - \Gamma_{kj,i} = 2\Gamma_{ij,k}.$$

Risulta perciò:

$$\Gamma_{ij,k} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij});$$

facendo una moltiplicazione contratta su ambo i membri per g^{kl}, si ottiene:

$$\Gamma_{ij,k} g^{kl} = \frac{1}{2} g^{kl}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}),$$

da cui:

$$\Gamma^l_{ij} = \frac{1}{2} g^{kl}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}).$$

Quindi, se una connessione simmetrica associata alla metrica esiste, è necessariamente di questa forma. Per provare l'esistenza, a partire dai coefficienti ora definiti, che, si è visto, costituiscono una connessione simmetrica, è sufficiente mostrare che vale l'identità di Ricci (10.42). Si ha infatti:

$$\begin{aligned}
\Gamma_{ki,j} \quad + \quad \Gamma_{kj,i} &= \Gamma^l_{ki} g_{lj} + \Gamma^l_{kj} g_{li} = && (10.44)\\
&= \frac{1}{2} g^{lh} g_{lj}(\partial_k g_{ih} + \partial_i g_{hk} - \partial_h g_{ki}) + \frac{1}{2} g^{lh} g_{li}(\partial_k g_{jh} + \partial_j g_{hk} - \partial_h g_{kj}) = && (10.45)\\
&= \frac{1}{2}\delta^h_j(\partial_k g_{ih} + \partial_i g_{hk} - \partial_h g_{ki}) + \frac{1}{2}\delta^h_i(\partial_k g_{jh} + \partial_j g_{hk} - \partial_h g_{kj}) = && (10.46)\\
&= \frac{1}{2}(\partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}) + \frac{1}{2}(\partial_k g_{ji} + \partial_j g_{ik} - \partial_i g_{kj}) = \partial_k g_{ij}, && (10.47)
\end{aligned}$$

ovvero quello che volevamo. Il teorema è completamente dimostrato. c.d.d.
La formula

$$\Gamma^k_{ij} \stackrel{\text{def}}{=} \frac{1}{2} g^{hk}(\partial_i g_{jh} + \partial_j g_{hi} - \partial_h g_{ij}) \tag{10.48}$$

prende il nome di 2^a **formula di Christoffel**. La connessione definita da questi coefficienti prende il nome di **connessione di Levi-Civita o riemanniana** della metrica g_{ij}.
Ora, facendo una moltiplicazione contratta su ambo i membri della (10.48) per g_{kl}, si ha, come visto nella dimostrazione del teorema precedente, la formula:

$$\Gamma_{ij,l} = \frac{1}{2}(\partial_i g_{jl} + \partial_j g_{li} - \partial_l g_{ij}), \tag{10.49}$$

chiamata 1^a **formula di Christoffel**, che definisce i simboli di Christoffel di prima specie della connessione di Levi-Civita.
Osservazione:
Nelle precedenti considerazioni abbiamo operato sempre in coordinate curvilinee generali, ma è chiaro che, in casi particolari, si possono usare **le coordinate cartesiane.** Ora si possono dare due diverse definizioni di coordinate cartesiane.

1) le coordinate $x^1, ..., x^n$ si dicono **cartesiane** se la matrice g_{ij} è euclidea in queste coordinate, cioè se

$$g_{ij} = \delta_{ij}; \tag{10.50}$$

2) le coordinate $x^1, ..., x^n$ sono dette **cartesiane** se i coefficienti Γ^k_{ij} di connessione sono identicamente nulli in queste coordinate:

$$\Gamma^k_{ij} \equiv 0, \qquad i, j, k = 1, ..., n \tag{10.51}$$

(o, in generale, se $\Gamma^k_{ij} = -\Gamma^k_{ji}$, per connessioni non simmetriche).
Qual'è la relazione tra queste due definizioni di coordinate cartesiane?
Come abbiamo detto all'inizio del paragrafo precedente la nozione di **connessione** e la nozione di **metrica riemanniana** non sono, in generale, in alcun modo legate fra loro: sono due strutture indipendenti introdotte sulla stessa varietà differenziabile M^n perciò si capisce che le espressioni (10.50) ed (10.51) definiscono nozioni differenti. Tuttavia, come risulta chiaro da quello che abbiamo visto in precedenza, se la connessione è associata alla metrica le due definizioni vengono a coincidere.

10.7.3 L'operatore divergenza

La connessione di Levi-Civita permette di definire l'operatore **divergenza** di un campo vettoriale su di una varietà M^n. Analogamente alla definizione che si dà nello spazio euclideo ordinario, come somma delle derivate parziali delle componenti del vettore fatte rispetto alle coordinate omonime: $\mathrm{div}\mathbf{v} = \sum_{i=1}^n \frac{\partial v^i}{\partial x^i}$, sulle varietà le derivate parziali vanno sostituite con le corrispondenti derivazioni covarianti. Avremo allora che la divergenza di un campo di vettori è data da un campo scalare così definito:

$$\nabla_i v^i = \partial_i v^i + v^k \Gamma^i_{ki}, \tag{10.52}$$

dove nei coefficienti di connessione Γ^i_{ki} essendoci due indici ripetuti, uno in basso e l'altro in alto, si deve intendere una sommatoria rispetto a tali indici i.

Nel caso di una varietà riemanniana per la quale è definito un tensore metrico g_{ik} non è difficile ottenere la forma esplicita della divergenza. Infatti, posto

$$g = \det g_{ij}(x^1, x^2, ..., x^n) \tag{10.53}$$

per la regola di derivazione dei determinanti (cfr. l'equazione (10.225) nell'APPENDICE 2 dove ci sono i dettagli del calcolo) si ha:

$$\partial_k g = \partial_k g_{ij} A^{ij},$$

dove $A^{ij} = cofg_{ij}$.
Ricordando che

$$g^{ij} = \frac{cofg_{ij}}{g},$$

si avrà

$$\partial_k g = g g^{ij} \partial_k g_{ij}. \tag{10.54}$$

Per l'identità di Ricci (10.42), si ottiene allora:

$$\partial_k g = g g^{ij} \partial_k g_{ij} = g g^{ij} \Gamma_{ki,j} + g g^{ij} \Gamma_{kj,i} = g \Gamma^i_{ki} + g \Gamma^j_{kj} = 2 g \Gamma^i_{ki},$$

da cui

$$\Gamma^i_{ki} = \frac{1}{2g} \partial_k g = \frac{1}{2} \partial_k \log |g| = \partial_k \log \sqrt{|g|} = \frac{1}{\sqrt{|g|}} \partial_k \sqrt{|g|}.$$

Sostituendo nella (10.52), perciò:

$$\nabla_i v^i = \partial_i v^i + v^k \frac{1}{\sqrt{|g|}} \partial_k \sqrt{|g|} = \frac{1}{\sqrt{|g|}} (\sqrt{|g|} \partial_i v^i + v^i \partial_i \sqrt{|g|}),$$

da cui, infine, la formula cercata della divergenza di v^i:

$$\nabla_i v^i = \frac{1}{\sqrt{|g|}} \partial_i (v^i \sqrt{|g|}). \tag{10.55}$$

10.8 Cenni di Calcolo delle variazioni

10.8.1 Il concetto di funzionale

Prima di affrontare lo studio delle geodetiche di una varietà riemanniana è opportuno fare qualche richiamo di Calcolo delle variazioni in quanto la definizione che daremo di geodetica è basata su di una equazione variazionale.
Nell'Analisi Matematica ordinaria si parla esplicitamente, di consueto, di *funzioni di punto* e di *funzioni d'insieme* (in particolare *funzioni di dominio*), ma si può dire che nelle applicazioni si incontrano svariati altri tipi di funzioni, cioè di corrispondenze fra elementi variabili e numeri.

Figura 10.2: Concetto di funzionale nel Calcolo delle variazioni.

Esempi:

i) *l'integrale definito* $\int_a^b f(x)\,dx$, con a e b fissi ed $a < b$, ha un valore che si può dire **"funzione" della funzione integranda** $f(x)$; esso, com'è noto, rappresenta l'area sottesa dalla curva $y = f(x)$, dall'asse x e dalle ordinate corrispondenti ai punti estremi a, b (Fig.10.2);

ii) presi nel piano (x, y) due punti $P_1(a_1, b_1)$ e $P_2(a_2, b_2)$, con $a_1 < a_2$, una linea regolare qualunque, di equazione $y = y(x)$, congiungente i punti P_1 e P_2, com'è noto, ha lunghezza uguale a

$$l[y] = \int_{a_1}^{a_2} \sqrt{1 + y'^2}\,dx.$$

Possiamo dire che questa lunghezza è **"funzione" della funzione** $y(x)$, e in questo caso, ci si può già porre un semplice problema di calcolo delle variazioni (Fig.10.3),

cioè: *nella famiglia* $[y]$ *costituita da tutte le linee* $y(x)$ *di estremi* P_1 *e* P_2, *continue nell'intervallo* $[a_1, a_2]$ *insieme alle loro derivate* $y'(x)$ *e soddisfacenti alle condizioni ai limiti:*

$$y(a_1) = b_1, \qquad y(a_2) = b_2,$$

determinare quella di **lunghezza minima.**
È evidente ch la linea che rende minimo $l[y]$ è costituita dal segmento P_1P_2.

Figura 10.3: Un semplice problema di minimo.

iii) il valore $y(x_1)$, in un dato punto x_1 di quella soluzione dell'equazione differenziale $y' = f(x, y)$ che verifica la condizione iniziale $y(x_0) = y_0$, è un numero che si può chiamare **"funzione"** della funzione $f(x, y)$.

È pertanto naturale porre il seguente concetto generale di funzione:
Dato un insieme E di elementi ξ di natura qualsiasi (numeri, punti, insiemi, funzioni di punto, funzioni di insieme, ecc.) se esiste una *una legge* che ad ogni elemento $\xi \in E$ faccia corrispondere un ben determinato numero u (reale o complesso), si dice che u è una *funzione reale o complessa* dell'elemento ξ, definita nell'insieme E.
Occorre però avvertire che, in questa accezione generale, è ormai invalso l'uso di sostituire alla parola *funzione* il termine **"funzionale"**: si dice cioè che u è un funzionale di ξ definito in E; l'elemento variabile ξ si chiama sovente **l'argomento** del funzionale e spesso si scrive $u = f[\xi]$.
Ciò premesso, è chiaro che nascono per i funzionali questioni analoghe a quelle che si considerano per le funzioni di punto; in particolare si può porre il problema della *ricerca dei massimi e dei minimi di un funzionale reale* (come abbiamo fatto più sopra nell'esempio ii), o più semplicemente, in via preliminare, della *ricerca delle condizioni di stazionarietà del detto funzionale*. Lo studio di tali problemi, costituisce quel capitolo dell'Analisi Matematica che si chiama *Calcolo delle variazioni* e che ha grande importanza sia teorica che applicativa. Si tratta di studi piuttosto difficili che da soli richiederebbero un intero corso. Noi qui ci limiteremo soltanto a richiamare alcune nozioni, fra le più elementari, necessarie per i nostri studi.[5]

10.8.2 Lemma fondamentale del Calcolo delle variazioni

Sia $\Phi(x)$ una funzione di classe C^0 definita nell'intervallo $[x_0, x_1]$ dell'asse x ed $\eta(x)$ una funzione pur essa definita in $[x_0, x_1]$ ed ivi di classe C^1, assoggettata o meno a delle condizioni agli estremi dell'intervallo.
Ha senso considerare l'integrale

$$J = \int_{x_0}^{x_1} \Phi(x)\eta(x)\, dx, \tag{10.56}$$

[5]Consigliamo al lettore di consultare qualche trattato di *Calcolo delle variazioni* per approfondire alcuni dei concetti che noi introduciamo necessariamente in maniera succinta. Si consiglia, a tal proposito, i seguenti due testi accessibili anche ad un principiante: Lev E. Elsgolts, *Equazioni differenziali e Calcolo delle variazioni*, Parte seconda, pag. 285 - 330, Editori Riuniti, Roma (1981) e Vladimir I. Smirnov, *Corso di Matematica Superiore*, Vol. IV, parte prima, pag. 204 - 312, Editori Riuniti, Roma (1980).
Per chi desideri una trattazione classica ma più rigorosa delle due precedenti, si consiglia: Edouard Goursat, *Cours d'Analyse Mathématique*, Tome III, Chap. XXXIV pag. 545 - 660, Gauthier-Villars, Paris (1949), translated from the original French edition by Dover Publications, Inc., as *A Course in Mathematical Analysis*, Vol.III, Part Two, pag. 229 - 352, New York (1964).
Una delle opere più moderne ed aggiornate sull'argomento è: Mariano Giaquinta e Stefan Hildebrandt, *Calculus of Variations*, Vol.I e Vol.II, Springer Verlag (2004).

il quale è *certamente nullo*, per *ogni* scelta della funzione $\eta(x)$, se è $\Phi(x) \equiv 0$. Ma vale anche il viceversa, sussiste infatti il seguente:
Lemma fondamentale del Calcolo delle variazioni:
Se, **per ogni scelta** *della funzione $\eta(x)$, soddisfacente alle condizioni*

$$\eta(x_0) = \eta(x_1) = 0, \tag{10.57}$$

si ha

$$\int_{x_0}^{x_1} \Phi(x)\eta(x)\, dx = 0, \tag{10.58}$$

è necessariamente

$$\Phi(x) \equiv 0. \tag{10.59}$$

Di questo teorema fondamentale ci limitiamo, per brevità, a darne l'enunciato, anche se la dimostrazione è piuttosto semplice. Il lettore può trovare tale dimostrazione nei testi citati nella nota a piè di pagina.
Del lemma precedente si può dare una generalizzazione immediata.
Lemma fondamentale generalizzato:
Supponiamo assegnate in uno stesso intervallo $[x_0, x_1]$ più funzioni continue $\Phi_1(x), \Phi_2(x), ..., \Phi_n(x)$, se è nullo l'integrale

$$J = \int_{x_0}^{x_1} [\Phi_1(x)\eta_1(x) + \Phi_2(x)\eta_2(x) + ... + \Phi_n(x)\eta_n(x)]\, dx = 0, \tag{10.60}$$

essendo $\eta_1(x), \eta_2(x), ..., \eta_n(x)$ **funzioni arbitrarie** *di classe C^1, nulle agli estremi dell'intervallo:*

$$\eta_i(x_0) = 0, \quad \eta_i(x_1) = 0, \quad (i = 1, 2, ..., n),$$

allora deve essere necessariamente:

$$\Phi_1(x) \equiv 0, \Phi_2(x) \equiv 0, ..., \Phi_n(x) \equiv 0, \tag{10.61}$$

in $[x_0, x_1]$.
Infatti, essendo le $\eta_i(x)$, $(i = 1, 2, ..., n)$ completamente arbitrarie, basta prenderle tutte nulle tranne una per ricadere nel caso precedente.
È chiaro che l'equivalenza delle condizioni di tipo locale (10.61) all'unica relazione di tipo integrale $J = 0$, è subordinata al fatto che le funzioni $\eta_1(x), \eta_2(x), ..., \eta_n(x)$ sono completamente arbitrarie.

10.8.3 Variazione prima di un integrale e metodo delle variazioni

Consideriamo la classe $\mathcal{C}$ delle funzioni $y(x)$ definite in un comune intervallo $[x_0, x_1]$ ivi continue insieme alle derivate prime e seconde, ed aventi valori prescritti agli estremi dell'intervallo. Immaginiamo poi assegnata una funzione $F(x, y, y')$ delle tre variabili x, y, y' definita in un certo dominio tridimensionale D dei tre argomenti detti.
Supponiamo inoltre che F sia sufficientemente regolare in D, ad esempio sia continua insieme con le sue derivate parziali prime e seconde. Detta $y(x)$ la generica funzione della classe $\mathcal{C}$, se il punto $[x, y(x), y'(x)]$ appartiene a D per ogni $x \in [x_0, x_1]$ ha senso considerare l'integrale

$$J[y] = \int_{x_0}^{x_1} F[x, y(x), y'(x)]\, dx, \tag{10.62}$$

il cui valore numerico dipende dalla funzione $y(x)$ che si è scelta in $\mathcal{C}$. Viene così definita una legge che ad ogni funzione $y(x)$ presa nella classe $\mathcal{C}$ fa corrispondere un ben determinato

numero reale, cioè, per quel che abbiamo detto alla fine del paragrafo 10.8.1, la (10.62) individua **un funzionale definito nella classe** $\mathcal{C}$.
Si tratta evidentemente di una generalizzazione del concetto ordinario di funzione, fungendo ora da variabile indipendente non una variabile numerica ma una funzione $y(x)$ nel suo complesso.
Per funzionali del tipo ora considerato si può porre un problema analogo a quello che si presenta per le funzioni ordinarie, cioè la ricerca dei **massimi e minimi assoluti** di $J[y]$ al variare della funzione $y(x)$, ovvero della curva $y = y(x)$, in $\mathcal{C}$. Più precisamente, fissata la classe $\mathcal{C}$ delle funzioni in cui si pensa definito il funzionale, ci si può domandare qual'è (e se esiste) la funzione che attribuisce a J il massimo o minimo valore, o più in generale, secondo una locuzione che preciseremo tra poco, lo rende **stazionario**. Naturalmente, come nel caso delle funzioni di punto, si può cominciare con la ricerca delle curve $y = y(x)$ che forniscono per $J[y]$ un *massimo ed un minimo relativo*. È opportuno però osservare che i funzionali del Calcolo delle Variazioni non sono generalmente continui (ma semicontinui), perciò per essi non si può applicare l'ordinario teorema di Weierstrass per la ricerca del massimo e del minimo assoluti di $J[y]$, ma una notevole sua generalizzazione.

Introduciamo ora il concetto di **variazione prima** del funzionale $J[y]$ ed un procedimento noto come **metodo delle variazioni**.
Sia $\overline{y}(x)$ una *particolare* funzione della classe $\mathcal{C}$. Ogni altra funzione $y(x)$ di $\mathcal{C}$ potrà mettersi nella forma

$$y(x) = \overline{y}(x) + \delta y(x), \tag{10.63}$$

$\delta y(x)$ essendo un'arbitraria funzione di classe C^2 in $[x_0, x_1]$, nulla agli estremi dell'intervallo, cioè tale che.

$$[\delta y(x)]_{x=x_0} = [\delta y(x)]_{x=x_1} = 0. \tag{10.64}$$

Allo scopo di passare da funzionali a funzioni ordinarie attribuiamo a $\delta y(x)$ la forma $\delta y(x) = \alpha\eta(x)$ dove la funzione $\eta(x)$ è un'arbitraria funzione di classe C^2 in $[x_0, x_1]$ che per la (10.64) si annulla agli estremi dell'intervallo, cioè:

$$\eta(x_0) = \eta(x_1) = 0,$$

ed α un parametro numerico infinitesimo soggetto alla condizione $-\varepsilon \leq \alpha \leq \varepsilon$, con $\varepsilon > 0$ arbitrariamente piccolo (Fig.10.4).
La (10.63) potrà dunque mettersi anche nella forma:

$$y(x, \alpha) = \overline{y}(x) + \alpha\eta(x), \tag{10.65}$$

che per $\alpha = 0$ si riduce alla $y(x, 0) \equiv \overline{y}(x)$.

Fissate le due funzioni $\overline{y}(x)$ ed $\eta(x)$ il funzionale $J[y]$, per la (10.62), si riduce ad una funzione ordinaria della variabile α:

$$J(\alpha) = \int_{x_0}^{x_1} F[x, \overline{y}(x) + \alpha\eta(x), \overline{y'}(x) + \alpha\eta'(x)]\, dx. \tag{10.66}$$

Naturalmente, per $\eta(x)$ fissata, $y(x) = \overline{y}(x) + \alpha\eta(x)$ non esaurisce generalmente, al variare di α, la classe $\mathcal{C}$. Calcoliamo ora la derivata di $J(\alpha)$ per $\alpha = 0$, e consideriamo la quantità

$$\delta J = \alpha \left[\frac{dJ(\alpha)}{d\alpha}\right]_{\alpha=0}. \tag{10.67}$$

Esplicitando la derivata sotto il segno di integrale si ottiene

$$\delta J = \alpha \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta'\right]_{\alpha=0} dx,$$

Figura 10.4: Metodo delle variazioni per un funzionale.

ovvero, integrando per parti sul secondo termine dell'integrale precedente, avremo:

$$\int_{x_0}^{x_1} \frac{\partial F}{\partial y'} \eta' dx = \left[\frac{\partial F}{\partial y'} \eta\right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} \frac{\partial F}{\partial y'} \eta dx.$$

Essendo $\eta(x_0) = \eta(x_1) = 0$, si avrà $\left[\frac{\partial F}{\partial y'} \eta\right]_{x_0}^{x_1} = 0$, perciò:

$$\int_{x_0}^{x_1} \frac{\partial F}{\partial y'} \eta' dx = -\int_{x_0}^{x_1} \frac{d}{dx} \frac{\partial F}{\partial y'} \eta dx,$$

ed in definitiva, tenuto conto che $\alpha\eta(x) = \delta y(x)$ si ottiene:

$$\delta J = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'}\right]_{y=\overline{y}, y'=\overline{y}'} \delta y(x) dx,$$

ovvero:

$$\delta J = -\int_{x_0}^{x_1} \left[\frac{d}{dx} \frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y}\right]_{y=\overline{y}, y'=\overline{y}'} \delta y(x) dx. \tag{10.68}$$

La quantità ora calcolata suol chiamarsi **la variazione prima del funzionale** $J[y]$ a partire da $\overline{y}(x)$ in corrispondenza alla variazione $\delta y(x) = \alpha\eta(x)$ della funzione indipendente. *Essa, nella forma* (10.67) *che la definisce, è manifestamente la spontanea generalizzazione del concetto di differenziale primo dell'analisi ordinaria.*

10.8.4 Equazione di Eulero associata ad un principio variazionale

Def.1[a]: *Si dice che una funzione* $y(x)$[6] *rende* **stazionario** *l'integrale* $J[y]$ *se, a partire da essa, risulta*

$$\delta J = 0 \tag{10.69}$$

per ogni scelta della variazione $\delta y(x)$.
È ovvio che tale condizione è verificata, in particolare, nel caso che $y(x)$ renda massimo o minimo l'integrale $J[y]$. Più in generale, la condizione globale (10.69) equivale, in virtù della (10.68) e del lemma fondamentale del Calcolo delle variazioni, alla condizione locale

$$\frac{d}{dx} \frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y} = 0, \tag{10.70}$$

cui deve soddisfare la $y(x)$.
Viceversa, ogni soluzione dell'equazione differenziale ora scritta rende ovviamente nulla δJ,

[6] Abbiamo tolto il soprassegno alla $\overline{y}$ ormai inutile.

sicché le condizioni (10.69) e (10.70) sono completamente equivalenti.
L'equazione (10.70) prende il nome di *equazione di Eulero (o euleriana) del principio variazionale* $\delta J = 0$.
Anche nel Calcolo delle Variazioni si presenta poi il problema di riconoscere sotto quali ulteriori condizioni sia assicurata l'esistenza di un massimo o di un minimo. Noi però non ci occuperemo di tale importante problema che comporterebbe l'analisi delle variazioni successive del funzionale $J[y]$, ma ci limiteremo a studiare le condizioni di stazionarietà, che, già da sole esauriscono un primo problema sul comportamento variazionale del funzionale. Il lettore tenga comunque presente che tali condizioni sufficienti sono state date sotto varie forme da Legendre, Jacobi, Weierstrass, ed altri e possono essere trovate nei testi citati.
Esplicitando la derivata totale la (10.70) si scrive

$$\frac{\partial^2 F}{\partial y'^2} y'' + \frac{\partial^2 F}{\partial y \partial y'} y' + \frac{\partial^2 F}{\partial x \partial y'} - \frac{\partial F}{\partial y} = 0. \tag{10.71}$$

Si tratta dunque di un'equazione differenziale ordinaria del 2° ordine nella funzione incognita $y(x)$, la cui soluzione è univocamente determinata non appena si tenga conto delle condizioni fissate agli estremi

$$y(x_0) = y_0, \quad y(x_1) = y_1, \quad [\delta y(x)]_{x=x_0} = [\delta y(x)]_{x=x_1} = 0. \tag{10.72}$$

Aggiungiamo ancora qualche osservazione sulla terminologia che spesso viene usata nel Calcolo delle Variazioni, ponendo le seguenti definizioni.
Def.2^a: *Si dice che una funzione $y(x)$ (ovvero la curva $y = y(x)$) è* **un'estremante** *del funzionale $J[y]$ se è di massimo o di minimo relativo.*
Def.3^a: *Le curve integrali dell'equazione di Eulero* (10.70) *si chiamano invece* **le curve estremali** *del funzionale $J[y]$.*
Si può perciò dire che le *curve estremanti* vanno ricercate fra le *curve estremali* del funzionale $J[y]$.

10.8.5 Equazioni di Eulero nel caso di un funzionale dipendente da più argomenti

Le considerazioni precedenti si estendono subito a funzionali dipendenti da più argomenti indipendenti del tipo

$$J[y^1(x), \ldots, y^n(x)] = \int_{x_0}^{x_1} F[x, y^1(x), \ldots, y^n(x), y'^{\,1}(x), \ldots, y'^{\,n}(x)] dx, \tag{10.73}$$

il cui dominio di definizione sia la famiglia $\phi = \{\gamma_\alpha\}$ di tutte le curve reali di classe C^2 definite nell'intervallo $[x_0, x_1]$ giacenti sulla varietà M^n e passanti per gli stessi punti estremi P_0 e P_1, così definite:

$$\gamma_\alpha : y^i(x, \alpha) = f^i(x) + \alpha \eta^i(x), \quad i = 1, 2, \ldots, n, \tag{10.74}$$

dove $f^i(x), \eta^i(x) \in C^2[x_0, x_1]$ e $\eta^i(x_0) = \eta^i(x_1) = 0$ ed inoltre $\alpha \in [-\varepsilon, \varepsilon]$ è un parametro infinitesimo con $\varepsilon > 0$. Sussiste il seguente teorema che enunciamo senza dimostrazione che lo studioso interessato può comunque reperire nei libri già citati.

Teorema 10.4 (Teorema fondamentale del calcolo delle variazioni, o di Eulero) *Condizione necessaria affinché la curva $\gamma_0 \in \phi$, $\gamma_0 : y^i(x, 0) = f^i(x)$, renda stazionario il funzionale* (10.73) *ovvero verifichi la condizione*

$$\delta J \equiv \alpha \left[\frac{dJ}{d\alpha} \right]_{\alpha=0} = 0, \tag{10.75}$$

in corrispondenza a variazioni arbitrarie $\delta y^1(x), \ldots, \delta y^n(x)$, nulle agli estremi, delle funzioni indipendenti, è che essa sia una **curva estremale** *del funzionale, cioè una soluzione del sistema (di Eulero)*

$$\frac{d}{dx} \left(\frac{\partial F}{\partial y'^{\,i}} \right) - \frac{\partial F}{\partial y^i} = 0, \quad i = 1, 2, \ldots, n. \tag{10.76}$$

10.8.6 Digressione sul Principio di Hamilton

I richiami che seguono sul Principio variazionale di Hamilton, riferito ai *sistemi olonomi* nell'ambito della meccanica lagrangiana, pur non essendo essenziali per il seguito, sono molto utili per la nostra trattazione delle geodetiche di una varietà riemanniana che faremo nel prossimo paragrafo.
Dalle considerazioni fatte nel paragrafo precedente, si riconosce subito che le equazioni di moto di un generico *sistema olonomo* S a vincoli bilaterali privi di attrito, soggetto ad una sollecitazione conservativa, corrispondono alle equazioni di Eulero di un *principio variazionale*. Precisamente, dette q^h le coordinate lagrangiane di S, $\dot{q}^h$ le loro derivate rispetto al tempo, le equazioni di moto del sistema considerato (equazioni di Lagrange) si scrivono, com'è noto, nella forma:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}}{\partial q^h} = 0, \quad h = 1, 2, \ldots, n. \tag{10.77}$$

dove $\mathcal{L} = T - V$ è la cosiddetta *lagrangiana* del sistema, definita (nel caso dinamico) come differenza fra l'energia cinetica $T(t, q^h, \dot{q}^h)$ e l'energia potenziale $V(q^h)$, oppure per i cosiddetti *sistemi lagrangiani generali*, è una funzione assegnata $\mathcal{L}(t, q^h, \dot{q}^h)$ delle $q^h, \dot{q}^h$ e del tempo t. Il confronto con le (10.76) mette bene in evidenza che le (10.77) sono equivalenti al *Principio variazionale di Hamilton* :

$$\delta \int_{t_0}^{t_1} \mathcal{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) dt = 0. \tag{10.78}$$

Figura 10.5: Illustrazione grafica del Principio di Hamilton.

In termini più precisi il principio medesimo si enuncia così (vedi Fig.10.5)
Principio di Hamilton: *Se C_0 e C_1 sono due configurazioni possibili per S, fra tutti i moti che il sistema può compiere, compatibilmente con i vincoli nell'intervallo temporale $[t_0, t_1]$ per portarsi da C_0 a C_1, quello effettivo corrisponde alla n-pla $q^h(t)$ che rende stazionario l'integrale*

$$\int_{t_0}^{t_1} \mathcal{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) dt \quad (\textit{Azione Hamiltoniana}).$$

Seguendo l'analogia fra le equazioni di Eulero (10.76) e le equazioni di Lagrange (10.77) è possibile, in maniera piuttosto semplice, com'è noto allo studioso dal Corso di Meccanica Razionale, mettere in evidenza certi integrali primi tipici del sistema lagrangiano che possono agevolare l'integrazione, o comunque la discussione, del sistema stesso. Tali considerazioni valgono naturalmente anche nel caso generale in cui la funzione lagrangiana $\mathcal{L}$ è sostituita da una qualunque funzione assegnata $F[t, q^1(t), \ldots, q^n(t), \dot{q}^1(t), \ldots, \dot{q}^n(t)]$ dei propri argomenti.

A. **Integrali primi generalizzati delle quantità di moto**
Cominciamo col supporre che la funzione $\mathcal{L}$ non dipenda esplicitamente da una delle funzioni incognite, ad esempio q^k.
Allora è chiaro che sarà $\dfrac{\partial \mathcal{L}}{\partial q^k} = 0$ e quindi, in virtù della k-ma equazione (10.77),

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) = 0, \quad h = 1, 2, \ldots, n;$$

il sistema (10.77) ammette l'integrale primo

$$\frac{\partial \mathcal{L}}{\partial \dot{q}^h} = cost. \tag{10.79}$$

Di tali integrali se ne avranno tanti quante sono le q che non compaiono esplicitamente in $\mathcal{L}$ (*coordinate ignorabili*): essi si chiamano *integrali primi generalizzati delle quantità di moto.*[7]

B. **Integrale primo generalizzato dell'energia**
Un altro integrale primo notevole è il cosiddetto *integrale primo generalizzato dell'energia*, valido nel caso che la lagrangiana sia esplicitamente indipendente dal tempo t. Per metterlo in evidenza introduciamo la cosiddetta *hamiltoniana* funzione delle $2n+1$ variabili $(t, q^h, \dot{q}^h)$:

$$\mathcal{H} = \sum_{h=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}^h}\dot{q}^h - \mathcal{L}, \tag{10.80}$$

e calcoliamo la derivata totale rispetto ad t immaginando di sostituire alle q e $\dot{q}$ una qualunque soluzione del sistema (10.77). Avremo

$$\frac{d\mathcal{H}}{dt} = \sum_{h=1}^{n}\left(\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\dot{q}^h + \frac{\partial \mathcal{L}}{\partial \dot{q}^h}\ddot{q}^h - \frac{\partial \mathcal{L}}{\partial q^h}\dot{q}^h - \frac{\partial \mathcal{L}}{\partial \dot{q}^h}\ddot{q}^h\right) - \frac{\partial \mathcal{L}}{\partial t}$$

e semplificando si ottiene

$$\frac{d\mathcal{H}}{dt} = \sum_{h=1}^{n}\left(\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^h} - \frac{\partial \mathcal{L}}{\partial q^h}\right)\dot{q}^h - \frac{\partial \mathcal{L}}{\partial t}. \tag{10.81}$$

Se ora teniamo presente che devono valere le equazioni di Lagrange otteniamo l'identità

$$\frac{d\mathcal{H}}{dt} = -\frac{\partial \mathcal{L}}{\partial t}. \tag{10.82}$$

Se allora $\mathcal{L}$ non dipende esplicitamente dal tempo t, cioè $\mathcal{L}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$, sarà $\dfrac{\partial \mathcal{L}}{\partial t} = 0$ e perciò lungo ogni soluzione del sistema (10.77) risulterà:

$$\mathcal{H}\Big(q^1(t), \ldots, q^n(t), \dot{q}^1(t), \ldots, \dot{q}^n(t)\Big) = cost. \tag{10.83}$$

L'integrale primo ora scritto prende il nome, come già abbiamo detto, di *integrale generalizzato dell'energia*: la denominazione trae origine dalla dinamica dei sistemi olonomi. Infatti nel caso che la funzione lagrangiana $\mathcal{L}$ non dipenda esplicitamente dal tempo (sistemi a vincoli indipendenti dal tempo), l'energia cinetica T del sistema si riduce ad una

[7] La denominazione dipende dal fatto che, nel caso di un punto materiale libero P, adottando coordinate cartesiane risulta $\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z)$ e quindi $\dfrac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$ che è appunto la componente cartesiana della quantità di moto relativa ad x.

forma quadratica del tipo $T = \frac{1}{2}\sum_{h,k=1}^{n} a_{hk}(q)\dot{q}^h\dot{q}^k$. Per il teorema di Eulero sulle funzioni omogenee[8] sarà:

$$\sum_{h=1}^{n} \frac{\partial T}{\partial \dot{q}^h}\dot{q}^h = 2T. \tag{10.84}$$

In tal caso, per la (10.80), si ottiene

$$\mathcal{H} = \sum_{h=1}^{n} \frac{\partial T}{\partial \dot{q}^h}\dot{q}^h - T + V = T + V \quad \text{(Energia meccanica totale)}. \tag{10.85}$$

C. **Caso di una lagrangiana omogenea nelle $\dot{q}^h$**
Occupiamoci infine di due Lemmi importanti che ci consentiranno di semplificare il calcolo delle equazioni differenziali delle geodetiche di una varietà riemanniana.
Cominciamo col dimostrare il seguente:
Lemma 1:
Ammettiamo che la funzione $\mathcal{L}$ non dipenda esplicitamente dal tempo t, sia cioè $\mathcal{L}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ e sia inoltre omogenea di un certo grado $r \neq 1$ nelle variabili $\dot{q}^h$, allora il sistema differenziale di Eulero-Lagrange

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}}{\partial q^h} = 0, \quad (h = 1, 2, \ldots, n), \tag{10.86}$$

ammette l'integrale generalizzato dell'energia:

$$\mathcal{L}\Big(q^1(t), \ldots, q^n(t), \dot{q}^1(t), \ldots, \dot{q}^n(t)\Big) = cost. \tag{10.87}$$

Dimostrazione:
Per il teorema di Eulero sulle funzioni omogenee si avrà:

$$\sum_{h=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}^h}\dot{q}^h = r\mathcal{L}, \quad (r \neq 1)$$

e quindi, per la (10.80), avremo

$$\mathcal{H} = (r-1)\mathcal{L} \quad (r \neq 1)$$

ciò che riduce l'integrale primo generalizzato dell'energia (10.83) semplicemente a

$$\mathcal{L}\Big(q^1(t), \ldots, q^n(t), \dot{q}^1(t), \ldots, \dot{q}^n(t)\Big) = cost. \tag{10.88}$$

È facile riconoscere che vale anche il seguente:
Lemma 2:
Se la lagrangiana $\mathcal{L}$ non dipende esplicitamente dal tempo t ed inoltre sia omogenea di grado

[8] Una funzione di più variabili $f(x^1, x^2, ..., x^n)$ si dice *omogenea di grado r* se, essendo λ un parametro arbitrario è soddisfatta l'identità:

$$f(\lambda x^1, \lambda x^2, ..., \lambda x^n) = \lambda^r f(x^1, x^2, ..., x^n).$$

A tal proposito sussiste il seguente **teorema di Eulero**: *C.N.E.S. affinché $f(x^1, x^2, ..., x^n)$ sia omogenea di grado r è che sia soddisfatta l'identità:*

$$\sum_{h=1}^{n} \frac{\partial f}{\partial x^h}x^h = rf(x^1, x^2, ..., x^n).$$

$r \neq 1$ nelle variabili $\dot{q}^h$, e se $m \neq 0$ è un numero reale arbitrario, allora i sistemi differenziali di Eulero-Lagrange associati alle due lagrangiane $\mathcal{L}$ ed $\mathcal{L}^m$ sono equivalenti

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}}{\partial q^h} = 0, \Longleftrightarrow \frac{d}{dt}\left(\frac{\partial \mathcal{L}^m}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}^m}{\partial q^h} = 0 \quad (h = 1, 2, \ldots, n), \tag{10.89}$$

ammettono cioè le stesse soluzioni.
Dim.
Consideriamo il binomio lagrangiano

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}^m}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}^m}{\partial q^h}, \tag{10.90}$$

esplicitando le derivate che vi compaiono avremo:

$$\frac{\partial \mathcal{L}^m}{\partial \dot{q}^h} = m\mathcal{L}^{m-1}\frac{\partial \mathcal{L}}{\partial \dot{q}^h}; \qquad \frac{\partial \mathcal{L}^m}{\partial q^h} = m\mathcal{L}^{m-1}\frac{\partial \mathcal{L}}{\partial q^h},$$

perciò:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}^m}{\partial \dot{q}^h}\right) = \frac{d}{dt}\left[m\mathcal{L}^{m-1}\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right] = m(m-1)\mathcal{L}^{m-2}\frac{d\mathcal{L}}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) + m\mathcal{L}^{m-1}\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right).$$

Il binomio lagrangiano (10.90) si scriverà perciò così:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}^m}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}^m}{\partial q^h} \equiv m(m-1)\mathcal{L}^{m-2}\frac{d\mathcal{L}}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) + m\mathcal{L}^{m-1}\left[\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}}{\partial q^h}\right]. \tag{10.91}$$

Per le ipotesi fatte, in virtù del **Lemma 1**, avremo che

$$\mathcal{L}\Big(q^1(t), \ldots, q^n(t), \dot{q}^1(t), \ldots, \dot{q}^n(t)\Big) = cost.$$

è un integrale primo del 1° sistema che figura nella (10.89), il che implica $\dfrac{d\mathcal{L}}{dt} = 0$, ma allora l'identità (10.91) si riduce all'altra identità:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}^m}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}^m}{\partial q^h} \equiv m\mathcal{L}^{m-1}\left[\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) - \frac{\partial \mathcal{L}}{\partial q^h}\right]. \tag{10.92}$$

La (10.92) implica che se è soddisfatta la $\dfrac{d}{dt}\left(\dfrac{\partial \mathcal{L}}{\partial \dot{q}^h}\right) - \dfrac{\partial \mathcal{L}}{\partial q^h} = 0$ allora risulta anche $\dfrac{d}{dt}\left(\dfrac{\partial \mathcal{L}^m}{\partial \dot{q}^h}\right) - \dfrac{\partial \mathcal{L}^m}{\partial q^h} = 0$ e viceversa, cioè ogni soluzione del 1° sistema che figura nella (10.89) è anche soluzione del 2° sistema e viceversa. c.d.d

10.9 Le geodetiche di una varietà riemanniana

In uno spazio tridimensionale euclideo, *una retta è il cammino più breve tra due punti*; noi vogliamo ora generalizzare questo concetto alle varietà riemanniane: si tratta del concetto fondamentale di **geodetica**.
In effetti abbiamo già incontrato questo tipo notevole di linee, nel caso di una superficie, nel capitolo sul parallelismo di Levi-Civita, ed abbiamo anche illustrato alcune semplici proprietà geometriche delle *geodetiche* con degli esempi, il più importante dei quali è che: *le geodetiche di una sfera sono i suoi cerchi massimi.*

Ora dobbiamo affrontare in tutta la sua generalità e dal punto di vista analitico la nozione di **geodetica di una varietà riemanniana**.
Sia dunque M^n una varietà riemanniana ad n dimensioni di *metrica*

$$\mathbf{ds}^2 = g_{hk}(x)dx^h dx^k, \qquad (g_{hk} = g_{kh}), \quad (h, k = 1, 2, ..., n) \tag{10.93}$$

Il $ds = \sqrt{|\mathbf{ds}^2|}$ si chiama anche *l'elemento lineare* della varietà.
Consideriamo una qualunque n-pla di funzioni di classe C^2

$$x^i = x^i(\lambda), \qquad \lambda_1 \leq \lambda \leq \lambda_2, \quad (i = 1, 2, ..., n) \tag{10.94}$$

soggetta alla condizione di assumere per $\lambda = \lambda_1$, $\lambda = \lambda_2$ valori prescritti; consideriamo cioè un arco di curva regolare γ che, su M^n, unisca due punti prefissati P_1, P_2.[9]
Diamo ora la seguente definizione fondamentale (Fig.10.6):

Figura 10.6: Geodetiche di una varietà riemanniana.

Definizione: *Diconsi* **geodetiche** *della data varietà riemanniana* (M^n, g) *le linee di questa varietà che rendono* **stazionario (o addirittura minimo)** *l'integrale dell'elemento di linea fra due punti* P_1, P_2 *prefissati su* M^n*, verificanti cioè la seguente equazione variazionale*:

$$\delta \int_{P_1}^{P_2} ds = 0. \tag{10.95}$$

Se $\mathbf{ds^2} > 0$, cioè se *la varietà è propriamente riemanniana*, il suo elemento lineare sarà

$$ds = \sqrt{g_{hk}(x)dx^h dx^k}, \tag{10.96}$$

e perciò sulle curve γ, di equazioni parametriche (10.94), l'ascissa curvilinea sarà funzione del parametro λ e potremo perciò scrivere la derivata di $s(\lambda)$ rispetto a λ così

$$\frac{ds}{d\lambda} = \sqrt{g_{hk}(x)\dot{x}^h \dot{x}^k}, \tag{10.97}$$

avendo indicato, per semplicità di scrittura, $\dfrac{dx^h}{d\lambda} = \dot{x}^h$. L'integrale che figura nella (10.95) potrà allora scriversi equivalentemente come

$$\delta \int_{P_1}^{P_2} ds \Longleftrightarrow \delta \int_{\lambda_1}^{\lambda_2} \frac{ds}{d\lambda} d\lambda = \delta \int_{\lambda_1}^{\lambda_2} \sqrt{g_{hk}(x)\dot{x}^h \dot{x}^k} d\lambda = 0, \tag{10.98}$$

e se poniamo

$$\mathcal{L}(x, \dot{x}) \stackrel{\text{def}}{=} g_{hk}(x)\dot{x}^h \dot{x}^k, \tag{10.99}$$

[9] λ è un parametro arbitrario su γ che, volendo, può identificarsi o col tempo t o con l'ascissa curvilinea s su γ.

l'equazione variazionale (10.98) delle geodetiche diventa

$$\delta \int_{\lambda_1}^{\lambda_2} \mathcal{L}^{1/2} ds = 0. \tag{10.100}$$

La funzione $\mathcal{L}$, che possiamo interpretare come una lagrangiana, non dipende esplicitamente da λ ed è una funzione omogenea di grado 2 nelle $(\dot{x}^1, \dot{x}^2, ..., \dot{x}^n) \equiv (\dot{x}^h)$, dunque è applicabile il **Lemma 2**, che abbiamo dimostrato nel paragrafo precedente, e che ora ci consente di scrivere la (10.100) così

$$\delta \int_{\lambda_1}^{\lambda_2} \mathcal{L}^{1/2} ds = 0 \Longleftrightarrow \delta \int_{\lambda_1}^{\lambda_2} \mathcal{L} ds = 0. \tag{10.101}$$

Ciò significa che i sistemi di equazioni di Eulero-Lagrange associati alle precedenti equazioni variazionali (10.101), ammettono le stesse soluzioni; possiamo dunque concludere che: *tutte le soluzioni dell'equazione variazionale* $\delta \int_{P_1}^{P_2} ds = 0$, *che definisce le geodetiche di* M^n *sono esattamente le stesse della* $\delta \int_{\lambda_1}^{\lambda_2} \mathcal{L}^{1/2} ds = 0$, *cioè le geodetiche di* M^n *possono essere definite anche con la:*

$$\delta \int_{\lambda_1}^{\lambda_2} \mathcal{L} ds = 0. \tag{10.102}$$

Quest'ultima definizione ha un duplice vantaggio rispetto alla (10.95) ovvero alla (10.98):

1°) ci fa determinare le equazioni differenziali delle geodetiche con calcoli più semplici non figurando in essa la metrica sotto il segno di radice quadrata;

2°) con la (10.102) non c'è più la necessità di avere un $\mathbf{ds}^2 > 0$, come per la (10.98), in quanto la (10.102) ha senso qualunque sia $\mathbf{ds^2} \neq 0$.

Possiamo ora finalmente procedere alla determinazione delle equazioni differenziali delle geodetiche di M^n, utilizzando l'equazione variazionale (10.102): sappiamo che essa è equivalente al sistema di equazioni di Eulero- Lagrange:

$$\frac{d}{d\lambda}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right) - \frac{\partial \mathcal{L}}{\partial x^i} = 0, \quad (i = 1, 2, \ldots, n). \tag{10.103}$$

Calcoliamoci prima le derivate parziali che figurano nella (10.103), avremo:

$$\frac{\partial \mathcal{L}}{\partial \dot{x}^i} = \frac{\partial}{\partial \dot{x}^i}[g_{hk}(x)\dot{x}^h\dot{x}^k] = g_{hk}(x)\delta_i^h\dot{x}^k + g_{hk}(x)\dot{x}^h\delta_i^k = g_{ik}(x)\dot{x}^k + g_{hi}(x)\dot{x}^h = 2g_{ik}(x)\dot{x}^k, \tag{10.104}$$

$$\frac{\partial \mathcal{L}}{\partial x^i} = \frac{\partial}{\partial x^i}[g_{hk}(x)\dot{x}^h\dot{x}^k] = \partial_i g_{hk}(x)\dot{x}^h\dot{x}^k. \tag{10.105}$$

Per non appesantire troppo le scritture sopprimiamo nei termini contenenti il tensore metrico $g_{hk}(x)$ la x scrivendo più semplicemente g_{hk}, ricordandoci comunque che il tensore metrico è generalmente funzione di tutte le coordinate x^h. Introducendo ora le (10.104), (10.105) nelle (10.103) otteniamo:

$$\frac{d}{d\lambda}[2g_{ik}\dot{x}^k] - \partial_i g_{hk}\dot{x}^h\dot{x}^k = 0,$$

da cui sviluppando i calcoli:

$$2\partial_j g_{ik}\dot{x}^j\dot{x}^k + 2g_{ik}\ddot{x}^k - \partial_i g_{hk}\dot{x}^h\dot{x}^k = 0,$$

dividendo, membro a membro, per 2 e scrivendo come primo termine quello contenente la derivata seconda di x, cioè $g_{ik}\ddot{x}^k$ avremo:

$$g_{ik}\ddot{x}^k + \partial_j g_{ik}\dot{x}^j\dot{x}^k - \frac{1}{2}\partial_i g_{hk}\dot{x}^h\dot{x}^k = 0.$$

Sdoppiando il 2° termine si ottiene:

$$g_{ik}\ddot{x}^k + \frac{1}{2}\partial_j g_{ik}\dot{x}^j\dot{x}^k + \frac{1}{2}\partial_j g_{ik}\dot{x}^j\dot{x}^k - \frac{1}{2}\partial_i g_{hk}\dot{x}^h\dot{x}^k = 0,$$

se, infine, scambiamo j con k nel 3° termine e sostituiamo h con j nell'ultimo termine, la precedente equazione diventa:

$$g_{ik}\ddot{x}^k + \frac{1}{2}\Big(\partial_j g_{ik} + \partial_k g_{ij} - \partial_i g_{jk}\Big)\dot{x}^j\dot{x}^k = 0.$$

Nel secondo termine di questa equazione si riconosce il simbolo di Christoffel di 1ª specie

$$\Gamma_{jk,i} \stackrel{\text{def}}{=} \frac{1}{2}\Big(\partial_j g_{ik} + \partial_k g_{ij} - \partial_i g_{jk}\Big)$$

già incontrato e studiato nell'ambito del parallelismo di Levi-Civita.
La precedente equazione si scrive allora, più sinteticamente così:

$$g_{ik}\ddot{x}^k + \Gamma_{jk,i}\dot{x}^j\dot{x}^k = 0,$$

infine moltiplicando 1° e 2° membro per il tensore metrico g^{ih} sotto forma controvariante e tenendo conto della sua simmetria, con facili calcoli, si ottiene:

$$\ddot{x}^h + \Gamma^h_{jk}\dot{x}^j\dot{x}^k = 0 \Longleftrightarrow \frac{d^2x^h}{d\lambda^2} + \Gamma^h_{jk}\frac{dx^j}{d\lambda}\frac{dx^k}{d\lambda} = 0, \tag{10.106}$$

essendo le

$$\Gamma^h_{jk} \stackrel{\text{def}}{=} g^{ih}\Gamma_{jk,i} = \frac{1}{2}g^{ih}\Big(\partial_j g_{ki} + \partial_k g_{ji} - \partial_i g_{jk}\Big)$$

i simboli di Christoffel di 2^a specie o *coefficienti di connessione* della varietà M^n.
Le (10.106) sono le cercate *equazioni differenziali delle* **geodetiche** *della varietà riemanniana* M^n.
Provenendo da una funzione $\mathcal{L}$ che non dipende esplicitamente dal parametro λ ed è omogenea di 2° grado nelle $\dot{x}^h$, per il **Lemma 1** del paragrafo precedente, esse *ammettono l'integrale generalizzato dell'energia* nella forma

$$\mathcal{L} \equiv g_{ik}(x)\dot{x}^i\dot{x}^k = cost. \tag{10.107}$$

Se si sceglie come parametro l'ascissa curvilinea s, tale integrale primo, si semplifica nella

$$g_{ik}(x)\frac{dx^i}{ds}\frac{dx^k}{ds} = \frac{\mathbf{ds}^2}{\mathbf{ds}^2} = 1, \tag{10.108}$$

la costante d'integrazione riducendosi ad 1.
Osservazione:
È noto che per una curva γ a $\mathbf{ds}^2$ di segno costante, ad esempio positivo ($\mathbf{ds}^2 = ds^2 > 0$),[10] si

[10]Ricordiamo al lettore che già nell'ambito della Teoria della Relatività Ristretta, e precisamente nel paragrafo 5.4.2 definimmo un 4-vettore infinitesimo $\mathbf{ds} \equiv (dx^i)$ come uno spostamento da un dato evento ad un altro infinitamente vicino in $\mathcal{M}^4$ e la sua norma venne così definita

$$||\mathbf{ds}|| \stackrel{\text{def}}{=} \mathbf{ds}\cdot\mathbf{ds} \equiv \mathbf{ds}^2 = g_{ik}(x)dx^i dx^k \gtreqless 0, \qquad (i,k = 1,2,\ldots,n),$$

mentre il suo modulo, chiamato *elemento lineare*, fu definito così

$$ds \stackrel{\text{def}}{=} \sqrt{|\mathbf{ds}^2|} \geq 0.$$

Le stesse definizioni possiamo adottarle nel caso, che ora ci interessa, di uno *spostamento infinitesimo* $\mathbf{ds} \equiv (dx^i)$ su di una varietà riemanniana M^n.
Perciò, nel caso che $\mathbf{ds}^2 > 0$, possiamo scrivere la metrica nella forma scalare ds^2 senza alcuna ambiguità in quanto, in tal caso, $ds^2 = \mathbf{ds}^2$, ciò che non è vero se $\mathbf{ds}^2 < 0$: in quest'ultima circostanza è invece $ds^2 = -\mathbf{ds}^2 > 0$.

può introdurre un'ascissa curvilinea $s(\lambda)$, semplicemente ponendo

$$s(\lambda) = \int_{\lambda_1}^{\lambda} \sqrt{g_{ik}[x(\lambda)]\dot{x}^i\dot{x}^k}d\lambda,$$

con λ_1 fissato e λ variabile.
Poiché risulta $\dfrac{ds}{d\lambda} = \sqrt{g_{ik}[x(\lambda)]\dot{x}^i\dot{x}^k}$, la (10.107) ci assicura che lungo una geodetica si ha

$$\frac{ds}{d\lambda} = cost., \tag{10.109}$$

cioè $s = a\lambda + b$, con a, b costanti. Se si vuole che $s = 0$ corrisponda a $\lambda = 0$, deve essere $b = 0$ e perciò $s(\lambda) = a\lambda$, cioè $s(\lambda)$ è proporzionale a λ. Pertanto il parametro λ può sostanzialmente interpretarsi, per tali geodetiche (a $ds^2 \neq 0$) come ascissa curvilinea. D'altra parte se $s = a\lambda + b \Longleftrightarrow \dfrac{ds}{d\lambda} = a = cost. \neq 0$ si avrà:

$$\frac{dx^i}{d\lambda} = \frac{dx^i}{ds}\frac{ds}{d\lambda} = a\frac{dx^i}{ds}; \qquad \frac{d^2x^i}{d\lambda^2} = \frac{d}{d\lambda}\left(a\frac{dx^i}{ds}\right) = a\frac{d^2x^i}{ds^2}\frac{ds}{d\lambda} = a^2\frac{d^2x^i}{ds^2}, \tag{10.110}$$

e perciò le equazioni differenziali delle geodetiche (10.106) si potranno scrivere anche nella forma:

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{hk}\frac{dx^h}{ds}\frac{dx^k}{ds} = 0, \quad (i, h, k = 1, 2, ..., n). \tag{10.111}$$

Se poi il parametro λ si identifica col tempo t, $(\lambda = t)$, è facile constatare che, in questo caso, le equazioni delle geodetiche (10.106) o (10.111) diventano:

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{hk}\frac{dx^h}{dt}\frac{dx^k}{dt} = 0, \quad (i, h, k = 1, 2, ..., n). \tag{10.112}$$

10.9.1 Ulteriori proprietà delle geodetiche

La nozione di trasporto parallelo su di una varietà riemanniana che abbiamo studiato nel paragrafo 9.1.11, permette di dare una notevole e caratteristica proprietà delle geodetiche di una M^n. Si consideri infatti il sistema di equazioni differenziali che le definisce (10.106) e, posto

$$T^h = \frac{dx^h}{d\lambda}, \tag{10.113}$$

vettore tangente alla generica geodetica, il sistema (10.106) si scrive allora

$$\frac{dT^h}{d\lambda} + \Gamma^h_{jk}T^jT^k = 0. \quad (j, h, k = 1, 2, ..., n), \tag{10.114}$$

Tenendo conto della definizione di *derivata assoluta* avremo:

$$\frac{DT^h}{d\lambda} = 0, \quad (h = 1, 2, ..., n), \tag{10.115}$$

ovvero equivalentemente in termini del *differenziale assoluto*:

$$DT^h = 0, \quad (h = 1, 2, ..., n). \tag{10.116}$$

Si può dunque affermare, in base alla (9.58) o alla (9.60) viste a suo tempo, che il vettore tangente T^h si trasporta per parallelismo lungo la geodetica, ovvero, come si suole anche dire, *le geodetiche sono curve autoparallele* e ciò le caratterizza completamente. Possiamo dunque dare per esse anche la seguente (Fig.10.7):

Figura 10.7: Le geodetiche sono curve autoparallele.

Definizione: *Si definisce* **geodetica** *(o* **curva autoparallela***) di M^n una curva γ, di equazioni parametriche $x^h = x^h(\lambda)$ $(h = 1, 2, ..., n)$ il cui vettore tangente*

$$T^h = \frac{dx^h}{d\lambda}, \quad (h = 1, 2, ..., n)$$

viene **trasportato per parallelismo** *lungo la curva stessa, e perciò soddisfacente alla semplice condizione che sia nullo il differenziale totale del vettore tangente*

$$DT^h = 0.$$

Resta così messa in luce per le geodetiche di una M^n una tipica proprietà che la retta possiede nella geometria euclidea, e la nozione di derivata assoluta (o differenziale assoluto) ne acquista un notevole contenuto geometrico.
In tal caso infatti essendo $M^n \equiv \mathbb{R}^n$ e riferendoci a coordinate cartesiane sarà $\Gamma^i_{hk} \equiv 0$, e la (10.114) si riduce alla

$$\frac{dT^h}{d\lambda} = 0, \quad (h = 1, 2, ..., n), \tag{10.117}$$

cioè $T^h = cost.$; le geodetiche di $\mathbb{R}^n$ sono perciò delle rette.

10.9.2 Curve nulle e geodetiche nulle

La metrica di una varietà riemanniana, come sappiamo, è definita dall'elemento lineare (invariante per trasformazioni di coordinate)

$$\mathbf{ds}^2 = g_{ik}(x)dx^i dx^k, \qquad (i, k = 1, 2, \ldots, n) \tag{10.118}$$

cioè da una forma quadratica nei differenziali dx^i delle coordinate, che fornisce il quadrato dello spostamento infinitesimo $PP' \equiv \mathbf{ds}$ da un generico punto $P(x^i)$ della varietà ad un punto infinitamente vicino $P'(x^i + dx^i)$. Tale forma quadratica può essere *definita positiva* $(\mathbf{ds}^2 > 0)$, nel caso delle varietà *propriamente riemanniane* (ad esempio: uno spazio euclideo $\mathbb{R}^n$ di dimensione arbitraria oppure una generica superficie regolare S dello spazio euclideo 3-dimensionale $\mathbb{R}^3$, in particolare la superficie di una sfera $\mathbb{S}^2$); ma può essere anche *una forma quadratica indefinita* $\left(\mathbf{ds}^2 \gtreqless 0\right)$ come nei casi che più ci interessano dello spaziotempo pseudoeuclideo $\mathcal{M}^4$ di Minkowski, nell'ambito della Relatività Ristretta, e dello spaziotempo $\mathcal{V}^4$ della Relatività Generale del quale parleremo nel Cap.11.
Se poi $\mathbf{ds}^2 = 0$ *con differenziali dx^i non tutti nulli, lo spostamento* $\mathbf{ds} \equiv (dx^i)$ *è detto uno* **spostamento nullo**, *anche se i due punti P e P' non coincidono.* Per ogni spostamento infinitesimo dx^i che non sia nullo, si può introdurre un *indicatore* $\epsilon = \pm 1$ in modo che $\mathbf{ds}^2$ risulti sempre positivo, cioè tale che

$$\mathbf{ds}^2 = \epsilon g_{ik}(x)dx^i dx^k > 0, \qquad (i, k = 1, 2, \ldots, n) \tag{10.119}$$

Utilizzando tale *indicatore* la quantità ds sarà considerata *sempre positiva (o nulla)* e si potrà scrivere

$$ds = \sqrt{\mathbf{ds}^2} \geq 0.$$

Consideriamo ora un arco di curva regolare γ che, su M^n, unisca due punti prefissati P_1, P_2, parametrizzato con un arbitrario parametro λ, come abbiamo fatto all'inizio del paragrafo 10.9, cioè consideriamo una qualunque n- pla di funzioni di classe C^2

$$x^i = x^i(\lambda), \qquad \lambda_1 \leq \lambda \leq \lambda_2, \quad (i = 1, 2, ..., n) \tag{10.120}$$

soggetta alla condizione di assumere per $\lambda = \lambda_1$, $\lambda = \lambda_2$ valori prescritti.
Possiamo definire la *lunghezza dell'arco di curva* γ con la formula

$$s = \int_{\lambda_1}^{\lambda} \sqrt{\epsilon g_{ik}[x(\lambda)] \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda}} d\lambda = \int_{\lambda_1}^{\lambda} \sqrt{\left| g_{ik}[x(\lambda)] \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda} \right|} d\lambda \tag{10.121}$$

Se lungo un arco di curva risulta $\mathbf{ds}^2 = 0$ (come succede per un raggio di luce in Relatività) allora risulterà

$$g_{ik}(x) \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda} = 0, \tag{10.122}$$

e i due punti corrispondenti a $\lambda = \lambda_1$, $\lambda = \lambda_2$ risultano a distanza zero l'uno dall'altro, sebbene essi non siano coincidenti. *Una tale curva è detta* **nulla o minimale**. In tal caso non è possibile scegliere come parametro λ l'ascissa curvilinea s perché con questo parametro dev'essere verificata la

$$g_{ik}(x) \frac{dx^i}{ds} \frac{dx^k}{ds} = 1, \tag{10.123}$$

e non la (10.122).
Esempio 1:
Nello spazio di Minkowski con metrica

$$\mathbf{ds}^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2, \tag{10.124}$$

la curva definita dalle equazioni parametriche

$$\begin{cases} x^0 = \int r \, d\lambda \\ x^1 = \int r \, \cos\theta \, \cos\psi \, d\lambda \\ x^2 = \int r \, \cos\theta \, \sin\psi \, d\lambda \\ x^3 = \int r \, \sin\theta \, d\lambda \end{cases} \tag{10.125}$$

dove r, θ e ψ sono funzioni assegnate di λ, è una **curva reale nulla** come si può constatare facilmente.
È ovvio che in uno spazio propriamente riemanniano, la cui metrica è definita positiva, non esiste nessuna curva reale nulla. Invece in uno spazio pseudoriemanniano con metrica indefinita esistono curve, che in generale, consistono di porzioni lungo le quali l'indicatore $\epsilon = +1$, di porzioni lungo le quali l'indicatore è $\epsilon = -1$, e di porzioni nulle. La lunghezza della curva è allora la somma delle lunghezze di queste porzioni, la parte nulla portando contributo nullo al valore della lunghezza.
Eccettuato il caso di curve nulle, come parametro λ si può scegliere o il tempo t oppure l'ascissa curvilinea s misurata a partire da un punto fisso della curva. Dalla (10.119) si ha allora

$$g_{ik}(x) \frac{dx^i}{ds} \frac{dx^k}{ds} = \epsilon, \tag{10.126}$$

lungo ogni porzione di una curva che non sia nulla.
Caso delle geodetiche
Nel paragrafo 10.9 abbiamo visto che le geodetiche di una varietà M^n sono le curve soluzioni del sistema di equazioni differenziali (10.106), perciò, alla luce di quello che abbiamo detto riguardo alle curve nulle, possiamo dare la seguente:
Definizione: *Una curva di M^n si dice* **una geodetica nulla** *se è soluzione del sistema di equazioni differenziali*

$$\frac{d^2x^h}{d\lambda^2} + \Gamma^h_{jk}\frac{dx^j}{d\lambda}\frac{dx^k}{d\lambda} = 0, \qquad (j,h,k = 1,2,\ldots,n) \tag{10.127}$$

e verifica la condizione:

$$g_{ik}(x)\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} = 0.$$

Brevemente possiamo dire che si tratta di una **curva nulla** che allo stesso tempo sia anche **geodetica.** È quasi ovvio che una curva nulla non è necessariamente una geodetica nulla.
Per le *geodetiche nulle* la distanza fra due suoi punti infinitamente vicini è sempre nulla, inoltre dobbiamo ricordarci che per queste curve non si può considerare come parametro lungo la curva l'ascissa curvilinea s. Il significato fisico delle *geodetiche nulle* diventerà chiaro nel prossimo capitolo sulla Relatività Generale. Qui possiamo limitarci a dire che: *così come l'equazione di una generica geodetica descrive il moto di una particella in un campo gravitazionale, così l'equazione di una geodetica nulla descrive il moto di una particella di massa propria nulla (come un fotone) e perciò descrive la propagazione delle onde elettromagnetiche e della luce.*
Esempio 2:
Consideriamo ancora lo spazio di Minkowski con la metrica (10.124), vogliamo dimostrare che ogni *curva nulla* può essere rappresentata nella seguente forma parametrica:

$$\begin{cases} x^0 = \int \sin\psi\, d\lambda \\ x^1 = \int \sin\psi\, \sin\theta\, \cos\phi\, d\lambda \\ x^2 = \int \sin\psi\, \sin\theta\, \sin\phi\, d\lambda \\ x^3 = \int \sin\psi\, \cos\theta\, d\lambda \end{cases} \tag{10.128}$$

dove ψ, θ e ϕ sono funzioni assegnate di λ.
Soluzione:
Sostituendo le (10.128), per le coordinate x^0, x^1, x^2, x^3, nella metrica (10.124) dello spazio di Minkowski, otteniamo

$$\mathbf{ds}^2 = \sin^2\psi(1 - \sin^2\theta\, \cos^2\phi - \sin^2\theta\, \sin^2\phi - \cos^2\theta)d\lambda^2 = 0, \tag{10.129}$$

perciò la curva rappresentata dalle (10.128) è una *curva nulla*. Si può verificare anche che ogni curva soddisfacente alla (10.122) può essere rappresentata nella forma (10.128).
È interessante trovare le condizioni affinché tale curva sia una geodetica. Poiché, in uno spazio di Minkowski, utilizzando coordinate cartesiane, i simboli di Christoffel sono nulli e l'equazione delle geodetiche (10.127) si riduce alla $\frac{d^2x^h}{d\lambda^2} = 0$, $(h = 0,1,2,3)$. Per $h = 0$, considerando quindi x^0, abbiamo $\frac{dx^0}{d\lambda} = \sin\psi$, e perciò $\frac{d^2x^0}{d\lambda^2} = \cos\psi\frac{d\psi}{d\lambda} = 0$, la quale implica che ψ è indipendente da λ. Per x^3 si ottiene $\frac{dx^3}{d\lambda} = \sin\psi\cos\theta$, perciò $\frac{d^2x^3}{d\lambda^2} = -\sin\psi\sin\theta\frac{d\theta}{d\lambda} = 0$, che è soddisfatta se θ è indipendente da λ. Infine per x^1 si ottiene $\frac{dx^1}{d\lambda} = \sin\psi\sin\theta\cos\phi$, ed allora $\frac{d^2x^1}{d\lambda^2} = -\sin\psi\sin\theta\sin\phi\frac{d\phi}{d\lambda} = 0$, la quale implica che anche ϕ è indipendente da λ. Perciò

affinché la curva (10.128) sia una geodetica è necessario che le tre funzioni ψ, θ e ϕ devono essere indipendenti dal parametro λ. Integrando le (10.128) otteniamo le equazioni parametriche delle *geodetiche nulle* dello spazio di Minkowski:

$$\begin{cases} x^0 - x_0^0 = \sin\psi(\lambda - \lambda_0) \\ x^1 - x_0^1 = \sin\psi\ \sin\theta\ \cos\phi(\lambda - \lambda_0) \\ x^2 - x_0^2 = \sin\psi\ \sin\theta\ \sin\phi(\lambda - \lambda_0) \\ x^3 - x_0^3 = \sin\psi\ \cos\theta(\lambda - \lambda_0), \end{cases} \tag{10.130}$$

dove $x_0^0, x_0^1, x_0^2, x_0^3$ e λ_0 sono costanti d'integrazione.

10.10 Coordinate geodetiche

Sappiamo già che se la varietà è euclidea esistono certamente sistemi di coordinate che fanno assumere alla metrica $\mathbf{ds}^2$ la forma pitagorica (o pseudopitagorica) ossia che rendono le g_{ik} costanti in M^n. Se *la varietà è curva* ciò non è possibile; però si può sempre scegliere, in infiniti modi, un sistema di coordinate (x^i) in modo che in un punto prefissato $P_0 \equiv (x^i_{(0)})$ di M^n risultino nulle tutte le derivate parziali del tensore metrico, $[\partial_k g_{ij}]_{(P_0)} = 0$, o, in forma equivalente, siano nulli tutti i simboli di Christoffel, $[\Gamma^k_{ij}]_{(P_0)} = 0$, e si abbia quindi *coincidenza locale tra derivate covarianti e derivate parziali ordinarie.* Tali coordinate si chiamano *coordinate geodetiche o localmente cartesiane.*
È opportuno ricordare allo studente che nell'Osservazione posta alla fine del paragrafo 10.7.2 abbiamo dato due distinte definizioni di *coordinate cartesiane*: la prima riferita alla metrica e la seconda riferita alla connessione.
Il concetto di coordinate geodetiche estende opportunamente ad una varietà arbitraria il concetto di coordinate cartesiane. Si possono dare due distinte definizioni, una relativa alla metrica ed una relativa alla connessione.
Def.1: *Sia M^n una varietà riemanniana con metrica g_{ij}. Un sistema di coordinate locali $(x) = (x^k)$ tale che in un punto $P_0 \in M^n$ le g_{ij} soddisfano le condizioni*

$$[\partial_k g_{ij}]_{(P_0)} = 0, \tag{10.131}$$

viene chiamato un sistema di **coordinate geodetiche** *o* **localmente cartesiane** *per la metrica g_{ij} nel punto P_0.*
Def.2: *Sia invece M^n una varietà differenziabile nella quale sia data una connessione tramite i coefficienti Γ^k_{ij}. Un sistema di coordinate locali $(x) = (x^k)$ tale che in un punto $P_0 \in M^n$ le Γ^k_{ij} soddisfano le condizioni*

$$[\Gamma^k_{ij}]_{(P_0)} = 0, \tag{10.132}$$

viene chiamato un sistema di **coordinate geodetiche** *o* **localmente cartesiane** *per la connessione Γ^k_{ij} nel punto P_0.*
Se nella varietà riemanniana M^n è data la connessione di Levi-Civita, allora la condizione (10.131) è equivalente alla condizione (10.132). Ciò segue dalla seconda formula di Christoffel e dall'identità di Ricci. In tal caso le due definizioni sono coincidenti; si parlerà quindi, semplicemente, di *coordinate geodetiche* o *localmente cartesiane* nel punto P_0. Tale punto viene detto il **polo** del sistema di coordinate geodetiche.
Vale il seguente:

Teorema 10.5 *Per ogni punto P_0 di una varietà riemanniana M^n, munita della connessione di Levi-Civita, esiste sempre un sistema di coordinate geodetiche di cui P_0 è il polo.*

Dimostrazione. Consideriamo in M^n la connessione di Levi-Civita e siano Γ^k_{ij} i suoi coefficienti espressi in un certo sistema di coordinate (x). Consideriamo ora un altro sistema di coordinate $(\overline{x})$, definito dalla relazione

$$\overline{x}^i = x^i - x^i_{(P_0)} + \frac{1}{2}[\Gamma^i_{rs}]_{(P_0)}(x^r - x^r_{(P_0)})(x^s - x^s_{(P_0)}). \tag{10.133}$$

Derivando rispetto a x^j:

$$\begin{aligned}
\frac{\partial \overline{x}^i}{\partial x^j} &= \delta^i_j + \frac{1}{2}[\Gamma^i_{rs}]_{(P_0)}\delta^r_j(x^s - x^s_{(P_0)}) + \frac{1}{2}[\Gamma^i_{rs}]_{(P_0)}(x^r - x^r_{(P_0)})\delta^s_j = \\
&= \delta^i_j + \frac{1}{2}[\Gamma^i_{js}]_{(P_0)}(x^s - x^s_{(P_0)}) + \frac{1}{2}[\Gamma^i_{rj}]_{(P_0)}(x^r - x^r_{(P_0)}) = \\
&= \delta^i_j + [\Gamma^i_{jr}]_{(P_0)}(x^r - x^r_{(P_0)}).
\end{aligned} \tag{10.134}$$

Nel punto P_0, perciò, si ha

$$\left[\frac{\partial \overline{x}^i}{\partial x^j}\right]_{(P_0)} = \delta^i_j,$$

da cui

$$\det\left[\frac{\partial \overline{x}^i}{\partial x^j}\right]_{(P_0)} = 1 \neq 0,$$

per cui la trasformazione è invertibile. Inoltre, si ha

$$\frac{\partial \overline{x}^i}{\partial x^j}\frac{\partial x^j}{\partial \overline{x}^k} = \delta^i_k, \qquad \left[\frac{\partial \overline{x}^i}{\partial x^j}\right]_{(P_0)}\left[\frac{\partial x^j}{\partial \overline{x}^k}\right]_{(P_0)} = \delta^i_k, \qquad \delta^i_j\left[\frac{\partial x^j}{\partial \overline{x}^k}\right]_{(P_0)} = \delta^i_k,$$

per cui, anche

$$\left[\frac{\partial x^i}{\partial \overline{x}^k}\right]_{(P_0)} = \delta^i_k.$$

Ora, moltiplicando membro a membro la (10.134) per $\dfrac{\partial x^j}{\partial \overline{x}^k}$, si ha

$$\frac{\partial \overline{x}^i}{\partial x^j}\frac{\partial x^j}{\partial \overline{x}^k} = \delta^i_j\frac{\partial x^j}{\partial \overline{x}^k} + [\Gamma^i_{jr}]_{(P_0)}(x^r - x^r_{(P_0)})\frac{\partial x^j}{\partial \overline{x}^k},$$

che possiamo riscrivere come:

$$\delta^i_k = \frac{\partial x^i}{\partial \overline{x}^k} + [\Gamma^i_{jr}]_{(P_0)}(x^r - x^r_{(P_0)})\frac{\partial x^j}{\partial \overline{x}^k}.$$

Derivando quest'ultima rispetto a $\overline{x}^h$, otteniamo:

$$0 = \frac{\partial^2 x^i}{\partial \overline{x}^k \partial \overline{x}^h} + [\Gamma^i_{jr}]_{(P_0)}(x^r - x^r_{(P_0)})\frac{\partial^2 x^j}{\partial \overline{x}^k \partial \overline{x}^h} + [\Gamma^i_{jr}]_{(P_0)}\frac{\partial x^r}{\partial \overline{x}^h}\frac{\partial x^j}{\partial \overline{x}^k}.$$

Valutando quest'ultima equazione nel punto P_0 si trova:

$$0 = \left[\frac{\partial^2 x^i}{\partial \overline{x}^k \partial \overline{x}^h}\right]_{(P_0)} + [\Gamma^i_{jr}]_{(P_0)}\delta^r_h\delta^j_k,$$

ovvero

$$\left[\frac{\partial^2 x^i}{\partial \overline{x}^k \partial \overline{x}^h}\right]_{(P_0)} = -[\Gamma^i_{kh}]_{(P_0)}.$$

Adesso, dalla formula di trasformazione per i simboli di Christoffel

$$\overline{\Gamma}^i_{kh} = \frac{\partial \overline{x}^i}{\partial x^j}\frac{\partial x^r}{\partial \overline{x}^k}\frac{\partial x^s}{\partial \overline{x}^h}\Gamma^j_{rs} + \frac{\partial \overline{x}^i}{\partial x^j}\frac{\partial^2 x^j}{\partial \overline{x}^k \partial \overline{x}^h}$$

valutando i singoli termini della formula precedente nel punto P_0, si ottiene infine

$$[\overline{\Gamma}^i_{kh}]_{(P_0)} = \delta^i_j \delta^r_k \delta^s_h [\Gamma^j_{rs}]_{(P_0)} - \delta^i_j [\Gamma^j_{kh}]_{(P_0)} = [\Gamma^i_{kh}]_{(P_0)} - [\Gamma^i_{kh}]_{(P_0)} = 0.$$

■

Osserviamo che nel polo $P_0 \equiv (x^i_{(0)})$, per la (10.133), si ha $\overline{x}^i_{(0)} = 0$, il che significa che *il polo è anche l'origine delle coordinate.* Le coordinate geodetiche possiedono dunque la seguente importante

Proprietà: *Le derivate covarianti si riducono alle corrispondenti derivate parziali nel polo, poiché in questo punto tutti i simboli di Christoffel sono nulli.*

In precedenza abbiamo visto che se un tensore è nullo in un dato sistema di coordinate, esso è nullo in qualsiasi altro sistema di coordinate, dal momento che la legge di trasformazione dei tensori è lineare. La verifica di un'equazione tensoriale spesso comporta pesanti manipolazioni algebriche e, generalmente, il volume del lavoro si riduce notevolmente provando dapprima l'equazione nel polo di un sistema di coordinate geodetiche. Ne consegue che l'equazione è vera per tutti i sistemi di coordinate in questo punto. *Se poi questo punto è un generico punto della varietà M^n, l'equazione è verificata in tutti i punti della varietà stessa.*

Esercizio 1:
Illustriamo questo metodo dimostrando che la regola di Leibniz sulla derivata di un prodotto di funzioni è valida anche per le derivate covarianti.
Consideriamo, per esempio, il tensore covariante di rango 2:

$$T_{hi} \equiv \nabla_h(A_{ij}B^j) - (\nabla_h A_{ij})B^j - A_{ij}(\nabla_h B^j),$$

e scegliamo un sistema di coordinate geodetiche con polo in un punto P_0 di M^n, dove le derivate covarianti si riducono alle ordinarie derivate parziali. Dal momento che le derivate parziali, rispetto ad una qualsiasi variabile x, soddisfano la legge di Leibniz del prodotto, cioè

$$\frac{\partial}{\partial x}(\varphi\psi) = \frac{\partial \varphi}{\partial x}\psi + \varphi\frac{\partial \psi}{\partial x},$$

il suddetto tensore è zero in P_0 nel sistema di coordinate geodetiche, cioè[11]:

$$T_{hi} \equiv \nabla_h(A_{ij}B^j) - \Big[(\nabla_h A_{ij})B^j - A_{ij}(\nabla_h B^j)\Big] \overset{*}{=} \partial_h(A_{ij}B^j) - \Big[(\partial_h A_{ij})B^j - A_{ij}(\partial_h B^j)\Big] = 0.$$

Ma P_0 è un punto generico di M^n, di conseguenza il suddetto tensore è nullo in tutti i punti di M^n. Questo stabilisce la validità della legge del prodotto:

$$\nabla_h(A_{ij}B^j) = (\nabla_h A_{ij})B^j - A_{ij}(\nabla_h B^j).$$

Esercizio 2:
Dimostrare che nel polo P_0 di un sistema di coordinate geodetiche si ha:

$$\nabla_k \nabla_j A_i = \frac{\partial^2 A_i}{\partial x^j \partial x^k} - A_h \partial_k \Gamma^h_{ij}.$$

[11] Se fra i due membri di un'equazione interponiamo il simbolo di uguaglianza sormontato da una stellina ($\overset{*}{=}$) intendiamo che l'equazione è valida in un particolare sistema di coordinate geodetiche.

10.11 La curvatura

L'idea di **curvatura** è semplice e familiare nella geometria euclidea. *Una linea è curva se devia da una retta, ed una superficie è curva se essa devia da un piano.* Tuttavia è possibile scoprire se una superficie è curva per mezzo di operazioni puramente intrinseche fatte sulla superficie stessa *"senza riguardarla come immersa in uno spazio ambiente euclideo".* Pensiamo alla più semplice delle superficie curve: *la sfera.* Immaginiamo un essere bidimensionale che si muova sulla superficie sferica, e non possa percepire nulla al di fuori di essa. Le operazioni che tale essere può effettuare consistono solamente di misure di lunghezza e di angoli sulla superficie. Dalle proprietà delle geodetiche (viste come linee di minima distanza fra due punti dati) egli può costruire dei cerchi massimi sulla sua sfera. Se lui misura gli angoli di un triangolo sferico ABC formato da archi di cerchi massimi, troverà che la somma degli angoli è maggiore di un angolo piatto e minore di tre angoli piatti ($\pi < \alpha + \beta + \gamma < 3\pi$).[12] Questo risultato suggerirà ad esso (se conosce la geometria euclidea) che la superficie dove si trova non è un piano (Fig.10.8).

Figura 10.8: La nozione di curvatura.

Questo semplice esempio illustra il nostro punto di vista nel discutere la curvatura di una varietà riemanniana. La curvatura è concepita come qualcosa di intrinseco allo spazio in considerazione (alla varietà), e non qualcosa da misurare confrontando uno spazio con un altro. Perciò noi non dobbiamo pensare ad una varietà riemanniana come necessariamente immersa in uno spazio euclideo, come eravamo tentati di fare quando dovevamo discutere la geometria di una sfera.
Nel creare una teoria geometrica di una varietà riemanniana M^n dobbiamo generalizzare i nostri concetti familiari in due modi:

1) *dobbiamo passare da 3 ad n dimensioni;*

2) *dobbiamo considerare la possibilità di una curvatura intrinseca, come si fa per una superficie dello spazio euclideo 3-dimensionale.*

Definiamo dapprima ciò che intendiamo per *varietà riemanniana piatta.*
Definizione 1: *Una varietà riemanniana è detta* **piatta** *se è possibile scegliere un sistema di coordinate per il quale la metrica si riduce a*

$$\mathbf{ds}^2 = \sum_{i=1}^{n} \epsilon_i (dx^i)^2, \tag{10.135}$$

essendo ϵ_i uguale o a $+1$ o a -1.
Definizione 2: *Una varietà riemanniana non piatta, cioè non soddisfacente la precedente definizione, è*

[12]Una dimostrazione elementare di questa relazione, il lettore può trovarla nell'articolo di Francesco Borghero, *Sui fondamenti della trigonometria sferica*, **Archimede**, Le Monnier, Firenze, Fasc. 2-3, (1987), pag.110-123.

detta **curva**.[13]
Il nostro prossimo compito è quello di sviluppare un test che ci dica *se una data varietà riemanniana* è piatta o curva. La questione è allora:
È possibile o no, mediante una trasformazione di coordinate, ridurre il ds^2 *alla forma* (10.135)*?* La risposta a questo quesito è contenuta in ciò che segue.

10.11.1 Non commutabilità di due derivate covarianti successive: Tensore di Riemann

Consideriamo una varietà differenziabile M^n sulla quale sia definita una connessione $\Gamma^k_{ij}(x)$ generalmente non simmetrica, cioè tale che $\Gamma^k_{ij}(x) \neq \Gamma^k_{ji}(x)$, ovvero sia non nullo il cosiddetto **tensore di torsione** $\Omega^k_{ij} \equiv \Gamma^k_{ij}(x) - \Gamma^k_{ji}(x) \neq 0$.[14]
Dato un tensore $T^{i_1 \dots i_p}_{j_1 \dots j_q}$ qualsiasi su M^n, sappiamo che la sua derivata covariante $\nabla_\alpha T^{i_1 \dots i_p}_{j_1 \dots j_q}$ è ancora un tensore con un indice di covarianza in più, del quale possiamo ancora calcolare la derivata covariante $\nabla_\beta(\nabla_\alpha T^{i_1 \dots i_p}_{j_1 \dots j_q})$. Ci chiediamo se le due operazioni sono commutabili, ovvero se risulta valida l'uguaglianza $\nabla_\beta(\nabla_\alpha T^{i_1 \dots i_p}_{j_1 \dots j_q}) = \nabla_\alpha(\nabla_\beta T^{i_1 \dots i_p}_{j_1 \dots j_q})$, oppure no.
Cominciamo dal caso più semplice in cui il tensore è di rango zero, cioè si riduce ad un campo scalare $\varphi(x^i)$, e consideriamo il suo gradiente: grad$\varphi \equiv \partial_l \varphi \equiv \nabla_l \varphi$.
Del campo di vettori $\nabla_l \varphi$ calcoliamo ulteriormente la derivata covariante $\nabla_m(\nabla_l \varphi)$, e poi scambiamo gli indici m ed l, avremo:

$$\nabla_l(\nabla_m \varphi) = \nabla_l(\partial_m \varphi) = \partial_{lm}\varphi - \partial_h \varphi \Gamma^h_{ml},$$

$$\nabla_m(\nabla_l \varphi) = \nabla_m(\partial_l \varphi) = \partial_{ml}\varphi - \partial_h \varphi \Gamma^h_{lm},$$

da cui sottraendo membro a membro:

$$\nabla_l(\nabla_m \varphi) - \nabla_m(\nabla_l \varphi) = \partial_h \varphi \Omega^h_{lm} = \nabla_h \varphi \Omega^h_{lm}, \tag{10.136}$$

dove $\Omega^h_{lm} \equiv \Gamma^h_{lm}(x) - \Gamma^h_{ml}(x) \neq 0$ è, come abbiamo detto, **il tensore di torsione**. Questo tensore è identicamente nullo soltanto se **la connessione è simmetrica**: $\Omega^h_{lm} \equiv 0$. In questo caso si ha:

$$\nabla_l(\nabla_m \varphi) = \nabla_m(\nabla_l \varphi), \tag{10.137}$$

cioè *se la connessione è simmetrica due derivate covarianti successive di un campo scalare commutano. Tale circostanza tuttavia non si verifica più quando le due derivazioni covarianti successive, anziché su uno scalare, operano su un generico campo di vettori o di tensori.* Consideriamo ad esempio un qualsiasi campo di vettori controvarianti v^i, ed eseguiamo su di esso due derivazioni covarianti successive, otteniamo:

$$\nabla_l(\nabla_m v^i) = \partial_l(\nabla_m v^i) + \nabla_m v^h \Gamma^i_{hl} - \nabla_h v^i \Gamma^h_{ml},$$

$$\nabla_m(\nabla_l v^i) = \partial_m(\nabla_l v^i) + \nabla_l v^h \Gamma^i_{hm} - \nabla_h v^i \Gamma^h_{lm},$$

[13] Dovrebbe essere chiaro che una varietà riemanniana M^n è piatta soltanto se è possibile ridurre il $\mathbf{ds}^2$ alla forma pitagorica (o pseudopitagorica) (10.135).
[14] Il tensore di torsione è identicamente nullo ($\Omega^k_{ij} \equiv 0$) se e solo se la connessione è simmetrica, cioè $\Gamma^k_{ij}(x) = \Gamma^k_{ji}(x)$.

da cui sottraendo membro a membro:

$$
\begin{aligned}
&\nabla_l(\nabla_m v^i) - \nabla_m(\nabla_l v^i) = \\
&= \partial_l(\nabla_m v^i) + \nabla_m v^h \Gamma^i_{hl} - \partial_m(\nabla_l v^i) - \nabla_l v^h \Gamma^i_{hm} + \nabla_h v^i \Omega^h_{lm} = \\
&= \partial_l(\partial_m v^i + v^k \Gamma^i_{km}) + (\partial_m v^h + v^k \Gamma^h_{km})\Gamma^i_{hl} - \\
&- \partial_m(\partial_l v^i + v^k \Gamma^i_{kl}) - (\partial_l v^h + v^k \Gamma^h_{kl})\Gamma^i_{hm} + \nabla_h v^i \Omega^h_{lm} = \\
&= \partial_{lm} v^i + \partial_l v^k \Gamma^i_{km} + v^k \partial_l \Gamma^i_{km} + \partial_m v^h \Gamma^i_{hl} + v^k \Gamma^h_{km} \Gamma^i_{hl} - \\
&- \partial_{ml} v^i - \partial_m v^k \Gamma^i_{kl} - v^k \partial_m \Gamma^i_{kl} - \partial_l v^h \Gamma^i_{hm} - v^k \Gamma^h_{kl} \Gamma^i_{hm} + \nabla_h v^i \Omega^h_{lm} = \\
&= v^k(\partial_l \Gamma^i_{km} - \partial_m \Gamma^i_{kl} + \Gamma^h_{km}\Gamma^i_{hl} - \Gamma^h_{kl}\Gamma^i_{hm}) + \nabla_h v^i \Omega^h_{lm}. \qquad (10.138)
\end{aligned}
$$

Si ha, infine:

$$\nabla_l(\nabla_m v^i) - \nabla_m(\nabla_l v^i) = \nabla_h v^i \Omega^h_{lm} + v^k R^i{}_{klm} \neq 0, \quad \forall v^k, \qquad (10.139)$$

avendo posto

$$R^i{}_{klm} \stackrel{\text{def}}{=} \partial_l \Gamma^i_{km} - \partial_m \Gamma^i_{kl} + \Gamma^i_{hl}\Gamma^h_{km} - \Gamma^i_{hm}\Gamma^h_{kl}. \qquad (10.140)$$

Nel caso in cui il tensore di torsione sia nullo ($\Omega^h_{lm} = 0$) la formula (10.139) si riduce a:

$$\nabla_l(\nabla_m v^i) - \nabla_m(\nabla_l v^i) = v^k R^i{}_{klm} \neq 0, \quad \forall v^k. \qquad (10.141)$$

In maniera analoga, per un campo di vettori covariante v_i si ottiene:

$$\nabla_l(\nabla_m v_i) - \nabla_m(\nabla_l v_i) = -v_k R^k{}_{iml} \neq 0, \quad \forall v_k. \qquad (10.142)$$

Le (10.141) e (10.142) prendono il nome di **formule di commutazione di Ricci**.
Esse mostrano anzitutto che *l'ordine di due derivazioni covarianti successive non è permutabile anche nel caso di connessioni simmetriche.* Inoltre, poiché i primi membri delle (10.141), (10.142) rappresentano dei tensori tripli, il termine $R^i{}_{klm}$ che figura al 2° membro, rappresenta, in base al criterio generalizzato di tensorialità, un tensore di rango 4 con tre indici di covarianza ed uno di controvarianza.
Il tensore definito dalla (10.140) è denominato **tensore di Riemann-Christoffel** o semplicemente **tensore di Riemann** o anche **tensore di curvatura** della varietà M^n (per una ragione che vedremo in seguito) ed è d'importanza fondamentale sia nella geometria riemanniana sia nella Teoria della Relatività Generale. Ne stabiliremo fra poco alcune delle proprietà più importanti, prima però facciamo qualche osservazione (di natura *"mnemotecnica"*) sulla sua definizione e struttura per ricordarlo agevolmente.
Osservazione 1:
Per ricordare la (10.140) dobbiamo osservare che il 2° membro è così strutturato: contiene 4 termini con segni alternati $+, -, +, -$ di cui i primi due sono derivate parziali dei simboli di Christoffel di 2ª specie, mentre il terzo e quarto termine sono formati ciascuno da somme di n prodotti fatti con simboli di Christoffel di 2ª specie. Precisamente: il 1°primo termine è la derivata $+\partial_l \Gamma^i_{km}$, col segno $+$, calcolata rispetto alla variabile x^l, il cui indice l è *il terzo indice* che figura al 1° membro in $R^i{}_{klm}$; il 2° termine è la derivata $-\partial_m \Gamma^i_{kl}$, col segno $-$, calcolata rispetto alla variabile x^m il cui indice m è *il quarto indice* che figura in $R^i{}_{klm}$.[15]
Per formare il 3° termine, cioè $+\Gamma^i_{hl}\Gamma^h_{km}$, immaginiamo di *"sdoppiare"* il simbolo di Christoffel Γ^i_{km} che figura nel 1° termine, così $\Gamma^i \Gamma_{km}$ e poi porre in basso su Γ^i, un indice muto h e l'indice

[15]Per la definizione del tensore di Riemann abbiamo adottato le notazioni usate da Landau e Lifschitz nel loro libro *Teoria dei Campi*, Editori Riuniti, Roma, edizione del 2010, pag. 341. Dobbiamo però avvertire il lettore che alcuni autori definiscono come tensore di Riemann l'opposto di quello usato da Landau e Lifschitz, cioè: $R^i{}_{klm} \stackrel{\text{def}}{=} \partial_m \Gamma^i_{kl} - \partial_l \Gamma^i_{km} + \Gamma^i_{hm}\Gamma^h_{kl} - \Gamma^i_{hl}\Gamma^h_{km}$.

l della prima derivazione, l'indice muto h verrà ripetuto in alto su Γ_{km}, in modo da ottenere $+\Gamma^i_{hl}\Gamma^h_{km}$ con lo stesso segno $+$ della prima derivata $\partial_l \Gamma^i_{km}$.
Infine, l'ultimo termine del tensore di Riemann $-\Gamma^i_{hm}\Gamma^h_{kl}$ si costruisce in maniera analoga al precedente col segno concorde con quello del secondo termine che figura al 2° membro della (10.140).

Per concludere, osserviamo ancora che se la data connessione è quella di Levi-Civita essa, per il teorema fondamentale della geometria riemanniana, è necessariamente simmetrica ed i simboli di Christoffel $\Gamma^k_{ij}(x)$ sono esprimibili mediante il tensore metrico $g_{ij}(x)$, in tal caso *il tensore di Riemann* (10.140) *contiene le derivate parziali del tensore metrico fino al 2° ordine: queste ultime compaiono linearmente, mentre le derivate prime vi figurano in maniera quadratica.* Analoghi risultati si ottengono se anziché derivare covariantemente un campo di vettori controvarianti, deriviamo un campo di vettori covarianti, come abbiamo già accennato, oppure, più in generale, se calcoliamo le derivate covarianti successive di un tensore qualunque: in tali casi si trova sempre il tensore di Riemann-Christoffel con calcoli molto più laboriosi. Per esempio, per un campo di vettori covarianti ed $\Omega^h_{lm} \neq 0$ si giunge alla relazione:

$$\nabla_l(\nabla_m v_j) - \nabla_m(\nabla_l v_j) = \nabla_h v_j \Omega^h_{lm} - v_k R^k{}_{jlm},$$

mentre per un tensore misto del 2° ordine si ottiene:

$$\nabla_l(\nabla_m v^i_j) - \nabla_m(\nabla_l v^i_j) = \nabla_h v^i_j \Omega^h_{lm} + v^k_j R^i{}_{klm} - v^i_k R^k{}_{jlm},$$

generalizzabili facilmente al caso di un tensore $T^{i_1\ldots i_p}_{j_1\ldots j_q}$ qualunque:

$$\begin{aligned}\nabla_l(\nabla_m T^{i_1\ldots i_p}_{j_1\ldots j_q}) \quad &- \quad \nabla_m(\nabla_l T^{i_1\ldots i_p}_{j_1\ldots j_q}) = \nabla_h T^{i_1\ldots i_p}_{j_1\ldots j_q}\Omega^h_{lm} + \\ &+ \quad \sum_{s=1}^{p} T^{i_1\ldots i_{s-1}k\, i_{s+1}\ldots i_p}_{j_1\ldots j_q} R^{i_s}{}_{klm} - \sum_{s=1}^{q} T^{i_1\ldots i_p}_{j_1\ldots j_{s-1}k\, j_{s+1}\ldots j_q} R^k{}_{j_s lm}.\end{aligned}$$

È ora evidente che la derivazione covariante non è, in generale, un'operazione commutativa. Questo rimane vero anche se la connessione è simmetrica, ovvero se risulta $\Omega^k_{ij} \equiv 0$; in tal caso, per un tensore qualsiasi $T^{i_1\ldots i_p}_{j_1\ldots j_q}$, la formula precedente si semplifica:

$$\begin{aligned}\nabla_l(\nabla_m T^{i_1\ldots i_p}_{j_1\ldots j_q}) \quad &- \quad \nabla_m(\nabla_l T^{i_1\ldots i_p}_{j_1\ldots j_q}) = \\ &= \quad \sum_{s=1}^{p} T^{i_1\ldots i_{s-1}k\, i_{s+1}\ldots i_p}_{j_1\ldots j_q} R^{i_s}{}_{klm} - \sum_{s=1}^{q} T^{i_1\ldots i_p}_{j_1\ldots j_{s-1}k\, j_{s+1}\ldots j_q} R^k{}_{j_s lm},\end{aligned}$$

ma l'operazione resta comunque non commutativa.
Osservazione 2:
Come già osservammo, quando in uno spazio euclideo si adottano coordinate cartesiane, tutti i simboli di Christoffel sono nulli, *e quindi si annulla il tensore di Riemann:* $R^i{}_{klm} = 0$. Naturalmente tale circostanza non è legata (come lo è invece per i simboli di Christoffel) a particolari scelte di coordinate, perché $R^i{}_{klm}$ è un tensore mentre i simboli di Christoffel non hanno carattere tensoriale. Si può dimostrare che vale anche il viceversa, cioè: *l'annullarsi del tensore di Riemann garantisce che la varietà M^n sia euclidea.*

10.11.2 Forma totalmente covariante del tensore di Riemann

Gioverà, per il seguito, procurarsi anche la forma totalmente covariante del tensore di Riemann. A tal fine basta abbassare l'indice i nella sua forma mista (10.140) utilizzando il tensore metrico g_{ik}, avremo:

$$R_{iklm} = g_{si} R^s{}_{klm}, \tag{10.143}$$

ovvero:

$$R_{iklm} = g_{si}R^s{}_{klm} = g_{si}(\partial_l\Gamma^s_{km} - \partial_m\Gamma^s_{kl} + \Gamma^s_{nl}\Gamma^n_{km} - \Gamma^s_{nm}\Gamma^n_{kl}), \tag{10.144}$$

che possiamo scrivere ancora nella forma:

$$R_{iklm} = g_{si}\partial_l\Gamma^s_{km} - g_{si}\partial_m\Gamma^s_{kl} + g_{si}(\Gamma^s_{nl}\Gamma^n_{km} - \Gamma^s_{nm}\Gamma^n_{kl}). \tag{10.145}$$

Applicando la regola di Leibniz sulla differenziazione di un prodotto, il 1° termine al secondo membro della (10.145) si può scrivere come una differenza

$$g_{si}\partial_l\Gamma^s_{km} = \partial_l(g_{si}\Gamma^s_{km}) - \partial_l g_{si}\Gamma^s_{km}.$$

La stessa cosa possiamo farla sul 2° termine, cioè

$$g_{si}\partial_m\Gamma^s_{kl} = \partial_m(g_{si}\Gamma^s_{kl}) - \partial_m g_{si}\Gamma^s_{kl}.$$

La (10.145) può allora riscriversi così:

$$R_{iklm} = [\partial_l(g_{si}\Gamma^s_{km}) - \partial_l g_{si}\Gamma^s_{km}] - [\partial_m(g_{si}\Gamma^s_{kl}) - \partial_m g_{si}\Gamma^s_{kl}] + g_{si}(\Gamma^s_{nl}\Gamma^n_{km} - \Gamma^s_{nm}\Gamma^n_{kl}),$$

introducendo i simboli di Christoffel di 1ª specie si ha: $g_{si}\Gamma^s_{km} = \Gamma_{km,i}$, $g_{si}\Gamma^s_{kl} = \Gamma_{kl,i}$, e la precedente equazione diventa

$$R_{iklm} = \partial_l\Gamma_{km,i} - \partial_l g_{si}\Gamma^s_{km} - \partial_m\Gamma_{kl,i} + \partial_m g_{si}\Gamma^s_{kl} + g_{si}(\Gamma^s_{nl}\Gamma^n_{km} - \Gamma^s_{nm}\Gamma^n_{kl}), \tag{10.146}$$

applicando l'identità di Ricci alle derivate $\partial_l g_{si}$, $\partial_m g_{si}$:

$$\partial_l g_{si} = \Gamma_{ls,i} + \Gamma_{li,s}, \qquad \partial_m g_{si} = \Gamma_{ms,i} + \Gamma_{mi,s},$$

e semplificando opportunamente avremo la seguente espressione per il tensore di Riemann:

$$R_{iklm} = \partial_l\Gamma_{km,i} - \partial_m\Gamma_{kl,i} + \Gamma_{mi,s}\Gamma^s_{kl} - \Gamma_{li,s}\Gamma^s_{km}. \tag{10.147}$$

Ora applichiamo la 1^a formula di Christoffel ai simboli di 1^a specie che figurano nelle prime due derivate al 2° membro della (10.147) e trasformiamo i simboli di 1^a specie in quelli di 2^a specie negli ultimi due termini; semplificando poi opportunamente avremo infine la seguente notevole espressione per il tensore di Riemann in forma completamente covariante:

$$R_{iklm} = \frac{1}{2}(\partial_{im}g_{kl} + \partial_{kl}g_{im} - \partial_{il}g_{km} - \partial_{km}g_{il}) + g_{rs}(\Gamma^r_{im}\Gamma^s_{kl} - \Gamma^r_{il}\Gamma^s_{km}). \tag{10.148}$$

Nella (10.148) l'ultima espressione in parentesi contiene soltanto simboli di Christoffel di 2ª specie, essa può ovviamente trasformarsi in una espressione contenente soltanto simboli di Christoffel di 1ª specie; dopo semplici calcoli si trova che:

$$g_{rs}(\Gamma^r_{im}\Gamma^s_{kl} - \Gamma^r_{il}\Gamma^s_{km}) = g^{sp}(\Gamma_{im,s}\Gamma_{kl,p} - \Gamma_{il,s}\Gamma_{km,p}), \tag{10.149}$$

e perciò la (10.149) può essere riscritta così:

$$R_{iklm} = \frac{1}{2}(\partial_{im}g_{kl} + \partial_{kl}g_{im} - \partial_{il}g_{km} - \partial_{km}g_{il}) + g^{sp}(\Gamma_{im,s}\Gamma_{kl,p} - \Gamma_{il,s}\Gamma_{km,p}). \tag{10.150}$$

10.11.3 Proprietà di simmetria ed antisimmetria del tensore di Riemann

Dalle due forme (10.148), (10.150) totalmente covarianti del tensore di Riemann resta messo in evidenza che esso *dipende dalle componenti del tensore metrico g_{ik} e dalle loro derivate prime e seconde.* Notiamo però esplicitamente come *la dipendenza di R_{iklm} dalle derivate seconde delle g_{ik}* **sia lineare**, *ciò che non è per le derivate prime che vi compaiono invece* **quadraticamente**.
Le due espressioni precedenti per R_{iklm} mostrano che *il tensore di Riemann è* **antisimmetrico** *sia rispetto allo scambio dei due indici della prima coppia, sia rispetto allo scambio dei due indici della seconda coppia, mentre è* **simmetrico** *rispetto allo scambio della prima coppia con la seconda*:

$$\begin{cases} R_{iklm} = -R_{kilm} \\ R_{iklm} = -R_{ikml} \\ R_{iklm} = R_{lmik}. \end{cases} \tag{10.151}$$

Inoltre esso gode di **una proprietà di simmetria ciclica** in virtù della quale *la somma delle componenti che si ottengono permutando ciclicamente tre indici (ad es. gli ultimi tre) è nulla.*
Infatti, usando la forma (10.148) (oppure la (10.150)), del tensore di Riemann R_{iklm}, mantenendo fisso il primo indice i, facendo una permutazione ciclica sugli ultimi tre indici klm e sommando membro a membro si ottiene facilmente:

$$R_{iklm} + R_{ilmk} + R_{imkl} = 0. \tag{10.152}$$

Dopo aver stabilito tali proprietà possiamo senz'altro affermare che il numero delle componenti indipendenti del tensore di Riemann è minore di n^4, cioè del numero delle disposizioni con ripetizione di n oggetti a 4 a 4. Per determinarlo con esattezza cominciamo con l'osservare che, per le due prime proprietà di antisimmetria espresse nella (10.151) *sono nulle tutte le componenti che hanno uguali fra di loro i primi due indici o gli ultimi due.* Possiamo allora ripartire tutte le componenti rimanenti in 3 insiemi ponendo:
nel 1° tutte quelle che hanno soltanto due indici diversi;
nel 2° quelle che hanno tre indici diversi;
nel 3° quelle che hanno tutti gli indici diversi.
Le componenti appartenenti al 1° insieme sono tutte del tipo:

$$R_{ilil}, \quad con\ i \neq l$$

poiché si ha

$$R_{iill} = 0, \quad R_{liil} = -R_{ilil}, \quad R_{lili} = R_{ilil}$$

il loro numero è ovviamente $\binom{n}{2}$, cioè tante quante le combinazioni semplici di n oggetti ad 2 a 2.
Quelle del 2° insieme possono tutte indicarsi con

$$R_{ikli}, \quad R_{iklk}, \quad R_{ilkl}$$

essendo

$$R_{ikil} = -R_{ikli}, \quad R_{ilki} = R_{ikli}, \quad R_{iikl} = 0,$$

il loro numero è $3\binom{n}{3}$.
Infine quelle del 3° insieme sono del tipo

$$R_{iklm}, \quad R_{ilmk}, \quad R_{imkl}$$

di cui però soltanto due sono indipendenti dovendo esse verificare la proprietà ciclica (10.152). Perciò il loro numero sarà $2\binom{n}{4}$. In definitiva *il numero totale delle componenti indipendenti del tensore di Riemann è allora*:

$$N = \binom{n}{2} + 3\binom{n}{3} + 2\binom{n}{4} = \frac{n^2(n^2-1)}{12}. \tag{10.153}$$

Per esempio il tensore di Riemann associato ad una superficie $(n = 2)$ ha un'unica componente indipendente, R_{1212}; mentre il tensore di Riemann associato ad una varietà differenziabile M^3 o ad una M^4 ha rispettivamente 6 o 20 componenti indipendenti.

10.11.4 Identità di Bianchi

Dopo aver esaminato le relazioni algebriche che sussistono per le componenti del tensore di Riemann, passiamo a stabilire tra esse degli importanti *legami differenzialei*, noti col nome di **identità di Bianchi**.
A tal proposito dimostriamo il seguente teorema fondamentale.
Teorema:
Dato il tensore di Riemann, in una varietà M^N ad N dimensioni, sotto forma mista $R^i{}_{klm}$, se ne calcoli la derivata covariante $\nabla_n R^i{}_{klm}$ e poi, lasciando fissi i primi due indici i, k e facendo una permutazione circolare sugli indici n, l, m si calcolino le analoghe derivate covarianti $\nabla_l R^i{}_{kmn}$, $\nabla_m R^i{}_{knl}$. I tre tensori misti del 5° ordine che si ottengono godono della proprietà che la loro somma è identicamente nulla, cioè:

$$\nabla_n R^i{}_{klm} + \nabla_l R^i{}_{kmn} + \nabla_m R^i{}_{knl} = 0, \qquad (i, k, l, n, m = 1, 2, ..., N). \tag{10.154}$$

Sono queste le importantissime **identità di Bianchi**.
Dim.
Partiamo dall'espressione esplicita del tensore di Riemann in forma mista:

$$R^i{}_{klm} \stackrel{\text{def}}{=} \partial_l \Gamma^i_{km} - \partial_m \Gamma^i_{kl} + \Gamma^i_{rl}\Gamma^r_{km} - \Gamma^i_{rm}\Gamma^r_{kl},$$

ed utilizziamo un sistema di coordinate geodetiche di polo un punto arbitrario P della data varietà riemanniana, avremo:

$$(\Gamma^i_{km})_P \stackrel{*}{=} 0, \qquad (\partial_l \Gamma^i_{km})_P \stackrel{*}{\neq} 0. \tag{10.155}$$

Il tensore di Riemann, nel dato sistema di coordinate geodetiche, si riduce all'espressione

$$R^i{}_{klm} \stackrel{*}{=} \partial_l \Gamma^i_{km} - \partial_m \Gamma^i_{kl}. \tag{10.156}$$

Derivando parzialmente la (10.156) rispetto ad x^j otteniamo:

$$\partial_j R^i{}_{klm} \stackrel{*}{=} \partial_{jl} \Gamma^i_{km} - \partial_{jm} \Gamma^i_{kl}, \tag{10.157}$$

operando ora una permutazione circolare sugli indici j, l, m lasciando fissi gli indici i, k, avremo altre due relazioni analoghe alla (10.157), cioè otterremo le tre relazioni, inclusa la (10.157):

$$\begin{cases} \partial_j R^i{}_{klm} \stackrel{*}{=} \partial_{jl} \Gamma^i_{km} - \partial_{jm} \Gamma^i_{kl} \\ \partial_l R^i{}_{kmj} \stackrel{*}{=} \partial_{lm} \Gamma^i_{kj} - \partial_{lj} \Gamma^i_{km} \\ \partial_m R^i{}_{kjl} \stackrel{*}{=} \partial_{mj} \Gamma^i_{kl} - \partial_{ml} \Gamma^i_{kj}, \end{cases} \tag{10.158}$$

tutte valide nel generico punto P della varietà nel sistema di coordinate geodetiche fissato.
Sommando membro a membro le tre relazioni (10.158) si ottiene:

$$\partial_j R^i{}_{klm} + \partial_l R^i{}_{kmj} + \partial_m R^i{}_{kjl} \stackrel{*}{=} 0. \tag{10.159}$$

Ma nel particolare sistema di coordinate geodetiche da noi scelto, le derivate parziali ordinarie coincidono con le corrispondenti derivate covarianti, perciò la (10.159) equivale alla:

$$\nabla_j R^i{}_{klm} + \nabla_l R^i{}_{kmj} + \nabla_m R^i{}_{kjl} \stackrel{*}{=} 0. \tag{10.160}$$

Il primo membro della (10.160) è un tensore del 5° ordine che è nullo nel punto P nel sistema di coordinate geodetiche scelto. Essendo però un'equazione tensoriale la (10.160) vale in P in tutti i possibili sistemi di coordinate, e poiché P è un punto generico della varietà riemanniana M^n, tale identità è valida in tutta la varietà, cioè si ha:

$$\nabla_j R^i{}_{klm} + \nabla_l R^i{}_{kmj} + \nabla_m R^i{}_{kjl} = 0, \qquad \forall P \in M^n. \tag{10.161}$$

La (10.161), essendo una identità tensoriale, racchiude in sé un certo numero di identità differenziali, e si tratta di relazioni cicliche fra le derivate covarianti del tensore di Riemann. Queste identità differenziali di cui gode il tensore di Riemann, sono note come **identità di Bianchi.**[16] Esse sono di fondamentale importanza sia nella geometria riemanniana sia nelle Teoria della Relatività Generale: *senza di esse le equazioni di campo di Einstein (equazioni gravitazionali) non si sarebbero potute formulare!*

10.11.5 Tensore di Ricci e scalare di curvatura

Dal tensore di Riemann, per contrazione può ottenersi un *tensore doppio simmetrico*. Facciamo osservare che affinché la contrazione non dia risultato nullo dovrà saturarsi l'indice in alto della prima coppia con un indice in basso della seconda. Precisamente, *noi converremo che la saturazione avvenga tra il 1° e 3° indice del tensore scritto sotto la forma mista* $R^j{}_{ijk}$ con cui lo abbiamo definito, avremo:[17]

$$R_{ik} \stackrel{\text{def}}{=} R^j{}_{ijk}. \tag{10.162}$$

Possiamo anche definirlo nelle due seguenti forme equivalenti:

$$R_{ik} \stackrel{\text{def}}{=} \delta^j_l R^l{}_{ijk}; \quad R_{ik} \stackrel{\text{def}}{=} g^{lj} R_{lijk} = R^j{}_{ijk} \tag{10.163}$$

l'ultima delle quali utilizza il tensore di Riemann nella sua espressione completamente covariante ed il tensore metrico che consente di innalzare un indice.
È facile riconoscere che il tensore R_{ik} definito dalle precedenti relazioni è un *tensore doppio simmetrico covariante*, infatti si ha

$$R_{ik} \stackrel{\text{def}}{=} R^j{}_{ijk} = g^{jl} R_{lijk} = g^{jl} R_{jkli} = R^l{}_{kli} = R_{ki}, \tag{10.164}$$

cioè:

$$R_{ik} = R_{ki}. \tag{10.165}$$

Questo tensore, definito dalla (10.162), va sotto il nome di **tensore contratto di curvatura o tensore contratto di Riemann** o più semplicemente **tensore di Ricci**, e come R_{lijk} è fondamentale sia nella geometria riemanniana che in Relatività Generale.
Sono utili anche le seguenti forme del tensore di Ricci, completamente controvariante la prima e mista la seconda:

$$R^{ik} = g^{il} g^{km} R_{lm}; \qquad R^i{}_k = g^{il} R_{lk}. \tag{10.166}$$

Eseguendo un'ulteriore contrazione si ottiene il cosiddetto **scalare di curvatura o scalare di Ricci** della varietà M^n:

$$R^k{}_k = g^{ik} R_{ik} = R_i{}^i = R \quad (\text{traccia di } R^i{}_k). \tag{10.167}$$

[16]Queste identità furono scoperte, per la prima volta, dal Prof. Ernesto Padova che le pubblicò nella nota *Sulle deformazioni infinitesime*, Rend. della Regia Accademia dei Lincei (4), vol.V, 1° Semestre 1889, pag.176. Il Padova le comunicò verbalmente al Prof. Gregorio Ricci-Curbastro che però se ne dimenticò. Il Prof. Luigi Bianchi, dell'Università di Pisa, le riscoprì nel 1902 e ne pubblicò la dimostrazione, conseguita mediante calcolo diretto (ibidem, (5), vol.XI, 1° Semestre 1902, pagg. 3-7). È interessante osservare che sia il Ricci sia il Levi-Civita furono entrambi allievi di Ernesto Padova.

[17]Osserviamo che alcuni autori definiscono il tensore di Ricci facendo una contrazione fra il 1° e 4° indice nel tensore di Riemann, si ottiene così un tensore di segno opposto rispetto a quello da noi definito seguendo Landau-Lifschitz.

Osserviamo ancora, come abbiamo accennato all'inizio del paragrafo, che se sul tensore di Riemann $R^j{}_{lik}$ si opera una contrazione fra il 1° ed il 2° indice si ottiene un tensore identicamente nullo, infatti si ha:

$$R^l{}_{lik} = g^{lj} R_{jlik} = -g^{lj} R_{ljik} = -R^j{}_{jik},$$

essendo j un indice muto, esso può essere sostituito nell'ultimo termine con l, la precedente relazione può essere dunque riscritta così:

$$R^l{}_{lik} = -R^l{}_{lik},$$

e ciò implica

$$R^l{}_{lik} = 0.$$

Un risultato analogo si ottiene contraendo il 3° e 4° indice.
Scriviamo ora in modo esplicito il **tensore di Ricci** a partire dal tensore misto di Riemann definito dalla (10.140) riscritta sostituendo gli indici i con j, k con i ed m con k:

$$R^j{}_{ilk} \stackrel{\text{def}}{=} \partial_l \Gamma^j_{ik} - \partial_k \Gamma^j_{il} + \Gamma^j_{ml} \Gamma^m_{ik} - \Gamma^j_{nk} \Gamma^n_{il}, \tag{10.168}$$

operando una contrazione fra il 1° e 3° indice, avremo:

$$R_{ik} = R^l{}_{ilk} \stackrel{\text{def}}{=} \partial_l \Gamma^l_{ik} - \partial_k \Gamma^l_{il} + \Gamma^l_{ml} \Gamma^m_{ik} - \Gamma^l_{nk} \Gamma^n_{il}. \tag{10.169}$$

Osserviamo che nel 2° e 3° termine del secondo membro della (10.169) figurano dei simboli di Christoffel con due indici ripetuti uno in alto e l'altro in basso, cioè Γ^l_{il} e Γ^l_{ml} per i quali è sottintesa una sommatoria rispetto a tali indici. È perciò necessario aprire una digressione per vedere come tali simboli di Christoffel possono essere calcolati.

Digressione:
Come abbiamo detto più volte, supponiamo che il tensore metrico della varietà riemanniana sulla quale operiamo non sia degenere, sia cioè

$$g = det\, g_{ik}(x) \neq 0. \tag{10.170}$$

Tale determinante è ovviamente una funzione di più variabili

$$g = g(x^1, x^2, ..., x^n) \tag{10.171}$$

ed è noto che le derivate parziali di g rispetto ad una determinata variabile x^j sono date dalla seguente formula:[18]

$$\frac{\partial g}{\partial x^j} = g g^{ik} \frac{\partial g_{ik}}{\partial x^j}, \tag{10.172}$$

che con le notazioni più volte usate si può scrivere, più semplicemente, così:

$$\partial_j g = g g^{ik} \partial_j g_{ik}. \tag{10.173}$$

Inoltre, il differenziale di g è dato da:

$$dg = \partial_j g dx^j = g g^{ik} \partial_j g_{ik} dx^j = g g^{ik} dg_{ik}. \tag{10.174}$$

Dalla (10.173) è possibile ricavare, in modo piuttosto semplice, i simboli di Christoffel di 2ª specie con due indici uguali del tipo Γ^k_{jk}. Avremo infatti per la (10.173) ed utilizzando l'identità di Ricci:

$$\partial_j g = g g^{ik} \partial_j g_{ik} = g g^{ik} (\Gamma_{ji,k} + \Gamma_{jk,i}) = g(\Gamma^i_{ji} + \Gamma^k_{jk}) = 2 g \Gamma^k_{jk},$$

[18] Per comodità del lettore, forniamo la dimostrazione della formula (10.172) nell'APPENDICE 2 di questo capitolo al paragrafo 10.13.

da cui la formula cercata:

$$\Gamma^k_{jk} = \frac{1}{2g}\frac{\partial g}{\partial x^j} \equiv \frac{1}{2g}\partial_j g. \tag{10.175}$$

È facile verificare che oltre alla (10.175) valgono le seguenti altre uguaglianze che ci saranno molto utili in seguito, nell'ambito della Relatività Generale, e, la cui semplice verifica lasciamo al lettore come utile esercizio:

$$\Gamma^k_{jk} = \frac{1}{2|g|}\frac{\partial |g|}{\partial x^j} = \frac{1}{2}\frac{\partial}{\partial x^j} log\,|g| = \frac{\partial}{\partial x^j} log\sqrt{|g|} = \frac{1}{\sqrt{|g|}}\frac{\partial\sqrt{|g|}}{\partial x^j}. \tag{10.176}$$

Tenendo conto di queste formule, possiamo concludere dicendo che il tensore di Ricci (10.169) assume la forma:

$$R_{ik} = \partial_l \Gamma^l_{ik} - \partial_{ik} log\sqrt{|g|} + \partial_m log\sqrt{|g|}\Gamma^m_{ik} - \Gamma^l_{nk}\Gamma^n_{il}. \tag{10.177}$$

che è quella più utilizzata nella pratica.

10.11.6 Tensore di Einstein e sua divergenza

Faremo ora vedere che manipolando in maniera opportuna le identità di Bianchi si dimostra l'esistenza di un tensore doppio simmetrico, essenziale per la formulazione delle equazioni di campo della Relatività Generale, il cosiddetto **tensore di Einstein** G^{ik} chiamato anche doverosamente **tensore di Einstein- Levi Civita** in quanto il grande matematico italiano contribuì in modo determinante alla sua corretta comprensione e formulazione. Tale tensore si esprime come *un'opportuna combinazione fatta con il tensore di Ricci R_{ik}, col tensore metrico g_{ik} e con lo scalare di curvatura R.*
A tal fine dimostriamo il seguente teorema fondamentale:

Teorema 10.6 *Le identità di Bianchi hanno come conseguenza necessaria l'esistenza del seguente tensore doppio simmetrico:*

$$G^{ik} \stackrel{def}{=} R^{ik} - \frac{1}{2}Rg^{ik}, \qquad (i,k = 1,2,....,n) \tag{10.178}$$

avente la proprietà di avere divergenza nulla, cioè verificante l'identità:

$$\nabla_i G^{ik} = 0. \tag{10.179}$$

Dimostrazione·
Consideriamo l'identità di Bianchi esprimendo il tensore di Riemann-Christoffel nella forma mista di tipo $(2,2)$ cioè con due indici di covarianza e due indici di controvarianza:

$$\nabla_l R^{in}{}_{kj} + \nabla_k R^{in}{}_{jl} + \nabla_j R^{in}{}_{lk} = 0, \tag{10.180}$$

ed eseguiamo una contrazione sugli indici i,k, si avrà:

$$\nabla_l R^{in}{}_{ij} + \nabla_i R^{in}{}_{jl} + \nabla_j R^{in}{}_{li} = 0. \tag{10.181}$$

Il tensore di Riemann che figura nella prima derivata covariante, nella (10.181), è ovviamente uguale al tensore di Ricci in forma mista (perché si ha una contrazione sul 1° e 3° indice), cioè $R^{in}{}_{ij} = R^n{}_j$; il tensore di Riemann che figura nella seconda derivata covariante, per l'antisimmetria rispetto allo scambio tra il 1° ed il 2° indice della prima coppia e analogamente per gli indici della seconda coppia, si può scrivere $R^{in}{}_{jl} = R^{ni}{}_{lj}$. Infine, il tensore di Riemann che figura nel terzo termine della (10.181), per l'antisimmetria rispetto allo scambio del 1° col 2° indice della

seconda coppia di indici, e tenendo ancora conto della definizione del tensore di Ricci, può scriversi così $R^{in}{}_{li} = -R^{in}{}_{il} = -R^{n}{}_{l}$. La (10.181) diventa allora:

$$\nabla_l R^n{}_j + \nabla_i R^{ni}{}_{lj} - \nabla_j R^n{}_l = 0. \tag{10.182}$$

Facciamo ora una contrazione su l ed n, ed introduciamo ancora, nel secondo termine della precedente identità, il tensore di Ricci, mentre sul terzo termine introduciamo lo scalare di curvatura R, si ha:

$$\nabla_l R^l{}_j + \nabla_i R^i{}_j - \nabla_j R = 0. \tag{10.183}$$

Essendo, nel primo termine della (10.183) l'indice l muto e lo stesso dicasi dell'indice i nel secondo termine, questi due termini sono uguali; se li sommiamo e dividiamo primo e secondo membro della relazione ottenuta per 2, avremo:

$$\nabla_i R^i{}_j - \frac{1}{2}\nabla_j R = 0. \tag{10.184}$$

Utilizzando il tensore misto di Kronecker δ^i_j, il termine $\nabla_j R$ possiamo riscriverlo così $\nabla_j R = \delta^i_j \nabla_i R$. Se ora ricordiamo che il tensore misto di Kronecker si identifica col tensore metrico in forma mista, cioè $\delta^i_j = g^i_j$ e applichiamo il teorema di Ricci sulla derivazione covariante possiamo scrivere:

$$\nabla_j R = \delta^i_j \nabla_i R = \nabla_i(\delta^i_j R) = \nabla_i(g^i{}_j R). \tag{10.185}$$

La (10.184) prende allora la forma definitiva:

$$\nabla_i(R^i{}_j - \frac{1}{2}g^i{}_j R) = 0. \tag{10.186}$$

Questa identità mette in luce *l'esistenza* di un tensore misto del 2° ordine che indicheremo così:

$$G^i_j \stackrel{\text{def}}{=} R^i{}_j - \frac{1}{2}g^i{}_j R. \tag{10.187}$$

È questo l'importantissimo **tensore di Einstein** scritto sotto forma mista. È ovvio che esso può essere scritto sia in forma covariante sia in forma controvariante, utilizzando in maniera opportuna il tensore metrico della varietà riemanniana M^n sulla quale stiamo operando, ed è manifestamente **simmetrico**:

$$G_{ik} \stackrel{\text{def}}{=} R_{ik} - \frac{1}{2}g_{ik}R, \qquad G^{ik} \stackrel{\text{def}}{=} R^{ik} - \frac{1}{2}g^{ik}R. \tag{10.188}$$

La (10.186) oltre che mettere in luce l'esistenza di questo tensore ne fornisce anche una notevole proprietà, e cioè:
la divergenza del tensore di Einstein è identicamente nulla, ovvero esso è un tensore conservativo verificante cioè le identità:

$$\nabla_i G^i_k \equiv 0, \qquad \nabla_i G^{ik} \equiv 0. \tag{10.189}$$

Il teorema è così completamete dimostrato.[19]

[19]Facciamo le seguenti due osservazioni:

1) Le precedenti identità possono essere generalizzate, tenuto conto del teorema di Ricci. Si avrà

$$\nabla_i(AG^i{}_k + Bg^i{}_k) = 0, \qquad \nabla_i(AG^{ik} + Bg^{ik}) = 0,$$

essendo A e B due costanti arbitrarie.

2) Se si vuole scrivere la divergenza del tensore G_{ik} nella sua forma covariante si dovrebbe scrivere $\nabla_i G_{ik} = 0$ che però è inconsistente con le notazioni che abbiamo adottato, si procede perciò così:

$$\nabla_i G^i_k = \nabla_i(g^{ih}G_{hk}) = g^{ih}(\nabla_i G_{hk}) = 0.$$

10.11.7 Relazione tra tensore di Riemann e parallelismo

Nel paragrafo 9.1.11 sottolineammo la circostanza che il trasporto parallelo su una varietà riemanniana M^n dipende generalmente dalla curva γ di trasporto. Si presenta allora spontaneo il problema di caratterizzare le varietà M^n nelle quali, *comunque venga scelto il vettore di partenza, il trasporto parallelo non dipende dal cammino γ.*
Tale condizione equivale alla possibilità di determinare univocamente in M^n un campo di vettori $v_i(x)$ riducentesi, in un particolare punto $\overline{P}$, ad un vettore $\overline{\mathbf{v}}$ arbitrariamente prescelto e soddisfacente alle equazioni di parallelismo:

$$Dv_i \equiv \nabla_j v_i dx^j \equiv dv_i - \Gamma^l_{ij} v_l dx^j = 0, \qquad (i, j, l = 1, 2, ..., n) \tag{10.190}$$

per ogni scelta del cammino elementare di trasporto dx^j. Data l'arbitrarietà delle dx^j, le equazioni (10.190) sono equivalenti al seguente sistema di n^2 equazioni alle derivate parziali nelle n incognite v_i:

$$\nabla_j v_i \equiv \frac{\partial v_i}{\partial x^j} - \Gamma^l_{ij} v_l = 0, \qquad (i, j, l = 1, 2, ..., n). \tag{10.191}$$

La sovrabbondanza delle equazioni rispetto alle incognite rende generalmente impossibile il sistema (10.191), a meno che non siano soddisfatte certe condizioni di compatibilità. Per ottenerle nel modo più rapido si osservi che le (10.191) implicano:

$$\nabla_h \nabla_j v_i = 0 \qquad e \qquad \nabla_j \nabla_h v_i = 0, \tag{10.192}$$

e quindi per differenza (si tenga presente la (10.142)):

$$\nabla_h \nabla_j v_i - \nabla_j \nabla_h v_i = -v_k R^k{}_{ijh} = 0, \tag{10.193}$$

valida in ogni punto di M^n. In particolare essa è valida in $\overline{P}$ dove, per l'arbitrarietà di $\overline{v}_i$, essa implica $R^k{}_{ijh}(\overline{P}) = 0$. Tenuto poi conto che anche $\overline{P}$ è completamente arbitrario, se ne deduce, per la compatibilità delle (10.191), che:

$$R^k{}_{ijh} = 0, \tag{10.194}$$

in tutta M^n.
Viceversa, come si potrebbe dimostrare[20], la (10.194) supposta verificata ovunque, è anche sufficiente per l'integrabilità del sistema (10.191), il quale diventa così atto a determinare univocamente il campo di vettori v_i non appena se ne assegni, ad arbitrio, la determinazione in un particolare punto $\overline{P}$. Resta così provato che *l'annullarsi identico del tensore di Riemann in una M^n, o in una sua regione, è condizione necessaria e sufficiente per l'indipendenza della operazione di trasporto parallelo dal cammino di trasporto.*
Si potrebbe anche dimostrare, ma vi rinunciamo per brevità, *che la suddetta indipendenza è caratteristica delle varietà euclidee.*[21] Può pertanto concludersi che **l'annullarsi identico del tensore di Riemann è caratteristico delle varietà euclidee.** Possiamo anzi affermare che: *carattere localmente euclideo di M^n, annullamento del tensore di Riemann ed indipendenza del trasporto parallelo dalla curva di trasporto sono circostanze concomitanti.* Non possiamo però, per ragioni di tempo e di spazio, soffermarci su tali questioni del più alto significato speculativo e concettuale, rimandando per esse ai trattati di Geometria differenziale e di Geometria riemanniana.

[20] Una dimostrazione si può trovare nel classico libro di Élie Cartan, *Leçons sur la Géométrie des espaces de Riemann*, Editions Gauthier-Villars, Paris, 1928, pagg. 53-58.

[21] Il lettore, desideroso dei dettagli, può trovare la dimostrazione per es. nel classico trattato di Paul Appell, *Traité de Mécanique Rationelle,* tome V, Chap. IV, n. 88, pag.115, Gauthier-Villars et Cie Éditeurs, Paris (1926).

10.12 APPENDICE 1: Un approccio alternativo alla derivazione covariante

10.12.1 Derivazione covariante di un campo tensoriale in $\mathbb{R}^n$

Tenuto conto delle considerazioni che abbiamo fatto nel paragrafo 10.1, all'inizio di questo capitolo, vogliamo ora introdurre il concetto di derivata covariante senza usare la nozione di parallelismo ma, in modo diretto e costruttivo, dapprima nello spazio euclideo $\mathbb{R}^n$ e poi in forma assiomatica in una varietà differenziabile arbitraria M^n.[22]
Il nostro scopo è dunque capire come differenziare invariantemente i campi tensoriali in coordinate curvilinee. Ciò è necessario perché le coordinate locali su di una arbitraria varietà differenziabile M^n sono *quasi sempre* curvilinee. Assumiamo, per il momento, che sia $M^n = \mathbb{R}^n$ e poniamoci il problema di cercare un'operazione, che denoteremo con ∇, avente le seguenti proprietà:

1. in coordinate cartesiane in $\mathbb{R}^n$, l'operazione ∇ deve coincidere con la derivata parziale ordinaria:
$$\nabla \equiv \{\partial_i\} \equiv \left\{\frac{\partial}{\partial x^i}\right\};$$
2. l'operazione ∇ deve essere tensoriale, cioè se T è un campo tensoriale su $\mathbb{R}^n$ (riferito a coordinate curvilinee), ∇T deve essere ancora un tensore.

Cominciamo con degli esempi.
Consideriamo in $\mathbb{R}^n$ un campo vettoriale controvariante T^i; sia (x) **un sistema di coordinate cartesiane** e (x') **un sistema di coordinate arbitrario**. Applichiamo le condizioni 1 e 2 sopra imposte a ∇. Nel sistema (x), essendo esso cartesiano, avremo

$$\nabla_j T^i \equiv \frac{\partial T^i}{\partial x^j}, \tag{10.195}$$

mentre nel sistema arbitrario (x') avremo (per la condizione 2):

$$\nabla_{j'} T^{i'} = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\nabla_j T^i. \tag{10.196}$$

Il problema consiste perciò nel trovare la forma esplicita dell'operazione ∇ e calcolare le componenti $\nabla_{j'} T^{i'}$ in un arbitrario sistema di coordinate. Un calcolo diretto ci dà:

$$\begin{aligned}
\nabla_{j'} T^{i'} &= \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\nabla_j T^i = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial}{\partial x^j}\left(\frac{\partial x^i}{\partial x^{k'}}T^{k'}\right) = \\
&= \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^i}{\partial x^{k'}}\frac{\partial T^{k'}}{\partial x^{\alpha'}}\frac{\partial x^{\alpha'}}{\partial x^j} + \frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}T^{k'}\frac{\partial}{\partial x^j}\left(\frac{\partial x^i}{\partial x^{k'}}\right) = \\
&= \delta^{i'}_{k'}\delta^{\alpha'}_{j'}\frac{\partial T^{k'}}{\partial x^{\alpha'}} + T^{k'}\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{h'}}\frac{\partial x^{h'}}{\partial x^j} = \\
&= \frac{\partial T^{i'}}{\partial x^{j'}} + T^{k'}\frac{\partial x^{i'}}{\partial x^i}\delta^{h'}_{j'}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{h'}} = \frac{\partial T^{i'}}{\partial x^{j'}} + T^{k'}\frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}}.
\end{aligned}$$

Perciò, in definitiva:

$$\nabla_{j'} T^{i'} = \frac{\partial T^{i'}}{\partial x^{j'}} + T^{k'}\Gamma^{i'}_{k'j'}, \tag{10.197}$$

[22] Nella trattazione che segue ci siamo ispirati all'approccio adottato nel libro di A. Mischchenko and A. Fomenko, *A Course of Differential Geometry and Topology*, Mir Pubblishers Moscow, 1988, Cap. 5, pag. 322-336.

avendo posto, per definizione:

$$\Gamma^{i'}_{k'j'} \stackrel{\text{def}}{=} \frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}}. \tag{10.198}$$

Abbiamo così ottenuto le funzioni $\Gamma^{i'}_{k'j'}$, che misurano la *"deviazione"* di ∇ dall'ordinaria derivazione parziale su un campo vettoriale controvariante.
Consideriamo adesso un campo di covettori T_i in $\mathbb{R}^n$ e vediamo come ∇ agisce su di esso.
Come prima, nel sistema (x), essendo esso cartesiano, avremo.

$$\nabla_j T_i \equiv \frac{\partial T_i}{\partial x^j}, \tag{10.199}$$

mentre nel sistema arbitrario (x') avremo:

$$\nabla_{j'} T_{i'} = \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\nabla_j T_i. \tag{10.200}$$

Quindi, con un calcolo diretto:

$$\begin{aligned}
\nabla_{j'} T_{i'} &= \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\nabla_j T_i = \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial}{\partial x^j}\left(\frac{\partial x^{k'}}{\partial x^i}T_{k'}\right) = \\
&= \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^{k'}}{\partial x^i}\frac{\partial T_{k'}}{\partial x^{\alpha'}}\frac{\partial x^{\alpha'}}{\partial x^j} + \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}T_{k'}\frac{\partial}{\partial x^j}\left(\frac{\partial x^{k'}}{\partial x^i}\right) = \\
&= \delta^{k'}_{i'}\delta^{\alpha'}_{j'}\frac{\partial T_{k'}}{\partial x^{\alpha'}} + T_{k'}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial^2 x^{k'}}{\partial x^i\partial x^j} = \\
&= \frac{\partial T_{i'}}{\partial x^{j'}} + T_{k'}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial^2 x^{k'}}{\partial x^i\partial x^j}.
\end{aligned}$$

Perciò, in definitiva:

$$\nabla_{j'} T_{i'} = \frac{\partial T_{i'}}{\partial x^{j'}} + T_{k'}\widetilde{\Gamma}^{k'}_{i'j'}, \tag{10.201}$$

avendo posto, per definizione:

$$\widetilde{\Gamma}^{k'}_{i'j'} \stackrel{\text{def}}{=} \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial^2 x^{k'}}{\partial x^i\partial x^j}. \tag{10.202}$$

Le $\widetilde{\Gamma}^{k'}_{i'j'}$ così ottenute sono funzioni che misurano la *"deviazione"* di ∇ dalla derivazione parziale ordinaria su un campo di covettori.

Lemma 10.7 *Si ha la seguente uguaglianza:*

$$\widetilde{\Gamma}^{i'}_{k'j'} = -\Gamma^{i'}_{k'j'}.$$

Dimostrazione. Partiamo dall'ovvia identità

$$\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^i}{\partial x^{k'}} = \delta^{i'}_{k'};$$

derivando rispetto a $x^{j'}$ si ha:

$$\frac{\partial}{\partial x^{j'}}\left(\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^i}{\partial x^{k'}}\right) = \frac{\partial^2 x^{i'}}{\partial x^i\partial x^j}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^i}{\partial x^{k'}} + \frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}} = 0,$$

ovvero

$$\frac{\partial^2 x^{i'}}{\partial x^i\partial x^j}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^i}{\partial x^{k'}} + \frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}} = 0,$$

e per le (10.198), (10.202) si ottiene

$$\widetilde{\Gamma}^{i'}_{k'j'} + \Gamma^{i'}_{k'j'} = 0,$$

cioè la tesi. c.d.d.

Le funzioni $\Gamma^{i'}_{k'j'}$ vengono chiamate *simboli di Christoffel di seconda specie*; ricordando la (10.198) ed in virtù delle considerazioni precedenti, sono definite dalla duplice uguaglianza:

$$\Gamma^{i'}_{k'j'} = \frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{k'}\partial x^{j'}} = -\frac{\partial^2 x^{i'}}{\partial x^i \partial x^j}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^i}{\partial x^{k'}}. \tag{10.203}$$

In conclusione, possiamo affermare che l'azione di ∇ su un campo di vettori o di covettori (in coordinate curvilinee in $\mathbb{R}^n$) è della forma:

$$\nabla_{j'} T^{i'} = \frac{\partial T^{i'}}{\partial x^{j'}} + T^{k'}\Gamma^{i'}_{k'j'}, \tag{10.204}$$

$$\nabla_{j'} T_{i'} = \frac{\partial T_{i'}}{\partial x^{j'}} - T_{k'}\Gamma^{k'}_{i'j'}. \tag{10.205}$$

Mediante ragionamenti analoghi è possibile dimostrare che l'azione di ∇ sul tensore misto T^i_j è del tipo:

$$\nabla_{k'} T^{i'}_{j'} = \frac{\partial T^{i'}_{j'}}{\partial x^{k'}} + T^{p'}_{j'}\Gamma^{i'}_{p'k'} - T^{i'}_{p'}\Gamma^{p'}_{j'k'}. \tag{10.206}$$

Generalizzando ciò che abbiamo detto in precedenza, per un tensore arbitrario, possiamo enunciare il seguente risultato generale:

Teorema 10.8 *Sia $M^n \equiv \mathbb{R}^n$, (x) un sistema di coordinate cartesiane e (x') un arbitrario sistema di coordinate curvilinee. Esiste allora in $\mathbb{R}^n$ un'operazione tensoriale ∇ definita su un arbitrario campo tensoriale $T^{i_1\dots i_p}_{j_1\dots j_q}$ per mezzo della formula:*

$$\nabla_{\alpha'} T^{i'_1\dots i'_p}_{j'_1\dots j'_q} = \frac{\partial T^{i'_1\dots i'_p}_{j'_1\dots j'_q}}{\partial x^{\alpha'}} + \sum_{s=1}^{p} T^{i'_1\dots i'_{s-1} r' i'_{s+1}\dots i'_p}_{j'_1\dots j'_q}\Gamma^{i'_s}_{r'\alpha'} - \sum_{s=1}^{q} T^{i'_1\dots i'_p}_{j'_1\dots j'_{s-1} r' j'_{s+1}\dots j'_q}\Gamma^{r'}_{j'_s\alpha'}, \tag{10.207}$$

dove le $\Gamma^{i'}_{j'k'}$, per la (10.203), sono definite dalle relazioni:

$$\Gamma^{i'}_{j'k'} \stackrel{def}{=} \frac{\partial x^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{j'}\partial x^{k'}} = -\frac{\partial^2 x^{i'}}{\partial x^i \partial x^j}\frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^i}{\partial x^{k'}}. \tag{10.208}$$

L'operazione così introdotta prende il nome di **derivazione covariante** e il risultato di tale operazione viene chiamato **derivata covariante** del campo tensoriale considerato.
Si noti che la derivata covariante di un tensore di rango n dà luogo ad un tensore di rango $n+1$, con un indice di covarianza in più.

10.12.2 Carattere non tensoriale dei simboli di Christoffel di seconda specie

Vediamo ora di fare qualche considerazione sui simboli di Christoffel di seconda specie. Innanzitutto questi sono stati ricavati da una particolare trasformazione di coordinate, da cartesiane (x) a curvilinee (x'). Poiché contengono derivate seconde delle coordinate (x) rispetto alle (x'), è chiaro che se la trasformazione viene fatta ancora in coordinate cartesiane, i simboli di Christoffel di seconda specie sono tutti identicamente nulli.

È chiaro poi che, per il modo in cui sono stati definiti (cfr. le formule (10.198) e (10.208)), i simboli di Christoffel sono simmetrici rispetto ai due indici che figurano in basso.[23]
Mostriamo inoltre che essi non costituiscono un tensore. A tal fine vediamo come si comportano in una trasformazione del tutto generale di coordinate (da curvilinee a curvilinee) $(x') \to (x'')$. Consideriamo un tensore controvariante del 1° ordine le cui componenti nei sistemi di coordinate cartesiane (x) e curvilinee (x'), (x'') siano rispettivamente T^i, $T^{i'}$, $T^{i''}$; calcoliamo la derivata covariante di $T^{i'}$, abbiamo:

$$\nabla_{k'} T^{i'} = \frac{\partial T^{i'}}{\partial x^{k'}} + T^{p'} \Gamma^{i'}_{p'k'}, \tag{10.209}$$

in modo analogo la derivata covariante di $T^{i''}$ sarà:

$$\nabla_{k''} T^{i''} = \frac{\partial T^{i''}}{\partial x^{k''}} + T^{p''} \Gamma^{i''}_{p''k''}. \tag{10.210}$$

Ora trasformiamo il 2° membro di quest'ultima relazione tenendo conto che essendo $T^{i''}$ e $T^{p''}$ tensori, passando da un sistema di coordinate all'altro, si trasformano con legge tensoriale. Con un calcolo diretto ma un po' tedioso avremo:

$$\nabla_{k''} T^{i''} = \frac{\partial T^{i''}}{\partial x^{k''}} + T^{p''} \Gamma^{i''}_{p''k''} = \frac{\partial T^{i''}}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^{k''}} + \frac{\partial x^{p''}}{\partial x^{p'}} T^{p'} \Gamma^{i''}_{p''k''} =$$

$$= \frac{\partial}{\partial x^{k'}} \left(\frac{\partial x^{i''}}{\partial x^{i'}} T^{i'} \right) \frac{\partial x^{k'}}{\partial x^{k''}} + \frac{\partial x^{p''}}{\partial x^{p'}} T^{p'} \Gamma^{i''}_{p''k''} =$$

$$= \left[\frac{\partial^2 x^{i''}}{\partial x^{k'} \partial x^{i'}} T^{i'} + \frac{\partial x^{i''}}{\partial x^{i'}} \frac{\partial T^{i'}}{\partial x^{k'}} \right] \frac{\partial x^{k'}}{\partial x^{k''}} + \frac{\partial x^{p''}}{\partial x^{p'}} T^{p'} \Gamma^{i''}_{p''k''},$$

ed in definitiva:

$$\nabla_{k''} T^{i''} = \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} \frac{\partial T^{i'}}{\partial x^{k'}} + T^{i'} \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial^2 x^{i''}}{\partial x^{k'} \partial x^{i'}} + T^{p'} \frac{\partial x^{p''}}{\partial x^{p'}} \Gamma^{i''}_{p''k''}. \tag{10.211}$$

D'altra parte essendo $\nabla_{k''} T^{i''}$ un tensore misto del 2° ordine, nel passaggio dal sistema di coordinate (x') al sistema (x'') si trasforma con legge tensoriale, avremo perciò:

$$\nabla_{k''} T^{i''} = \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} \nabla_{k'} T^{i'} = \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} \left(\frac{\partial T^{i'}}{\partial x^{k'}} + T^{p'} \Gamma^{i'}_{p'k'} \right),$$

ovvero

$$\nabla_{k''} T^{i''} = \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} \frac{\partial T^{i'}}{\partial x^{k'}} + \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} T^{p'} \Gamma^{i'}_{p'k'}, \tag{10.212}$$

Confrontando la (10.212) con la (10.211) e cambiando in quest'ultima, nel secondo addendo del 2° membro l'indice muto i' in p' avremo:

$$T^{p'} \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} \Gamma^{i'}_{p'k'} = T^{p'} \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial^2 x^{i''}}{\partial x^{k'} \partial x^{p'}} + T^{p'} \frac{\partial x^{p''}}{\partial x^{p'}} \Gamma^{i''}_{p''k''}.$$

Poiché questa identità deve essere soddisfatta *qualunque* sia il campo tensoriale $T^{i'}$, si ottiene:

$$\frac{\partial x^{p''}}{\partial x^{p'}} \Gamma^{i''}_{p''k''} = \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial x^{i''}}{\partial x^{i'}} \Gamma^{i'}_{p'k'} - \frac{\partial x^{k'}}{\partial x^{k''}} \frac{\partial^2 x^{i''}}{\partial x^{k'} \partial x^{p'}},$$

[23] Questo perché supponiamo implicitamente che le trasformazioni di coordinate siano almeno di classe C^2, e che quindi sia valido il teorema d'inversione dell'ordine di derivazione.

da cui, moltiplicando ambo i membri per $\partial x^{p'}/\partial x^{j''}$ e contraendo su p':

$$\delta^{p''}_{j''}\Gamma^{i''}_{p''k''} = \frac{\partial x^{p'}}{\partial x^{j''}}\frac{\partial x^{k'}}{\partial x^{k''}}\frac{\partial x^{i''}}{\partial x^{i'}}\Gamma^{i'}_{p'k'} - \frac{\partial x^{p'}}{\partial x^{j''}}\frac{\partial x^{k'}}{\partial x^{k''}}\frac{\partial^2 x^{i''}}{\partial x^{k'}\partial x^{p'}}.$$

Infine ricordando il risultato del Lemma 10.7:

$$-\frac{\partial x^{p'}}{\partial x^{j''}}\frac{\partial x^{k'}}{\partial x^{k''}}\frac{\partial^2 x^{i''}}{\partial x^{k'}\partial x^{p'}} = \frac{\partial x^{i''}}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^{j''}\partial x^{k''}}, \tag{10.213}$$

otteniamo la cercata importante formula di trasformazione dei simboli di Christoffel di seconda specie nel passaggio dal sistema di coordinate curvilinee (x') all'altro sistema (x'') pur esso di coordinate curvilinee:

$$\Gamma^{i''}_{j''k''} = \frac{\partial x^{i''}}{\partial x^{i'}}\frac{\partial x^{j'}}{\partial x^{j''}}\frac{\partial x^{k'}}{\partial x^{k''}}\Gamma^{i'}_{j'k'} + \frac{\partial x^{i''}}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^{j''}\partial x^{k''}}. \tag{10.214}$$

Come si vede chiaramente dalla (10.214), i simboli di Christoffel di seconda specie non si trasformano, in generale, con legge tensoriale, ovvero non formano un tensore a meno che la trasformazione di coordinate sia lineare nel qual caso sono nulle tutte le derivate parziali seconde fatte sulle coordinate ed in conseguenza la legge di trasformazione diventa tensoriale.

10.12.3 Derivazione covariante di un campo tensoriale su di una varietà

Finora abbiamo dimostrato l'esistenza di un'operazione, chiamata derivazione covariante, ma soltanto nel caso che $M^n \equiv \mathbb{R}^n$, cioè quando M^n è munito di coordinate cartesiane. Ciò ci ha consentito di calcolare esplicitamente i simboli di Christoffel, che misurano la deviazione di ∇ dall'ordinaria derivazione parziale. Per fare questo abbiamo ammesso tacitamente che in $\mathbb{R}^n$ esistano coordinate cartesiane rispetto alle quali l'operazione ∇ coincide con l'ordinaria derivazione parziale.

Consideriamo ora un'arbitraria varietà differenziabile[24] M^n e definiamo l'operazione ∇ *assiomaticamente*, in modo tale che abbia le proprietà dimostrate per ∇ in $\mathbb{R}^n$.

Definizione:
Se esistono n^3 funzioni $\Gamma^k_{ij}(x)$ (a priori non simmetriche rispetto agli indici inferiori) che si trasformano per un generico cambiamento di coordinate (curvilinee) $(x) \to (x')$ secondo la legge:

$$\Gamma^{k'}_{i'j'} = \frac{\partial x^{k'}}{\partial x^k}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\Gamma^k_{ij} + \frac{\partial x^{k'}}{\partial x^k}\frac{\partial^2 x^k}{\partial x^{i'}\partial x^{j'}}, \tag{10.215}$$

allora si definisce l'operazione di derivazione covariante ∇ come quella che su un generico tensore $T^{i_1\dots i_p}_{j_1\dots j_q}$ agisce (in ogni sistema di coordinate) verificando la seguente relazione:

$$\nabla_\alpha T^{i_1\dots i_p}_{j_1\dots j_q} = \frac{\partial T^{i_1\dots i_p}_{j_1\dots j_q}}{\partial x^\alpha} + \sum_{s=1}^{p} T^{i_1\dots i_{s-1} r\, i_{s+1}\dots i_p}_{j_1\dots j_q}\Gamma^{i_s}_{r\alpha} - \sum_{s=1}^{q} T^{i_1\dots i_p}_{j_1\dots j_{s-1} r\, j_{s+1}\dots j_q}\Gamma^{r}_{j_s\alpha}. \tag{10.216}$$

In questa definizione il sistema (x) (non accentato) non è più cartesiano perché ora abbiamo a che fare con una varietà arbitraria M^n.
Si suole dire che l'operazione ∇ (o, equivalentemente, l'insieme delle funzioni $\Gamma^k_{ij}(x)$) definisce una *connessione* (affine) su M^n. Le funzioni $\Gamma^k_{ij}(x)$ continueranno ad essere chiamate *simboli di Christoffel di seconda specie*, oppure anche, in virtù del concetto di connessione, *coefficienti di connessione*.
L'esistenza di varietà sulle quali l'operazione ∇ è valida, è già stata provata, possiamo scegliere

[24] D'ora in avanti per varietà differenziabile intenderemo una varietà liscia, o di classe C^r, con $r \geq 2$.

infatti $M^n \equiv \mathbb{R}^n$; allora le $\Gamma^k_{ij}(x)$ sono simmetriche rispetto agli indici inferiori e sono esprimibili per mezzo delle (10.208).
L'insieme delle Γ^k_{ij} non forma un tensore, anche se esse si trasformano con legge tensoriale nel caso di una trasformazione lineare di coordinate. A partire dai coefficienti di connessione è possibile tuttavia definire un nuovo ente, avente carattere tensoriale, che chiameremo **tensore di torsione** della connessione Γ^k_{ij}:

$$\Omega^k_{ij} \stackrel{\text{def}}{=} \Gamma^k_{ij} - \Gamma^k_{ji}. \tag{10.217}$$

Il fatto che questo sia un tensore si può verificare facilmente utilizzando la formula (10.215) di trasformazione dei simboli di Christoffel di seconda specie. Se risulta $\Omega^k_{ij} \equiv 0$ si dice che la connessione definita dalle Γ^k_{ij} è *simmetrica*. La connessione introdotta in $\mathbb{R}^n$ è manifestamente simmetrica, ciò segue dall'espressione delle Γ^k_{ij} in termini delle derivate parziali prime e seconde delle coordinate come è chiaro dalla loro definizione espressa dalla (10.208).

Su di una arbitraria varietà M^n, i simboli di Christoffel non sono necessariamente esprimibili attraverso le derivate parziali prime e seconde delle coordinate; tale forma si ha soltanto se esistono localmente su M^n coordinate cartesiane.

La derivazione covariante gode di alcune proprietà notevoli. Si ha in proposito il seguente:

Teorema 10.9 *La derivazione covariante ∇ su una generica varietà M^n soddisfa le seguenti proprietà:*

- *l'operazione $\nabla \equiv \{\nabla_k\}$ è lineare;*
- *se $T^{i_1\dots i_p}_{j_1\dots j_q}$ è un generico campo tensoriale, l'insieme di funzioni $\nabla_\alpha T^{i_1\dots i_p}_{j_1\dots j_q}$ forma un campo tensoriale;*
- *se il campo tensoriale è scalare (cioè una funzione φ regolare su M^n), allora $\nabla_k\varphi = \partial_k\varphi$, ovvero $\nabla\varphi = \mathbf{grad}\,\varphi$;*
- *l'operazione ∇ soddisfa la seguente regola di Leibniz:*

$$\nabla_\alpha(T^{i_1\dots i_p}_{j_1\dots j_q} R^{k_1\dots k_r}_{h_1\dots h_s}) = (\nabla_\alpha T^{i_1\dots i_p}_{j_1\dots j_q}) R^{k_1\dots k_r}_{h_1\dots h_s} + T^{i_1\dots i_p}_{j_1\dots j_q}(\nabla_\alpha R^{k_1\dots k_r}_{h_1\dots h_s}), \tag{10.218}$$

  dove $T^{i_1\dots i_p}_{j_1\dots j_q}$ e $R^{k_1\dots k_r}_{h_1\dots h_s}$ sono campi tensoriali arbitrari;
- *l'operazione ∇ è permutabile con l'operazione di contrazione, ovvero, dato un qualsiasi campo tensoriale $T^{i_1\dots i_p}_{j_1\dots j_q}$ è possibile eseguire prima la derivazione covariante $\nabla_\alpha T^{i_1\dots i_p}_{j_1\dots j_q}$ e poi contrarre due indici di tipo diverso (non α), oppure il viceversa.*

10.13 APPENDICE 2: Derivata di un determinante

Dimostriamo dapprima il teorema che ci consente di calcolare la derivata di un determinante i cui elementi $g_{ik}(x)$ sono funzioni di una sola variabile x; estenderemo poi il risultato al caso di un determinante i cui elementi sono funzioni di più variabili.
Teorema:
Sia $(g_{ik}(x))$ una matrice quadrata di ordine n i cui n^2 elementi siano funzioni derivabili di una variabile x, è chiaro che il $\det(g_{ik}) = g$ sarà anch'esso funzione derivabile di x, avremo allora:

$$g(x) = \begin{vmatrix} g_{11}(x) & g_{12}(x) & \dots & g_{1n}(x) \\ g_{21}(x) & g_{22}(x) & \dots & g_{2n}(x) \\ \dots & \dots & \dots & \dots \\ g_{n1}(x) & g_{n2}(x) & \dots & g_{nn}(x) \end{vmatrix} \tag{10.219}$$

$$g'(x) = \begin{vmatrix} g'_{11}(x) & \dots & g'_{1n}(x) \\ g_{21}(x) & \dots & g_{2n}(x) \\ \dots & \dots & \dots \\ g_{n1}(x) & \dots & g_{nn}(x) \end{vmatrix} + \begin{vmatrix} g_{11}(x) & \dots & g_{1n}(x) \\ g'_{21}(x) & \dots & g'_{2n}(x) \\ \dots & \dots & \dots \\ g_{n1}(x) & \dots & g_{nn}(x) \end{vmatrix} + \dots + \begin{vmatrix} g_{11}(x) & \dots & g_{1n}(x) \\ g_{21}(x) & \dots & g_{2n}(x) \\ \dots & \dots & \dots \\ g'_{n1}(x) & \dots & g'_{nn}(x) \end{vmatrix} \tag{10.220}$$

cioè la derivata di un determinante di ordine n è data dalla somma degli n determinanti che si ottengono dal dato sostituendo agli elementi della $1^a, 2^a, \dots\dots, n^a$ riga (oppure agli elementi della $1^a, 2^a, \dots\dots, n^a$ colonna) di esso le loro rispettive derivate.

Dimostrazione. Infatti dalla definizione di determinante si ha[25]:

$$g(x) = \varepsilon^{k_1 k_2 \cdots k_n} g_{1k_1} g_{2k_2} \cdots g_{nk_n}, \tag{10.221}$$

dove $\varepsilon^{k_1 k_2 \cdots k_n}$ è il cosiddetto *indicatore o simbolo di Levi- Civita*, così definito:

$$\varepsilon^{k_1 k_2 \cdots k_n} = \begin{cases} +1, & \text{se } k_1 k_2 \cdots k_n \text{ è una permutazione di classe pari della permutazione} \\ & \text{naturale che è quella strettamente crescente degli stessi indici;} \\ -1, & \text{se } k_1 k_2 \cdots k_n \text{ è una permutazione di classe dispari della permutazione} \\ & \text{naturale degli stessi indici;} \\ 0, & \text{se almeno due degli indici } k_1 k_2 \cdots k_n \text{ sono uguali.} \end{cases}$$

Derivando rispetto ad x si ottiene il risultato voluto:

$$g'(x) = \varepsilon^{k_1 k_2 \cdots k_n} g'_{1k_1} g_{2k_2} \cdots g_{nk_n} + \varepsilon^{k_1 k_2 \cdots k_n} g_{1k_1} g'_{2k_2} \cdots g_{nk_n} + \dots \varepsilon^{k_1 k_2 \cdots k_n} g_{1k_1} g_{2k_2} \cdots g'_{nk_n}. \tag{10.222}$$

La formula precedente può essere messa in una forma molto espressiva che ora ci proponiamo di dimostrare. Osserviamo che ciascuno degli n determinanti dello sviluppo (10.222) può ottenersi applicando il 1° teorema di Laplace sviluppando ciascuno di essi secondo gli elementi della riga che contiene le derivate, avremo:

$$g'(x) = (g'_{11} A^{11} + \dots g'_{1n} A^{1n}) + (g'_{21} A^{21} + \dots g'_{2n} A^{2n}) + \dots (g'_{n1} A^{n1} + \dots + g'_{nn} A^{nn}), \tag{10.223}$$

avendo indicato con A^{ik} il cofattore (o complemento algebrico) dell'elemento g_{ik} nel determinante $g(x)$. Per la convenzione di sommazione sui secondi indici potremo scrivere:

$$g'(x) = g'_{1k} A^{1k} + g'_{2k} A^{2k} + \dots + g'_{nk} A^{nk},$$

ovvero, con una doppia somma su i e k, più sinteticamente così:

$$g'(x) = g'_{ik} A^{ik}, \tag{10.224}$$

che è la formula cercata.

■

Se ora supponiamo che gli elementi g_{ik} del dato determinante siano funzioni di più variabili indipendenti, cioè:

$$g_{ik} = g_{ik}(x^1, x^2, \dots, x^n),$$

avremo anche che il determinante $g = det(g_{ik})$ è funzione delle stesse variabili

$$g \equiv det(g_{ik}) = g(x^1, x^2, \dots, x^n).$$

[25]Tale formula si può trovare per es. nell'ottimo libro di Enrico Bompiani *Geometria Analitica*, Editrice "Lo Scaffale" - Diffusione del libro, Roma (1961), pagg.22-28.

Derivando allora parzialmente rispetto alla variabile x^j, per la (10.224), avremo:

$$\frac{\partial g}{\partial x^j} = \frac{\partial g_{ik}}{\partial x^j} A^{ik}. \tag{10.225}$$

Introducendo gli elementi reciproci degli g_{ik} nel dato determinante, avremo per definizione

$$g^{ik} = \frac{A^{ik}}{g},$$

dalla quale si ricavano i complementi algebrici:

$$A^{ik} = g g^{ik}, \tag{10.226}$$

per cui la (10.225) assume la forma definitiva, molto utile in pratica:

$$\frac{\partial g}{\partial x^j} = g g^{ik} \frac{\partial g_{ik}}{\partial x^j}. \tag{10.227}$$

A completamento delle formule precedenti calcoliamo il differenziale di g, avremo per definizione:

$$dg = \frac{\partial g}{\partial x^j} dx^j = \partial_j g dx^j, \tag{10.228}$$

e tenuto conto delle (10.225), (10.226) si ottiene:

$$dg = \frac{\partial g_{ik}}{\partial x^j} A^{ik} dx^j = dg_{ik} A^{ik} = g g^{ik} dg_{ik}. \tag{10.229}$$

Questa formula verrà utilizzata nel paragrafo 11.16 del Cap.11 nella deduzione variazionale delle equazioni gravitazionali di Einstein.

Capitolo 11

FONDAMENTI DELLA TEORIA EINSTEINIANA DELLA GRAVITAZIONE

Newton, perdonami; tu hai trovato la sola via che,
ai tuoi tempi, fosse possibile per un uomo di altissimo intelletto
e potere creativo. I concetti che tu hai creato guidano ancora oggi
il nostro pensiero nel campo della fisica, anche se ora noi sappiamo
che dovranno essere sostituiti con altri assai più discosti dalla sfera
della nostra esperienza immediata, se si vorrà raggiungere una
conoscenza più profonda dei rapporti fra le cose.
(Albert Einstein, *Autobiografia scientifica*, Boringhieri, Torino, 1979, pag. 24.)

Einstein, nel costruire la sua teoria della gravitazione, non cercava di spiegare
qualche risultato osservativo. Al contrario. Tutto il suo modo di procedere tendeva alla ricerca di una
teoria bella, una teoria come l'avrebbe scelta la Natura.
Era guidato solo dal requisito che la teoria avesse la bellezza e l'eleganza che ci si
aspetta di trovare in una descrizione fondamentale della Natura. Il suo lavoro
muoveva esclusivamente dall'idea di come la Natura dovrebbe essere e non dalla
necessità di dar conto di certi risultati sperimentali.
(Paul A. M. Dirac, *La bellezza come metodo*, Indiana Editore srl, Milano, 2013, pag. 175.)

In questo capitolo e nei due seguenti si espongono le idee fondamentali della Teoria della Relatività Generale così come il suo unico fondatore, Albert Einstein, l'ha formulata nel 1915. Da allora ragguardevoli contributi sono stati apportati da moltissimi eccellenti scienziati; molti oggetti sono stati scoperti nel campo dell'Astronomia e dell'Astrofisica: quasar, pulsar, buchi neri, lenti gravitazionali, radiazione cosmica di fondo, modelli cosmologici, ecc. Essi hanno tutti a che fare con la costruzione teorica einsteiniana della Relatività Generale. Non abbiamo certamente la pretesa, in un libro per principianti, di discutere tali questioni la cui trattazione completa riempie svariati volumi. Confidiamo che la nostra trattazione elementare della Relatività Generale ne metta in luce almeno le idee fondamentali.

11.1 Introduzione

La meccanica di Newton e la sua teoria della gravitazione hanno dato in campo astronomico risultati così meravigliosi che veramente c'è da chiedersi come mai Einstein abbia sentito la necessità di una **nuova teoria della gravitazione**. Infatti prima del 1915, anno in cui Einstein presentò la sua monumentale formulazione relativistica di questa teoria - chiamata **Relatività Generale** - poche teorie fisiche potevano essere paragonate alla teoria della gravitazione di Newton, in quanto a precisione delle loro previsioni. Il moto e la posizione dei pianeti e dei satelliti, entro il sistema solare, le eclissi del Sole e della Luna, il moto delle comete, ecc. sono previsti dalla meccanica celeste classica con precisione a dir poco sbalorditiva, inoltre tali previsioni sono in accordo con i dati osservativi entro pochi secondi d'arco su intervalli temporali di parecchi anni. La scoperta del pianeta Nettuno, fatta con soli calcoli, da parte del matematico ed astronomo francese Urbain Jean-Joseph Le Verrier fu un vero trionfo della precisione della teoria newtoniana.[1] Tuttavia la teoria di Newton si rivelò non del tutto esatta: il moto di precessione del perielio dei pianeti più vicini al Sole, previsti dalla teoria newtoniana, si discosta leggermente dai valori osservati. Nel caso del pianeta Mercurio, tale scostamento è maggiore che per gli altri pianeti ed ammonta a circa $43''$ per secolo e non è assolutamente spiegato dalla meccanica newtoniana. Questa piccola deviazione fu scoperta attraverso calcoli molto laboriosi sia da Le Verrier nel 1845, sia dall'astronomo e matematico statunitense Simon Newcomb nel 1882, che ricalcolò tale deviazione.
La spiegazione della precessione del perielio di Mercurio fu uno dei primi successi della teoria relativistica della gravitazione di Einstein.
Ci dobbiamo perciò chiedere quali siano i motivi che hanno portato Einstein alla formulazione di una nuova teoria della gravitazione: universalmente nota col nome di **Relatività Generale**.
Conviene subito osservare che la Relatività Ristretta è nata per due necessità fondamentali:

1) rimuovere l'apparente incompatibilità tra il fenomeno della propagazione della luce ed il Principio di Relatività di Galileo;

2) eliminare il disaccordo tra i risultati di alcune esperienze ottiche (soprattutto l'esperimento di Michelson-Morley) e le previsioni teoriche desunte dai principi della fisica classica.

La Relatività Generale invece è sorta per tutt'altre esigenze che ora esamineremo brevemente.
Da un lato essa intendeva rimuovere dall'antica teoria newtoniana alcuni aspetti poco soddisfacenti, non dal punto di vista del controllo sperimentale che era pressoché perfetto, come abbiamo già detto in precedenza, ma piuttosto da un punto di vista speculativo, come vedremo fra poco.
Da un altro lato la Relatività Generale si proponeva di dare una formulazione più unitaria delle leggi fisiche: intento giustificato dalla convinzione, ogni giorno sempre più confermata dall'esperienza, di una interdipendenza tra tutti i fenomeni fisici. Un tale processo di unificazione si era già iniziato con la Relatività Ristretta, che aveva fuso - ad opera di Hermann Minkowski (1908) - in un unico ente, lo **spaziotempo**, le nozioni, prima separate, di **spazio** e di **tempo**, riducendo simultaneamente il numero degli enti assoluti.
Naturalmente la Relatività Generale accetta come acquisita tale fusione spazio-temporale, inserendola in un più vasto programma di unificazione di cui ci occuperemo nei paragrafi seguenti.
Dal punto di vista storico, dobbiamo osservare che il lavoro fondamentale sulla **Teoria della Relatività Generale** fu presentato da Einstein, alla Sezione fisico-matematica dell'Accademia Prussiana delle Scienze, il 25 Novembre del 1915 e fu ricevuto dalla redazione della prestigiosa

[1] Nettuno venne individuato nella notte del 23 Settembre 1846 all'Osservatorio Astronomico di Berlino dall'astronomo Joahnn Gottfried Galle su indicazione di Le Verrier. La scoperta fu fatta, quasi contemporaneamente, dall'astronomo inglese John Adams a Cambridge.

rivista tedesca "Annalen der Physik" il 20 Marzo 1916; venne pubblicato col titolo: *Die Grundlage der Allgemeinen Relativitätstheorie*, Annalen der Physik, **(4), 49**, 1916, pp. 769-822.[2] Questo lavoro è una pietra miliare nella storia della fisica e forse uno dei vertici più alti raggiunti finora dall'ingegno umano!
Per l'esattezza storica, dobbiamo osservare che appena cinque giorni prima, il 20 Novembre 1915, David Hilbert presentò all'Accademia delle Scienze di Gottinga una memoria che conteneva le stesse equazioni gravitazionali dedotte da Einstein, con l'aggiunta però di una restrizione superflua sul tensore energia-impulso. David Hilbert, da grandissimo matematico quale era, dedusse le equazioni gravitazionali generalizzando il principio variazionale di Hamilton, ben noto in meccanica classica, utilizzando una opportuna lagrangiana. Einstein invece pervenne a tali equazioni, dopo diversi tentativi ed errori, durati circa 10 anni di duro lavoro - avvalendosi anche della collaborazione del suo amico matematico Marcel Grossman - guidato sempre però dal suo *incomparabile ed infallibile intuito fisico*. Possiamo affermare comunque che non fu un caso che essi pervennero, quasi contemporaneamente, alle stesse equazioni gravitazionali in quanto fra i due grandi scienziati c'era stata una fitta corrispondenza epistolare su tali questioni. A tal proposito, il matematico Felix Klein sostiene che: "fra i due non ci possa essere questione di priorità, dal momento che essi seguivano linee di pensiero talmente differenti che la stessa compatibilità dei risultati non apparve subito sicura". E Abraham Pais soggiunge: *sono convinto che Einstein sia stato l'unico creatore* **della teoria fisica** *della Relatività Generale e che la scoperta della fondamentale equazione gravitazionale debba essere attribuita sia a lui che a Hilbert* [3].
Nello stesso anno 1916 in cui fu pubblicato il lavoro fondamentale sulla Relatività Generale, Einstein pubblicò anche un'esposizione divulgativa sia della Relatività Ristretta sia della Relatività Generale di notevole pregio didattico, la cui lettura è fortemente raccomandata allo studente principiante[4].

11.2 Programma della teoria einsteiniana

11.2.1 Le critiche di Einstein alla teoria newtoniana e programma della teoria einsteiniana

La critica profonda che Einstein fece - anche sulla base delle idee filosofiche (e fisiche) espresse dal fisico e filosofo austriaco Ernest Mach - della teoria newtoniana della gravitazione, si può sintetizzare in tre punti, che rappresentano insieme il programma base di una nuova teoria della gravitazione[5].

1. **Eliminazione delle azioni gravitazionali istantanee a distanza**.
   È proprio sulla legge di gravitazione universale che si fonda la prima obiezione alla teoria newtoniana.
   Accettate le idee della Relatività Ristretta, **la quale pone un limite superiore alla velocità di propagazione di ogni azione fisica**, anche la forza newtoniana, concepita come

[2]Di questo lavoro fondamentale abbiamo una traduzione italiana dovuta ad Aldo Pratelli dal titolo: *I fondamenti della teoria della Relatività Generale*, ed inserita nel volume *Cinquant'anni di Relatività* a cura di Mario Pantaleo, Editrice Universitaria, Firenze 1955, pp. 509-559; nuova edizione Giunti Barbera, Firenze 1980, pp. 479-504. Questa memoria è stata anche riprodotta nel volume: Albert Einstein, *Opere scelte* a cura di Enrico Bellone, Bollati Boringhieri, 1988, pp. 282-343.

[3]Abraham Pais, *Sottile è il Signore....,La scienza e la vita di Albert Einstein*, Boringhieri, Torino,1986, pp. 283-284. Questo libro è considerato una delle migliori biografie su Albert Einstein, se ne consiglia perciò vivamente la lettura

[4]Il titolo originale del libro è *Über die spezielle und allgemeine Relativitätstheorie (gemeinverständlich)*, la cui traduzione italiana a cura di Virginia Geymonat, dal titolo *Relatività, esposizione divulgativa,* fu pubblicata dall'Editore Boringhieri (Torino), 1967 e ripetutamente ristampata anche recentemente.

[5]In ciò che segue ci atteniamo all'impostazione data da Carlo Cattaneo in *Introduzione alla Teoria einsteiniana della Gravitazione*, Vol. I, Libreria Eredi Virgilio Veschi, Università degli Studi di Roma, Anno Accademico 1960-61.

un'azione immediata a distanza, deve essere esclusa per principio. A tale obiezione non sfugge neppure l'equazione di campo della gravitazione newtoniana, rappresentata dall'equazione di Poisson, della quale abbiamo discusso nel Cap. 1 paragrafo 1.3.11:

$$\Delta_2 \Phi = 4\pi G \mu(P), \tag{11.1}$$

dove Φ denota il potenziale gravitazionale, G costante newtoniana della gravitazione e $\mu(P)$ la densità di massa. Poiché in tale equazione la variabile temporale interviene al più come parametro ed è quindi implicita la propagazione istantanea del campo gravitazionale generato da ogni distribuzione di materia.
Si tratta dunque di una vera obiezione di principio, che imporrebbe almeno una modifica della legge della gravitazione universale. Ma, come vedremo la teoria einsteiniana darà una interpretazione ben più radicale della gravitazione.

2. **Abolizione di ogni riferimento privilegiato. Principio generale di relatività (o Principio di covarianza generale)**.
C'è nella meccanica newtoniana un altro elemento concettualmente insoddisfacente. È ben noto che l'equazione di moto di una particella (seconda legge della dinamica)

$$\frac{d}{dt}(m\mathbf{v}) = \mathbf{f}, \tag{11.2}$$

anche nella interpretazione che ad essa si dà in Relatività Ristretta, deve ritenersi valida quando si adotti un sistema di riferimento inerziale o galileiano, cioè un riferimento solidale con le stelle fisse ovvero in moto traslatorio rettilineo uniforme rispetto ad esse. Non appena si scelga un riferimento diverso, per esempio un riferimento in rotazione rispetto alle stelle fisse, o linearmente accelerato, la legge di moto cessa di essere valida, o, se si vuole conservare la sua validità formale, occorre ammettere che in un tale riferimento sono presenti, oltre alle **"forze effettive" $\mathbf{f}_{eff.}$**, i cosiddetti **campi di forza "apparenti" (di trascinamento $\mathbf{f}_\tau$ e di Coriolis $\mathbf{f}_{Cor}$)**.
Malgrado tale estensione formale della (11.2) ad arbitrari sistemi di riferimento, è ben noto che in meccanica classica un sistema inerziale è assolutamente indistinguibile da qualsiasi altro sistema inerziale (**Principio Galileiano di Relatività**), è invece possibile effettuare delle convenienti esperienze dinamiche per accertarsi se ci si trova in un sistema inerziale ovvero in uno non inerziale. Infatti se ci troviamo, per esempio, su di una piattaforma priva di attrito, ruotante con velocità angolare costante, e poniamo, in un dato istante, una massa puntiforme in un punto della piattaforma, non appartenente all'asse di rotazione, con velocità nulla, negli istanti successivi vedremo tale massa muoversi radialmente verso l'esterno. Analogamente se poniamo, in un punto arbitrario di un riferimento linearmente accelerato, una particella con velocità nulla in un dato istante, negli istanti successivi vedremo che tale particella si mette in moto spontaneamente con una accelerazione vettorialmente opposta a quella spettante al sistema di riferimento. Se la stessa esperienza si effettuasse in un riferimento inerziale, la data massa puntiforme continuerebbe a rimanere in quiete nella stessa posizione iniziale, rispettando così il Principio d'inerzia di Galileo.
Perciò di fronte alla diversità sperimentalmente riconoscibile fra riferimenti galileiani e non galileiani, due atteggiamenti mentali sono possibili; o la si accetta semplicemente come un dato di fatto, ovvero ci si pone la domanda: perché un riferimento non galileiano si comporta in modo diverso da uno galileiano ?
La Fisica Classica e la Relatività Ristretta assumono il primo atteggiamento. La Relatività Generale sceglie invece il secondo atteggiamento e si chiede: *Concomitanti con questa diversità di comportamento, dei sistemi inerziali e non inerziali, quali altre diverse circostanze fisiche sono presenti e sperimentalmente osservabili, negli uni e negli altri ?*

Una risposta, abbastanza verosimile è la seguente: l'unica ma grandiosa differenza è che negli uni l'enorme, sterminata quantità di materia costituente le "stelle fisse" si trova sensibilmente in quiete o in moto traslatorio uniforme. Negli altri le masse medesime appaiono dotate di accelerazioni, anche grandissime. Come allora non ritenere che sia proprio il diverso stato di moto di questo enorme sistema di masse ad influire diversamente sulle leggi fisiche nei due riferimenti ? Tale punto di vista, che sul piano speculativo risale al fisico e filosofo viennese Ernest Mach ed in parte anche al grande matematico francese Henri Poincaré, venne accolto da Einstein ed influì profondamente sul suo pensiero epistemologico. Con ciò si deve intendere che nessun riferimento è a priori privilegiato, ma che in esso le leggi fisiche sono condizionate dalla presenza e dal moto della materia dell'universo rispetto al riferimento medesimo. Si arriva così alla formulazione del Principio Generale di Relatività che enunceremo nel prossimo paragrafo.

3. **Unificazione concettuale dei campi gravitazionali "veri" e "apparenti" (Principio di Equivalenza).**
L'esperienza dimostra che le proprietà dei **campi di forza apparenti** sono in tutto e per tutto riconducibili a quelle dei **campi gravitazionali**. In particolare, **in entrambi i casi, il moto di una particella libera non dipende dal valore della sua massa.** *Se si vuole eliminare l'esistenza di riferimenti privilegiati, quindi, i campi apparenti in un riferimento non inerziale devono essere considerati come veri e propri campi gravitazionali.*
Accettato il fatto che le masse dell'universo ed il loro moto condizionano in qualche modo le leggi fisiche, nella teoria gravitazionale classica appare un fatto abbastanza sorprendente. L'influenza che le masse dell'universo hanno sui fenomeni fisici, per esempio sui fenomeni di moto, sembra manifestarsi in duplice modo. Da un lato ci sono le masse del sistema solare, che provocano direttamente un campo gravitazionale che si manifesta attraverso la legge di gravitazione universale oppure per mezzo dell'equazione di campo di Poisson; dall'altro vi sono altre masse assai più lontane, che sembrano influenzare la stessa legge di moto creando, nei riferimenti non inerziali, i **"campi apparenti"**.
Se si tiene conto del grande rapporto tra la distanza delle stelle più vicine al Sole e l'analoga distanza del pianeta più lontano, sembrerebbe quasi potersi affermare che i campi gravitazionali "veri" sono provocati dalle masse "vicine" mentre i "campi apparenti" sono provocati dalle masse "lontane". Ma si tratta veramente di due influenze di natura diversa?
Benché suggerito dall'osservazione diretta, questo modo di concepire le cose appare concettualmente insoddisfacente, perché certamente non può avere carattere universale una formulazione che debba distinguere masse "vicine" da masse "lontane". Appare quindi desiderabile, sul piano concettuale, una unificazione di queste due influenze in una influenza più generale; cioè, in altri termini, una unificazione dei campi gravitazionali "veri" con i "campi apparenti".
Del resto che tra campi gravitazionali "veri" e "campi apparenti" vi sia una stretta connessione può riconoscersi anche con la seguente constatazione.
Una proprietà importante dei "campi apparenti" è che il moto di una particella libera in tale campo risulta, a parità di condizioni iniziali, **indipendente dal valore della sua massa**. Come è noto la stessa proprietà vale per i campi gravitazionali "veri": il moto di un grave nel vuoto, così come il moto di un pianeta soggetto all'attrazione newtoniana da parte del Sole e degli altri pianeti, è **indipendente dalla sua massa**. Il fatto, come ben si sa, **è conseguenza dell'identità tra massa inerziale e massa gravitazionale.**[6]

[6]Forse è qui opportuno fare un'osservazione apparentemente marginale, ma che verosimilmente ha avuto una qualche influenza nella formazione della Relatività Generale. Esiste in Meccanica Classica un'altra categoria di movimenti che risultano *indipendenti dalla massa mobile: si tratta dei moti che avvengono sotto l'azione esclusiva di vincoli privi di attrito, siano essi fissi o dipendenti dal tempo.* Si può dunque dire che la notevole proprietà è comune ai tre tipi

Questa identità è stata dimostrata sperimentalmente in laboratorio, con una precisione meravigliosa, dalle celebri esperienze del fisico ungherese Loránd Eötvos (1889) e ripetuta ancora da Eötvos e dai suoi collaboratori Pekar e Fekete (1908) aumentandone la precisione. Infine l'esperienza fu ripetuta nel 1970 da due gruppi di fisici: il primo, dell'Università di Princeton, guidato dall'eminente fisico sperimentale Robert Dicke e il secondo, dell'Università statale di Mosca, guidato da Vladimir Braginski[7]. Essi ottennero miglioramenti consistenti in quanto a precisione.
L'identità fra massa inerziale e massa gravitazionale accomuna perciò "campi veri" con "campi apparenti". Naturalmente fra le due specie di campi ci sono anche delle differenze. Una differenza è questa: un "campo apparente" può essere eliminato in blocco mediante il cambiamento del sistema di riferimento; la medesima eliminazione non è invece possibile, in generale, per un campo gravitazionale "vero", per esempio il campo gravitazionale del Sole o della Terra. Si osservi però che se non è possibile l'eliminazione in blocco di tutto un campo gravitazionale "vero" è però possibile la sua eliminazione in piccolo. Basta pensare ad una cabina liberamente gravitante nello spazio intorno alla Terra e priva di moto rotatorio: entro tale cabina il campo gravitazionale "vero" della Terra risulta del tutto eliminato.
L'eliminabilità locale dei campi "veri" indica la loro stretta parentela con i "campi apparenti", parentela che non può essere annullata dalla loro non eliminabilità globale. Tale presunzione può essere avvalorata dalla seguente analogia elettromagnetica.
Se entro un riferimento galileiano si hanno cariche elettriche in moto di insieme traslatorio uniforme, si manifesta un campo elettromagnetico che può essere totalmente eliminato semplicemente passando al riferimento galileiano in cui le cariche appaiono in quiete. Ma tale eliminazione non è più possibile, in blocco, se le cariche anziché di moto traslatorio, si muovono in modo disordinato. Appare dunque naturale, a conclusione di queste considerazioni, porre tra i punti fondamentali del programma della Relatività Generale anche la unificazione dei campi gravitazionali "veri" e dei campi gravitazionali "apparenti". A tal principio suol darsi il nome di **Principio di equivalenza** che enunceremo esplicitamente nel paragrafo seguente.
Questi sono, nelle loro linee essenziali, i punti salienti del programma einsteiniano, gli scopi principali che la teoria si propone.

11.2.2 Gli assiomi della Relatività Generale

Le considerazioni del paragrafo precedente indussero Einstein a fondare la nuova teoria della gravitazione sui due principi: il **Principio generale di relatività** ed il **Principio di equivalenza**. A questi due principi cardine possiamo aggiungerne però altri tre che implicitamente hanno influenzato e guidato il pensiero di Einstein. Questi cinque assiomi sono i seguenti:

1. **Principio Generale di Relatività (o Principio di Covarianza Generale)**;
2. **Principio di Equivalenza (di Gravità ed Inerzia)**;
3. **Principio di Mach (o di Mach-Poincaré)**;
4. **Principio di Trascrizione**;

di campi: campi gravitazionali "veri", campi "apparenti," e campi "vincolari". Tale avvicinamento suggerisce che l'unificazione dei campi "veri" e dei campi "apparenti" possa farsi, forse, mediante opportuni campi "vincolari". Tale era in sostanza l'idea di Hertz, nel suo tentativo, che pur non ebbe grande fortuna, di eliminare dalla Meccanica il concetto di forza. Ma in altra forma l'idea di Hertz sembra veramente rinascere in Relatività Generale, dove **il campo gravitazionale, verrà sostituito, come vedremo, da opportuni "vincoli" dello spaziotempo.**

[7]Un resoconto interessante di questi esperimenti può trovarsi nel libro: Clifford M. Will, *Einstein aveva ragione?*, Cap.2, pag.24, Bollati Boringhieri, Torino (1989).

5. **Principio di Corrispondenza**.

Enunciamoli esplicitamente uno per uno, in maniera articolata su più punti, ciascuno dei quali coglie un particolare aspetto del contenuto e del significato di ciascun assioma illuminandone la portata.

1. **Principio Generale di Relatività di Einstein**:
   I) Le leggi della fisica debbono essere di natura tale che le si possano applicare a sistemi di riferimento comunque in moto;
   Principio di Covarianza Generale:
   II) Le leggi generali della natura debbono potersi esprimere mediante equazioni che valgano per tutti i sistemi di coordinate, siano cioè covarianti rispetto a qualunque sostituzione (covarianti in modo generale).[8]

2. **Il Principio di Equivalenza** scaturisce da una notevole proprietà già nota nella meccanica classica e risalente a Galileo e cioè:
   Proprietà fondamentale del campo gravitazionale:
   Il campo gravitazionale ha la proprietà di imprimere a tutti i corpi la stessa accelerazione indipendentemente dalle loro dimensioni e dalla loro costituzione fisica.
   Detto in altre parole, ciò dipende dal fatto che:

   $$\textbf{massa inerziale} = \textbf{massa gravitazionale.}$$

   Questa proprietà del campo gravitazionale implica alcune circostanze, che sintetizziamo in 4 enunciati, che nel loro insieme costituiscono il: **Principio di Equivalenza (di Gravità ed Inerzia)**:

   I) Le forze gravitazionali e le forze inerziali (o "apparenti") sono completamente equivalenti dal punto di vista fisico, cioè esse sono di identica natura e conseguentemente è impossibile distinguere localmente le une dalle altre con esperimenti fisici;

   II) Non esistono esperimenti locali che possano distinguere un riferimento non ruotante in caduta libera in un campo gravitazionale da un riferimento in moto uniforme in assenza di campo gravitazionale;

   III) Un riferimento uniformemente accelerato, rispetto ad un riferimento inerziale (della Relatività Ristretta), è localmente identico ad un riferimento in quiete in un campo gravitazionale;

   IV) Per ogni regione spaziotemporale infinitamente piccola (così piccola, cioè, che in essa la variazione spaziale e temporale della gravità possa venire trascurata) esiste sempre un sistema di coordinate K_0, nel quale è assente ogni effetto della gravità, sia sul movimento dei punti materiali, che su qualunque altro fenomeno fisico (es.: cabina in caduta libera, ossia liberamente gravitante).

   In breve: **è sempre possibile eliminare qualunque campo gravitazionale in regioni di universo infinitamente piccole.**[9]

Queste affermazioni riposano ora su basi sperimentali molto sicure per opera di: Eötvös (1889); Eötvös, Pekar, Fekete (1908); Dicke e Braginski (1970).

[8] I precedenti due Principi sono stati enunciati da Albert Einstein in *I fondamenti della teoria della Relatività Generale* (1916), rispettivamente a pag.511 e a pag.514 della traduzione italiana, inserita nel volume già citato *Cinquant'anni di Relatività*, Editrice Universitaria, Firenze, 1955.

[9] Il contenuto del punto *IV* è stato sostanzialmente enunciato da Einstein (1916, pag.515), in una forma leggermente diversa ma concettualmente equivalente: *Per regioni 4-dimensionali infinitamente piccole, se le coordinate sono scelte convenientemente, rimane valida la Teoria della Relatività Ristretta.*

Osservazione di Landau-Lifschitz
".....I campi ai quali sono equivalenti sistemi di riferimento non inerziali non sono del tutto identici ai campi gravitazionali "reali", esistenti anche nei sistemi inerziali".
Esiste fra di essi una differenzia sostanziale per ciò che concerne le loro proprietà all'infinito. Ad una distanza infinita dai corpi che lo generano, **un campo gravitazionale reale** tende sempre a zero. Al contrario, **i campi "apparenti"** ai quali sono equivalenti sistemi di riferimento non inerziali, crescono illimitatamente all'infinito, oppure, al massimo, tendono ad un valore finito. Così, le forze centrifughe, che nascono in un sistema di riferimento ruotante (con velocità angolare costante), tendono all'infinito quando ci si allontana dall'asse di rotazione; il campo equivalente ad un sistema di riferimento in moto rettilineo accelerato è identico in tutto lo spazio compreso l'infinito.
I campi equivalenti a sistemi di riferimento non inerziali scompaiono passando ad un riferimento inerziale. Contrariamente a ciò, i campi gravitazionali reali (esistenti anche in un sistema inerziale) non si possono eliminare globalmente mediante un'opportuna scelta del sistema di riferimento. La sola cosa che si possa ottenere con una scelta adeguata del sistema di riferimento, è di eliminare il campo gravitazionale in una regione dello spazio sufficientemente piccola perché in essa il campo possa essere considerato uniforme."

3. **Principio di Mach**: Il contenuto di questo principio è articolato nei seguenti tre punti:

   I) La distribuzione di materia-energia nell'Universo determina completamente la sua Geometria;

   II) La metrica dello spaziotempo è determinata dalla distribuzione della materia-energia nell'Universo;

   III) Soltanto nel caso di un Universo, assolutamente ipotetico, che fosse del tutto vuoto di materia-energia, e quindi privo di ogni campo gravitazionale vero, lo spaziotempo si ridurrebbe ad uno spazio piatto pseudoeuclideo di Minkowski.

4. **Principio di Trascrizione (o Regola di Trascrizione)**:

   I) Ogni relazione fisica a carattere locale algebrico valida in Relatività Ristretta (RR), si trasporta inalterata in Relatività Generale (RG);

   II) Ogni legge fisica differenziale che involga grandezze tensoriali che già esistono in RR, e che in questa sia formulata in coordinate cartesiane (galileiane), si trascrive in RG, in coordinate generali (gaussiane), con la semplice sostituzione della derivazione parziale ordinaria ∂_i con la derivazione covariante ∇_i e della differenziazione ordinaria d con la differenziazione assoluta D:

$$\partial_i \longrightarrow \nabla_i\,, \;\; d \longrightarrow D. \tag{11.3}$$

   Tale principio si può induttivamente giustificare paragonando **la constatazione d'ordine fisico** della eliminabilità locale di un qualsiasi campo gravitazionale mediante la scelta di un opportuno riferimento (cabina liberamente gravitante), **alla constatazione di carattere matematico** della esistenza di *coordinate geodetiche* (localmente galileiane), rispetto alle quali la derivazione covariante si identifica con la derivazione parziale ordinaria. Nella cabina gravitante, che si può pensare come un riferimento localmente galileiano, è plausibile accettare le leggi differenziali della RR senza alcuna modifica; **il passaggio a coordinate del tutto generali** si farà poi con la materiale sostituzione indicata dalle (11.3).

Veniamo ora all'ultimo principio a cui è opportuno far precedere le seguenti considerazioni:
La Fisica Matematica si occupa, in generale, della formulazione di modelli matematici per lo

studio di svariati fenomeni fisici, ed accanto ad ogni modello deve specificare il suo ambito di validità, inoltre quando si costruisce una nuova teoria, più ampia di altre teorie precedenti, essa deve soddisfare al requisito che in un limite appropriato deve ridursi alle teorie già note. Anche la teoria einsteiniana della gravitazione deve soddisfare ai precedenti requisiti, possiamo perciò enunciare il quinto ed ultimo assioma così:

5. **Principio di Corrispondenza**: *La Relatività Generale*, considerata come modello matematico relativistico della gravitazione deve, in assenza di gravitazione ($G = 0$) avere come limite **la Relatività Ristretta** e nel caso di campi gravitazionali deboli e basse velocità (paragonate a quella della luce, $v/c \to 0$) avere come limite **la teoria gravitazionale newtoniana**. Inoltre quest'ultima per $G = 0$ deve ridursi alla **meccanica newtoniana in assenza di gravitazione** e questa deve potersi ottenere come limite per $c \to \infty$ dalla **Relatività Ristretta.**

Concludendo, è opportuno, per il seguito, tenere presente esplicitamente la seguente importante osservazione sulla propagazione della luce conseguenza degli assiomi precedenti:
Osservazione di Einstein[10]
"Il **principio della costanza della velocità della luce nel vuoto** *dev'essere modificato, in quanto si constata facilmente che se la luce si propaga in linea retta con velocità costante rispetto ad un riferimento galileiano $\mathcal{K}$, la sua traiettoria rispetto ad un sistema di riferimento $\mathcal{K}'$, in moto traslatorio uniformemente accelerato rispetto a $\mathcal{K}$, deve essere in generale curvilinea".*

11.3 Necessità della geometria riemanniana

Come sappiamo dalla Relatività Ristretta, in un sistema di riferimento inerziale K, in coordinate cartesiane (o galileiane) l'intervallo spaziotemporale fra due eventi infinitamente vicini è dato dalla:

$$\mathbf{ds}^2 \equiv \eta_{ik} dx^i dx^k = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \tag{11.4}$$

Nel passaggio a qualsiasi altro sistema inerziale, applicando le trasformazioni di Lorentz, l'intervallo, com'è noto, conserva la stessa espressione algebrica (è un invariante) pur essendo espresso nelle nuove coordinate. Se però si passa ad un sistema di riferimento non inerziale K', il $\mathbf{ds}^2$ non è più esprimibile come somma algebrica dei quadrati dei differenziali delle quattro coordinate $(x'^i) = (x'^0, x'^1, x'^2, x'^3)$, come in (11.4). Esempi di queste trasformazioni sono il passaggio ad un sistema di riferimento in moto traslatorio uniformemente accelerato rispetto al riferimento inerziale K, oppure ad un sistema in moto rotatorio uniforme rispetto a K.

1) Nel caso del passaggio ad un sistema di riferimento in moto traslatorio uniformemente accelerato rispetto a un riferimento galileiano, supponiamo, più precisamente, che il sistema non inerziale sia costituito da una terna cartesiana $K' \equiv O'x'y'z'$ equiorientata con la terna inerziale $K \equiv Oxyz$ ed in moto rispetto ad essa con accelerazione costante, di modulo a, parallelamente all'asse z talchè la trasformazione delle coordinate è del tipo:

$$\begin{cases} x = x' + x_0, \\ y = y' + y_0, \\ z = z' + z_0 + \dfrac{1}{2}at^2, \end{cases} \tag{11.5}$$

[10]Questa osservazione di Einstein si trova alla fine del paragrafo 2, pag. 512, della traduzione italiana del lavoro fondamentale del 1916: *I fondamenti della teoria della Relatività Generale* già citato, all'inizio di questo capitolo, nel paragrafo 11.1.

essendo x_0, y_0, z_0, le coordinate dell'origine $O'(x' = y' = z' = 0)$ del riferimento K' nell'istante iniziale $t = 0$. In virtù della trasformazione (11.5), il $\mathbf{ds}^2$ espresso dalla (11.4), si trasforma nella seguente espressione:

$$\mathbf{ds}'^2 = (c^2 - a^2t^2)dt^2 - dx'^2 - dy'^2 - dz'^2 - 2at\, dz'\, dt. \tag{11.6}$$

Qualunque possa essere la relazione fra i tempi in K e K' questa forma quadratica non potrà mai ridursi alla somma algebrica dei quadrati dei differenziali delle coordinate.

2) Nel secondo caso, supponiamo che la terna $O'x'y'z'$ sia inizialmente sovrapposta alla $Oxyz$ e che ruoti con velocità angolare costante ω attorno all'asse $Oz \equiv O'z'$, la trasformazione delle coordinate è allora data da:

$$\begin{cases} x = x' \cos \omega t - y' \sin \omega t, \\ y = x' \sin \omega t + y' \cos \omega t, \\ z = z'. \end{cases} \tag{11.7}$$

L'intervallo spaziotemporale assume la forma:

$$\mathbf{ds}'^2 = [c^2 - \omega^2(x'^2 + y'^2)]dt^2 - (dx'^2 + dy'^2 + dz'^2) + 2\omega dt(y'dx' - x'dy'), \tag{11.8}$$

ed anche in questo caso esso non può mai ridursi alla forma minkowskiana. Di conseguenza in un sistema di riferimento non inerziale l'intervallo spaziotemporale sarà una forma quadratica che include termini che sono prodotti dei differenziali delle coordinate. D'altra parte, è ovvio che se si opera un'arbitraria trasformazione differenziabile ed invertibile di coordinate:

$$x^i = x^i(x'^k), \tag{11.9}$$

si avrà

$$dx^i = \frac{\partial x^i}{\partial x'^j} dx'^j, \qquad dx^k = \frac{\partial x^k}{\partial x'^h} dx'^h. \tag{11.10}$$

Se sostituiamo queste espressioni nella forma quadratica minkowskiana (11.4) avremo

$$\mathbf{ds}'^2 = \eta_{ik} \frac{\partial x^i}{\partial x'^j} \frac{\partial x^k}{\partial x'^h} dx'^j dx'^h. \tag{11.11}$$

Se poi poniamo

$$g'_{jh} = \eta_{ik} \frac{\partial x^i}{\partial x'^j} \frac{\partial x^k}{\partial x'^h}, \quad (g'_{jh} = g'_{hj}) \tag{11.12}$$

le dieci funzioni g'_{jh} saranno, in generale, funzioni di tutte le coordinate x'^i. Possiamo allora concludere dicendo che: *in un sistema di riferimento non inerziale l'intervallo spaziotemporale sarà una forma quadratica generale dei differenziali delle coordinate spaziali x^1, x^2, x^3 e della coordinata temporale x^0 cioè esso definisce una* **metrica riemanniana**:

$$\mathbf{ds}^2 = g_{ik} dx^i dx^k, \quad (g_{ik} = g_{ki}). \tag{11.13}$$

Applicando una trasformazione generale del tipo (11.9) alla (11.13) si ritrova un'espressione dello stesso tipo, si tratta dunque di *una forma invariante* per tali trasformazioni generali. Siccome, per il Principio di Equivalenza, *un sistema non inerziale è equivalente ad un campo gravitazionale* è del tutto naturale, tenuto conto delle considerazioni precedenti, convincerci che la geometria appropriata per lo studio dei campi gravitazionali sia **la geometria riemanniana.** Questa idea sarà sviluppata nei prossimi paragrafi.

11.4 Lo spaziotempo $\mathcal{V}^4$ della Relatività generale

Nei primi due principi, **di Relatività Generale e di Equivalenza,** non c'è ancora traccia della **luminosa idea geometrica** che permise ad Einstein di realizzare, d'un sol colpo, le richieste del suo programma, **illuminando in maniera semplice e geniale il fenomeno della gravitazione.** La visione geometrica di Einstein scaturisce invece sostanzialmente dal **Principio di Mach** e da una evoluzione che egli perseguì seguendo l'idea originale di **spaziotempo** dovuta al matematico lituano Hermann Minkowski (1908), già professore di Einstein all'epoca in cui era studente al Politecnico di Zurigo.
Infatti un risultato essenziale conseguito da Minkowski e incorporato in seguito nella RR, *consiste nell'aver trasportato* **l'assoluto** *dallo* **spazio** *e dal* **tempo**, *com'erano concepiti nella meccanica classica e cioè in maniera del tutto indipendente l'uno dall'altro, allo* **spaziotempo**, *cioè all'insieme di tutti gli eventi costituenti l'Universo nel suo divenire, munito di una metrica opportuna.*
In Relatività Ristretta **lo spaziotempo** è uno **spazio pseudoeuclideo (detto spazio di Minkowski $\mathcal{M}^4$)** nel quale, adottando coordinate galileiane, il divario spaziotemporale tra due eventi infinitamente vicini, $\mathcal{A} \equiv (x^0, x^1, x^2, x^3)$ e $\mathcal{B} \equiv (x^0 + dx^0, x^1 + dx^1, x^2 + dx^2, x^3 + dx^3)$, è espresso (cfr. i paragrafi 5.1 e 5.2 del Cap.5) da una forma quadratica indefinita del tipo:

$$\mathbf{ds}^2 \equiv \eta_{ik} dx^i dx^k = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2, \qquad (11.14)$$

dove $x^1 = x, x^2 = y, x^3 = z$ sono le coordinate spaziali ed $x^0 = ct$ la coordinata temporale dell'evento $\mathcal{A}$.
Lo spazio di Minkowski $\mathcal{M}^4$ non porta soltanto il vantaggio di un comodo e sintetico linguaggio geometrico, ma fornisce un potente strumento per costruire teorie fisiche soddisfacenti ai due postulati della Relatività Ristretta.
Infatti, come si è già visto in RR, **una teoria fisica che sia formulata in termini geometrici intrinseci, vettoriali o tensoriali di $\mathcal{M}^4$ gode della proprietà di essere espressa in una medesima forma in tutti i riferimenti galileiani, soddisfacendo in tal modo automaticamente al Principio Einsteiniano di Relatività.**
Anche in Relatività Generale sono gli **eventi** ad avere un significato assoluto, e la loro totalità costituisce ancora naturalmente una varietà a quattro dimensioni. Se teniamo conto delle considerazioni fatte nel paragrafo precedente possiamo porre la seguente
Definizione:
Dicesi spaziotempo della Relatività Generale - *e lo indicheremo con $\mathcal{V}^4 \equiv (\mathcal{E}^4, g_{ik})$ - l'insieme di tutti i possibili eventi $\mathcal{E}^4$ che accadono nel nostro universo e caratterizzati ciascuno - in un arbitrario sistema di coordinate generali (gaussiane) - dalla quaterna $(x^i) \equiv (x^0, x^1, x^2, x^3)$, con $x^0 = ct$* **coordinata temporale** *e $x^\alpha, \alpha = 1, 2, 3,$* **coordinate spaziali**, *munito di una generica* **"metrica riemanniana"**

$$\mathbf{ds}^2 = g_{ik}(x) dx^i dx^k \lesseqgtr 0, \quad (i, k = 0, 1, 2, 3), \qquad (11.15)$$

definita dal tensore simmetrico $g_{ik}(x) \equiv g_{ik}(x^0, x^1, x^2, x^3)$.[11]
Spesso la metrica verrà scritta in forma più esplicita separando i termini nel seguente modo:

$$\mathbf{ds}^2 = g_{00}(dx^0)^2 + 2g_{0\alpha} dx^0 dx^\alpha + g_{\alpha\beta} dx^\alpha dx^\beta \qquad (\alpha, \beta = 1, 2, 3), \qquad (11.16)$$

con le $g_{00}, g_{0\alpha}, g_{\alpha\beta}$ beninteso dipendenti generalmente da tutte e quattro le coordinate spaziotemporali (x^0, x^1, x^2, x^3).
Ma che cosa ha in comune la varietà $\mathcal{V}^4$ della Relatività Generale con uno spazio $\mathcal{M}^4$ di

[11] Avvertiamo il lettore che, siccome la forma quadratica che esprime la metrica dello spaziotempo non è definita in segno (indefinita), alcuni autori ritengono che si debba parlare più propriamente di *metrica pseudo-riemmaniana*, riservando la parola *riemanniana* al caso in cui la metrica sia definita positiva. Noi, per semplicità, continueremo ad usare il termine *metrica riemanniana* in senso generale.

Minkowski, e in che cosa ne differisce?
Secondo il punto di vista della Relatività Generale, e precisamente per il punto III del Principio di Mach, la varietà $\mathcal{V}^4$ dovrebbe anzitutto ridursi ad uno spazio pseudoeuclideo di Minkowski soltanto nel caso di un Universo, assolutamente ipotetico, che fosse del tutto vuoto di materia, quindi *privo di ogni campo gravitazionale vero.*
Se allora si tiene presente che, almeno localmente, ogni campo gravitazionale può essere eliminato adottando un opportuno riferimento fisico (cabina liberamente gravitante)(*punto IV del Principio di Equivalenza*), appare spontaneo ammettere che $\mathcal{V}^4$ abbia, localmente, la struttura di uno spaziotempo di Minkowski $\mathcal{M}^4$. Ciò equivale ad ammettere che in ogni punto P di $\mathcal{V}^4$ *sia lo spazio tangente $T_P(\mathcal{V}^4)$ ad avere struttura minkowskiana.* In altre parole, ciò significa pensare $\mathcal{V}^4$ come una **varietà riemanniana** dotata in ogni punto di una metrica che in coordinate generali è espressa dalla (11.15), il cui tensore metrico $g_{ik}(x)$ è costituito da funzioni delle coordinate generali x^i del generico punto P della varietà, ma che avrà, in ogni punto di $\mathcal{V}^4$, la stessa struttura algebrica della metrica di Minkowski.
In altri termini la presenza di materia-energia nell'Universo, in virtù del Principio di Mach, ha per effetto di rendere lo spaziotempo una varietà riemanniana generalmente curva (tensore di Riemann $R^i{}_{jkl} \neq 0$).
D'altra parte, *un campo gravitazionale reale* non può essere *totalmente eliminato* con nessuna trasformazione di coordinate. Ciò significa che in presenza di un campo gravitazionale, lo spaziotempo $\mathcal{V}^4 \equiv (\mathcal{E}^4, g_{ik})$ è tale che le g_{ik}, che ne determinano la metrica, non possono essere ridotte, con trasformazioni di coordinate, alla forma galileiana (o minkowskiana) contemporaneamente in tutto lo spazio. Un tale spaziotempo è detto **curvo**, contrariamente allo spaziotempo **piatto** dove questa riduzione è possibile.
È tuttavia possibile, con una appropriata trasformazione delle coordinate, ridurre le g_{ik} ad una forma galileiana in un punto arbitrario dello spaziotempo $\mathcal{V}^4 \equiv (\mathcal{E}^4, g_{ik})$: basta ridurre a forma diagonale una forma quadratica a coefficienti costanti (i valori delle g_{ik} nel punto dato). Chiameremo un tale sistema di coordinate **sistema galileiano** per il punto dato.
Ciò è consistente anche dal punto di vista matematico, in quanto in una varietà riemanniana è sempre possibile, come sappiamo, trovare, per ogni punto della varietà, un sistema di coordinate, le cosiddette ***coordinate geodetiche o localmente cartesiane*** che gode della proprietà che, in tali punti, tutti i simboli di Christoffel di 2ª specie sono nulli come nel caso di coordinate cartesiane, e perciò in coordinate geodetiche la derivazione covariante si riduce alla derivazione parziale ordinaria con enorme vantaggio per il calcolo. Notiamo che, ridotta alla forma diagonale nel punto dato, la matrice delle g_{ik} ha un autovalore positivo (+1) e tre autovalori negativi (-1,-1,-1) (l'insieme dei segni di questi autovalori costituisce la *segnatura* della matrice). Ne risulta, in particolare, che *il determinante $g = det(g_{ik})$ della metrica, in uno spaziotempo reale è sempre negativo*[12]:

$$g < 0. \tag{11.17}$$

Il fatto che $\mathcal{V}^4$ ha, in ogni suo punto, una metrica del tipo di Minkowski permette, come in RR, di classificare i vettori v^i e le direzioni uscenti da un generico punto di $\mathcal{V}^4$ in tre categorie: vettori (e direzioni) **del genere tempo**, a norma positiva $g_{ik}v^iv^k > 0$; vettori (e direzioni) **del genere spazio**, a norma negativa $g_{ik}v^iv^k < 0$; vettori (e direzioni) **del genere luce**, a norma nulla $g_{ik}v^iv^k = 0$. Anche ora come in RR, il cono dei vettori del genere luce, con vertice in un evento $\mathcal{E}$, separa gli eventi che per $\mathcal{E}$ sono definitivamente passati da quelli futuri e da quelli presenti.
Anche ora *la storia* di una particella materiale non potrà che essere rappresentata da una linea oraria a $\mathbf{ds}^2$ positivo (che potremo perciò indicare più semplicemente con ds^2) e questo ds^2 servirà ancora a definire *l'elemento di tempo proprio $d\tau$* della particella

$$d\tau^2 = \frac{1}{c^2}ds^2 \quad \Longleftrightarrow \quad d\tau = \frac{1}{c}ds. \tag{11.18}$$

[12]Ciò sarà dimostrato in maniera più formale nel paragrafo 11.8.

La struttura localmente minkowskiana di $\mathcal{V}^4$ ci assicura ancora che tutte le grandezze che in RR hanno una rappresentazione locale, mediante scalari o vettori o tensori, godono della stessa proprietà anche in RG.
A questo punto è spontaneo chiedersi: *Qual'è la reciproca influenza tra la struttura della varietà $\mathcal{V}^4$ ed i fenomeni fisici che vi si svolgono?*
Come vedremo, la risposta che Einstein dà a questa domanda **tramuta la gravitazione in un fatto puramente geometrico dello spaziotempo.**
Dividiamo la precedente domanda in due parti, alle quali risponderemo separatamente:

1) *Come influisce la curvatura di $\mathcal{V}^4$ sui fenomeni che vi si svolgono?*

2) *Come, viceversa, la distribuzione ed il moto della materia, più generalmente ogni manifestazione energetica, influiscono sulla curvatura dello spaziotempo?*

Risponderemo a queste domande dettagliatamente, almeno dal punto di vista concettuale, nei due paragrafi successivi, 11.5, 11.6.

11.5 Influenza della struttura di $\mathcal{V}^4$ sui fenomeni fisici

Proponiamoci ora di rispondere alla prima delle due domande poste alla fine del paragrafo precedente, cioè: *Come influisce la curvatura dello spaziotempo $\mathcal{V}^4$ sui fenomeni che vi si svolgono?*
Supponiamo, in qualche modo, conosciuta la struttura riemanniana di $\mathcal{V}^4$, cioè supponiamo che in ogni suo punto sia assegnato il tensore metrico $g_{ik}(x)$. Consideriamo ora una legge fisica che in RR abbia una certa formulazione tensoriale, di carattere algebrico o differenziale, in ogni punto dello spazio di Minkowski $\mathcal{M}^4$. *Come dovrà modificarsi la formulazione di tale legge per il fatto che il fenomeno si svolge entro $\mathcal{V}^4$ anziché in $\mathcal{M}^4$, cioè in uno spazio curvo anziché piatto?*
Per rispondere a tale domanda è opportuno soffermarsi su un punto chiave, di cui è necessario tenere conto se si vogliono estendere i concetti della Relatività Ristretta a quella Generale.
Sappiamo, per programma, che nella teoria della Relatività sia Ristretta che Generale, qualsiasi operazione su enti assoluti (tensori) deve produrre enti assoluti; sappiamo anche che in una generica varietà riemanniana le operazioni ordinarie di derivazione e di differenziazione non sono assolute, nel senso che, a partire da enti assoluti, non producono enti assoluti (tensori). In una qualunque varietà riemanniana le corrispondenti operazioni tensoriali sono *la derivazione covariante e la differenziazione assoluta*, che godono inoltre della notevole proprietà che applicate usando, anziché coordinate generali, *coordinate geodetiche*, si riducono alle corrispondenti operazioni ordinarie. Quindi, ogni legge differenziale che involge grandezze tensoriali esistenti in Relatività Ristretta, e che in questa sia formulata in coordinate cartesiane, va trascritta, in Relatività Generale, in virtù del *Principio di Trascrizione (o Regola di Trascrizione)* - e utilizzando coordinate generali - con la semplice sostituzione della derivazione parziale ordinaria ∂_i con la derivazione covariante ∇_i e della differenziazione ordinaria d con la differenziazione assoluta D:

$$\partial_i \to \nabla_i, \qquad d \to D. \tag{11.19}$$

Tale *Regola di Trascrizione* si può induttivamente giustificare, in base al *Principio di equivalenza*, paragonando la constatazione d'ordine fisico della eliminabilità locale di un qualsiasi campo gravitazionale mediante la scelta di un opportuno riferimento (cabina liberamente gravitante), alla constatazione di carattere matematico della esistenza di *coordinate geodetiche* (localmente galileiane), rispetto alle quali la derivazione covariante si identifica con la derivazione parziale ordinaria. Nella cabina gravitante, che si può pensare come un riferimento localmente galileiano, è plausibile accettare le leggi differenziali della RR senza alcuna modifica; il passaggio a coordinate del tutto generali si farà poi con la materiale sostituzione indicata dalle (11.19).

Ad illustrazione di ciò che abbiamo detto, ci limiteremo ad applicare il *Principio di Trascrizione* a tre casi molto importanti.

1) Come primo esempio vediamo come si trascrive in RG la legge di moto di una particella materiale libera.
Detto $d\tau$ l'intervallo elementare di tempo proprio della particella, che è un invariante per definizione come in RR, noi sappiamo che se ξ^i sono coordinate galileiane, si chiama 4-velocità della particella il 4-vettore **U** le cui componenti controvarianti sono

$$U^i = \frac{d\xi^i}{d\tau}, \quad (i = 0, 1, 2, 3) \tag{11.20}$$

che risulta tangente alla sua linea oraria, ed ha norma costante uguale a $||\mathbf{U}|| \equiv \mathbf{U} \cdot \mathbf{U} = \eta_{ik} U^i U^k = c^2 > 0$.
Ciò premesso, ricordiamo che in RR la legge di moto di una particella, supposta completamente libera da ogni forza, può scriversi così:

$$\frac{dU^i}{d\tau} = 0 \quad \Longleftrightarrow \quad \frac{d^2\xi^i}{d\tau^2} = 0 \quad (i = 0, 1, 2, 3). \tag{11.21}$$

Tale equazione, in una forma o nell'altra, traduce ovviamente **la legge d'inerzia**, ed esprime in sostanza che la linea oraria della particella in $\mathcal{M}^4$ è una retta, cioè **una geodetica** di $\mathcal{M}^4$.
Secondo Einstein, e precisamente secondo il *Principio di Equivalenza*, il fatto che ogni campo gravitazionale è localmente eliminabile (cabina liberamente gravitante) *permette di considerare, in tale cabina assunta quale sistema di coordinate, come assolutamente libero anche un punto materiale soggetto all'azione di un campo gravitazionale*.
Ne segue che, per il *Principio di Trascrizione*, la legge di moto di un punto materiale liberamente gravitante si potrà ottenere dalla legge di moto (11.21) espressa in coordinate galileiane ξ^i, esprimendola ora in coordinate curvilinee generali x^i e sostituendo semplicemente la derivazione ordinaria con la derivazione assoluta:

$$\frac{DU^i}{d\tau} \equiv \nabla_k U^i \frac{dx^k}{d\tau} = 0 \Longleftrightarrow \frac{d^2x^i}{d\tau^2} + \Gamma^i_{hk} \frac{dx^h}{d\tau}\frac{dx^k}{d\tau} = 0 \quad (i, h, k = 0, 1, 2, 3), \tag{11.22}$$

dove la 4-velocità U^i, in analogia con la (11.20) della RR, s'intende definita, in coordinate generali, con la $U^i = \dfrac{dx^i}{d\tau}$.
Tenuto conto che $ds = cd\tau$ e ricordando le equazioni che abbiamo stabilito nei paragrafi (10.8) e (10.9) sulle geodetiche di una varietà riemanniana, riconosciamo che la (11.22), rappresentante la linea oraria di una particella liberamente gravitante, è ancora una **geodetica dello spaziotempo**, ma non più una retta poiché la $\mathcal{V}^4$ è una varietà curva. Ne segue che anche la traiettoria della particella nello spazio fisico tridimensionale non sarà generalmente una retta, bensì una linea curva.
Osserviamo che questo risultato è scaturito dalla semplice applicazione formale del *Principio di Trascrizione*: quest'ultimo, se applicato correttamente, costituisce dunque un potente mezzo euristico di scoperta di leggi fisiche relative ai campi gravitazionali.
Ecco così come la sola curvatura dello spaziotempo può produrre effetti analoghi a quelli prodotti da un campo gravitazionale senza che forze gravitazionali esplicite, in senso newtoniano, debbano intervenire a produrre tali effetti.
È notevole, a questo proposito, il fatto che anche una piccola curvatura dello spaziotempo può dar luogo a geodetiche anch'esse poco incurvate, che però nello spazio fisico 3-

dimensionale possono dar luogo a traiettorie spaziali molto incurvate[13].
Come si vede, nella teoria della gravitazione einsteiniana **il campo gravitazionale come entità fisica è scomparso**; *infatti la particella gravitante viene trattata come in RR viene trattato un punto materiale libero da ogni forza.* **Ciò che classicamente era il campo di forze gravitazionali è ora implicitamente contenuto nella struttura geometrica dello spaziotempo $\mathcal{V}^4$, e cioè in sostanza nei coefficienti di connessione** $\Gamma^i_{hk}(x)$ *ovvero, se la connessione deriva dalla metrica $g_{ik}(x)$ di $\mathcal{V}^4$, da questo tensore.*
Dunque **le forze gravitazionali come entità fisiche vengono sostituite da entità geometriche, le** $\Gamma^i_{hk}(x)$, *e siccome queste ultime, tramite le formule di Christoffel, si esprimono per mezzo delle derivate parziali delle $g_{ik}(x)$, è naturale (in analogia con quello che si fa in meccanica classica) chiamare le $g_{ik}(x)$* **potenziali gravitazionali**; tale denominazione sarà anche giustificata in seguito in modo più circostanziato.

2) Come secondo esempio, consideriamo non più una semplice particella ma un sistema continuo costituito da **materia disgregata.**[14] Come si fa in RR, anche in Relatività Generale, se indichiamo con μ_0 *la densità propria di massa propria* del sistema e con $U^i = \dfrac{dx^i}{d\tau}$ (oppure alternativamente con $u^i = \dfrac{dx^i}{ds}$) la 4-velocità della generica particella, la distribuzione ed il moto della materia sono rappresentabili, col cosiddetto **tensore energetico o tensore energia-impulso**[15], manifestamente simmetrico:

$$T^{ik} = \mu_0 U^i U^k = c^2 \mu_0 u^i u^k. \tag{11.23}$$

Ciò premesso, **la legge di evoluzione del sistema**, che in RR e in coordinate galileiane è espressa, come vedremo nel paragrafo 11.17, dall'equazione tensoriale

$$\partial_k T^{ik} = 0, \tag{11.24}$$

secondo *il Principio di Trascrizione* si traduce, in coordinate generali, con la semplice sostituzione della derivata parziale ordinaria con la derivazione covariante, cioè:

$$\nabla_k T^{ik} = 0. \tag{11.25}$$

Osserviamo che, più in generale, la (11.24), e perciò anche la (11.25), esprime **la legge di evoluzione di un generico sistema chiuso**, caratterizzato da un tensore energetico T^{ik}, anche quando questo non abbia la struttura (11.23).

3) Come terzo esempio, applichiamo la regola di trascrizione alle equazioni di Maxwell. Osserviamo, a tal fine, che le grandezze tensoriali F^{ik} (*campo elettromagnetico*) e J^i (*4-densità*

[13]Per renderci conto di come ciò possa accadere utilizziamo la seguente analogia: si consideri, nello spazio ordinario, un'elica circolare di passo molto grande rispetto al raggio del cilindro sul quale è tracciata, la cui curvatura, com' è noto, è $\kappa_{elica} = \dfrac{\sin^2\theta}{R}$, dove θ è l'angolo d'inclinazione dell'elica rispetto all'asse del cilindro il cui raggio è R. La proiezione dell'elica su di un piano ortogonale all'asse del cilindro è una circonferenza che, nelle ipotesi poste, ha curvatura $\kappa_{circonf.} = \dfrac{1}{R}$ tanto più grande di quella dell'elica quanto è minore θ, cioè quanto è maggiore il suo passo.

[14]Dicesi **materia disgregata o materia incoerente non interagente o polvere** un sistema di particelle materiali, per il quale siano trascurabili, rispetto alle forze di volume, le forze di contatto, cioè in sostanza gli sforzi.

[15]Si noti che con la definizione (11.23) il tensore simmetrico T^{ik} ha le dimensioni di una **densità di energia** essendo le u^i adimensionali. Si può anche definire il tensore energia-impulso così $T^{ik} = \mu_0 u^i u^k$ (come fanno, per esempio, Adler-Bazin-Schiffer) ed in questo caso T^{ik} ha le dimensioni di una **densità di massa**. Il tensore energia-impulso verrà studiato nei dettagli nel paragrafo 11.17.

di corrente) conservano significato in ciascun punto di $\mathcal{V}^4$, inoltre le equazioni omogenee (1^a coppia delle equazioni di Maxwell) assumono la forma[16]:

$$\nabla_i F_{kl} + \nabla_k F_{li} + \nabla_l F_{ik} = 0, \quad (i, k, l = 0, 1, 2, 3). \tag{11.26}$$

Mentre per le equazioni non omogenee (2ª coppia delle equazioni di Maxwell)[17] si ha:

$$\nabla_k F^{ik} = J^i, \quad (i, k = 0, 1, 2, 3). \tag{11.27}$$

Osserviamo che quest'ultima equazione può anche assumere la stessa forma che ha in RR, cioè espressa per mezzo delle derivate parziali ordinarie; ma ciò è dovuto ad un fatto puramente casuale, derivante dalla sua struttura particolare e dalle proprietà di antisimmetria del tensore F^{ik} che, nell'esplicitare le derivate covarianti, permettono la eliminazione di tutti i termini contenenti simboli di Christoffel.

Con ciò che precede resta messo in evidenza, in modo qualitativo, come si manifesti l'influenza della curvatura dello spaziotempo $\mathcal{V}^4$ sui vari fenomeni fisici, senza l'intervento esplicito di un campo gravitazionale in senso newtoniano.

11.6 Influenza della materia e del suo moto sulla struttura di $\mathcal{V}^4$.

Veniamo ora a rispondere alla seconda domanda che ponemmo alla fine del paragrafo 11.3, ovvero: *Come la distribuzione ed il moto della materia, più generalmente ogni manifestazione energetica, influiscono sulla curvatura dello spaziotempo?*
Secondo la concezione einsteiniana, ispirata al *Principio di Mach*, la materia - ed in generale ogni manifestazione energetica - influisce sui fenomeni fisici in modo, per così dire indiretto, agendo essa direttamente sulla struttura geometrica dello spaziotempo e provocandone un minore o maggiore incurvamento, a seconda della sua minore o maggiore concentrazione.
Questo in termini vaghi. *Ma qual'è il meccanismo esatto secondo il quale avviene tale influenza sulla struttura di $\mathcal{V}^4$?*
Si è visto che il moto di una particella liberamente gravitante, in virtù del Principio di Trascrizione, è determinato completamente dalla struttura geometrica di $\mathcal{V}^4$, e quindi dai coefficienti di connessione $\Gamma^i_{hk}(x)$ ovvero dal tensore metrico cioè dai **potenziali gravitazionali** $g_{ik}(x)$. Appare allora ben naturale che, come classicamente, la distribuzione della materia influisce sul potenziale gravitazionale Φ attraverso **l'equazione di Poisson**, così in RG la distribuzione ed il moto della materia debbono influire sui potenziali gravitazionali $g_{ik}(x)$ e quindi sulla struttura di $\mathcal{V}^4$.
Per stabilire matematicamente tale influenza, ossia per scrivere *le equazioni differenziali del campo gravitazionale*, possiamo presumere che Einstein abbia tenuto presente, in maniera più o meno esplicita, i cinque assiomi enunciati nel paragrafo 11.2.2, ma che più specificamente e pragmaticamente, si sia ispirato **ai seguenti criteri** che da tali assiomi scaturiscono:

1) L'influenza della materia e delle altre manifestazioni energetiche sul tensore fondamentale $g_{ik}(x)$, deve esprimersi mediante equazioni differenziali alle derivate parziali del 2° ordine nelle $g_{ik}(x)$ (**le cosiddette equazioni di campo o equazioni gravitazionali**), lineari nelle derivate seconde. Tutto ciò in analogia con l'equazione di Poisson

$$\Delta_2 \Phi = 4\pi G \mu(P),$$

per il potenziale gravitazionale newtoniano, inoltre quest'ultima si dovrebbe poter ottenere, in un limite appropriato, dalle stesse equazioni di campo;

[16] Cfr. Cap.7, paragrafo 7.7, formula (7.67).

[17] Cfr. Cap.7, paragrafo 7.7, formula (7.84), a meno del fattore costante $\frac{4\pi}{c}$.

2) Tali equazioni siano formalmente invarianti rispetto al più generale cambiamento di coordinate, siano perciò espresse in forma tensoriale (in virtù del *Principio Generale di Relatività*);

3) Poiché il *tensore energia-impulso* T^{ik} è l'analogo della *densità di massa o della densità di energia* della teoria newtoniana, esso dovrebbe essere la causa ovvero la sorgente del campo gravitazionale (in virtù del *Principio di Mach*);

4) Le equazioni stesse devono essere tali che, ove per ipotesi manchi in tutto lo spazio ogni presenza di materia ed ogni altra manifestazione energetica, nel qual caso si avrà $T^{ik} = 0$, esse riducano $\mathcal{V}^4$ ad uno spazio pseudoeuclideo di Minkowski $\mathcal{M}^4$;

5) Esse infine siano tali che le leggi
$$\nabla_k T^{ik} = 0$$
di evoluzione della materia (o di altro campo fisico che riempia lo spazio), leggi che per ora noi abbiamo ricavato dalle analoghe (11.24) della RR applicando la *Regola di Trascrizione*, ne discendano come conseguenza logica necessaria.

In base ai precedenti criteri, si può anzitutto osservare che la cercata relazione differenziale non potrà ridursi ad un'unica equazione scalare perché, se in fisica classica, per l'immediata propagazione a distanza delle azioni gravitazionali, la distribuzione di un mezzo materiale può a sufficienza venire descritta da un campo scalare quale è la densità di massa $\mu(x, y, z, t)$, in RR e quindi anche in RG, ove non si può prescindere dal suo moto, essa viene descritta da un tensore doppio T^{ik}, il tensore energetico. Perciò volendo costruire qualcosa di simile all'equazione di Poisson, converrà considerare quest'ultima come espressione, eventualmente approssimata, di una delle equazioni che andiamo cercando, facente parte di un'equazione tensoriale che abbia a secondo membro proprio il tensore energetico e sia perciò del tipo:
$$S^{ik} = T^{ik},$$
dove il tensore materiale T^{ik} è la *sorgente del campo gravitazionale* ed S^{ik} è un tensore simmetrico del 2° ordine, al pari di T^{ik}, *rappresentante la geometria dello spaziotempo* $\mathcal{V}^4$ e perciò costruito con i *potenziali gravitazionali* g^{ik}, costituenti il tensore metrico, e con le sue derivate prime e seconde. Per mantenere l'analogia qualitativa con la teoria newtoniana della gravitazione, per quanto detto nel precedente 1° criterio, chiediamo che nell'espressione di S^{ik} le derivate di g^{ik} compaiano al 1° e al 2° ordine e queste ultime vi figurino linearmente. Allora le equazioni di campo formeranno un'equazione tensoriale avente perciò la stessa forma in qualunque sistema di coordinate.
Osserviamo che *scrivendo le equazioni di campo in una forma che è valida in tutti i sistemi di riferimento, ammettiamo che in RG non ci sono riferimenti privilegiati.*
Le equazioni di campo saranno studiate in dettaglio e scritte in forma esplicita nei prossimi paragrafi 11.15.1, 11.15.2, 11.15.3, 11.16.
Vogliamo terminare questo paragrafo sintetizzando ciò che finora abbiamo detto, con un interessante ed espressivo aforisma del famoso relativista americano John Archibald Wheeler: **"Spacetime tells matter how to move, matter tells spacetime how to curve".**

11.7 Grandezze assolute e grandezze relative. Concetto di riferimento fisico.

Nel paragrafo 11.5, mediante *la Regola di Trascrizione*, abbiamo ottenuto una formulazione 4-dimensionale nello spaziotempo $\mathcal{V}^4$, perciò avente carattere assoluto, ossia indipendente dal

sistema di riferimento adottato: 1) della legge di moto di una particella liberamente gravitante, 2) della legge di evoluzione di un sistema continuo costituito da materia disgregata e più generalmente di un sistema chiuso[18], 3) e infine delle equazioni di Maxwell.
Occorre però osservare che se la descrizione di un fenomeno e delle sue leggi direttamente in termini assoluti 4-dimensionali è perfettamente soddisfacente da un punto di vista concettuale, essa non può bastare. Chi osserva un fenomeno naturale non può fare a meno, nel descrivere il fenomeno osservato, di adottare un particolare riferimento fisico, operazione questa che comporta di per sé necessariamente la separazione della nozione di spazio da quella di tempo. Di qui la necessità di porre in relazione le grandezze assolute 4- dimensionali che figurano nelle leggi con grandezze tradizionali 3- dimensionali che costituiscono l'oggetto immediato delle loro misure. Così, per esempio, il fisico dovrà sapere come dal 4-tensore (ente assoluto) F^{ik} campo elettromagnetico egli possa desumere il campo elettrico **E** e il campo magnetico **H**, 3-dimensionali; e come dal 4-vettore **J** egli possa ricavare la ordinaria densità di corrente **j** e la densità di carica ρ. Queste grandezze **E**, **H**, **j**, ρ dipendono in modo essenziale dal riferimento adottato, sono cioè, come suol dirsi, **grandezze essenzialmente relative.** Per il fisico si pone dunque la questione di desumere le **grandezze fisiche relative** 3-dimensionali dalle **grandezze fisiche assolute** 4-dimensionali. Cioè, egli si può ragionevolmente domandare se per le stesse leggi, di cui conosce **la formulazione assoluta** nello spaziotempo $\mathcal{V}^4$, non sia possibile anche **una formulazione relativa**, fatta naturalmente rispetto al più generale riferimento fisico. Ora ciò è effettivamente possibile, anche se la stragrande maggioranza dei testi sulla Relatività Generale non ne fa alcun cenno, sviluppando quasi esclusivamente la trattazione 4-dimensionale nello spaziotempo e tornando in maniera occasionale (non sistematica) alla considerazione delle corrispondenti grandezze relative 3-dimensionali. Una **formulazione relativa delle leggi fisiche**, perfettamente coerente dal punto di vista logico e matematico, è stata formulata, quasi contemporaneamente ed indipendentemente dal russo A. L. Zelmanof (1956)[19] e dall'italiano Carlo Cattaneo (1958-1963)[20] e successivamente ripresa anche dal danese Christian Møller[21] seguendo questi autori. La trattazione completa di tale formulazione non è né ovvia né banale ed esce fuori dagli scopi di questo libro. A noi basta aver accennato all'esistenza della problematica in questione, e perciò ci limiteremo ad illustrare alcuni concetti utili per la nostra esposizione seguendo Cattaneo.[22]

---

[18]Ricordiamo che si dicono *sistemi chiusi* quei sistemi materiali caratterizzati da un tensore energetico $T_i{}^k$ (non necessariamente simmetrico) soddisfacente alle equazioni $\partial_k T_i{}^k = 0$ che sono le equazioni di evoluzione del sistema medesimo, interpretabili come *equazioni di conservazione dell'energia, e della quantità di moto*, non appena si attribuiscono alle componenti del tensore $T_i{}^k$ uno specifico significato relativo: $T_0{}^0$ = densità relativa di energia, $\frac{1}{c}T_\alpha{}^0$= densità di quantità di moto, $cT_0{}^\alpha$= densità di corrente di energia, $T_\alpha{}^\beta$ = densità di corrente di quantità di moto.

[19]A. L. Zelmanof (1956), *Dokl. Akad. Nauk* SSSR **107**, 815.

[20]C. Cattaneo (1958) *Nuovo Cim.*, **10**, 318; (1959a) *Nuovo Cim.*, **11**, 733; (1959b) *Nuovo Cim.*, **13**, 237; (1959c) *Atti Accad. Naz. Lincei Rc.* **27**, 54; (1959d) *Annali Mat. Pura e Appl.* **48**, 361; (1963) *Comp. Ren. Hebd. Sèanc. Acad. Sci. Paris* **256**, 3974.

[21]Christian Møller, *The Theory of Relativty*, The International Series of Monographs on Physics, Second Edition, Delhi, Oxford University Press, 1972, First Indian Impression 1974, Sixth impression 1990, Cap.10, paragrafo 10.1, pagg. 371-379.

[22]La trattazione completa della formulazione relativa delle leggi fisiche secondo Cattaneo è contenuta nei suoi lavori precedentemente citati, ma è esposta con chiarezza anche nel libro: Carlo Cattaneo, *Introduzione alla Teoria Einsteiniana della Gravitazione*, Vol.I, Cap.VII, Libreria Eredi Virgilio Veschi, Roma (Appunti dal *Corso di Fisica Matematica* Anno Accademico 1960-61, raccolti da Gaetano Caricato e Giorgio Ferrarese). Ulteriori complementi sono contenuti nell'altro volume: Carlo Cattaneo, *Introduzione alla Teoria Einsteiniana della Gravitazione*, Vol.II, Libreria Eredi Virgilio Veschi, Roma (Lezioni tenute all'Istituto Nazionale di Alta Matematica "Francesco Severi", Anno Accademico 1961-62, raccolte da Giorgio Ferrarese).
Il lettore che fosse interessato all'opera del Prof. Carlo Cattaneo può consultare utilmente la **Commemorazione** a lui dedicata pronunciata da Giorgio Ferrarese il 2 Maggio 1979 e pubblicata nel volume: *Scritti Matematici dedicati a Carlo Cattaneo nel decennale della morte*, **Rendiconti di Matematica e delle sue Applicazioni**, Università "La Sapienza", Centro stampa d'Ateneo, Roma, 1989.

Cominciamo col precisare che cosa si deve intendere in Relatività Generale per **riferimento fisico.** C'è un poco l'abitudine di identificare un riferimento fisico con *un sistema di coordinate.* In verità già in fisica classica si può convenientemente distinguere tra *spazio di riferimento* o, come più espressivamente si dice, *solido di riferimento*, e sistema di coordinate adottate entro il solido stesso. Quello che ha diritto al nome di *spazio fisico di riferimento* è il solido di riferimento: entro il solido stesso sono poi possibili infinite diverse scelte di sistemi di coordinate, cartesiane, polari, o curvilinee di specie qualunque.
In RR vengono di regola considerati soltanto *solidi di riferimento galileiani.* Scegliere un sistema di riferimento (con annesso sistema di orologi sincronizzati) equivale a scegliere nello spazio-tempo di Minkowski $\mathcal{M}^4$, un asse temporale ad arbitrio (scelto naturalmente tra le rette del genere tempo). Su questa retta con origine arbitraria, restano individuati, in una scala arbitraria, i successivi istanti. In ciascuno di questi istanti resta poi automaticamente determinato il sottospazio a tre dimensioni ortogonale all'asse temporale, sottospazio che rappresenta lo spazio fisico in quell'istante, o, se si vuole, il solido di riferimento galileiano.
Conviene dare maggiore concretezza allo spazio fisico pensandolo riempito di un'infinità di particelle ideali (*particelle di prova*) tra loro rigidamente collegate, così da costituire un vero riferimento fisico. Beninteso si tratta di particelle ideali che non turbano con la loro presenza il libero svolgersi dei fenomeni formando soltanto il riferimento necessario alla loro descrizione. Con tale ipotesi ogni retta parallela all'asse temporale si può interpretare come linea oraria, cioè come la storia, di una di queste ∞^3 particelle di riferimento.
In Relatività Generale un riferimento galileiano che riempie tutto lo spazio non esiste. Infatti, nel caso generale di un campo gravitazionale arbitrario e variabile la metrica dello spazio non solo non è euclidea ma varia inoltre col tempo. Ciò vuol dire che i rapporti fra le distanze geometriche dei vari punti variano col tempo. Ne segue che la disposizione di "*particelle di prova*", poste nel campo, non può restare inalterata in nessun sistema di coordinate.
Consideriamo un semplice esempio che spiega chiaramente perché è inevitabile che le proprietà non euclidee dello spazio compaiono quando si passa da un riferimento inerziale ad uno non inerziale (Fig.11.1).

Figura 11.1: Riferimento solidale ad un disco ruotante: intervento della geometria non euclidea.

Sia $K' \equiv O'x'y'z'$ un sistema di riferimento solidale con un disco circolare piatto che ruota uniformemente intorno all'asse $O'z'$ perpendicolare al disco. Sia $K \equiv Oxyz$ un sistema di riferimento inerziale con l'asse Oz coincidente con l'asse $O'z'$ del disco. Una circonferenza

nel piano xy del sistema K, di centro nell'origine delle coordinate, può essere considerata anche come una circonferenza nel piano $x'y'$ del sistema K'. Misurando la lunghezza della circonferenza e del suo diametro con un regolo rigido campione nel sistema K, otteniamo dei valori il cui rapporto è π, dato che nel sistema di riferimento inerziale la geometria è euclidea. Supponiamo ora che le stesse misure siano fatte con lo stesso regolo campione in quiete nel riferimento K'. Osservando questo processo dal riferimento K (in cui vale la Teoria della Relatività Ristretta), troviamo che il regolo applicato tangenzialmente alla circonferenza, muovendosi parallelamente alla sua velocità, subisce la contrazione di Lorentz, mentre resta inalterato quando è applicato radialmente.[23] Ne segue che la lunghezza della circonferenza ha un valore maggiore che in precedenza, mentre la lunghezza del diametro rimane invariata, ciò implica che il loro rapporto sarà maggiore di π. Poiché il rapporto tra la lunghezza della circonferenza ed il suo diametro non è uguale a π, è ovvio che se le distanze fra le particelle lungo il diametro restano invariate, devono necessariamente variare le distanze tra le particelle disposte sulla circonferenza, e viceversa. Perciò l'immobilità relativa di un sistema di particelle di prova, nella Relatività Generale è assolutamente impossibile. Questa circostanza cambia la nozione stessa di *sistema di riferimento in Relatività Generale* rispetto al suo significato in RR.
Il modo più semplice e più spontaneo di estendere alla RG la nozione di riferimento fisico è quello di **considerare ancora lo spazio riempito di una infinità di particelle ideali, comunque mobili, purché con regolarità, senza più imporre, beninteso, alcuna condizione di rigidità.** Ecco così come dalla nozione di **"solido di riferimento"** abituale in fisica classica ed in relatività ristretta, si giunge in modo del tutto naturale alla nozione generale di **"fluido di riferimento".** Scegliere un *fluido di riferimento* $\mathcal{S}$ equivale a scegliere nello spaziotempo $\mathcal{V}^4$, che ora è uno spazio curvo, *una congruenza* Γ *di linee del genere tempo, che sono poi le linee orarie delle particelle costituenti il fluido di riferimento.*[24]

Figura 11.2: Riferimento fluido in relatività generale.

Assegnare la congruenza Γ equivale poi ad assegnare un campo di vettori, che siano tangenti alle linee medesime ed orientati verso il futuro (Fig.11.2). Il campo più semplice di vettori tangenti alle linee di Γ che godono di tale proprietà è ovviamente il campo dei vettori unitari $\mathbf{u}$, che come sappiamo dalla RR, sono le 4- velocità adimensionali delle particelle del fluido.[25].
Con la scomparsa del riferimento galileiano viene anche meno in Relatività Generale, almeno a priori, la possibilità di associare ad un dato riferimento fisico *un tempo che abbia veramente un*

[23]Per essere precisi bisognerebbe considerare che il regolo applicato tangenzialmente alla circonferenza sia solidale con un opportuno sistema di riferimento inerziale K'' (*riferimento inerziale di moto incipiente*) che istante per istante si muova di moto TRU rispetto a K con velocità $\mathbf{v}$ uguale alla velocità periferica del riferimento ruotante K'.

[24]Einstein nella sua esposizione divulgativa della Relatività, tale **"fluido di riferimento"** lo chiama, con un'espressione pittoresca, **"mollusco di riferimento"**, si veda: A. Einstein, *Relatività, Esposizione divulgativa*, Boringhieri, Torino (1967), pag. 118.

[25]Ricordiamo al lettore che noi abbiamo definito la 4-velocità di una particella come un vettore le cui componenti, secondo una terna inerziale di riferimento, sono le derivate delle coordinate x^i, che individuano la posizione della particella, rispetto al tempo proprio τ: $\mathbf{U} = \frac{dx^i}{d\tau}$. Altri autori però derivano le coordinate rispetto all'elemento di linea s: $\mathbf{u} = \frac{dx^i}{ds}$; in tal caso il campo delle 4-velocità (adimensionali) è proprio il campo dei vettori unitari che definiscono la congruenza Γ.

contenuto fisico. Ad esso viene sostituito un **tempo coordinato** x^0 che si può pensare indicato da tanti orologi, legati alle varie particelle del fluido di riferimento. La mancanza di un riferimento galileiano rende vana la speranza di un accordo fisico coerente di tutti questi orologi: a questi non si richiederà altro requisito che quello *di non fermarsi mai e di soddisfare alla duplice condizione che a due eventi infinitamente vicini non soltanto corrispondano linee orarie di riferimento infinitamente vicine, ma anche due istanti, letti nei rispettivi orologi, infinitamente vicini.* Naturalmente è ovvio che coordinate così generali, come quelle individuate da un tale sistema di particelle ed orologi, benché perfettamente adatte a coordinare gli eventi del mondo, hanno *individualmente* ben poco significato fisico.
Una volta scelto nello spaziotempo $\mathcal{V}^4$ un riferimento fisico $\mathcal{S}$, diremo che un sistema di coordinate x^0, x^1, x^2, x^3 *è adattato al riferimento stesso* se le linee coordinate $x^0 = var.$ si identificano con le linee orarie (*del genere tempo*) delle particelle di riferimento, e se le linee $x^\alpha = var.$ sono del *genere spazio*. Tali coordinate si dicono anche *coordinate fisicamente ammissibili.* Converremo, per motivi dimensionali, che la coordinata x^0 dia *il tempo coordinato t* a meno del fattore costante c, velocità della luce nel vuoto: $x^0 = ct$. *Le coordinate spaziali* $(x^\alpha) \equiv (x^1, x^2, x^3)$ servono invece ad individuare ciascuna particella di riferimento, rimanendo costanti lungo la storia della particella stessa.
Il carattere di ammissibilità fisica di un sistema di coordinate si traduce, come si può subito verificare, nelle due disuguaglianze:

$$g_{00} > 0, \qquad g_{\alpha\beta} dx^\alpha dx^\beta < 0 \;\; (\alpha, \beta = 1, 2, 3), \tag{11.28}$$

per valori delle dx^α, dx^β non simultaneamente nulli, con tutte le condizioni che esse implicano. Infatti, considerato *il vettore del genere tempo* $(dx^0, 0, 0, 0) \equiv (dx^0 \neq 0, dx^\alpha = 0)$ che si ottiene congiungendo due punti infinitamente vicini di una linea coordinata $x^0 = var.$, e tenendo conto che esso ha norma positiva, si ha.

$$\mathbf{ds}^2 = g_{00}(dx^0)^2 > 0, \tag{11.29}$$

da cui segue la prima delle (11.28). Mentre se indichiamo con $(dx^i) = (0, dx^1, dx^2, dx^3) \equiv (dx^0 = 0, dx^\alpha \neq 0)$ un vettore infinitesimo, *del genere spazio*, che congiunge due punti infinitamente vicini di una varietà $x^0 = cost.$, si ha per esso:

$$\mathbf{ds}^2 = g_{ik} dx^i dx^k = g_{\alpha\beta} dx^\alpha dx^\beta < 0 \;\; (\alpha, \beta = 1, 2, 3), \tag{11.30}$$

la forma quadratica ternaria $g_{\alpha\beta} dx^\alpha dx^\beta$ risulta quindi definita negativa.
Una trasformazione arbitraria di coordinate, pur fisicamente ammissibili, comporta generalmente un cambiamento delle linee temporali x^0 e quindi un cambiamento del riferimento fisico. Perché ciò non accada, cioè affinché la trasformazione lasci ferme le linee temporali x^0, pur cambiando le altre, è necessario e sufficiente che essa sia del tipo:

$$x'^0 = x'^0(x^0, x^1, x^2, x^3), \qquad x'^\alpha = x'^\alpha(x^1, x^2, x^3). \tag{11.31}$$

Una tale trasformazione si può chiamare una *trasformazione interna* al riferimento fisico considerato. Essa può decomporsi nel prodotto di una *trasformazione puramente spaziale* delle sole coordinate spaziali

$$x'^0 = x^0, \qquad x'^\alpha = x'^\alpha(x^1, x^2, x^3), \tag{11.32}$$

(la quale lascia invariato il tempo coordinato e fa cambiare nome alle singole particelle di riferimento) per una *trasformazione puramente temporale* del tipo

$$x'^0 = x'^0(x^0, x^1, x^2, x^3), \qquad x'^\alpha = x^\alpha, \tag{11.33}$$

(la quale lascia invariato il nome delle singole particelle di riferimento ma cambia l'orologio coordinato di ogni particella di riferimento).
È implicito in ciò che precede che, mentre ad ogni sistema di coordinate fisicamente ammissibile corrisponde un ben determinato riferimento fisico $\mathcal{S}$, viceversa ad un dato riferimento fisico corrispondono infiniti sistemi di coordinate ammissibili, l'uno ottenibile dagli altri mediante trasformazioni del tipo (11.31).

11.8 Distanze, intervalli di tempo e metrica spaziale

Dalle considerazioni fatte nel paragrafo precedente segue che in RG la scelta di un *sistema di riferimento fluido* non è soggetta a limitazioni di sorta: il ruolo delle tre coordinate spaziali x^1, x^2, x^3 $(x^\alpha,\ \alpha = 1, 2, 3)$ può essere svolto da grandezze qualsiasi che definiscano la posizione dei corpi nello spazio (supposti puntiformi), e la coordinata temporale x^0 (*tempo coordinato*) può essere determinata da un orologio, solidale con ciascun corpo, che segni il tempo in modo arbitrario.[26]
Sorge qui spontaneo il problema: *come si possono determinare, a partire dai valori delle coordinate* $(x^i) \equiv (x^0, x^\alpha)$, *le distanze e gli intervalli di tempo reali?*
Determiniamo dapprima la relazione fra il **"tempo reale"**, che identificheremo con il **"tempo proprio"** di una particella, e indicheremo, come al solito, con τ, e la *"coordinata temporale"* x^0 che chiameremo invece **"tempo coordinato"**. A questo scopo, considereremo due eventi distinti infinitamente vicini, che avvengono nello stesso punto dello spazio fisico 3-dimensionale ad istanti diversi, definiti perciò da $dx^0 \neq 0$, *e* $dx^\alpha = 0$. Allora l'intervallo spazio-temporale tra questi due eventi non è altro, per la (11.18), che $ds = cd\tau$, dove $d\tau$ è **l'intervallo di tempo reale** tra i due eventi. Ponendo $dx^1 = dx^2 = dx^3 = 0$ nell'espressione generale (11.16) del $\mathbf{ds}^2$ che qui riscriviamo per comodità:

$$\mathbf{ds}^2 = g_{00}(x^i)(dx^0)^2 + 2g_{0\alpha}(x^i)dx^0dx^\alpha + g_{\alpha\beta}(x^i)dx^\alpha dx^\beta \qquad (\alpha, \beta = 1, 2, 3), \tag{11.34}$$

troviamo:

$$\mathbf{ds}^2 = c^2d\tau^2 = g_{00}(x^i)(dx^0)^2, \tag{11.35}$$

da cui, essendo $x^0 = ct$:

$$d\tau = \frac{1}{c}\sqrt{g_{00}(x^i)}dx^0 = \sqrt{g_{00}(x^i)}dt, \tag{11.36}$$

o ancora, per l'intervallo temporale tra due eventi verificantisi nello stesso punto dello spazio fisico 3-dimensionale, ma in istanti diversi:

$$\tau = \frac{1}{c}\int \sqrt{g_{00}(x^i)}dx^0. \tag{11.37}$$

Tali sono le relazioni che determinano gli *intervalli di tempo reali*, in un dato punto dello spazio fisico, in funzione del *"tempo coordinato"* x^0. Dalla (11.35) ne segue, come già abbiamo visto nel paragrafo precedente, che, essendo l'intervallo spazio-temporale positivo ($\mathbf{ds}^2 > 0$) perché del genere tempo, deve essere necessariamente:

$$g_{00}(x^i) > 0. \tag{11.38}$$

Occorre sottolineare la differenza tra il significato della condizione (11.38) ed il significato dell'altra condizione (11.17) determinata dai segni degli autovalori della matrice formata dagli

[26] In questo paragrafo seguiremo l'esposizione di Landau-Lifschitz, *Teoria dei campi*, Cap.X, § 84, Editori Riuniti, Roma (nuova edizione 2010), tenendo comunque conto di ciò che abbiamo detto nei due paragrafi precedenti sullo spaziotempo $\mathcal{V}^4$ e sui sistemi di riferimento.

elementi del tensore metrico g_{ik} calcolato in un determinato punto della varietà $\mathcal{V}^4$. Un tensore metrico che non soddisfa la (11.17) non può, in generale, corrispondere ad un campo gravitazionale reale, qualunque esso sia, non può cioè determinare la metrica di uno spaziotempo reale. Se d'altra parte non fosse soddisfatta la (11.38), ciò significherebbe solamente che il sistema di riferimento corrispondente non può essere realizzato da corpi reali: se invece viene soddisfatta la condizione sugli autovalori, con una trasformazione appropriata delle coordinate si può ottenere allora che g_{00} diventi positiva.[27]
Occupiamoci ora della determinazione **dell'intervallo spaziale** *$d\sigma$ fra due eventi infinitamente vicini.* In RR la distanza spaziale $d\sigma$ si può definire come l'intervallo fra due eventi infinitamente vicini aventi luogo in uno stesso istante. In RG non si può più generalmente adottare questo procedimento, cioè non si può definire $d\sigma$ semplicemente ponendo $dx^0 = 0$ nel $\mathbf{ds}^2$. Ciò è dovuto al fatto che il tempo proprio in un campo gravitazionale è legato al tempo coordinato x^0 da una relazione, la (11.36), che è diversa in diversi punti dello spazio (essendo g_{00} generalmente dipendente da tutte le coordinate x^i).
Per determinare $d\sigma$ procederemo ora come segue.

Figura 11.3: Segnale luminoso emesso da un punto B verso un punto A infinitamente vicino e riflesso istantaneamente.

Supponiamo che un segnale luminoso venga emesso da un punto B dello spazio di coordinate $x^\alpha + dx^\alpha$, verso un punto A infinitamente vicino, di coordinate x^α (Fig.11.3) e riflesso istantaneamente lungo lo stesso cammino. *Il tempo reale (o proprio) necessario perché avvenga l'intero processo (misurato nello stesso punto B)*[28] *è dato evidentemente dal rapporto fra il doppio della distanza $2d\sigma$ dei due punti e la velocità c della luce.* Scriviamo ora l'intervallo spaziotemporale nella forma (11.34), mettendo in evidenza i differenziali della coordinata temporale ed i differenziali delle coordinate spaziali. L'intervallo spaziotemporale fra i due eventi consistenti nella partenza del segnale luminoso dal punto B e nell'arrivo dello stesso segnale nel punto A, essendo del genere luce, è nullo cioè $\mathbf{ds}^2 = 0$. Risolvendo allora l'equazione di 2° grado

$$g_{00}(dx^0)^2 + 2g_{0\alpha}dx^0dx^\alpha + g_{\alpha\beta}dx^\alpha dx^\beta = 0, \tag{11.39}$$

[27]Un esempio che illustra tale possibilità si ha nel caso di un sistema di coordinate in rotazione. Lo studioso interessato a questo esempio può trovarlo in: L. Landau - E. Lifschitz, *Teoria dei campi*, Cap.X, § 89, Editori Riuniti, Roma (nuova edizione 2010).

[28]Ovviamente dobbiamo supporre che in B ci sia un osservatore munito di un orologio arbitrario che segni un intervallo di tempo dt a cui corrisponde l'intervallo di tempo coordinato $dx^0 = cdt$. E la stessa cosa dicasi per A.

rispetto a dx^0 troviamo due radici[29]:

$$(dx^0)_1 = \frac{1}{g_{00}}\Big(- g_{0\alpha}dx^\alpha - \sqrt{(g_{0\alpha}g_{0\beta} - g_{00}g_{\alpha\beta})dx^\alpha dx^\beta}\Big) \tag{11.40}$$
$$(dx^0)_2 = \frac{1}{g_{00}}\Big(- g_{0\alpha}dx^\alpha + \sqrt{(g_{0\alpha}g_{0\beta} - g_{00}g_{\alpha\beta})dx^\alpha dx^\beta}\Big)$$

che corrispondono alla propagazione del segnale nelle due direzioni, tra i punti A e B. Se x^0 è l'istante di arrivo del segnale in A, allora gli istanti di tempo coordinato della sua partenza da B e del suo ritorno in B saranno rispettivamente $x^0 + (dx^0)_1$ ed $x^0 + (dx^0)_2$. Nella figura schematica (Fig.11.3) le rette continue rappresentano le linee di universo corrispondenti alle coordinate date x^α e $x^\alpha + dx^\alpha$, e le rette tratteggiate rappresentano le linee di universo dei segnali.[30] È ovvio che l'intervallo di tempo coordinato totale tra l'emissione ed il ritorno del segnale nello stesso punto B è uguale a:

$$\Delta(dx^0) \equiv [x^0 + (dx^0)_2] - [x^0 + (dx^0)_1] \equiv (dx^0)_2 - (dx^0)_1 = \frac{2}{g_{00}}\sqrt{(g_{0\alpha}g_{0\beta} - g_{00}g_{\alpha\beta})dx^\alpha dx^\beta}. \tag{11.41}$$

L'intervallo di tempo proprio corrispondente si trova, conformemente alla (11.36), moltiplicando la precedente relazione per $\frac{\sqrt{g_{00}}}{c}$:

$$\Delta(d\tau) = \frac{\sqrt{g_{00}}}{c}\Delta(dx^0). \tag{11.42}$$

Perciò la distanza spaziale $d\sigma$ tra i due punti si ottiene moltiplicando la (11.42) per $\frac{c}{2}$:

$$d\sigma = \Delta(d\tau)\frac{c}{2} = \frac{1}{2}\sqrt{g_{00}}\Delta(dx^0).$$

In definitiva, tenendo conto della (11.41) troviamo:

$$d\sigma^2 = \Big(- g_{\alpha\beta} + \frac{g_{0\alpha}g_{0\beta}}{g_{00}}\Big)dx^\alpha dx^\beta. \tag{11.43}$$

È questa l'espressione cercata che determina *la distanza 3- dimensionale fra due punti infinitamente vicini in funzione degli incrementi delle coordinate spaziali*. Riscriviamo la formula ottenuta come segue:

$$d\sigma^2 = \gamma_{\alpha\beta}dx^\alpha dx^\beta, \tag{11.44}$$

dove

$$\gamma_{\alpha\beta} = -g_{\alpha\beta} + \frac{g_{0\alpha}g_{0\beta}}{g_{00}}, \qquad (\gamma_{\alpha\beta} = \gamma_{\beta\alpha}) \tag{11.45}$$

è *il tensore metrico tridimensionale che definisce la metrica puramente spaziale, cioè le proprietà geometriche dello spazio fisico reale.* La relazione (11.45) stabilisce il legame tra la metrica dello spazio

[29]Nelle due formule (11.40) abbiamo sostituito l'espressione $(g_{0\alpha}dx^\alpha)^2 = (g_{0\alpha}dx^\alpha)(g_{0\alpha}dx^\alpha)$ con l'altra equivalente $(g_{0\alpha}dx^\alpha)(g_{0\beta}dx^\beta) = g_{0\alpha}g_{0\beta}dx^\alpha dx^\beta$, essendo α e β indici muti.

[30]Nella (Fig.11.3) si è supposto che $(dx^0)_1 < 0$ e $(dx^0)_2 > 0$, ipotesi che non è affatto necessaria: $(dx^0)_1$ e $(dx^0)_2$ possono avere lo stesso segno. Il fatto che in questo caso il valore di $x^0(A)$ nell'istante di arrivo del segnale in A potrebbe essere inferiore al valore $x^0(B)$ nell'istante della sua partenza da B non è contraddittorio, perché non si suppone che gli orologi situati in punti diversi dello spazio siano in qualche modo sincronizzati.

reale e la metrica dello spaziotempo $\mathcal{V}^4$. ***Si osservi che il tensore metrico $\gamma_{\alpha\beta}$ che determina la geometria spaziale nel dato sistema di riferimento, non è, in generale, semplicemente uguale alla parte spaziale $g_{\alpha\beta}$ del tensore 4-dimensionale g_{ik}.***[31]
Bisogna tuttavia tener presente che le g_{ik} dipendono, in generale, da x^0, in modo che la metrica spaziale (11.44) varia pure col tempo. Per questa ragione non ha senso integrare $d\sigma$:[32] un tale integrale dipenderebbe dalla linea di universo d'integrazione tra i due punti spaziali dati. Di conseguenza **in RG la nozione di distanza determinata fra corpi non ha, in generale senso: essa ha solo un significato locale.** *L'unico caso in cui la distanza può essere definita anche in regioni finite dello spazio è quello dei sistemi di riferimento dove le g_{ik} non dipendono dal tempo, e di conseguenza, l'integrale $\int d\sigma$ su una curva spaziale acquista un significato determinato.*
La (11.44), essendo la metrica dello spazio fisico 3-dimensionale, deve essere *una forma quadratica definita positiva* nelle variabili dx^α, dx^β a coefficienti funzioni delle 4 variabili (x^0, x^α) $(\gamma_{ik} = \gamma_{ik}(x^i))$, perciò si avrà:

$$d\sigma^2 = \gamma_{\alpha\beta} dx^\alpha dx^\beta > 0. \tag{11.46}$$

Per un noto teorema di Sylvester sulle forme quadratiche definite positive[33] i coefficienti della (11.46) debbono soddisfare le condizioni:

$$\gamma_{11} > 0, \quad \begin{vmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{vmatrix} = \gamma = det(\gamma_{ik}) > 0. \tag{11.47}$$

Utilizzando le (11.45) è possibile esprimere le γ_{ik} mediante le g_{ik}; si trova con calcoli semplici ma laboriosi, che le condizioni (11.47) assumono la forma:

$$\begin{vmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{vmatrix} < 0, \quad \begin{vmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{vmatrix} > 0, \quad g = det(g_{ik}) = \begin{vmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{vmatrix} < 0. \tag{11.48}$$

Ricordando la (11.38), possiamo concludere col seguente risultato:
In RG le componenti del tensore metrico 4-dimensionale g_{ik}, in qualsiasi riferimento realizzabile con corpi reali, devono soddisfare sempre alle seguenti condizioni:

$$g_{00} > 0, \quad \begin{vmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{vmatrix} < 0, \quad \begin{vmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{vmatrix} > 0, \quad g = \begin{vmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{vmatrix} < 0, \tag{11.49}$$

dunque il determinante della metrica dello spaziotempo $\mathcal{V}^4$ è sempre negativo:

$$g = det(g_{ik}) < 0. \tag{11.50}$$

[31]Si ha qui un primo esempio di ciò che nel paragrafo 11.7 abbiamo discusso ritenendo essenziale, oltre che la formulazione assoluta 4-dimensionale della RG, anche una sua adeguata formulazione relativa 3-dimensionale. Con ciò si vuol dire (in questo caso specifico della determinazione delle distanze spaziali) che si potrebbe ritenere ingenuamente ed erroneamente che il tensore che determina la geometria dello spazio fisico reale sia semplicemente uguale alla parte spaziale $g_{\alpha\beta}$ del tensore 4-dimensionale g_{ik}, mentre abbiamo visto più sopra che il tensore corretto è $\gamma_{\alpha\beta}$ che, in generale, non coincide affatto con $g_{\alpha\beta}$. Si avrebbe coincidenza, a meno del segno, soltanto se risultasse $g_{0\alpha} = 0$; in seguito ci occuperemo di questo caso particolare in cui il sistema di coordinate adattato al riferimento fisico utilizzato si dice *tempo- ortogonale.*

[32]Non avrebbe senso infatti definire una distanza finita che varia nel tempo.

[33]Il Teorema di Sylvester a cui si allude, si enuncia così: *Condizione necessaria e sufficiente affinché una forma quadratica $f(x^1, x^2, ..., x^n) = a_{ik}x^i x^k$, in n variabili, a coefficienti costanti, sia definita positiva, è che i minori principali della matrice dei coefficienti siano tutti positivi.*

Osserviamo che la precedente disuguaglianza era già stata stabilita nel paragrafo 11.3 (formula (11.17)) con un ragionamento sintetico.
Concludiamo osservando che il tensore spaziale cambiato di segno, cioè $-\gamma_{\alpha\beta}$, è l'inverso del tensore controvariante tridimensionale $g^{\alpha\beta}$. Infatti se scriviamo l'uguaglianza 4-dimensionale $g^{ik}g_{kl} = \delta^i_l$ mettendo in evidenza i termini contenenti gli indici spaziali α, β e quelli contenenti l'indice temporale 0, avremo[34]:

$$\begin{cases} g^{\alpha 0}g_{0\mu} + g^{\alpha\beta}g_{\beta\mu} = \delta^\alpha_\mu, \\ g^{\alpha\beta}g_{\beta 0} + g^{\alpha 0}g_{00} = 0, \\ g^{0\beta}g_{\beta 0} + g^{00}g_{00} = 1. \end{cases} \tag{11.51}$$

Eliminando $g^{\alpha 0}$ fra la prima e la seconda delle precedenti relazioni otteniamo:

$$-g^{\alpha\beta}\gamma_{\beta\mu} = \delta^\alpha_\mu, \tag{11.52}$$

ciò che dimostra la nostra affermazione. Questo risultato può essere espresso in altre parole dicendo che le grandezze $-g^{\alpha\beta}$ formano un tensore metrico controvariante 3-dimensionale corrispondente alla metrica spaziale (11.46), cioè:

$$\gamma^{\alpha\beta} = -g^{\alpha\beta}. \tag{11.53}$$

Possiamo inoltre constatare, tenuto conto delle relazioni (11.47), (11.48), (11.49), (11.50), che i determinanti g e γ dei tensori g_{ik} e γ_{ik} sono legati fra di loro dalla semplice relazione:

$$g = -g_{00}\gamma. \tag{11.54}$$

Ricordando che $g_{00} > 0$, al fine di introdurre alcune considerazioni geometriche che faremo nei prossimi paragrafi è utile riscrivere la relazione (11.45), che stabilisce il legame tra la metrica dello spazio reale e la metrica dello spaziotempo $\mathcal{V}^4$, nel modo seguente:

$$\gamma_{\alpha\beta} = -g_{\alpha\beta} + \frac{g_{0\alpha}}{\sqrt{g_{00}}}\frac{g_{0\beta}}{\sqrt{g_{00}}}, \tag{11.55}$$

la quale suggerisce l'introduzione di un vettore tridimensionale $\mathbf{g}$ le cui componenti covarianti sono definite come segue[35]:

$$g_\alpha = \frac{g_{0\alpha}}{\sqrt{g_{00}}}, \qquad \left(g_\beta = \frac{g_{0\beta}}{\sqrt{g_{00}}}\right). \tag{11.56}$$

Con tale definizione, la (11.45) si scriverà in maniera più semplice:

$$\gamma_{\alpha\beta} = -g_{\alpha\beta} + g_\alpha g_\beta. \tag{11.57}$$

Considerando $\mathbf{g}$ come un vettore dello spazio avente la metrica (11.57), le sue componenti controvarianti saranno:

$$g^\alpha = \gamma^{\alpha\beta}g_\beta. \tag{11.58}$$

[34]Occorre stare attenti a *non commettere un errore piuttosto comune*. Nella relazione $g^{ik}g_{kl} = \delta^i_l$ l'indice muto di sommazione k, deve assumere necessariamente tutti i valori $0, 1, 2, 3$; tale formula sarebbe sbagliata se si dovesse scrivere così: $g^{i\alpha}g_{\alpha l} = \delta^i_l$ con $\alpha = 1, 2, 3$.

[35]Avvertiamo il lettore che nel libro di Landau-Lifschitz, più volte citato, le componenti covarianti del vettore 3- dimesionale $\mathbf{g}$ sono definite dalla relazione $g_\alpha = -\frac{g_{0\alpha}}{g_{00}}$, e perciò in maniera leggermente diversa da ciò che abbiamo fatto noi. La definizione da noi adottata (usata per esempio da Møller, Arzeliès e Cattaneo) gode della proprietà consistente nel fatto che le g_α hanno un preciso significato geometrico: esse coincidono con le componenti spaziali covarianti u_α del campo dei 4-vettori unitari (4-velocità adimensionali) $\mathbf{u}$ le cui componenti controvarianti sono $(u^0 = 1/\sqrt{g_{00}}, u^\alpha = 0)$ tangenti (ed orientati verso il futuro) alle linee del genere tempo della congruenza Γ costituente *il riferimento fisico adottato*, che abbiamo introdotto nel paragrafo 11.7.

Sostituendo in quest'ultima $\gamma^{\alpha\beta}$ con $-g^{\alpha\beta}$ come dalla (11.53) e tenendo conto della seconda delle (11.51), si ottiene:

$$g^{\alpha} = \sqrt{g_{00}}\, g^{0\alpha}. \tag{11.59}$$

Infine dalla terza uguaglianza delle (11.51), in virtù delle (11.56), (11.58) se ne trae quest'ultima:

$$g^{00} = \frac{1}{g_{00}}(1 - g_{\alpha} g^{\alpha}). \tag{11.60}$$

11.9 Direzione temporale e piattaforma spaziale associata.

Dedicheremo questo paragrafo ed il successivo all'introduzione di alcuni concetti, dovuti a Cattaneo[36], che ci conducono con ragionamenti strettamente geometrici alla stessa definizione della metrica spaziale $\gamma_{\alpha\beta}$ introdotta nel paragrafo precedente seguendo Landau-Lifschitz. Riattaccandoci alle considerazioni fatte nel paragrafo 11.7 ricordiamo che: un generico **riferimento fisico** è individuato univocamente da una congruenza Γ di linee del genere tempo, la quale, a sua volta, resta univocamente individuata da un campo di 4-vettori unitari $\mathbf{u}(x)$ tangenti alle linee medesime, ed orientati verso il futuro (Fig.11.4).[37]

Figura 11.4: Direzione temporale e piattaforma spaziale associata.

Osserviamo inoltre che ad una congruenza di linee del genere tempo ($\mathbf{ds}^2 > 0$) è sempre possibile, ed in infinite maniere, associare un sistema di coordinate fisicamente ammissibili x^i che ammetta come linee coordinate $x^0 = var.$ le linee di Γ. Ciò posto, in coordinate adattate, i vettori unitari $\mathbf{u}(x)$ hanno componenti controvarianti espressi da $(u^0 \neq 0, u^{\alpha} = 0)$, con $u^0 > 0$, dato che sono orientati verso il futuro, inoltre essendo unitari e del genere tempo la loro norma deve essere uguale ad 1, cioè:

$$||\mathbf{u}|| = \mathbf{u} \cdot \mathbf{u} = g_{ik} u^i u^k = g_{00}(u^0)^2 + 2 g_{0\alpha} u^0 u^{\alpha} + g_{\alpha\beta} u^{\alpha} u^{\beta} = 1. \tag{11.61}$$

Ed essendo $u^{\alpha} = 0$ per $\alpha = 1, 2, 3$, $u^0 > 0$ e $g_{00} > 0$ avremo $u^0 = 1/\sqrt{g_{00}}$. Perciò le componenti controvarianti del vettore unitario $\mathbf{u}(x)$, in coordinate adattate sono:

$$u^0 = \frac{1}{\sqrt{g_{00}}}, \qquad u^{\alpha} = 0. \tag{11.62}$$

[36]Si veda il testo già citato di Carlo Cattaneo, *Introduzione alla Teoria Einsteiniana della Gravitazione*, Vol.I, Cap.VII, Libreria Eredi Virgilio Veschi, Roma (Anno Accademico 1960-61).

[37]Osserviamo che il campo dei 4-vettori unitari $\mathbf{u}(x)$ è indicato da Cattaneo con $\mathbf{fl}(x)$; noi non usiamo questa notazione perché la lettera γ la riserviamo esclusivamente per indicare il fattore di Lorentz.

Le componenti covarianti di $\mathbf{u}$ si ottengono poi con la solita regola di abbassamento degli indici $u_i = g_{ik}u^k$ che, tenuto conto delle (11.62), ci da:

$$u_0 = \sqrt{g_{00}}, \qquad u_\alpha = \frac{g_{0\alpha}}{\sqrt{g_{00}}}, \qquad (\alpha = 1, 2, 3) \tag{11.63}$$

che possiamo sintetizzare nell'unica relazione:

$$u_i = \frac{g_{0i}}{\sqrt{g_{00}}} \qquad (i = 0, 1, 2, 3). \tag{11.64}$$

Osserviamo subito che le u_α definite dalla seconda delle (11.63) coincidono con le g_α date nel paragrafo precedente dalla (11.56).

La sottovarietà $x^0 = cost.$ si interpreta talvolta come spazio fisico all'istante $t = \dfrac{x^0}{c}$. Tale interpretazione non ha una grande importanza, dato lo scarso significato fisico della variabile temporale x^0. Ben più importante è la definizione seguente. In corrispondenza al generico evento $(x^i) \equiv (x)$ di $\mathcal{V}^4$ consideriamo lo spazio tangente T_x ad esso associato: in esso consideriamo il sottospazio Θ_x ad una dimensione tangente alla linea coordinata x^0; la sua direzione la chiameremo *direzione temporale* localmente associata al riferimento fisico scelto. Consideriamo inoltre il sottospazio Σ_x di T_x a tre dimensioni, ortogonale a Θ_x: lo chiameremo *piattaforma spaziale*, localmente associata al riferimento adottato. Θ_x e Σ_x sono due sottospazi di T_x tra loro *supplementari*[38], sicché lo spazio tangente T_x si può pensare decomposto nella loro somma:

$$T_x = \Theta_x \oplus \Sigma_x. \tag{11.65}$$

In conformità delle definizioni ora date, diremo *puramente temporali* i vettori di Θ_x e *puramente spaziali* i vettori di Σ_x.

In coordinate adattate i vettori puramente temporali sono caratterizzati dall'avere nulle le ultime tre componenti controvarianti (quelle aventi gli indici $\alpha = 1, 2, 3$, come mostrano le (11.62) per il vettore, anch'esso puramente temporale $\mathbf{u}$); i vettori puramente spaziali sono invece caratterizzati dall'avere nulla la prima componente covariante (quella di indice zero) come si riconosce facilmente imponendo la perpendicolarità a $\mathbf{u}$. Se indichiamo con $\mathbf{A}$ ed $\mathbf{N}$ due generici vettori rispettivamente temporale e spaziale avremo perciò:

$$A^0 \neq 0, \qquad A^\alpha = 0, \qquad (\alpha = 1, 2, 3) \tag{11.66}$$

$$N_0 = 0, \qquad N_\alpha \neq 0, \qquad (\alpha = 1, 2, 3). \tag{11.67}$$

Consideriamo ora un generico vettore $\mathbf{V}$ di T_x; com'è noto esso è decomponibile univocamente nella somma di due vettori $\mathbf{A}$ ed $\mathbf{N}$ rispettivamente appartenenti a Θ_x (cioè parallelo a $\mathbf{u}$) e a Σ_x (cioè normale a $\mathbf{u}$). Precisamente si ha, come si può subito controllare[39]:

$$\mathbf{A} = (\mathbf{V} \cdot \mathbf{u})\mathbf{u}, \qquad \mathbf{N} = \mathbf{V} - (\mathbf{V} \cdot \mathbf{u})\mathbf{u}, \tag{11.68}$$

[38] Ricordiamo che due sottospazi V_p e V_q, di dimensioni rispettive $p < n$ e $q < n$, di un dato spazio vettoriale E_n, di dimensione n, si dicono *supplementari* quando essi sono *disgiunti* (cioè non hanno altri vettori in comune che il vettore nullo) e la somma delle loro dimensioni eguaglia la dimensione di E_n, cioè $n = p + q$.
Osserviamo che le basi di due sottospazi fra loro supplementari costituiscono complessivamente una base di E_n, inoltre sussiste il seguente importante
Teorema: *Dati due sottospazi supplementari V_p e V_q, ogni vettore di E_n può* **univocamente decomporsi** *nella somma di due vettori appartenenti l'uno a V_p, l'altro a V_q.*

[39] Infatti si deve avere $\mathbf{V} = \mathbf{A} + \mathbf{N}$, ed essendo $\mathbf{A}$ parallelo a $\mathbf{u}$, esso potrà scriversi nella forma $\mathbf{A} = \lambda\mathbf{u}$ con λ reale, si avrà allora $\mathbf{V} = \lambda\mathbf{u} + \mathbf{N}$. Moltiplicando scalarmente 1° e 2° membro di quest'ultima relazione per $\mathbf{u}$, avremo: $\mathbf{V} \cdot \mathbf{u} = \lambda\mathbf{u} \cdot \mathbf{u} + \mathbf{N} \cdot \mathbf{u}$. Ricordando che $\mathbf{u} \cdot \mathbf{u} = 1$, ed $\mathbf{N} \cdot \mathbf{u} = 0$, si avrà $\lambda = \mathbf{V} \cdot \mathbf{u}$, perciò $\mathbf{A} = (\mathbf{V} \cdot \mathbf{u})\mathbf{u}$, ed infine $\mathbf{N} = \mathbf{V} - (\mathbf{V} \cdot \mathbf{u})\mathbf{u}$.

ovvero, tenuto conto che $\mathbf{V} \cdot \mathbf{u} = g_{jh} V^j u^h = V^j u_j$, in termini di componenti naturali controvarianti (cioè secondo la base indotta in T_x dal sistema di coordinate adattato x^i) avremo:

$$A^k = (g_{jh} V^j u^h) u^k = u_j u^k V^j, \qquad N^k = V^k - (V^j u_j) u^k. \tag{11.69}$$

Le corrispondenti componenti covarianti si esprimono così:

$$A_i = g_{ik} A^k = u_i u_j V^j, \qquad N_i = g_{ik} N^k = (g_{ik} - u_i u_k) V^k = \Gamma_{ik} V^k, \tag{11.70}$$

avendo posto:

$$\Gamma_{ik} \stackrel{\text{def}}{=} g_{ik} - u_i u_k. \tag{11.71}$$

Siffatta decomposizione di $\mathbf{V}$ nella somma di un vettore puramente temporale e di un vettore puramente spaziale verrà chiamata *decomposizione naturale* di $\mathbf{V}$.
I due vettori componenti $\mathbf{A}$ ed $\mathbf{N}$ si diranno rispettivamente *la proiezione temporale* e la *la proiezione spaziale* di $\mathbf{V}$ e possono essere indicati simbolicamente così:

$$\mathcal{P}_\Theta(V_i) = A_i, \qquad \mathcal{P}_\Sigma(V_i) = N_i, \tag{11.72}$$

$\mathcal{P}$ indicando genericamente un operatore di proiezione. Possiamo verificare che le ultime tre componenti controvarianti del vettore temporale $\mathbf{A}$ sono nulle, mentre del vettore spaziale $\mathbf{N}$ è nulla senz'altro la prima componente covariante. Segue infatti dalle (11.62), (11.63):

$$\begin{cases} A^\alpha = u^\alpha u_k V^k = 0, \\ N_0 = (g_{0k} - u_0 u_k) V^k = 0. \end{cases} \tag{11.73}$$

La (11.71), tenuta presente l'espressione esplicita (11.64) di u_i, mostra che le componenti covarianti significative del tensore doppio simmetrico Γ_{ik} sono soltanto le $\Gamma_{\alpha\beta}$, essendo nulle le rimanenti Γ_{0k}, cioè $\Gamma_{ik} \equiv \Gamma_{\alpha\beta}$. Ciò si esprime dicendo che il tensore Γ_{ik} è totalmente spaziale. Viceversa, il tensore doppio simmetrico $u^i u^k$ ha nulle tutte le componenti controvarianti diverse da $u^0 u^0$. Ciò si esprime dicendo che esso è un tensore totalmente temporale.
Le (11.70) mettono in evidenza come le due proiezioni, temporale e spaziale, di un vettore V^i si ottengono rispettivamente operando su di esso per composizione, mediante il tensore $u_i u_k$ o il tensore Γ_{ik}, ai quali si addice pertanto il nome di *proiettore temporale* e *proiettore spaziale*.
Ai tensori medesimi compete anche un'altra funzione: suol chiamarsi *norma temporale* e *norma spaziale* del vettore $\mathbf{V}$ rispetto al vettore temporale di riferimento $\mathbf{u}$, la norma spaziotemporale dei vettori $\mathbf{A}$ ed $\mathbf{N}$ rispettivamente:

$$\begin{cases} ||\mathbf{V}||_\Theta \stackrel{\text{def}}{=} ||\mathbf{A}|| = g_{ik} A^i A^k \\ \\ ||\mathbf{V}||_\Sigma \stackrel{\text{def}}{=} ||\mathbf{N}|| = g_{ik} N^i N^k. \end{cases} \tag{11.74}$$

Vediamo come possono essere trasformati i secondi membri delle (11.74) tenendo conto delle (11.69) e che $g_{ik} u^i u^k = u_i u^i = 1$. Per la prima avremo:

$$g_{ik} A^i A^k = g_{ik} (u_j u^i V^j)(u_h u^k V^h) = (g_{ik} u^i u^k)(u_j V^j)(u_h V^h) = u_i u_k V^i V^k,$$

e per la seconda:

$$g_{ik} N^i N^k = N^i N_i = [V^i - (V^h u_h) u^i] \Gamma_{ik} V^k = \Gamma_{ik} V^i V^k - (g_{ik} - u_i u_k) V^k V^h u_h u^i.$$

ed essendo $u_i u_k V^i V^k = (u_i V^i)^2$, $\Gamma_{ik} = \Gamma_{\alpha\beta}$, a conti fatti si ha:

$$\begin{cases} ||\mathbf{V}||_\Theta \stackrel{\text{def}}{=} ||\mathbf{A}|| = g_{ik} A^i A^k = u_i u_k V^i V^k \\ \\ ||\mathbf{V}||_\Sigma \stackrel{\text{def}}{=} ||\mathbf{N}|| = g_{ik} N^i N^k = \Gamma_{ik} V^i V^k = \Gamma_{\alpha\beta} V^\alpha V^\beta. \end{cases} \tag{11.75}$$

Da tali formule si rileva nuovamente l'intervento sistematico dei due tensori $u_i u_k$ e Γ_{ik} i quali, del vettore $\mathbf{V}$, mettono in evidenza, rispettivamente, il contenuto temporale ed il contenuto spaziale. Evidentemente ciascuno di essi compie nelle (11.70) una funzione diversa che nelle (11.75): in queste ultime essi compiono un ufficio metrico, mentre nelle (11.70), come si è già detto, operando per semplice composizione, esse hanno per effetto di proiettare il vettore $\mathbf{V}$ rispettivamente su Θ_x e su Σ_x. In relazione a tale nuovo ufficio i due tensori $u_i u_k$ e Γ_{ik} si diranno anche, rispettivamente, *tensore metrico temporale* e *tensore metrico spaziale*.
Dalle (11.75) si desume infine che *la norma temporale di un vettore è sempre positiva* e che *la sua norma spaziale è sempre negativa*, infatti la prima delle (11.75) può essere scritta nella forma

$$u_i u_k V^i V^k = (u_i V^i)(u_k V^k) = (u_i V^i)^2 > 0,$$

mentre la seconda diventa

$$\Gamma_{ik} V^i V^k = \Gamma_{\alpha\beta} V^\alpha V^\beta = g_{\alpha\beta} V^\alpha V^\beta - (u_\alpha V^\alpha)^2 < 0.$$

La forma quadratica $\Gamma_{ik} V^i V^k = \Gamma_{\alpha\beta} V^\alpha V^\beta$ differendo per il termine negativo $-(u_\alpha V^\alpha)^2$ dalla forma quadratica definita negativa $g_{\alpha\beta} V^\alpha V^\beta$ (in virtù della (11.28)) è essa stessa definita negativa.
Osservazione sul tensore metrico $\Gamma_{ik} = \Gamma_{\alpha\beta}$.
Osserviamo che il tensore puramente spaziale definito in questo paragrafo dalla (11.71)

$$\Gamma_{\alpha\beta} = g_{\alpha\beta} - u_\alpha u_\beta$$

differisce dal tensore, pur esso puramente spaziale,

$$\gamma_{\alpha\beta} = -g_{\alpha\beta} + g_\alpha g_\beta$$

definito nel paragrafo precedente dalla (11.57) soltanto per il segno, cioè:

$$\Gamma_{\alpha\beta} = -\gamma_{\alpha\beta}, \tag{11.76}$$

e ciò in virtù delle (11.56) e (11.63) che identificano g_α con u_α.
Questo implica ovviamente che, come abbiamo visto più sopra, essendo

$$\Gamma_{\alpha\beta} dx^\alpha dx^\beta = -\gamma_{\alpha\beta} dx^\alpha dx^\beta, \tag{11.77}$$

una forma quadratica *definita negativa*, sarà invece *definita positiva* la forma quadratica ottenuta dalla precedente cambiandole segno, cioè la:

$$\gamma_{\alpha\beta} dx^\alpha dx^\beta. \tag{11.78}$$

Questo fatto verrà subito utilizzato all'inizio del prossimo paragrafo per definire *la prima grandezza relativa standard*.

11.10 Grandezze relative standard

Tenuto conto di ciò che si è detto alla fine del paragrafo precedente, consideriamo in $\mathcal{V}^4$ due eventi qualsiasi infinitamente vicini, caratterizzati dalle coordinate x^i, $x^i + dx^i$, possiamo definire per essi le seguenti grandezze:

1. **Quadrato della distanza spaziale standard** fra i due eventi[40]:

$$d\sigma^2 \stackrel{\text{def}}{=} \gamma_{\alpha\beta} dx^\alpha dx^\beta > 0, \tag{11.79}$$

[40]Si ricordi che la **distanza spaziale** fra due eventi infinitamente vicini è già stata calcolata nel paragrafo 11.8 e coincide proprio con quella che qui viene chiamata **distanza spaziale standard.**

2. **Intervallo di tempo relativo standard** tra i due eventi:

$$dT \stackrel{\text{def}}{=} \frac{1}{c} u_i dx^i. \tag{11.80}$$

Entrambe le quantità $d\sigma^2$ e dT ora introdotte non hanno carattere assoluto, ma dipendono dal riferimento fisico considerato, cioè dal campo di vettori $\mathbf{u}(x)$ che lo definisce. Tuttavia esse hanno un ruolo essenziale quando si voglia dare una corretta formulazione relativa delle leggi fisiche.
Per giustificare le definizioni introdotte, cominciamo con l'osservare che, nel caso di uno spazio piatto di Minkowski, riferito a coordinate galileiane, $x^0 = ct, x^1, x^2, x^3$ le (11.79) e (11.80) danno rispettivamente il quadrato dell'ordinaria distanza tra i due eventi, $(dx^1)^2 + (dx^2)^2 + (dx^3)^2 = \sum_{\alpha=1}^{3}(dx^\alpha)^2 = \delta_{\alpha\beta}dx^\alpha dx^\beta$ e l'intervallo di tempo dt, misurati nel riferimento galileiano adottato. Infatti essendo allora

$$\mathbf{ds}^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2,$$

cioè

$$g_{00} = 1, \quad g_{\alpha\alpha} = -1, \quad g_{ik} = 0, \ \ per \ \ i \neq k.$$

le (11.64) danno

$$u_i = \frac{g_{0i}}{\sqrt{g_{00}}} = \delta_{0i}, \qquad \gamma_{\alpha\beta} = \delta_{\alpha\beta}$$

ciò che riduce le (11.79) e (11.80) proprio a

$$d\sigma^2 = \delta_{\alpha\beta}dx^\alpha dx^\beta, \qquad dT = \frac{1}{c}u_0 dx^0 + u_\alpha dx^\alpha = \frac{1}{c}dx^0 = dt.$$

Si osservi che $d\sigma^2$ e dT sono *invarianti* di fronte ad ogni cambiamento di coordinate che non alteri il campo di vettori $\mathbf{u}(x)$, che sia cioè interno al riferimento fisico prefissato. Questa proprietà d'invarianza è ovviamente essenziale per una corretta formulazione relativa. Come in fisica classica le equazioni che traducono, in forma relativa ad un prefissato solido di riferimento $\mathcal{R}$, le leggi fisiche si scrivono in forma invariantiva rispetto ad ogni cambiamento di coordinate interno ad $\mathcal{R}$, esprimendole mediante vettori o tensori, così in Relatività Generale una corretta formulazione relativa delle leggi fisiche deve essere invariante rispetto al fluido di riferimento, invariante cioè rispetto a tutte le trasformazioni di coordinate, interne al riferimento fisico cioè del tipo (11.31), che mutano in sé la congruenza fondamentale Γ.
Con l'introduzione delle due precedenti grandezze standard, la metrica spaziotemporale può in ogni caso scriversi:

$$\mathbf{ds}^2 = c^2dT^2 - g_{\alpha\beta}dx^\alpha dx^\beta = c^2dT^2 - d\sigma^2, \tag{11.81}$$

formula che in sostanza decompone la norma spaziotemporale del vettore infinitesimo dx^i nella differenza fra la sua norma temporale c^2dT^2 e la sua norma spaziale $d\sigma^2$. Infatti si ha, aggiungendo e sottraendo $u_iu_kdx^idx^k$ alla metrica spaziotemporale:

$$\mathbf{ds}^2 = g_{ik}dx^idx^k + u_iu_kdx^idx^k - u_iu_kdx^idx^k = (u_idx^i)(u_kdx^k) - (-g_{ik} + u_iu_k)dx^idx^k =$$
$$= (u_idx^i)^2 - \gamma_{ik}dx^idx^k = c^2dT^2 - \gamma_{\alpha\beta}dx^\alpha dx^\beta = c^2dT^2 - d\sigma^2.$$

Ciò posto, limitiamoci a dare qualche nozione di Cinematica di una particella. Ricordiamo innanzitutto che una linea di $\mathcal{V}^4$ si interpreta come linea oraria di una particella materiale se essa è del genere tempo (la linea oraria di un fotone è invece necessariamente del genere luce). Si presenta allora spontanea la seguente definizione: *una particella P si dice in quiete nel riferimento fisico $\mathcal{R}$ se la sua linea oraria coincide con una linea della congruenza Γ.* In tal caso il

vettore infinitesimo dP è sempre parallelo a $\mathbf{u}(x)$ e quindi $dx^\alpha \equiv 0$.
Se la linea oraria di P non coincide con una delle linee della congruenza Γ si dice invece che P è in moto nel riferimento considerato: risulta allora $dx^\alpha \neq 0$.
Siano

$$x^i = x^i(\lambda), \tag{11.82}$$

le equazioni della linea oraria di P, con λ un qualunque parametro. Se si considera l'intervallo di tempo relativo standard ad esso associato,

$$dT = \frac{1}{c} u_i dx^i = \frac{1}{c} \mathbf{u} \cdot dP, \tag{11.83}$$

poiché dP è del genere tempo, o al più del genere luce, se la data particella è un fotone, il prodotto scalare $\mathbf{u} \cdot dP$ risulta diverso da zero in quanto anche $\mathbf{u}$ è del genere tempo. Ha senso allora pensare $\lambda = \lambda(T)$ e scrivere le (11.82) nella forma

$$x^i = x^i(T). \tag{11.84}$$

Facciamo osservare che, mentre la (11.84) ha senso incondizionato, non ha invece senso esprimere le coordinate x^i in funzione del tempo proprio τ, $x^i = x^i(\tau)$, tutte le volte che la linea oraria è del genere luce (fotone), perché in tal caso $d\tau = 0$.
Chiameremo *4-velocità relativa* della particella il vettore $\dfrac{dx^i}{dT}$ e invece *velocità relativa standard* la sua proiezione su Σ_x:

$$v_i = \Gamma_{ij} \frac{dx^j}{dT} = -\gamma_{i\beta} \frac{dx^\beta}{dT}, \qquad (v_0 = 0). \tag{11.85}$$

Si tratta, per definizione, di un vettore di Σ_x, cioè di un vettore a tre componenti. Volendolo considerare anche come vettore di T_x, ciò che è perfettamente lecito, occorre introdurre una prima componente covariante nulla, $v_0 = 0$ (la corrispondente componente controvariante sarà invece tale che risulti $u_i v^i = 0$, perché essendo u_i del genere tempo e v^i del genere spazio, sono ortogonali; inoltre si ha $v^0 = -\dfrac{u_\alpha v^\alpha}{u_0}$). Per le componenti controvarianti si ha

$$v^\alpha = -\gamma^\alpha_\beta \frac{dx^\beta}{dT} = -(-\delta^\alpha_\beta + u^\alpha u_\beta) \frac{dx^\beta}{dT} = \delta^\alpha_\beta \frac{dx^\beta}{dT}$$

cioè:

$$v^\alpha = \frac{dx^\alpha}{dT}. \tag{11.86}$$

La (11.85) si può allora scrivere anche così:

$$v_\alpha = \gamma_{\alpha\beta} v^\beta. \tag{11.87}$$

Naturalmente chiameremo *norma o quadrato della velocità relativa standard* la quantità positiva

$$v^2 = \gamma_{\alpha\beta} v^\alpha v^\beta = \frac{d\sigma^2}{dT^2}, \tag{11.88}$$

che, per quanto si è detto, ha senso anche per un fotone. Facendo intervenire l'intervallo di tempo proprio $d\tau^2 = \dfrac{1}{c^2} ds^2$ e la decomposizione (11.81) si ricava:

$$\left(\frac{d\tau}{dT}\right)^2 = 1 - \frac{v^2}{c^2}, \tag{11.89}$$

dalla quale, per $d\tau \neq 0$, si avrà:

$$\frac{dT}{d\tau} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{11.90}$$

relazione formalmente identica al fattore di Lorentz della Relatività Ristretta.
Si noti come il rapporto $\frac{dT}{d\tau}$ sia direttamente influenzato soltanto dal moto della particella e non dal campo gravitazionale (che interviene indirettamente, nella definizione delle misure standard, di lunghezza e di tempo).
Per un fotone ($d\tau = 0$) risulta, in base alla (11.89)

$$v = c \tag{11.91}$$

perciò: *l'adozione delle misure standard, di spazio e di tempo, attribuisce alla velocità della luce nel vuoto il valore universale c anche in presenza di un campo gravitazionale.*
In termini sostanzialmente equivalenti si può dire che: *l'adozione delle misure standard equivale ad assumere la velocità della luce nel vuoto, presente o no il campo gravitazionale, come* **la velocità campione per eccellenza, invariabile per definizione.** *In sostanza, così facendo, si estende, nell'ambito della R.G., ad ogni possibile riferimento fisico ciò che in R.R. è limitato ai soli riferimenti galileiani.*
Osservazione 1:
Si osservi che generalmente *dT non è un differenziale esatto*, cioè la forma differenziale $u_i dx^i$, in generale, non è integrabile, sicché non risulta possibile definire nel riferimento $\mathcal{R}$ un unico tempo standard, o, come suol dirsi, in $\mathcal{R}$ *non è possibile la sincronizzazione degli orologi standard.*
La cosa sarà invece possibile nel caso che le u_i si identifichino con le derivate parziali di una funzione scalare $f(x)(u_i = \partial_i f)$ per il che occorre e basta che sia $\Omega_{ij} = \partial_i u_j - \partial_j u_i \equiv 0$. In tal caso la (11.80) si integra immediatamente e si avrà $T = \frac{1}{c} f(x)$. Si osservi che, in generale, Ω_{ij} è un tensore antisimmetrico che si identifica con il rotore del campo di vettori unitari $\mathbf{u}$, cioè $\Omega_{ij} \equiv \mathrm{rot}\mathbf{u}$.
Per una coppia di eventi, x^i, $x^i + dx^i$, appartenenti alla storia di una medesima particella materiale (quindi a $\mathbf{ds}^2 > 0$) restano definiti tre intervalli temporali: $dt = \frac{dx^0}{c}$, **intervallo di tempo coordinato** (che di per sé non ha un vero contenuto fisico); dT, **intervallo di tempo relativo standard**; $d\tau$ **intervallo di tempo proprio**. Tra dT e $d\tau$ intercorre la relazione (11.90), mentre tra dt e dT intercorre quest'altra:

$$\frac{dt}{dT} = \frac{1 - u_\alpha \frac{v^\alpha}{c}}{\sqrt{g_{00}}}. \tag{11.92}$$

Per riconoscere la validità di questa seconda relazione basta osservare che la (11.80) si può scrivere $dT = \frac{1}{c} u_0 dx^0 + \frac{1}{c} u_\alpha dx^\alpha = u_0 dt + \frac{1}{c} u_\alpha \frac{dx^\alpha}{dT} dT$, ovvero

$$dT\left(1 - u_\alpha \frac{v^\alpha}{c}\right) = u_0 dt,$$

da cui la (11.92).
Osservazione 2:
Se si usano distanze ed intervalli temporali non standard in coordinate non adattate ma del tutto generali, le cose vanno in maniera molto diversa: in particolare la velocità di un fotone (ovvero di un segnale luminoso) in un campo gravitazionale, non avrà un valore costante ma dipenderà sia dalla posizione sia dalla direzione di propagazione del raggio luminoso, come vedremo nel paragrafo 11.12.
Terminiamo questo paragrafo definendo, per una generica particella materiale, alcune altre

grandezze relative.
Tenendo conto delle (11.85), (11.86), (11.90), ricordando inoltre le espressioni della 4-velocità $U^j = \frac{dx^j}{d\tau} = \frac{dx^j}{dT}\frac{dT}{d\tau}$ e della 4-quantità di moto $P^j = m_0 U^j = m_0 \frac{dx^j}{dT}\frac{dT}{d\tau}$, possiamo definire la:
Quantità di moto standard

$$p_i = \mathcal{P}_\Sigma(P_i) = -\gamma_{ij}P^j = -\gamma_{i\alpha}P^\alpha = -\gamma_{i\alpha}(m_0\frac{dx^\alpha}{dT}\frac{dT}{d\tau}) = mv_i \qquad (essendo\ \gamma_{i0} = 0). \tag{11.93}$$

Si tratta, per definizione, di un vettore spaziale a tre componenti, appartenente a $\Sigma_x(v_0 = 0, p_0 = 0)$. Si noti come, in base alle precedenti definizioni di 4-velocità relativa e di velocità relativa standard, p_α possa mettersi nella forma classica mv_α, a patto di introdurre la seguente altra definizione.
Dicesi *Massa relativa standard*

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{11.94}$$

La relazione tra massa relativa e massa propria si presenta così formalmente, la stessa che in Relatività Ristretta.
Definiremo infine:
Energia materiale relativa standard

$$E = cu_iP^i = m_0cu_i\frac{dx^i}{d\tau} = m_0c^2\frac{dT}{d\tau} = mc^2. \tag{11.95}$$

Ad E competono evidentemente anche le seguenti altre espressioni:

$$E = cu^iP_i = cu^0P_0 = -c\frac{P_0}{u_0}. \tag{11.96}$$

Si osservi che, essendo $\mathcal{P}_\Theta(P_i) = \frac{E}{c}u_i$, la quantità $\frac{E}{c}$ non differisce dalla proiezione ortogonale di P_i secondo $\mathbf{u}$. L'ultima espressione che compare in (11.95) mette poi in evidenza, proprio come in RR, l'identificazione sostanziale della massa (relativa) e dell'energia (relativa) materiale. Conviene rimarcare esplicitamente come tutte le grandezze introdotte abbiano carattere *invariantivo* di fronte ad un generico cambiamento di coordinate *interno* al riferimento fisico adottato.

11.11 Concetto di simultaneità in Relatività Generale

Nella Teoria della Relatività Ristretta abbiamo dimostrato che è possibile introdurre il concetto di **simultaneità** di due eventi in maniera globale e non ambigua in ogni sistema di riferimento inerziale, previa sincronizzazione degli orologi situati in ciascun punto del riferimento, ed abbiamo constatato che tale concetto non è assoluto ma è relativo ad ogni sistema di riferimento inerziale. In Relatività Generale non è generalmente possibile introdurre una nozione analoga in punti arbitrari di un qualunque sistema di riferimento $\mathcal{R}$: qui la nozione di **simultaneità** è un po' più delicata e complessa. Noi introduciamo questo concetto in un modo che riteniamo piuttosto semplice ed intuitivo, utilizzando la nozione di *intervallo di tempo relativo standard* di Cattaneo e seguendo in parte Møller.[41]

[41]Molti dei testi moderni sulla Relatività Generale, stranamente ed inspiegabilmente, non fanno cenno al *concetto di simultaneità*. A nostra conoscenza, ci sono almeno tre trattati che se ne occupano seriamente sono: 1) il già, più volte citato, Landau-Lifschitz, *Teoria dei campi*, Editori Riuniti, Roma 2010, §84, pag. 310-311; 2) Christian Møller, *The Theory of Relativty*, The International Series of Monogaphs on Physics, Second Edition, Delhi, Oxford University Press, 1972, First Indian Impression 1974, Sixth impression 1990, paragrafi 9.16, 10.2, in particolare pagg. 376, 377, 378; 3) Henry Arzeliès, *Relativité Généralisée, Gravitation*, Fascicule I, Chap.IV, pag. 89, Gauthier- Villars, Paris, 1961.

Se definiamo entro il riferimento $\mathcal{R}$ un ben determinato sistema di coordinate, potremmo formalmente dire, in maniera ingenua, che due eventi sono simultanei quando essi hanno lo stesso valore della variabile temporale t. Tuttavia questa "*simultaneità coordinata*" ha differenti significati in diversi sistemi di coordinate entro $\mathcal{R}$.
Intuitivamente, sembrerebbe più soddisfacente se fosse possibile definire la *"simultaneità fra due eventi"* in maniera tale da dipendere soltanto dal sistema di riferimento. Per eventi che accadono in punti infinitamente vicini ciò è effettivamente possibile utilizzando il concetto di **intervallo di tempo relativo standard** che abbiamo introdotto nel paragrafo precedente. A tal fine poniamo la seguente:
Definizione: *Diciamo che due eventi P e P' di coordinate spaziotemporali (x^i) e $(x^i + dx^i)$ rispettivamente, sono* **simultanei**, *nel dato sistema di riferimento, se è nullo l'intervallo di tempo relativo standard che li separa, cioè:*

$$dT = 0. \tag{11.97}$$

Poiché, come sappiamo, $dT = \frac{1}{c} u_i dx^i$ è invariante di fronte ad ogni cambiamento di coordinate che non alteri il campo di vettori $\mathbf{u}(x)$, che sia cioè interno al riferimento fisico prefissato, la simultaneità definita in questo modo, che chiameremo **"simultaneità standard"**, dipende soltanto dal sistema di riferimento.
Tuttavia, se si tenta di estendere questa definizione ad eventi che sono spazialmente distanti unendoli da una curva e usando per ogni sua porzione infinitesima l'espressione $dT = \frac{1}{c} u_i dx^i = 0$, si trova che la simultaneità ottenuta in questo modo viene a dipendere dalla curva che connette i due eventi. D'altra parte dT generalmente non è un differenziale esatto, cioè la forma differenziale lineare $u_i dx^i$ non è in generale integrabile, il che significa che dT non è il differenziale totale di una variabile temporale T avente carattere globale. Perciò in un qualsiasi sistema di riferimento non è possibile definire globalmente la simultaneità standard per una coppia arbitraria di eventi.
Vediamo ora come si può esprimere in termini dei differenziali delle coordinate la relazione $dT = 0$ che definisce la simultaneità standard, avremo

$$u_i dx^i = u_0 dx^0 + u_\alpha dx^\alpha = 0,$$

e tenendo conto delle (11.63) si ottiene:

$$dx^0 = -\frac{g_{0\alpha}}{g_{00}} dx^\alpha. \tag{11.98}$$

Questo risultato coincide esattamente con la definizione di simultaneità data da Landau-Lifschitz, utilizzando argomentazioni completamente diverse, che riproduciamo qui sotto per comodità del lettore.

Consideriamo, come nel paragrafo 11.8, il processo di propagazione di un segnale luminoso fra due punti infinitamente vicini A e B, rappresentato con un diagramma spaziotemporale nella Fig.11.5. Precisamente supponiamo che un segnale luminoso venga emesso da un punto B dello spazio di coordinate $x^\alpha + dx^\alpha$, all'istante $x^0 + (dx^0)_1$, venga ricevuto da A all'istante x^0, e riflesso istantaneamente lungo lo stesso cammino e ricevuto da B all'istante $x^0 + (dx^0)_2$.
Landau-Lifschitz scrivono:
"*Bisogna considerare come* **simultaneo** *con l'istante x^0 nel punto A l'indicazione dell'orologio in B nell'istante tra la partenza e il ritorno del segnale in questo punto, cioè l'istante*

$$x^0 + dx^0 = x^0 + \frac{1}{2}[(dx^0)_1 + (dx^0)_2]. \tag{11.99}$$

Figura 11.5: Simultaneità di due eventi in Relatività Generale.

Sostituendo, nella precedente relazione le (11.40)*, troviamo la differenza dei valori del "tempo x^0" per due eventi simultanei aventi luogo in due punti infinitamente vicini nella forma"*

$$dx^0 = -\frac{g_{0\alpha}}{g_{00}} dx^\alpha. \tag{11.100}$$

In altri termini:
Si deve considerare come **simultaneo** *con l'istante x^0 segnato dall'orologio in A l'indicazione dell'orologio in B, diciamola x^0_*, definita dalla semisomma tra l'istante di partenza e l'istante di ritorno del segnale in B:*

$$x^0_* = \frac{1}{2}\{[x^0 + (dx^0)_1] + [x^0 + (dx^0)_2]\} = x^0 + \frac{1}{2}[(dx^0)_1 + (dx^0)_2]. \tag{11.101}$$

Come si vede la (11.100), ottenuta da Landau-Lifschitz, come conseguenza della definizione da loro data di simultaneità, coincide esattamente con la (11.98) che abbiamo più sopra ottenuto come conseguenza immediata della nostra definizione di *simultaneità standard*.
Perciò la (11.100) è la relazione che permette di **sincronizzare** gli orologi in qualsiasi volume infinitesimo di spazio. Continuando una tale sincronizzazione con altri punti, si possono sincronizzare gli orologi, cioè determinare la simultaneità di eventi, lungo una qualsiasi linea non chiusa. Per quanto riguarda la sincronizzazione degli orologi lungo un contorno chiuso, essa è, in generale, impossibile. Infatti descrivendo il contorno e ritornando nel punto di partenza, si otterrebbe per dx^0 un valore diverso da zero. A maggior ragione è impossibile una sincronizzazione univoca degli orologi in tutto lo spazio. Fanno eccezione i riferimenti fisici che ammettono un sistema di coordinate adattate in cui sono nulle tutte le componenti $g_{0\alpha}$ ($g_{0\alpha} = 0$) della metrica.[42] Questi sistemi si dicono **riferimenti tempo-ortogonali** perché sono tali che le linee orarie $x^0 = var.$ sono ortogonali alle varietà $x^0 = cost.$ In tal caso, si dice che la congruenza Γ, che definisce il riferimento fisico, è una **congruenza normale**, cioè tale da ammettere una famiglia di ∞^1 ipersuperficie le cui traiettorie ortogonali coincidono con le linee di Γ.
È necessario sottolineare che l'impossibilità della sincronizzazione di tutti gli orologi è una proprietà dei sistemi di riferimento arbitrari, e non dello spaziotempo come tale. In qualsiasi campo gravitazionale si può sempre scegliere (e perfino in infiniti modi) un sistema di riferimento in modo tale da annullare identicamente i tre coefficienti $g_{0\alpha}$ della metrica e, di conseguenza, da rendere possibile la sincronizzzione totale degli orologi.[43]

[42]Sono da includere qui anche i casi in cui le $g_{0\alpha}$ si possono annullare con una semplice trasformazione della coordinata temporale, a prescindere dalla scelta del sistema di oggetti che servono alla determinazione delle coordinate spaziali.
[43]Si veda Landau-Lifschitz, *Teoria dei campi*, Editori Riuniti, Roma 2010, § 97, pag. 375.

Già in Relatività Ristretta lo scorrere del tempo reale è differente per orologi in moto relativo. In Relatività Generale, invece, il tempo reale scorre in modo differente nei diversi punti dello spazio anche in uno stesso sistema di riferimento. Ciò vuol dire che l'intervallo di tempo proprio tra due eventi che accadono nello stesso punto dello spazio e l'intervallo di tempo proprio tra due eventi contemporanei ai primi in un altro punto dello spazio sono, in generale, differenti l'uno dall'altro.
Osservazione 1:
Un'altro procedimento di sincronizzazione che porta allo stesso risultato dei due precedenti è il seguente.[44]
Si sceglie preliminarmente un'origine comune per la regolazione degli orologi procedendo così: siano ancora A e B i due punti infinitamente vicini, considerati in precedenza, in cui noi collochiamo due orologi *qualunque.* Un segnale luminoso effettua il percorso di andata e ritorno BAB. Prendiamo come *evento origine per i due orologi* il passaggio del segnale nel punto A. Per questo passaggio l'orologio $\mathcal{O}_A$ *segna zero per convenzione.* Da parte sua l'orologio $\mathcal{O}_B$ segna:

$$\text{Partenza da B}: \quad (dx^0)_1;$$

$$\text{Ritorno in B}: \quad (dx^0)_2;$$

$(dx^0)_1$ e $(dx^0)_2$ essendo rispettivamente i tempi di andata e ritorno del segnale luminoso calcolati nel paragrafo 11.8 a partire dal $\mathbf{ds}^2$ supposto conosciuto.
Per dare un senso alla **nozione di simultaneità** di avvenimenti che si producono nei punti distinti A e B dello spazio, bisogna definire l'istante letto in B di un evento che accade in A.

Figura 11.6: Ancora sul concetto di simultaneità di due eventi in Relatività Generale.

Nel diagramma spazio-temporale di Fig.11.6, le linee orarie di A e B sono rappresentate da due rette parallele a tratto pieno, mentre le linee orarie dei segnali luminosi sono rappresentate da linee tratteggiate oblique. Si vede chiaramente che nel tempo che il segnale luminoso passa da B ad A, la sua linea oraria interseca la linea oraria di A nel punto rappresentativo $\mathcal{A}_0$, il cui orologio segna $x^0 = 0$. Il punto rappresentativo di B, non rimane in quiete come il punto B ma, siccome il tempo scorre, esso si muove sulla sua linea oraria verso l'alto raggiungendo la posizione $\mathcal{B}_*$, che è punto medio del segmento avente per estremi i punti rappresentativi $\mathcal{B}_1 \equiv (dx^0)_1$, $\mathcal{B}_2 \equiv (dx^0)_2$.
Dato che i percorsi di andata e ritorno del segnale luminoso sono del tutto simmetrici quando il segnale raggiunge $\mathcal{A}_0$ il punto rappresentativo di B raggiunge il punto medio $\mathcal{B}_*$ del segmento di linea oraria $\mathcal{B}_1\mathcal{B}_2$. Invece quando il segnale di ritorno raggiunge B, il suo punto rappresentativo

[44]Nelle considerazioni che seguono abbiamo adottato il punto di vista di Henry Arzeliès, *Relativité Généralisée, Gravitation*, Fascicule I, pag. 89, Gauthier-Villars, Paris, 1961.

raggiunge, sulla sua linea oraria, la posizione $\mathcal{B}_2$. Se allora teniamo conto del fatto che le coordinate del punto medio di un segmento di cui si conoscono le coordinate degli estremi, sono dati dalla semisomma di queste ultime, possiamo concludere dicendo che:
Per definizione, noi poniamo che il passaggio del segnale in A (il cui orologio $\mathcal{O}_A$ segna lo zero) si effettua, per l'orologio $\mathcal{O}_B$, all'istante x_*^0 dato dalla semisomma dei tempi di andata e ritorno del segnale in B:

$$x_*^0 = \frac{1}{2}[(dx^0)_1 + (dx^0)_2]. \tag{11.102}$$

Ciò *definisce* per l'orologio $\mathcal{O}_B$ *la simultaneità di eventi che si producono in punti infinitamente vicini.* Con questa *convenzione*, del tutto naturale, la scelta di un evento origine non permette in generale d'attribuire la lettura zero ai due orologi. Se uno degli orologi, per esempio $\mathcal{O}_A$ segna zero per convenzione, noi dobbiamo attribuire all'altro, nel nostro caso $\mathcal{O}_B$, la lettura (11.102), che tenendo conto delle (11.40), possiamo scrivere così:

$$x_*^0 = \frac{1}{2}[(dx^0)_1 + (dx^0)_2] = -\frac{g_{0\alpha}}{g_{00}}dx^\alpha, \tag{11.103}$$

cioè ancora la (11.100).
Se il campo di gravitazione è costante, questa quantità, per due dati punti A e B infinitamente vicini, è indipendente dal tempo; *essa è nulla in un riferimento tempo-ortogonale* in quanto in tale sistema si ha $g_{0\alpha} = 0$. Si vede dunque che, *per un evento scelto come origine, le indicazioni dei diversi orologi saranno differenti; noi non possiamo dunque attribuire l'istante zero che ad uno solo fra essi.*
Osservazione 2:
In un riferimento galileiano $\mathcal{K}_0$ (Relatività Ristretta) un dato evento si produce allo stesso istante per tutti gli orologi di $\mathcal{K}_0$, poiché questi orologi sono sincronizzati. È dunque inutile precisare a quale orologio di $\mathcal{K}_0$ ci si riferisce. Nella maggior parte dei ragionamenti si utilizza l'orologio locale, vale a dire l'orologio fisso nel punto in cui si verifica l'evento; ma ciò non è per niente necessario, la scelta dell'orologio essendo indifferente.
In un riferimento arbitrario $\mathcal{R}$, le cose si presentano meno semplicemente. L'ora di un evento differisce a seconda dell'orologio di riferimento, perché gli orologi di $\mathcal{R}$ non sono sincronizzati. *In assenza di qualcosa di meglio, noi assumiamo sempre che l'ora è letta su di un orologio locale.* Ma allora, in generale, *dire che due eventi sono simultanei non significa che gli orologi situati nei luoghi dove essi si verificano misurino lo stesso tempo!* In altri termini, generalmente, **due eventi simultanei non si producono alla stessa ora!**

11.12 Velocità della luce e velocità di una particella in Relatività Generale

In un generico riferimento $\mathcal{R}$, consideriamo un segnale luminoso che parte da un punto $A(x^\alpha)$ dello spazio 3-dimensionale all'istante t, ed arriva al punto infinitamente vicino $B(x^\alpha + dx^\alpha)$ all'istante $t + dt$, questi istanti essendo letti su due orologi situati rispettivamente in A e B. La legge di propagazione della luce, come si è detto più volte, è definita dalla relazione:

$$ds^2 = g_{ik}dx^i dx^k = 0. \tag{11.104}$$

Riscriviamola sotto la forma (11.39), cioè:

$$g_{00}(dx^0)^2 + 2g_{0\alpha}dx^0 dx^\alpha + g_{\alpha\beta}dx^\alpha dx^\beta = 0. \tag{11.105}$$

Facendo ora intervenire l'intervallo spaziale dato dalla (11.57), tenuto conto della (11.56), la precedente equazione diventa:

$$g_{00}(dx^0)^2 + 2\sqrt{g_{00}}g_\alpha dx^0 dx^\alpha - \gamma_{\alpha\beta}dx^\alpha dx^\beta + g_\alpha g_\beta dx^\alpha dx^\beta = 0, \tag{11.106}$$

che possiamo riscrivere così:

$$g_{00}(dx^0)^2 + 2\sqrt{g_{00}}g_\alpha dx^0 dx^\alpha - d\sigma^2 + (g_\alpha dx^\alpha)^2 = 0, \tag{11.107}$$

Definiamo ora la **velocità spaziale della luce (ovvero di un fotone)**, nel riferimento $\mathcal{R}$, per mezzo delle sue componenti controvarianti:

$$v_l^\alpha = \frac{dx^\alpha}{dt}, \tag{11.108}$$

il suo modulo sarà:

$$v_l = \frac{d\sigma}{dt} = \sqrt{\gamma_{\alpha\beta} v_l^\alpha v_l^\beta}. \tag{11.109}$$

Si ha dunque che: *In un riferimento arbitrario, la velocità di un segnale luminoso dipende in generale dal punto in cui si considera ed in questo punto dalla direzione di propagazione.*
Per ottenere la sua espressione, dividiamo ambo i membri della (11.107) per dt^2 e teniamo conto delle (11.108) e (11.109), si avrà:

$$g_{00}c^2 + 2c\sqrt{g_{00}}g_\alpha v_l^\alpha - v_l^2 + (g_\alpha v_l^\alpha)^2 = 0, \tag{11.110}$$

perciò:

$$v_l^2 = (c\sqrt{g_{00}} + g_\alpha v_l^\alpha)^2. \tag{11.111}$$

La precedente equazione ammette due soluzioni opposte delle quali quella col segno meno è da scartare perché fisicamente inammissibile, perciò l'unica soluzione è:

$$v_l = c\sqrt{g_{00}} + g_\alpha v_l^\alpha. \tag{11.112}$$

Come si vede, il modulo della velocità di propagazione della luce, dipende dalla sua direzione tramite le componenti controvarianti v_l^α. Una direzione qualunque è definita dai suoi parametri direttori

$$\theta^\alpha = \frac{dx^\alpha}{d\sigma}, \qquad \gamma_{\alpha\beta}\theta^\alpha\theta^\beta = 1. \tag{11.113}$$

possiamo anche scrivere

$$\theta^\alpha = \frac{dx^\alpha}{d\sigma} = \frac{dx^\alpha}{dt}\frac{dt}{d\sigma} = \frac{\frac{dx^\alpha}{dt}}{\frac{d\sigma}{dt}} = \frac{v_l^\alpha}{v_l}, \tag{11.114}$$

da cui la relazione:

$$v_l = \frac{v_l^\alpha}{\theta^\alpha}, \tag{11.115}$$

che ci fornisce il modulo della velocità secondo la direzione θ^α in funzione delle sue componenti e dei parametri direttori. Per la (11.115), la (11.112) può essere trasformata in quest'altra formula:

$$v_l = \frac{c\sqrt{g_{00}}}{1 - g_\alpha\theta^\alpha} = \frac{cg_{00}}{\sqrt{g_{00}} - g_{0\alpha}\theta^\alpha}, \tag{11.116}$$

che fornisce la velocità della luce in funzione dei coefficienti g_{00} e $g_{0\alpha}$ della metrica, e dei parametri direttori della sua direzione di propagazione. Nel caso particolare in cui $g_{0\alpha} = 0 \Leftrightarrow g_\alpha = 0$ si avrà la formula notevolmente semplificata:

$$v_{luce} = c\sqrt{g_{00}}. \tag{11.117}$$

Si può dimostrare che ciò si verifica nel caso in cui la congruenza Γ, del genere tempo, costituente il riferimento fisico non è una generica congruenza ma una **congruenza normale** Γ_n, cioè tale

da ammettere una famiglia ∞^1 di ipersuperfici le cui traiettorie ortogonali coincidono con le linee di Γ_n. In tal caso, il particolare riferimento fisico definito dalla Γ_n è un **riferimento tempo-ortogonale** la cui definizione è stata data nel paragrafo precedente.[45]
Un ragionamento analogo a quello fatto più sopra si può ripetere per determinare la velocità di una qualsiasi particella, in tal caso però il divario spaziotemporale è del genere tempo e perciò $ds^2 > 0$. Indicando con v_p la velocità della particella, si può ripetere il ragionamento precedente sostituendo nella (11.105) il segno $=$ col segno $>$. Si otterrà.

$$v_p < c\sqrt{g_{00}} + g_\alpha v_p^\alpha. \tag{11.118}$$

11.13 Moto di una particella liberamente gravitante

Ci occupiamo ancora, data l'importanza dell'argomento, del moto di una particella soggetta soltanto all'azione di un campo gravitazionale.
L'equazione differenziale del moto di tale particella, la (11.22), è già stata ricavata in maniera piuttosto formale nel paragrafo 11.5, applicando il *Principio di Trascrizione*. Ora vogliamo ricavarla in modo diretto come una istruttiva applicazione del *Principio di Equivalenza*.
Consideriamo, a tal fine, una particella che si muova liberamente sotto l'azione di forze puramente gravitazionali. Secondo il **Principio di equivalenza** (punti *III e IV)* esiste un sistema di coordinate ξ^i liberamente gravitante *nel quale è assente ogni effetto della gravità e perciò equivalente ad un sistema inerziale* nel quale le equazioni di moto della particella sono quelle di una retta (*linea oraria*) nello spaziotempo di Minkowski $\mathcal{M}^4$ cioè:

$$\frac{dU^i}{d\tau} = 0 \Longleftrightarrow \frac{d^2\xi^i}{d\tau^2} = 0, \qquad (i = 0, 1, 2, 3) \tag{11.119}$$

dove $U^i = \dfrac{d\xi^i}{d\tau}$ è la 4-velocità della particella, $d\tau = \dfrac{ds}{c}$ è l'intervallo elementare di tempo proprio e

$$\mathbf{ds}^2 = \eta_{hk} d\xi^h \xi^k \qquad (h, k = 0, 1, 2, 3) \tag{11.120}$$

è *il tensore metrico dello spaziotempo di Minkowski.*
Supponiamo ora di usare un altro sistema di coordinate x^i, che può essere un sistema cartesiano in quiete nel laboratorio, o più generalmente un sistema di coordinate curvilinee, oppure accelerato, rotante o ciò che si vuole. Le coordinate ξ^i del sistema liberamente gravitante saranno funzioni invertibili delle x^i ed il 1° membro dell'equazione (11.119) diventa

$$\frac{d}{d\tau}\left(\frac{d\xi^i}{d\tau}\right) = \frac{d}{d\tau}\left(\frac{\partial\xi^i}{\partial x^k}\frac{dx^k}{d\tau}\right) = \frac{\partial\xi^i}{\partial x^k}\frac{d^2x^k}{d\tau^2} + \frac{\partial^2\xi^i}{\partial x^h \partial x^k}\frac{dx^h}{d\tau}\frac{dx^k}{d\tau}.$$

Perciò la (11.119) si trasforma nella

$$\frac{\partial\xi^i}{\partial x^k}\frac{d^2x^k}{d\tau^2} + \frac{\partial^2\xi^i}{\partial x^h \partial x^k}\frac{dx^h}{d\tau}\frac{dx^k}{d\tau} = 0,$$

moltiplicando quest'ultima, membro a membro, per $\dfrac{\partial x^l}{\partial\xi^i}$ e sommando rispetto ad i, avremo

$$\frac{\partial x^l}{\partial\xi^i}\frac{\partial\xi^i}{\partial x^k}\frac{d^2x^k}{d\tau^2} + \frac{\partial x^l}{\partial\xi^i}\frac{\partial^2\xi^i}{\partial x^h \partial x^k}\frac{dx^h}{d\tau}\frac{dx^k}{d\tau} = 0,$$

[45] Il lettore desideroso di approfondire lo studio delle congruenze normali e dei *sistemi tempo-ortogonali* può consultare utilmente il libro, già citato, del Prof. Carlo Cattaneo, *Introduzione alla teoria einsteiniana della gravitazione*, vol. I, pag. 172-180, Università degli Studi di Roma, Libreria Eredi Virgilio Veschi, Roma, 1961.

e siccome $\dfrac{\partial x^l}{\partial \xi^i}\dfrac{\partial \xi^i}{\partial x^k} = \delta^l_k$, la precedente equazione diventa:

$$\frac{d^2x^l}{d\tau^2} + \frac{\partial x^l}{\partial \xi^i}\frac{\partial^2 \xi^i}{\partial x^h \partial x^k}\frac{dx^h}{d\tau}\frac{dx^k}{d\tau} = 0, \qquad (i,l,h,k = 0,1,2,3), \tag{11.121}$$

cioè:

$$\frac{d^2x^l}{d\tau^2} + \gamma^l_{hk}\frac{dx^h}{d\tau}\frac{dx^k}{d\tau} = 0, \qquad (l,h,k = 0,1,2,3), \tag{11.122}$$

avendo posto:

$$\gamma^l_{hk} \stackrel{\text{def}}{=} \frac{\partial x^l}{\partial \xi^i}\frac{\partial^2 \xi^i}{\partial x^h \partial x^k}, \quad (\gamma^l_{hk} = \gamma^l_{kh}) \qquad (l,h,k = 0,1,2,3) \tag{11.123}$$

che chiameremo *connessione affine* relativa alla trasformazione di coordinate $\xi \to x$; si osservi che tale connessione è *simmetrica*.
Nella (11.122) il $d\tau$ può essere sostituito da ds, in quanto è $ds = cd\tau$ ed il fattore c a calcoli fatti non figura più nell'equazione finale. Perciò la (11.122) è equivalente alla

$$\frac{d^2x^l}{ds^2} + \gamma^l_{hk}\frac{dx^h}{ds}\frac{dx^k}{ds} = 0 \qquad (l,h,k = 0,1,2,3). \tag{11.124}$$

Osserviamo che l'equazione (11.124) ha la stessa struttura formale dell'equazione differenziale (10.111) delle geodetiche di una varietà riemanniana

$$\frac{d^2x^l}{ds^2} + \Gamma^l_{hk}\frac{dx^h}{ds}\frac{dx^k}{ds} = 0, \qquad (l,h,k = 0,1,2,3), \tag{11.125}$$

ma ne differisce per il fatto che in quest'ultima compaiono i simboli di Christoffel di 2^a specie Γ^i_{hk} relativi alla connessione di Levi-Civita proveniente da un opportuno tensore metrico g_{ik}, mentre nella (11.124) compaiono i simboli γ^i_{hk} della connessione affine definita poc'anzi dalla (11.123). Che relazione c'è fra le due connessioni? Non è difficile dimostrare che in effetti esse coincidono, cioè risulta

$$\gamma^i_{hk} = \Gamma^i_{hk}, \qquad (l,h,k = 0,1,2,3), \tag{11.126}$$

purché il tensore metrico che definisce le Γ^i_{hk} provenga dal $\mathbf{ds}^2$ dato dalla (11.120) tenuto conto della trasformazione delle coordinate $\xi \to x$ come ora vedremo.
Infatti la metrica (11.120) può essere espressa per mezzo delle nuove coordinate x^i, scrivendo i differenziali delle vecchie coordinate ξ^i per mezzo delle nuove, cioè: $d\xi^h = \dfrac{\partial \xi^h}{\partial x^i}dx^i$, $d\xi^k = \dfrac{\partial \xi^k}{\partial x^j}dx^j$. Queste espressioni sostituite nella (11.120) ci danno:

$$\mathbf{ds}^2 = \eta_{hk}d\xi^h d\xi^k = \eta_{hk}\frac{\partial \xi^h}{\partial x^i}dx^i\frac{\partial \xi^k}{\partial x^j}dx^j = \eta_{hk}\frac{\partial \xi^h}{\partial x^i}\frac{\partial \xi^k}{\partial x^j}dx^i dx^j$$

ovvero

$$\mathbf{ds}^2 = g_{ij}(x)dx^i\,dx^j, \tag{11.127}$$

dove $g_{ij}(x)$ è il tensore metrico dello spaziotempo nelle nuove coordinate definito da

$$g_{ij}(x) \stackrel{\text{def}}{=} \eta_{hk}\frac{\partial \xi^h}{\partial x^i}\frac{\partial \xi^k}{\partial x^j}, \qquad (g_{ij} = g_{ji}). \tag{11.128}$$

Se ora moltiplichiamo la (11.123) membro a membro per $\dfrac{\partial \xi^m}{\partial x^l}$ e sommiamo rispetto all'indice l, avremo:

$$\frac{\partial \xi^m}{\partial x^l}\gamma^l_{hk} = \frac{\partial \xi^m}{\partial x^l}\frac{\partial x^l}{\partial \xi^i}\frac{\partial^2 \xi^i}{\partial x^h \partial x^k} = \delta^m_i\frac{\partial^2 \xi^i}{\partial x^h \partial x^k} = \frac{\partial^2 \xi^m}{\partial x^h \partial x^k},$$

che riscriviamo così:

$$\frac{\partial^2 \xi^m}{\partial x^h \partial x^k} = \frac{\partial \xi^m}{\partial x^l} \gamma^l_{hk}. \tag{11.129}$$

Deriviamo ora membro a membro la metrica (11.128) rispetto ad x^m:

$$\partial_m g_{ij} = \eta_{hk} \frac{\partial^2 \xi^h}{\partial x^m \partial x^i} \frac{\partial \xi^k}{\partial x^j} + \eta_{hk} \frac{\partial \xi^h}{\partial x^i} \frac{\partial^2 \xi^k}{\partial x^m \partial x^j} = \eta_{hk} \frac{\partial \xi^h}{\partial x^l} \gamma^l_{mi} \frac{\partial \xi^k}{\partial x^j} + \eta_{hk} \frac{\partial \xi^h}{\partial x^i} \frac{\partial \xi^k}{\partial x^l} \gamma^l_{mj}$$

cioè

$$\partial_m g_{ij} = g_{lj} \gamma^l_{mi} + g_{il} \gamma^l_{mj} = \gamma_{mi,j} + \gamma_{mj,i} \tag{11.130}$$

relazione simile all'identità di Ricci. Se sulla (11.130) permutiamo circolarmente gli indici m, i, j otteniamo le seguenti tre identità

$$\partial_m g_{ij} = \gamma_{mi,j} + \gamma_{mj,i}$$

$$\partial_i g_{jm} = \gamma_{ij,m} + \gamma_{im,j}$$

$$\partial_j g_{mi} = \gamma_{jm,i} + \gamma_{ji,m}.$$

Sommando le prime due relazioni ottenute e sottraendo la terza, con un calcolo esattamente analogo a quello fatto a suo tempo per dimostrare la 1ª formula di Christoffel, si ottiene:

$$\gamma_{mi,j} \equiv \frac{\partial x^l}{\partial \xi^i} \frac{\partial^2 \xi^i}{\partial x^h \partial x^k} = \frac{1}{2}(\partial_h g_{ki} + \partial_k g_{hi} - \partial_i g_{hk}),$$

ciò che comporta

$$\gamma^l_{hk} = \frac{1}{2} g^{li} (\partial_h g_{ki} + \partial_k g_{hi} - \partial_i g_{hk}),$$

Ma questa è proprio la connessione di Levi-Civita relativa alla metrica (11.128), possiamo perciò concludere che

$$\gamma^l_{hk} \equiv \Gamma^l_{hk}, \tag{11.131}$$

e dunque la (11.124) si identifica con l'equazione differenziale delle geodetiche (11.125) della varietà riemanniana (spaziotempo) di metrica (11.128). c.d.d.
Possiamo perciò concludere che:
Una particella liberamente gravitante si muove su una geodetica dello spaziotempo $\mathcal{V}^4$.
La precedente trattazione del moto di una particella in caduta libera ha dunque mostrato che **chi determina "la forza gravitazionale" è la connessione** Γ^l_{hk}, e siccome, per la formula precedente, essa si esprime per mezzo delle derivate delle componenti g_{ij} del tensore metrico, è lecito chiamare queste funzioni **potenziali gravitazionali**, in analogia con la terminologia usata nella teoria newtoniana nella quale le componenti cartesiane di un campo di forze conservative sono date dalle derivate parziali del corrispondente potenziale.
Osservazione 1:
Dobbiamo osservare che le Γ^l_{hk} rappresentano la connessione metrica di g_{ij} che è ancora una metrica piatta ma non la metrica di Minkowski η_{hk}. I termini aggiuntivi, rispetto alla (11.119), che appaiono nella (11.125) e contengono le Γ^l_{hk} sono interpretabili come i termini che caratterizzano la **"forza inerziale" o "forza apparente"** che si manifesta in un sistema di riferimento non inerziale. Allora il *Principio di Equivalenza* richiede che le forze gravitazionali, come le forze inerziali siano date da una appropriata connessione Γ^l_{hk}. In questo caso non possiamo più mantenere uno spaziotempo piatto, perché altrimenti non ci sarebbe distinzione dal caso non gravitazionale. La generalizzazione **più semplice** è quella di ritenere le Γ^l_{hk} come coefficienti della connessione metrica **di una metrica non piatta**.
Osservazione 2:

Per una particella di **massa nulla**, come il **fotone**,[46] le equazioni di moto rispetto ad un sistema liberamente gravitante (cioè in caduta libera) sono le stesse della (11.119) eccetto per il fatto che come variabile indipendente non è più possibile scegliere il tempo proprio τ, perché per particelle prive di massa, il 1° membro (e quindi anche il 2°) della (11.120) è uguale a zero.[47] Al posto di τ (o di s) possiamo usare come parametro un parametro arbitrario, per es. ξ^0, che possiamo indicare con σ, $(\sigma \equiv \xi^0)$, cosicché le (11.119) e (11.120) diventano rispettivamente

$$\frac{d^2\xi^i}{d\sigma^2} = 0,$$

$$\eta_{hk}\frac{d\xi^h}{d\sigma}\frac{d\xi^k}{d\sigma} = 0.$$

Seguendo lo stesso ragionamento fatto in precedenza, troviamo che le equazioni di moto di un fotone in un arbitrario campo gravitazionale ed in un arbitrario sistema di coordinate sono:

$$\frac{d^2x^l}{d\sigma^2} + \Gamma^l_{hk}\frac{dx^h}{d\sigma}\frac{dx^k}{d\sigma} = 0, \tag{11.132}$$

con

$$g_{ij}\frac{dx^i}{d\sigma}\frac{dx^j}{d\sigma} = 0, \tag{11.133}$$

dove Γ^l_{hk} e g_{ij} sono dati, come prima, rispettivamente dalle (11.123), (11.128).
Conclusione:
Concludendo, osserviamo che la nostra trattazione del moto di una particella in caduta libera ha mostrato che *il campo che determina la forza gravitazionale è la* **connessione affine** Γ^l_{hk}, *mentre, per ciò che abbiamo detto all'inizio del paragrafo 11.8, l'intervallo di tempo proprio fra due eventi infinitamente vicini è determinato dal coefficiente* $g_{00}(x)$ *del* **tensore metrico**.

11.14 Gravità come fenomeno metrico. Limite newtoniano per le geodetiche

In questo paragrafo ci proponiamo di dimostrare, attraverso un appropriato procedimento di approssimazione, che *l'effetto del campo gravitazionale può essere descritto usando le equazioni differenziali delle geodetiche di una particella liberamente gravitante e consentendo al tensore metrico di differire di poco dal tensore di Minkowski.* Il successo di questo procedimento ci servirà come giustificazione *per collegare la forza di gravità (come concetto fisico) con la natura non-Minkowskiana (riemanniana) dello spaziotempo e per usare le equazioni delle geodetiche come equazioni di moto per particelle libere in un campo gravitazionale.*
Poiché **"misure fisiche ordinarie"** *suggeriscono che la geometria del mondo reale è, con buona approssimazione, pressoché Euclidea, la metrica dello spaziotempo deve essere in effetti molto vicina a quella di Minkowski. Tuttavia è proprio questa differenza molto piccola fra la metrica reale e quella di Minkowski che noi desideriamo far vedere essere l'agente degli effetti gravitazionali.*
Perciò considereremo un tensore metrico **indipendente dal tempo** della forma

$$g_{ik}(x^\alpha) = \eta_{ik} + \epsilon h_{ik}(x^\alpha), \quad (i,k = 0,1,2,3), \quad (\alpha = 1,2,3) \tag{11.134}$$

[46] Per molto tempo, oltre al **fotone**, anche il **neutrino** è stato considerato come una particella priva di massa, ma le evidenze sperimentali più recenti dicono che si tratta di una particella di massa estremamente piccola ma non nulla. Anche i cosiddetti **gluoni** sono ritenuti particelle di massa nulla.

[47] Il 1° membro della (11.120), per particelle di **massa nulla** è uguale a zero, cioè $\mathbf{ds}^2 = 0$, perché tali particelle necessariamente viaggiano alla velocità della luce c, com'è noto dalla Relatività Ristretta.

dove η_{ik} è il tensore metrico di Minkowski, h_{ik} un tensore simmetrico arbitrario indipendente dal tempo, ϵ una costante molto piccola, tale che $|\epsilon h_{ik}| \ll 1$. Il tensore ϵh_{ik} rappresenta perciò una perturbazione molto piccola **indipendente dal tempo (approssimazione di campo debole)**, la quale è dovuta alla presenza di un **corpo gravitante**; inoltre ϵh_{ik} tende a zero allontanandosi dal corpo. Chiameremo questa metrica g_{ik} una metrica vicina a quella di Minkowski (o metrica quasi-Minkowskiana).
Per mostrare che **il termine ϵh_{ik} è veramente l'agente delle forze gravitazionali**, applicheremo le approssimazioni di cui sopra alle equazioni di moto delle geodetiche dello spazio riemanniano la cui metrica sia definita dal tensore espresso dalla (11.134). Successivamente per mostrare l'intima connessione con la teoria classica newtoniana, supporremo che il modulo v della velocità ordinaria della particella lungo la geodetica sia molto minore di c ($v \ll c$), o equivalentemente, che la velocità romeriana $\beta = \dfrac{v}{c}$ sia molto piccola ($\beta \ll 1$) (**basse velocità**); perciò nei nostri calcoli approssimati tratteniamo soltanto termini del 1° ordine in ϵ e β, trascurando termini del 2° ordine, quali $\epsilon^2, \beta^2, \epsilon\beta$, o di ordine superiore (**approssimazione di campo debole e basse velocità**). Usando le coordinate della RR ed il tensore metrico (11.134), otteniamo immediatamente la metrica (detta anche l'elemento di linea)[48]

$$\mathbf{ds}^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 + \epsilon h_{ik} dx^i dx^k, \quad (x^0 = ct). \tag{11.135}$$

Tenendo conto che $\dfrac{(dx^1)^2 + (dx^2)^2 + (dx^3)^2}{dt^2} = \dfrac{dl^2}{dt^2} = v^2$, è il quadrato della 3-velocità della particella, avremo:

$$\left(\frac{\mathbf{ds}}{dt}\right)^2 = c^2 - v^2 + \epsilon h_{ik} \frac{dx^i}{dt}\frac{dx^k}{dt} = c^2\left(1 - \beta^2 + \epsilon h_{ik}\frac{dx^i}{dx^0}\frac{dx^k}{dx^0}\right), \tag{11.136}$$

che possiamo riscrivere più esplicitamente

$$\left(\frac{\mathbf{ds}}{dt}\right)^2 = c^2\left(1 - \beta^2 + \epsilon h_{00} + 2\epsilon h_{0\alpha}\frac{dx^\alpha}{dx^0} + \epsilon h_{\alpha\beta}\frac{dx^\alpha}{dx^0}\frac{dx^\beta}{dx^0}\right). \tag{11.137}$$

Ricordando che le componenti v^α della velocità $\mathbf{v}$ di una particella sono definite come $v^\alpha = \dfrac{dx^\alpha}{dt} = c\dfrac{dx^\alpha}{dx^0}$ e perciò $\dfrac{dx^\alpha}{dx^0} = \dfrac{v^\alpha}{c} \lesssim \beta$ risulta plausibile **l'approssimazione di moto lento**

$$\frac{dx^\alpha}{dx^0} \simeq \beta \quad \Longleftrightarrow \quad dx^\alpha \simeq \beta dx^0. \tag{11.138}$$

Perciò al 1° ordine in ϵ e β, cioè trascurando termini del 2° ordine o di ordine superiore in ϵ e β, la (11.137) diventa

$$\left(\frac{\mathbf{ds}}{dt}\right)^2 \simeq c^2(1 + \epsilon h_{00}). \tag{11.139}$$

Ora applichiamo la stessa approssimazione alle equazioni differenziali delle geodetiche dello spaziotempo che, come sappiamo, sono le equazioni di moto di una particella liberamente gravitante, cioè le (11.125) del paragrafo 11.13 che, cambiando gli indici l in i, e h in j, possiamo riscrivere così:

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0 \qquad (i, h, k = 0, 1, 2, 3), \tag{11.140}$$

[48] Si badi bene che dire che **la metrica è indipendente dal tempo** significa che le componenti del tensore metrico sono indipendenti da x^0 cioè $g_{ik}(x^\alpha)$ ovvero $\partial_0 g_{ik} = 0$ (**campo stazionario**). Tutto ciò non significa però che nell'espressione (11.135) del $\mathbf{ds}^2$ manchino i termini contenenti il differenziale dx^0 della variabile temporale, in quanto in essa generalmente compaiono tutti e quattro i differenziali dx^0, dx^1, dx^2, dx^3 delle variabili spazio-temporali, a meno che non si facciano particolari ipotesi di simmetria sul corpo gravitante come si fa nella trattazione del problema di Schwarzschild che vedremo in seguito.

o più esplicitamente così:

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{00}\left(\frac{dx^0}{ds}\right)^2 + 2\Gamma^i_{0\alpha}\frac{dx^0}{ds}\frac{dx^\alpha}{ds} + \Gamma^i_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0 \qquad (i = 0, 1, 2, 3). \tag{11.141}$$

Tenuto conto che i simboli di Christoffel di 2ª specie si possono esprimere in funzione di quelli di 1^a specie e ricordando che la metrica g_{ik} è definita dalla (11.134) con il tensore di Minkowski η_{ik} costante nello spazio e nel tempo, avremo

$$\Gamma^i_{jk} = g^{im}\Gamma_{jk,m} \approx \frac{1}{2}g^{im}\left(\partial_j g_{km} + \partial_k g_{mj} - \partial_m g_{jk}\right) = \frac{1}{2}\left(\eta^{im} - \epsilon h^{im}\right)\epsilon\left(\partial_j h_{km} + \partial_k h_{mj} - \partial_m h_{jk}\right), \tag{11.142}$$

dove il tensore metrico sotto forma controvariante è $g^{im} = \eta^{im} - \epsilon h^{im}$.[49] Perciò tutte le Γ^i_{jk} contengono il fattore ϵ. D'altra parte i prodotti del tipo $\frac{dx^j}{ds}\frac{dx^k}{ds}$ che figurano nella (11.140), tenuto conto della (11.139), diventano

$$\frac{dx^j}{ds}\frac{dx^k}{ds} = \frac{dx^j}{dt}\frac{dx^k}{dt}\left(\frac{dt}{ds}\right)^2 \approx \frac{dx^j}{dt}\frac{dx^k}{dt}\frac{1}{c^2(1+\epsilon h_{00})}. \tag{11.143}$$

Se né j, né k è zero ($j \neq 0$, $k \neq 0$, $j = \alpha$, $k = \beta$), avremo:

$$\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt}\frac{1}{c^2(1+\epsilon h_{00})} = \frac{v^\alpha}{c}\frac{v^\beta}{c}\frac{1}{1+\epsilon h_{00}} \approx \beta^2\frac{1}{1+\epsilon h_{00}} \tag{11.144}$$

che è di ordine β^2, perciò trascurabile.
Se soltanto uno degli indici j o k è diverso da zero, per esempio $j \neq 0$ ($j = \alpha, k = 0, 1, 2, 3$) l'espressione (11.143) è o di ordine β se $k = 0$ o di ordine β^2 se $k \neq 0$, ed essendo, come detto più sopra, Γ^i_{jk} di ordine ϵ, il prodotto $\Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds}$ è di ordine $\epsilon\beta$ o superiore, e può essere trascurato nella nostra approssimazione. Le stesse considerazioni valgono se è solo $k \neq 0$ ($k = \alpha, j = 0, 1, 2, 3$).
Le equazioni differenziali (11.141) delle geodetiche, al 1° ordine in ϵ e β, sono allora:

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{00}\left(\frac{dx^0}{ds}\right)^2 = 0 \qquad (i = 0, 1, 2, 3). \tag{11.145}$$

Per quanto riguarda il 1° termine della (11.145), possiamo trasformarlo nel seguente modo:

$$\frac{d^2x^i}{ds^2} = \frac{d}{ds}\left(\frac{dx^i}{ds}\right) = \frac{d}{dt}\left(\frac{dx^i}{ds}\right)\frac{dt}{ds} = \frac{d}{dt}\left[\frac{dx^i}{dt}\frac{dt}{ds}\right]\frac{dt}{ds} =$$

$$= \left[\frac{d^2x^i}{dt^2}\frac{dt}{ds} + \frac{dx^i}{dt}\frac{d}{dt}\left(\frac{dt}{ds}\right)\right]\frac{dt}{ds} = \frac{d^2x^i}{dt^2}\left(\frac{dt}{ds}\right)^2,$$

[49] Avvertiamo il lettore che nell'approssimazione di campo debole l'innalzamento e l'abbassamento degli indici per enti tensoriali si effettua con il tensore metrico η_{ik} o η^{ik} e non con g_{ik} o g^{ik}. Siccome abbiamo definito $g_{ik} = \eta_{ik} + \epsilon h_{ik}$ con $|\epsilon h_{ik}| \ll 1$, ed essendo $\eta_{ik} = \eta^{ik}$, l'espressione del tensore metrico in forma controvariante sarà, nello stesso ordine di approssimazione, analoga a quella del tensore metrico covariante, cioè $\eta^{ik} + \epsilon\widetilde{h}^{ik}$ con $|\epsilon\widetilde{h}_{ik}| \ll 1$. Vogliamo però che, nell'approsimazione di campo debole, sia $g_{ik}g^{il} \approx \delta^l_k$, ma $g_{ik}g^{il} = (\eta_{ik} + \epsilon h_{ik})(\eta^{il} + \epsilon\widetilde{h}^{il}) = \delta^l_k + \epsilon\widetilde{h}^l_k + \epsilon h^l_k + \epsilon^2 h_{ik}\widetilde{h}^{il}$. L'ultimo termine è trascurabile perché del 2° ordine in ϵ, dunque deve essere $\widetilde{h}^l_k = -h^l_k$ affinché sia $g_{ik}g^{il} \approx \delta^l_k$. Se nella $\widetilde{h}^l_k = -h^l_k$ innalziamo l'indice inferiore, cambiando nome agli indici, avremo $\widetilde{h}^{ik} = -h^{ik}$, dunque si ha $g^{ik} = \eta^{ik} - \epsilon h^{ik}$. c.d.d.

essendo $\frac{d}{dt}\left(\frac{dt}{ds}\right) = 0$ perché, per la (11.139), $\frac{dt}{ds}$ è indipendente dal tempo e perciò è nulla la sua derivata temporale. Dunque, per ciò che precede, avremo

$$\frac{d^2x^i}{ds^2} \simeq \frac{d^2x^i}{dt^2}\frac{1}{c^2(1+\epsilon h_{00})}. \tag{11.146}$$

e l'equazione (11.145) delle geodetiche si riduce alla

$$\frac{d^2x^i}{dt^2}\frac{1}{c^2(1+\epsilon h_{00})} + \Gamma^i_{00}\frac{1}{1+\epsilon h_{00}} = 0 \qquad (i = 0,1,2,3).$$

ed in definitiva alla

$$\frac{d^2x^i}{dt^2} + c^2\Gamma^i_{00} = 0 \qquad (i = 0,1,2,3). \tag{11.147}$$

Osservazione:
Un procedimento alternativo a quello precedente per semplificare ulteriormente la (11.145) è il seguente.
Ricordiamo che $ds = cd\tau = c\sqrt{1-\beta^2}dt$, siccome trascuriamo termini del 2° ordine in β avremo $d\tau \simeq dt$ e perciò $ds \simeq cdt$. Con questa ulteriore approssimazione, essendo $dx^0 = cdt$, le due derivate che compaiono nella (11.145) diventano

$$\frac{d^2x^i}{ds^2} = \frac{d}{ds}\left(\frac{dx^i}{ds}\right) \simeq \frac{1}{c}\frac{d}{dt}\left(\frac{1}{c}\frac{dx^i}{dt}\right) = \frac{1}{c^2}\frac{d^2x^i}{dt^2}$$

e

$$\left(\frac{dx^0}{ds}\right)^2 \simeq 1$$

Perciò la (11.145) si riduce ancora alla (11.147).
Per semplificare ulteriormente la (11.147) calcoliamo i simboli di Christoffel Γ^i_{00} esplicitamente, abbiamo

$$\Gamma^i_{00} = g^{ir}\Gamma_{00,r}$$

dove

$$\Gamma_{00,r} = \frac{1}{2}\Big(\partial_0 g_{0r} + \partial_0 g_{r0} - \partial_r g_{00}\Big).$$

Poiché, per la nostra ipotesi di indipendenza della metrica dal tempo, g_{ik} è indipendente da x^0 ed η_{ik} è costante e perciò indipendente da tutte le variabili x^i, l'espressione precedente si riduce a

$$\Gamma_{00,r} = -\frac{1}{2}\partial_r g_{00} = -\frac{1}{2}\epsilon\partial_r h_{00}. \tag{11.148}$$

Innalzando l'indice r per ottenere i simboli di Christoffel di 2ª specie e trascurando i termini di ordine ϵ^2 o superiore, otteniamo.

$$\Gamma^i_{00} = g^{ir}\Gamma_{00,r} = g^{ir}\Big(-\frac{1}{2}\epsilon\partial_r h_{00}\Big) = \Big(\eta^{ir} - \epsilon h^{ir}\Big)\Big(-\frac{1}{2}\epsilon\partial_r h_{00}\Big) =$$

$$= -\frac{1}{2}\epsilon\,\eta^{ir}\partial_r h_{00} + \frac{1}{2}\epsilon^2 h^{ir}\partial_r h_{00},$$

cioè

$$\Gamma^i_{00} \simeq -\frac{1}{2}\epsilon\,\eta^{ir}\partial_r h_{00}. \tag{11.149}$$

Le (11.149) ci forniscono quattro simboli di Christoffel per $(i = 0, 1, 2, 3)$ che ora esamineremo uno alla volta.
Per $i = 0$ otteniamo:

$$\Gamma^0_{00} = -\frac{1}{2}\epsilon(\eta^{00}\partial_0 h_{00} + \eta^{0\alpha}\partial_\alpha h_{00}) = 0, \tag{11.150}$$

per l'assunzione di indipendenza della metrica dal tempo e per il fatto che $\eta^{0\alpha} = 0$, per $\alpha = 1, 2, 3$; inoltre è anche ovviamente $\frac{d^2x^0}{dt^2} = \frac{d}{dt}\left(\frac{dx^0}{dt}\right) = \frac{dc}{dt} = 0$, perciò la prima delle equazioni differenziali delle geodetiche (11.147) si riduce all'identità 0=0.
Per $i = \alpha = 1, 2, 3$, abbiamo

$$\Gamma^\alpha_{00} = -\frac{1}{2}\epsilon\,\eta^{\alpha r}\partial_r h_{00} = -\frac{1}{2}\epsilon\Big(\eta^{\alpha 0}\partial_0 h_{00} + \eta^{\alpha 1}\partial_1 h_{00} + \eta^{\alpha 2}\partial_2 h_{00} + \eta^{\alpha 3}\partial_3 h_{00}\Big).$$

Se $\alpha = 1$ entro parentesi rimane solo il termine $\eta^{11}\partial_1 h_{00} = -\partial_1 h_{00}$, se $\alpha = 2$ rimane solo il termine $\eta^{22}\partial_2 h_{00} = -\partial_2 h_{00}$, se $\alpha = 3$ rimane solo il termine $\eta^{33}\partial_3 h_{00} = -\partial_3 h_{00}$, perciò avremo

$$\Gamma^\alpha_{00} = \frac{1}{2}\epsilon\,\partial_\alpha h_{00}. \tag{11.151}$$

Usando queste approssimazioni per i simboli di Christoffel, possiamo scrivere l'equazione delle geodetiche (11.147) come

$$\frac{d^2x^\alpha}{dt^2} = -\frac{1}{2}c^2\epsilon\,\partial_\alpha h_{00} \qquad (\alpha = 1, 2, 3). \tag{11.152}$$

In notazione vettoriale 3-dimensionale, indicando con $\mathbf{x}$ il 3-vettore di posizione della particella, la (11.152) può scriversi

$$\frac{d^2\mathbf{x}}{dt^2} = -\mathrm{grad}\Big(\frac{1}{2}c^2\epsilon\,h_{00}\Big). \tag{11.153}$$

In questa equazione riconosciamo **la legge del moto di Newton** *di una particella mobile in un campo gravitazionale classico dipendente da un potenziale scalare*[50]

$$\frac{d^2\mathbf{x}}{dt^2} = -\mathrm{grad}\,\Phi, \tag{11.154}$$

purché si identifichi il potenziale scalare Φ con l'argomento del gradiente nella (11.153), cioè:

$$\Phi = \frac{1}{2}c^2\epsilon\,h_{00}. \tag{11.155}$$

Viceversa, dato il potenziale newtoniano classico Φ, il moto di una particella liberamente gravitante si svolgerà su di una geodetica 4-dimensionale dello spaziotempo, se il coefficiente $g_{00} = \eta_{00} + \epsilon h_{00} = 1 + \epsilon h_{00}$ del tensore metrico ha la forma

$$g_{00} \simeq 1 + \frac{2\Phi}{c^2}. \tag{11.156}$$

Le altre componenti del tensore metrico non entrano nella nostra approssimazione, eccetto che attraverso l'ipotesi che esse sono indipendenti dal tempo e prossimi ad essere Minkowskiane.
Riassumiamo i precedenti risultati:
Se ignoriamo termini del 2° ordine o di ordine maggiore in ϵ e β **(campo debole indipendente dal tempo e basse velocità)**, *allora l'equazione delle geodetiche* (11.140), *o* (11.141), *di una particella liberamente gravitante* **(relazione puramente geometrica)** *è equivalente alla* 2^a *legge di Newton* (11.154)

[50] Si ricordi che, nell'ambito della Meccanica Classica, il moto di una particella mobile in un campo gravitazionale, generato per esempio da una massa puntiforme M, è indipendente dalla massa m della particella.

(relazione puramente meccanica), *purché la componente g_{00} del tensore metrico soddisfi la relazione* (11.156).
Questa equivalenza fornisce una giustificazione piuttosto buona per l'uso di una metrica non-minkowskiana per descrivere un campo gravitazionale e per l'uso dell'equazione delle geodetiche come equazioni di moto nel relativo spazio riemanniano. Inoltre la relazione approssimata (11.156) per il termine g_{00} del tensore metrico sarà essa stessa utile per molte ragioni:
1) Ci permette di correlare certe costanti che appaiono formalmente nei successivi sviluppi della teoria relativistica a quantità classiche familiari, come la massa e la costante newtoniana G di gravitazione;
2) Possiamo verificare, quando risolveremo le equazioni gravitazionali di Einstein, che la componente g_{00} è consistente col nostro risultato approssimato;
3) Possiamo predire **il redshift delle linee spettrali** in un campo gravitazionale, usando soltanto i risultati precedenti nell'ambito della RR; ciò sarà fatto nel Cap.13, paragrafo 13.4.

11.15 Le equazioni gravitazionali di Einstein

11.15.1 Genesi euristica delle equazioni gravitazionali

Nel paragrafo 11.6, al quale rimandiamo il lettore per riprendere il discorso ivi impostato, abbiamo riconosciuto che le equazioni di campo della RG devono costituire una sorta di *equazione di bilancio fra geometria e materia* e perciò devono avere sinteticamente la forma:

$$S^{ik} = T^{ik}. \tag{11.157}$$

Queste equazioni determineranno, almeno in via di principio nel caso che siano risolubili, la metrica g^{ik} dello spaziotempo a partire da una data distribuzione delle sorgenti rappresentate dal tensore energetico T^{ik}. La metrica non viene data a priori ma viene determinata dalla distribuzione materiale che noi abbiamo considerato. Possiamo cioè dire, col Principio di Mach, che: *la metrica dello spaziotempo $\mathcal{V}^4$ è determinata dalla distribuzione della materia-energia nell'Universo.*
Questo principio ha giocato un'importante ruolo **euristico**[51] negli sforzi di Einstein per arrivare ad una formulazione corretta della Relatività Generale[52].
Come ora riconosceremo, seguendo appunto un procedimento euristico, il complesso delle condizioni imposte dai cinque criteri esposti nel paragrafo 11.6, determina univocamente la scelta delle equazioni di campo che andiamo cercando.
Siccome S^{ik} è il tensore rappresentativo della geometria dello spaziotempo, in un primo approccio Einstein postulò che esso fosse proporzionale al tensore di Ricci, che non solo rende conto della curvatura di $\mathcal{V}^4$, ma soddisfa altresì al 1° criterio del paragrafo 11.6 in quanto contiene le derivate prime e seconde del tensore metrico g^{ik} ed inoltre dipende linearmente da queste ultime in analogia con l'equazione di Poisson. Perciò Einstein postulò che fosse $S^{ik} = AR^{ik}$, $(A = cost. \neq 0)$, in tal caso l'equazione tensoriale di campo sarebbe

$$R^{ik} = \kappa T^{ik}, \tag{11.158}$$

con $\kappa = \dfrac{1}{A}$ costante.
Nelle regioni dello spazio prive di materia abbiamo $T^{ik} = 0$ e conseguentemente le precedenti equazioni si ridurrebbero alle

$$R^{ik} = 0. \tag{11.159}$$

[51] **L'euristica** è una parola che deriva dal greco **"eurisco"** che significa **"scopro o trovo"**, è una parte **dell'epistemologia** ovvero del **metodo scientifico.**

[52] Per una visione storica del percorso di pensiero seguito da Einstein per formulare la Relatività Generale si veda il già citato: Abraham Pais, *Sottile è il Signore....,La scienza e la vita di Albert Einstein*, Boringhieri, Torino,1986, pp.283-284 e Gerald Holton, *La Lezione di Einstein*, Parte seconda, cap. 6 e 7, Feltrinelli Editore, Milano, 1997.

Le (11.159) sarebbero dunque le equazioni del campo gravitazionale per regioni spaziali vuote, cioè del tutto prive di materia. Questa conclusione si rivelerà a posteriori corretta anche se dedotta da un falso presupposto costituito dalla (11.158) come ora riconosceremo.
Infatti, Einstein si rese conto, successivamente, che le (11.158) non possono essere le corrette equazioni di campo, e perciò devono essere modificate, a causa delle seguenti difficoltà.
Il tensore energetico obbedisce in RR, quando si usano riferimenti inerziali, alla legge di conservazione

$$\partial_k T^{ik} = 0. \tag{11.160}$$

Se si considerano coordinate generali, questa relazione deve essere scritta, in virtù del *Principio di Trascrizione*, nella forma

$$\nabla_k T^{ik} = 0. \tag{11.161}$$

In accordo col *Principio di equivalenza* noi dobbiamo richiedere la validità dell'equazione (11.161) anche in uno spazio riemanniano che descrive il campo gravitazionale in RG, ne segue allora, per la (11.158), che il tensore di Ricci dovrebbe soddisfare la relazione

$$\nabla_k R^{ik} = 0, \tag{11.162}$$

avere cioè divergenza nulla. Ciò però non è generalmente vero, a meno che le (11.162) vengano interpretate come 4 ulteriori equazioni, in aggiunta alle (11.158), per le 10 componenti del tensore metrico g_{ik}, conducenti però ad una sovraddeterminazione per le stesse g_{ik}.
Purtroppo Einstein, all'epoca in cui lavorava alla ricerca delle equazioni di campo, non conosceva le *identità di Bianchi* [53] utilizzando le quali - come abbiamo visto a suo tempo - non è difficile scoprire che esiste un tensore:

$$G^{ik} \equiv R^{ik} - \frac{1}{2} g^{ik} R, \tag{11.163}$$

costruito in maniera univoca con il tensore di Ricci R^{ik}, lo scalare di curvatura R ed il tensore metrico g^{ik}, dipendente perciò esclusivamente dalla struttura geometrica di $\mathcal{V}^4$, che gode della notevole proprietà di avere - quali che siano R^{ik}, R, e g^{ik} - divergenza nulla, cioè è identicamente

$$\nabla_k G^{ik} = 0. \tag{11.164}$$

Le 4 equazioni (11.164), non impongono affatto, come le (11.162), ulteriori condizioni al tensore metrico g^{ik} proprio perché sono identicamente soddisfatte. A questo importantissimo tensore Einstein è pervenuto dopo molteplici tentativi ed errori che gli sono costati molti anni di duro lavoro. Il tensore G^{ik} fu successivamente chiamato **tensore di Einstein** o di **Einstein - Levi-Civita** in quanto fu proprio Levi-Civita a chiarire la questione utilizzando le identità di Bianchi che lui conosceva molto bene.
Con la conoscenza del tensore di Einstein G^{ik}, si può intanto sostituire più correttamente la (11.162) con la (11.164) e l'errata equazione di campo (11.158) con la sua corretta formulazione:

$$R^{ik} - \frac{1}{2} R g^{ik} = \kappa T^{ik}. \tag{11.165}$$

Se con **G** e **T** indichiamo sinteticamente il tensore di Einstein ed il tensore energia-impulso, la (11.165) può scriversi più concisamente

$$G^{ik} = \kappa T^{ik} \quad \text{ovvero} \quad \mathbf{G} = \kappa \mathbf{T}. \tag{11.166}$$

È questa la celebre *equazione gravitazionale di Einstein*, chiamata anche *equazione di campo della Relatività Generale* o semplicemente *equazione di Einstein*; in essa il coefficiente κ si chiama *costante*

[53] Cfr. Cap.10, paragrafo 10.11.4.

gravitazionale einsteiniana.
Dimostreremo in seguito che risulta $\kappa = 8\pi G/c^4$, dove G è la costante classica della gravitazione newtoniana.[54]
L'equazione gravitazionale (11.165) non è la più generale possibile, essa può essere generalizzata aggiungendo il cosiddetto **termine cosmologico** nel seguente modo.
Basta tener presente che, per il teorema di Ricci sulla derivazione covariante del tensore metrico, risulta $\nabla_j g^{ik} = 0$ per cui, contraendo sugli indici j, k, si ha anche $\nabla_k g^{ik} = 0$, dunque l'identità (11.164) si generalizza nella seguente relazione:

$$\nabla_k (AG^{ik} + Bg^{ik}) = 0, \tag{11.167}$$

dove A e B sono degli scalari invarianti.
Se ne deduce che il tensore S^{ik}, che figura nel 1° membro della (11.157), deve avere la seguente espressione

$$S^{ik} = AG^{ik} + Bg^{ik}, \tag{11.168}$$

essendo A e B delle costanti opportune. La (11.157) prende quindi la forma

$$R^{ik} - \frac{1}{2} R g^{ik} + \Lambda g^{ik} = \kappa T^{ik}, \quad (i, k = 0, 1, 2, 3) \tag{11.169}$$

avendo posto $\Lambda = B/A$ e $\kappa = 1/A$, con $A \neq 0$.
È questa *la più generale equazione gravitazionale* contenente **il termine cosmologico** Λg^{ik}, dove Λ è chiamata la **costante cosmologica** (di dimensione cm^{-2}) ed ha un valore molto piccolo ($|\Lambda| \leq 10^{-55}\,\mathrm{cm}^{-2}$), generalmente trascurabile se si considerano piccole porzioni di Universo, mentre ha notevole importanza in questioni cosmologiche, cioè quando si considera l'Universo nella sua globalità. Ritorneremo nel seguito con più dettagli sulle questioni che riguardano la costante cosmologica ed il relativo temine cosmologico che figura nella (11.169).
Osserviamo infine che l'equazione tensoriale (11.165) o la più generale (11.169) costituisce, dal punto di vista matematico, un sistema di 16 equazioni differenziali alle derivate parziali, cioè tante quante sono le disposizioni con ripetizione di 4 oggetti a 2 a 2 ($D'_{4,2} = 4^2 = 16$), in quanto gli indici i, k assumono 4 valori distinti $i, k = 0, 1, 2, 3$. Di queste, però, solo 10 sono effettivamente indipendenti, tante quante sono le componenti indipendenti dei tensori simmetrici del 2° ordine che vi figurano e cioè tante quante sono le funzioni incognite g^{ik} (i cosiddetti *potenziali gravitazionali*), e cioè tante quante sono le combinazioni con ripetizione di 4 oggetti a 2 a 2 ($C'_{4,2} = \binom{4+2-1}{2} = \binom{5}{2} = 10$).
Naturalmente si suppone che sia assegnata la distribuzione ed il moto della materia dell'Universo compendiati nel tensore energetico T^{ik}.

11.15.2 Il teorema di Cartan

Come abbiamo più volte ripetuto Einstein è pervenuto alle equazioni gravitazionali (11.165), (11.169) dopo circa dieci anni di tentativi, errori e duro lavoro, senza avere un piano matematicamente prestabilito[55].
Nel paragrafo precedente abbiamo visto che il tensore simmetrico

$$S^{ik} = R^{ik} - \frac{1}{2} R g^{ik} + \Lambda g^{ik}$$

[54] Avvertiamo il lettore che nella letteratura sulla Relatività Generale, talvolta si trova per κ, anziché il valore riportato più sopra, cioè $\kappa = 8\pi G/c^4$, il valore $\kappa = 8\pi G/c^2$ che differisce dal precedente, per un fattore c^2 al denominatore. Ciò dipende dal fatto che il tensore energia-impulso, nel primo caso, è stato definito con la $T^{ik} = \mu_0 U^i U^k = c^2 \mu_0 u^i u^k$ avente le dimensioni di una **densità di energia**, mentre nel secondo caso è stato definito con la $T^{ik} = \mu_0 u^i u^k$ che ha invece le dimensioni di una **densità di massa**.
[55] Si consulti il libro già citato di A. Pais, o anche Gerald Holton, *La Lezione di Einstein*, Feltrinelli, Milano 1997.

soddisfa sostanzialmente ai 5 criteri, visti nel paragrafo 11.6, che Einstein seguì per scrivere le equazioni del campo gravitazionale. *Ma chi ci assicura che esso sia l'unico tensore che risolve il nostro problema?* La risposta a questa domanda è stata data, nel 1922, dal grande matematico francese Èlie Cartan, sei anni dopo la pubblicazione del lavoro fondamentale di Einstein, col seguente notevole:

Teorema 11.1 (di Cartan)[56]*:*

1) *Il più generale tensore simmetrico S^{ik}, del 2° ordine, che sia funzione soltanto del tensore metrico g^{ik}, delle sue derivate prime e dipenda linearmente dalle sue derivate seconde è della forma*

$$S^{ik} = aR^{ik} + bg^{ik}R + cg^{ik}, \tag{11.170}$$

dove R^{ik} è il tensore di Ricci, R è lo scalare di curvatura ed a, b e c sono delle costanti (con $a \neq 0$).

2) *Se si richiede inoltre che S^{ik} sia a divergenza nulla (o conservativo), ovvero che risulti $\nabla_k S^{ik} = 0$, allora esso è necessariamente della forma:*

$$S^{ik} = A\left(R^{ik} - \frac{1}{2}Rg^{ik}\right) + Bg^{ik}, \tag{11.171}$$

con A e B costanti ed $A \neq 0$.

La richiesta del punto 2) del teorema di Cartan, cioè che il tensore S^{ik} che rappresenta la geometria dello spaziotempo $\mathcal{V}^4$, sia a divergenza nulla, è necessaria se vogliamo che la legge di conservazione (11.161) sia una conseguenza logica delle equazioni gravitazionali (11.169), come viene richiesto nel 5° criterio del paragrafo 11.6.
Osserviamo che il coefficiente a che si trova nel 1° termine del 2° membro della (11.170) non può essere nullo perché in tal caso si avrebbe $S^{ik} = (bR + c)g^{ik}$ e, per la (11.157), il tensore energetico T^{ik} sarebbe semplicemente proporzionale al tensore metrico g^{ik}, ciò che generalmente non si verifica.
Sostituendo l'espressione (11.171) del tensore S^{ik} nel 1° membro della (11.157) otteniamo l'equazione gravitazionale (11.169): possiamo dunque concludere che quest'ultima è ora posta su basi matematicamente più solide di quelle fornite dai ragionamenti euristici di Einstein grazie all'applicazione del teorema di Cartan!
Come abbiamo detto, l'equazione gravitazionale (11.169) essendo un'equazione tensoriale è equivalente ad un sistema di 10 equazioni alle derivate parziali, non lineari, di difficilissima soluzione, ed è perciò notevole che *la prima soluzione esatta* sia stata trovata dal matematico ed astronomo tedesco Karl Schwarzschild nel 1916 quasi 2 mesi dopo la pubblicazione dell'articolo di Einstein. Appena Einstein lesse il lavoro di Schwarzschild le scrisse immediatamente dicendogli: *Ho letto il suo lavoro con il massimo interesse. Non mi sarei mai aspettato che si potesse determinare la soluzione esatta del problema in modo così semplice. La formulazione analitica mi sembra splendida.* (Lettera del 9 Gennaio 1916).

11.15.3 Una forma alternativa dell'equazione gravitazionale

L'equazione di campo di Einstein, che finora abbiamo scritto in forma controvariante, essendo un'equazione tensoriale, può essere scritta ovviamente anche in forma covariante:

$$R_{ik} - \frac{1}{2}Rg_{ik} + \Lambda g_{ik} = \kappa T_{ik}, \tag{11.172}$$

[56]Questo teorema, di difficile dimostrazione, è stato pubblicato da Èlie Cartan col titolo: *Sur les équations de la gravitation d'Einstein* sul Journal des Math. Pures et Appliquées, 1, pp. 141-203, Paris (1922).

oppure in forma mista:

$$R^i_k - \frac{1}{2}\delta^i_k R + \Lambda\delta^i_k = \kappa T^i_k. \tag{11.173}$$

Da quest'ultima, mediante una contrazione sugli indici (i, k), tenuto conto che $\delta^i_i = 4$, otteniamo l'espressione dello scalare di curvatura nella forma:

$$R = 4\Lambda - \kappa T, \tag{11.174}$$

dove $T = T^i_i$ prende il nome di *invariante lineare o scalare materiale o scalare di Max Von Laue*. Sostituendo questo risultato nella (11.172) si ottiene

$$R_{ik} - 2g_{ik}\Lambda + \frac{1}{2}g_{ik}\kappa T + \Lambda g_{ik} = \kappa T_{ik}, \tag{11.175}$$

dalla quale ne segue una forma alternativa per l'equazione gravitazionale:

$$R_{ik} - \Lambda g_{ik} = \kappa\left(T_{ik} - \frac{1}{2}g_{ik}T\right), \tag{11.176}$$

scritta in forma covariante, che talvolta può essere comodo utilizzare.

11.15.4 Digressione: Cenno sulle densità scalari e densità tensoriali

In questo paragrafo facciamo una breve digressione introducendo i concetti di **densità scalare** e di **densità tensoriale** che utilizzeremo nel paragrafo successivo per dare una formulazione variazionale delle equazioni gravitazionali.[57]
Ci sono quantità $Q(x^1, x^2, ..., x^n) \equiv Q(x^i)$, funzioni di più variabili, che in una trasformazione di coordinate $x \to x'$ si trasformano nel modo seguente:

$$Q(x'^{\,i}) = J^W Q(x^i), \tag{11.177}$$

dove $J = \left|\frac{\partial x}{\partial x'}\right| = det\left(\frac{\partial x}{\partial x'}\right)$ denota lo jacobiano della trasformazione dal sistema di coordinate x^i al sistema di coordinate $x'^{\,i}$, e W *è un numero intero positivo o negativo* ($W \in \mathbb{Z}$). Una quantità Q che si trasforma secondo la formula precedente si chiama una **densità scalare di peso** W.
Un esempio di densità scalare di peso $W = 2$ è fornito dal determinante della matrice i cui elementi sono le componenti T_{ik} di un tensore covariante del 2° ordine. La legge di trasformazione di un tale tensore può essere riguardata come un'equazione matriciale:

$$T_{i'k'} = \left(\frac{\partial x^i}{\partial x^{i'}}\right)(T_{ik})\left(\frac{\partial x^k}{\partial x^{k'}}\right), \tag{11.178}$$

utilizzando la regola per il calcolo del determinante di un prodotto di matrici, troviamo che

$$det\, T_{i'k'} = \left|\frac{\partial x}{\partial x'}\right|^2 det\, T_{ik}. \tag{11.179}$$

Perciò possiamo dire, con locuzione abbreviata, che il determinante di un tensore covariante di rango 2 è una densità scalare di peso $W = 2$.
Esercizio:
Un altro esempio, che proponiamo come esercizio, è il seguente. Consideriamo un tensore

[57] Sul concetto di densità tensoriale il lettore può consultare con profitto il libro di Achille Papapetrou, *Lectures on General Relativity*, D. Reidel Publishing Company, Dordrecht-Holland, 1974, Chapter 1, pagg. 11-14. Un'altra buona referenza è Moshe Carmeli, *Classical fields: General Relativity and Gauge Theory*, John Wiley and Sons, New York, 1982, Chapter 2, pagg. 35-38.

covariante antisimmetrico di ordine $n = 4$: T_{ijhk}. Le componenti non nulle di questo tensore sono ovviamente uguali a $\pm T_{1234}$. Ci domandiamo: come cambia la componente T_{1234} per una generica trasformazione di coordinate $x \to x'$?
La risposta è contenuta nella formula di trasformazione tensoriale:

$$T'_{1234} = \frac{\partial x^i}{\partial x'^1}\frac{\partial x^j}{\partial x'^2}\frac{\partial x^h}{\partial x'^3}\frac{\partial x^k}{\partial x'^4}T_{ijhk}.$$

La somma al secondo membro è estesa a tutte le componenti non nulle del tensore T_{ijhk}. Il loro numero è $4! = 24$, 12 delle quali uguali a T_{1234} (se $ijhk$ è una permutazione di classe pari della permutazione fondamentale 1234) e le altre 12 uguali a $-T_{1234}$ (se $ijhk$ è una permutazione di classe dispari della permutazione fondamentale.) Si può constatare che il risultato finale è:

$$T'_{1234} = JT_{ijhk},$$

con $J = det\left(\frac{\partial x}{\partial x'}\right)$.
Più in generale, per un tensore antisimmetrico covariante di ordine n si avrà:

$$T'_{12....n} = JT_{12....n},$$

perciò, essendo l'esponente di J uguale ad 1, la componente $T_{12....n}$ è una densità scalare di ordine n e peso $W = 1$.
Il prodotto di un tensore di tipo (p, q) ed ordine $n = p + q$ per una densità scalare di peso W è per definizione una **densità tensoriale di tipo** (p, q)**, ordine** $n = p + q$ **e peso** W. La formula di trasformazione per una **densità tensoriale** si deduce dalle formule di trasformazione di un tensore e dalla formula di trasformazione di una densità scalare.
Più generalmente possiamo dare la seguente:
Definizione:
Dicesi **densità tensoriale** *una quantità $\mathcal{T}^{i...}_{k...}$ che si trasforma come un tensore, con la differenza che nella sua legge di trasformazione compare un fattore, che è lo jacobiano della trasformazione di coordinate, elevato ad una certa potenza.*
Avremo perciò che una **densità tensoriale** $\mathcal{T}^{i...}_{k...}$ si trasforma secondo la seguente legge:

$$\mathcal{T}'^{\,i...}_{\;k...} = \left|\frac{\partial x}{\partial x'}\right|^W \frac{\partial x'^{\,i}}{\partial x^j}.....\frac{\partial x^h}{\partial x'^{\,k}}.....\mathcal{T}^{j...}_{h...} \tag{11.180}$$

ovvero utilizzando l'abuso di notazione, che abbiamo sempre usato, cioè di mettere l'apice sopra gli indici anziché sopra le lettere che rappresentano la densità tensoriale o una generica coordinata:

$$\mathcal{T}^{i'...}_{k'...} = \left|\frac{\partial x}{\partial x'}\right|^W \frac{\partial x^{i'}}{\partial x^i}.....\frac{\partial x^k}{\partial x^{k'}}.....\mathcal{T}^{i...}_{k...}. \tag{11.181}$$

Nelle precedenti equazioni $\left|\frac{\partial x}{\partial x'}\right| = det\left(\frac{\partial x}{\partial x'}\right)$ denota lo jacobiano della trasformazione dal sistema di coordinate x^i al sistema di coordinate $x^{i'}$, e *W è un numero intero positivo o negativo, chiamato il* **peso** *della densità tensoriale ($W \in \mathbb{Z}$).* Se il numero degli indici in alto di $\mathcal{T}^{i...}_{k...}$ è p e quello degli indici in basso è q, con $p + q = n$, si dice che *la densità tensoriale è di tipo (p, q), di ordine n e di peso W.*
Una densità tensoriale di *ordine 1* è chiamata *densità vettoriale*, ed una densità di *ordine 0* è ovviamente una *densità scalare*. I tensori ordinari sono *densità tensoriali* di **peso 0**.
Possiamo ora applicare la (11.179) al tensore metrico g_{ik}: si trova allora che per una generica trasformazione di coordinate $x \to x'$, si ha

$$g' = \left|\frac{\partial x}{\partial x'}\right|^2 g = J^2 g, \tag{11.182}$$

dove $g = det g_{ik}$ e $g' = det g_{i'k'}$, con $J = \left|\frac{\partial x}{\partial x'}\right| = det\left(\frac{\partial x}{\partial x'}\right)$.
Dalla (11.182) si deduce che g e g' hanno sempre lo stesso segno, essendo $J^2 > 0$. Dalla teoria del calcolo differenziale ed integrale per funzioni di più variabili, è noto che per una generica trasformazione di coordinate della forma

$$x^{i'} = x^{i'}(x^i), \quad (i, i' = 0, 1, 2, 3), \tag{11.183}$$

l'elemento di volume $d^4x = dx^0 dx^1 dx^2 dx^3$ si trasforma nel modo seguente

$$d^4x' = |J'| d^4x, \tag{11.184}$$

dove $|J'| = |J^{-1}| = \left|det\left(\frac{\partial x'}{\partial x}\right)\right|$ è il modulo dello jacobiano della trasformazione (11.183) che è l'inversa della trasformazione $x^i = x^i(x^{i'})$; queste trasformazioni hanno ovviamente jacobiani che sono l'uno l'inverso dell'altro, cioè $J' = J^{-1}$ ovvero $J = (J')^{-1}$.
Se riscriviamo la (11.182) nella forma

$$|J| = \frac{\sqrt{|g'|}}{\sqrt{|g|}},$$

ovvero

$$|J'| = |J^{-1}| = \frac{\sqrt{|g|}}{\sqrt{|g'|}},$$

la (11.184) diventa allora:

$$\sqrt{|g'|} d^4x' = \sqrt{|g|} d^4x, \tag{11.185}$$

e siccome in Relatività Generale $g < 0$, $g' < 0$, possiamo riscrivere la (11.185) così:

$$\sqrt{-g'}\, d^4x' = \sqrt{-g}\, d^4x = Invariante\ per\ trasf.\ di\ coord. \tag{11.186}$$

Questa è l'espressione **dell'elemento di volume 4-dimensionale dello spaziotempo $\mathcal{V}^4$, invariante per una generica trasformazione di coordinate.**
L'equazione (11.182) possiamo utilizzarla nella legge di trasformazione delle densità tensoriali (11.181) riscrivendo quest'ultima nella forma:

$$(-g')^{-W/2} \mathcal{T}^{i'\ldots}_{k'\ldots} = \frac{\partial x^{i'}}{\partial x^i} \ldots\ldots \frac{\partial x^k}{\partial x^{k'}} \ldots\ldots (-g)^{-W/2} \mathcal{T}^{i\ldots}_{k\ldots}. \tag{11.187}$$

Si vede che la quantità

$$(-g)^{-W/2} \mathcal{T}^{i\ldots}_{k\ldots}\ , \tag{11.188}$$

segue la regola di trasformazione di un tensore ordinario $T^{i\ldots}_{k\ldots}$.
Conseguentemente una densità tensoriale di peso W, moltiplicata per il fattore $(-g)^{-W/2}$, è un tensore ordinario, ovvero una densità tensoriale di peso 0. Allo stesso modo un tensore ordinario, se moltiplicato per $(-g)^{W/2}$, diventa una densità tensoriale di peso W. in particolare la quantità $\sqrt{-g} T^{i\ldots}_{k\ldots}$ è una densità tensoriale di peso $W = 1$ se $T^{i\ldots}_{k\ldots}$ è un tensore ordinario.
In generale le regole dell'algebra tensoriale possono essere generalizzate alle densità tensoriali. Abbiamo allora le seguenti regole:

1. Una combinazione lineare di densità tensoriali dello stesso tipo e peso, assegnate nello stesso punto dello spaziotempo $\mathcal{V}^4$, è uguale ad una nuova densità tensoriale dello stesso tipo e peso, nello stesso punto.

2. Il prodotto diretto di due densità tensoriali di ordine m ed n e pesi W_1 e W_2 rispettivamente, definite nello stesso punto dello spaziotempo, produce una nuova densità tensoriale di ordine $m+n$ e peso $W = W_1 + W_2$, nello stesso punto.

3. La contrazione di due indici su di una densità tensoriale mista di ordine n e peso W, produce una nuova densità tensoriale di ordine $n-2$ e peso W.

Come conseguenza delle regole precedenti ne segue che una densità tensoriale $\mathcal{T}^{i\cdots}_{k\cdots}$ di peso W può essere espressa come prodotto di un tensore ordinario $T^{i\cdots}_{k\cdots}$ e di una densità scalare di peso W come avevamo già posto, per definizione, all'inizio delle nostre considerazioni. Ne segue che innalzando od abbassando degli indici, per mezzo del tensore metrico g_{ik}, di una densità tensoriale essa non cambia peso. Inoltre, le proprietà di simmetria delle densità tensoriali sono definite esattamente allo stesso modo che per i tensori ordinari.

11.16 Deduzione variazionale delle equazioni gravitazionali

Nei paragrafi precedenti abbiamo ricavato le equazioni gravitazionali dapprima con un procedimento euristico, più o meno approssimativo, poi con più precisione utilizzando il teorema di Cartan; in questo paragrafo ricaveremo le stesse equazioni in maniera perfettamente rigorosa utilizzando un principio variazionale dovuto ad Hilbert (1915).
Com'è noto le leggi della meccanica classica si possono compendiare in un solo principio variazionale, il cosiddetto
Principio di Hamilton:
Per un sistema materiale, a vincoli perfetti bilaterali, sottoposto ad una sollecitazione conservativa, il generico moto naturale è caratterizzato come quello che rende stazionaria l'azione hamiltoniana $S = \int_{t_0}^{t_1} \mathcal{L} dt$, *cioè tale che:*

$$\delta \int_{t_0}^{t_1} \mathcal{L}(q^i, \dot{q}^i, t)\, dt = 0, \qquad (i = 1, 2, ..., n)$$

rispetto a tutti i moti variati sincroni fra le stesse configurazioni estreme.
Si noti che la *lagrangiana o potenziale cinetico* $\mathcal{L}$ è uguale alla somma tra *l'energia cinetica* del sistema ed *il potenziale* delle forze attive ad esso applicate:

$$\mathcal{L} = T + U$$

ed è invariante per tutte le trasformazioni di coordinate.
In Relatività Generale si ammette che valga la seguente proprietà:
Le equazioni differenziali del campo gravitazionale si possono dedurre da un principio analogo a quello di Hamilton, ammettendo cioè che sia nulla la variazione di una opportuna **funzione universale (l'integrale d'azione)**, *costruito con una opportuna lagrangiana* $\mathcal{L}$, *la quale sia invariante rispetto a tutte le trasformazioni di coordinate dello spaziotempo.*
Tale lagrangiana $\mathcal{L}$, relativa, diremo così, all'unità di volume dello spaziotempo $\mathcal{V}^4$, la immagineremo, in analogia col Principio di Hamilton, decomposta in due parti:
una prima parte $\mathcal{L}_{\mathcal{G}}$ che rappresenti **il contributo spaziale e temporale (e quindi quello del campo gravitazionale) e che sia uno scalare invariante**; più una seconda parte $\mathcal{L}_{\mathcal{M}}$ che **tenga conto delle azioni fisiche che si esercitano nel campo (gravitazione esclusa) e quindi dovute alla presenza della materia-energia**: tale funzione sarà dunque esprimibile nella forma:

$$\mathcal{L} = \mathcal{L}_{\mathcal{G}} + \mathcal{L}_{\mathcal{M}}. \tag{11.189}$$

Discutiamo ora in maniera più dettagliata l'introduzione di queste due funzioni.

1) Per determinare la **lagrangiana gravitazionale** $\mathcal{L}_{\mathcal{G}}$ dobbiamo ricordarci che le equazioni di campo che andiamo cercando devono essere equazioni differenziali alle derivate parziali del 2° ordine nei potenziali gravitazionali incogniti g_{ik} (componenti del tensore metrico di $\mathcal{V}^4$), contenenti le derivate prime di g_{ik} e le derivate seconde linearmente. Poiché le equazioni di campo ottenute da una lagrangiana contengono usualmente derivate di un ordine superiore a quelle che figurano nella lagrangiana stessa, ne segue che $\mathcal{L}_{\mathcal{G}}$ dovrebbe contenere il tensore metrico g_{ik} e solamente le sue derivate prime al primo ordine. In altri termini, $\mathcal{L}_{\mathcal{G}}$ dovrebbe includere il tensore metrico g_{ik} ed i coefficienti di connessione, o simboli di Christoffel di 2^a specie Γ^j_{ik} che dipendono appunto linearmente dalle $\partial_h g_{ik}$.
Come sappiamo, dalla geometria riemanniana, l'unico scalare costruito col tensore metrico e con le sue derivate è lo *lo scalare di curvatura* R di Ricci, il quale gode della proprietà di contenere le derivate prime (non linearmente) del tensore metrico, le sue derivate seconde linearmente ed il tensore metrico stesso. Perciò se noi usiamo R come lagrangiana del campo gravitazionale, in virtù della linearità delle derivate seconde di g_{ik}, i termini contenenti tali derivate non contribuiscono a produrre termini contenenti derivate di g_{ik} di ordine superiore a 2. Tenuto conto di tutto ciò possiamo dunque porre:

$$\mathcal{L}_{\mathcal{G}} = R, \tag{11.190}$$

inoltre, siccome l'integrale d'azione relativo ad $\mathcal{L}_{\mathcal{G}}$ deve essere esteso ad una regione 4-dimensionale di $\mathcal{V}^4$, detta Ω una porzione dello spaziotempo, *l'elemento invariante di volume*, per la (11.186) del paragrafo precedente, è dato da

$$d\Omega = \sqrt{|g|}\, d^4x,$$

ed essendo $g = \det(g_{ij}) < 0$ avremo

$$d\Omega = \sqrt{-g}\, d^4x. \tag{11.191}$$

avendo indicato con d^4x la quantità $dx^0dx^1dx^2dx^3$; $d\Omega$ è la cosiddetta *forma di volume riemanniana*. Siamo ora in grado di scrivere l'integrale d'azione relativo ad $\mathcal{L}_{\mathcal{G}}$ nella forma

$$\int_\Omega \mathcal{L}_{\mathcal{G}} d\Omega = \int_\Omega R d\Omega = \int_\Omega \sqrt{-g}\, R\, d^4x. \tag{11.192}$$

2) La seconda parte della lagrangiana $\mathcal{L}$ che vogliamo costruire deve tener conto, come abbiamo già detto, delle azioni fisiche che si esercitano nel campo (**gravitazione esclusa**) e quindi dovute alla presenza della materia-energia, possiamo dunque identificarla, a meno di una costante κ, con *l'invariante lineare* $T = T^i_i = g^{ik}T_{ik}$ del *tensore energia-impulso*:

$$\mathcal{L}_{\mathcal{M}} = \kappa T, \tag{11.193}$$

ed il corrispondente integrale d'azione sarà:

$$\int_\Omega \mathcal{L}_{\mathcal{M}} d\Omega = \int_\Omega \kappa T d\Omega = \int_\Omega \sqrt{-g}\, \kappa T\, d^4x. \tag{11.194}$$

La lagrangiana totale è allora:

$$\mathcal{L} = \mathcal{L}_{\mathcal{G}} + \mathcal{L}_{\mathcal{M}} = R + \kappa T. \tag{11.195}$$

Se si vuole che nelle equazioni gravitazionali figuri anche il cosiddetto *termine cosmologico* è necessario aggiungere, alla lagrangiana, un'altra costante che, per comodità di calcolo, indicheremo con -2Λ, essendo Λ la cosiddetta *costante cosmologica*[58]. Dunque la più generale

[58] Il segno della costante cosmologica Λ è a priori indeterminato, potendo essere positivo o negativo. Per il momento non discuteremo il significato fisico di tale costante, riservandoci di ritornare sulla questione, in seguito, nel capitolo dedicato alla Cosmologia.

lagrangiana sarà

$$\mathcal{L} = R - 2\Lambda + \kappa T, \tag{11.196}$$

ed il nostro principio variazionale potrà essere enunciato nel seguente modo:
Principio variazionale di Hilbert:
È nulla la variazione dell'integrale d'azione

$$\delta \int_\Omega \mathcal{L}\, d\Omega = \delta \int_\Omega (R - 2\Lambda + \kappa T)\, d\Omega = 0, \tag{11.197}$$

esteso ad un campo Ω dello spaziotempo $\mathcal{V}^4$, quando i potenziali gravitazionali g_{ik} o g^{ik}, e le loro derivate prime, subiscono variazioni arbitrarie in ogni punto di Ω, le quali siano nulle sull'ipersuperficie $\Sigma \equiv \partial\Omega$ costituente il bordo del campo Ω, cioè

$$\delta g^{ik} \text{ arbitrarie in } \Omega, \quad (\delta g^{ik})_\Sigma = 0 \tag{11.198}$$

$$\delta\Big(\frac{\partial g^{ik}}{\partial x^j}\Big) \text{ arbitrarie in } \Omega, \quad \delta\Big(\frac{\partial g^{ik}}{\partial x^j}\Big)_\Sigma = 0. \tag{11.199}$$

La funzione $\mathcal{L}$ definita dalla (11.196) è evidentemente invariante rispetto a qualunque trasformazione di coordinate dello spaziotempo, perché è una somma di quantità invarianti.
Ora, come dal principio di Hamilton si deducono le equazioni dinamiche di Lagrange, così coi noti procedimenti del calcolo delle variazioni, dal *Principio di Hilbert* dedurremo le equazioni di campo della Relatività Generale, le quali saranno covarianti rispetto a tutte le trasformazioni di coordinate.
A tal fine, essendo lo scalare di curvatura R così definito:

$$R = g^{ik} R_{ik}, \tag{11.200}$$

ed essendo l'invariante lineare $T = g^{ik} T_{ik}$, tenuto conto della (11.191), l'equazione variazionale (11.197) diventa:

$$\delta \int_\Omega \Big\{ g^{ik} R_{ik} \sqrt{-g} - 2\Lambda \sqrt{-g} + \kappa g^{ik} T_{ik} \sqrt{-g} \Big\}\, d^4x = 0, \tag{11.201}$$

dalla quale, essendo l'operatore variazionale δ permutabile col segno d'integrale otteniamo:

$$\int_\Omega \Big\{ \delta g^{ik} R_{ik} \sqrt{-g} + g^{ik} \delta R_{ik} \sqrt{-g} + g^{ik} R_{ik} \delta \sqrt{-g} - 2\Lambda \delta \sqrt{-g} + \kappa \delta (g^{ik} T_{ik} \sqrt{-g}) \Big\}\, d^4x = 0. \tag{11.202}$$

Calcoliamo dapprima $\delta\sqrt{-g}$ che figura nel 3°, 4° e 5° termine dell'integrando della (11.202), avremo

$$\delta\sqrt{-g} = \frac{\delta(-g)}{2\sqrt{-g}} = -\frac{\delta g}{2\sqrt{-g}}; \tag{11.203}$$

la variazione δg del determinante $g = det(g_{ik})$ della metrica, per la formula (10.229) stabilita nel paragrafo 10.13, del Cap. 10, è uguale alla moltiplicazione contratta fra il differenziale δg_{ik} di ciascuna componente del tensore metrico per il complemento algebrico corrispondente, cioè:

$$\delta g = \delta g_{ik} A^{ik}. \tag{11.204}$$

D'altra parte, per il tensore metrico sotto forma controvariante (elemento reciproco di g_{ik} in g) si ha $g^{ik} = \dfrac{A^{ik}}{g}$ da cui $A^{ik} = g\, g^{ik}$ e perciò la (11.204), potrà scriversi nella forma

$$\delta g = g g^{ik} \delta g_{ik}. \tag{11.205}$$

Inoltre si ha l'identità

$$g_{ik}g^{ik} = \delta^i_i = 4,$$

e perciò differenziando quest'ultima identità avremo $\delta g_{ik}g^{ik} + g_{ik}\delta g^{ik} = 0$, da cui

$$g^{ik}\delta g_{ik} = -g_{ik}\delta g^{ik}. \tag{11.206}$$

Sostituendo la (11.206) nella (11.205) otteniamo

$$\delta g = -g g_{ik}\delta g^{ik}. \tag{11.207}$$

La (11.203) diventa allora

$$\delta\sqrt{-g} = -\frac{1}{2}\frac{\delta g}{\sqrt{-g}} = -\frac{1}{2}\sqrt{-g}g_{ik}\delta g^{ik}. \tag{11.208}$$

In virtù della (11.208), tenuto conto della definizione dello scalare di curvatura $R = g^{ik}R_{ik}$, il 1°, 3° e 4° termine della (11.202) sommati ed integrati danno perciò:

$$\int_\Omega \left\{R_{ik} - \frac{1}{2}Rg_{ik} + \Lambda g_{ik}\right\}\sqrt{-g}\,\delta g^{ik}\,d^4x. \tag{11.209}$$

Calcoliamo ora il 2° termine della (11.202) che è il più difficile da calcolarsi dato che contiene la variazione δR_{ik} del tensore di Ricci, che come sappiamo ha un'espressione piuttosto complicata. Si ha infatti per definizione:

$$R_{ik} \equiv R^l_{.ilk} = \partial_l\Gamma^l_{ik} - \partial_k\Gamma^l_{il} + \Gamma^l_{ml}\Gamma^m_{ik} - \Gamma^l_{mk}\Gamma^m_{il}, \tag{11.210}$$

dunque essendo la variazione δ permutabile con la derivazione parziale ∂ si avrà:

$$\delta R_{ik} = \partial_l(\delta\Gamma^l_{ik}) - \partial_k(\delta\Gamma^l_{il}) + \delta\Gamma^l_{ml}\Gamma^m_{ik} + \Gamma^l_{ml}\delta\Gamma^m_{ik} - \delta\Gamma^l_{mk}\Gamma^m_{il} - \Gamma^l_{mk}\delta\Gamma^m_{il}. \tag{11.211}$$

Osserviamo ora che mentre i simboli di Christoffel Γ^l_{ik} non rappresentano, in generale, dei tensori, le loro variazioni $\delta\Gamma^l_{ik}$ sono invece sempre dei tensori di tipo $(1,2)$, essendo la differenza di due connessioni. Perciò, in un generico cambiamento di coordinate, le $\delta\Gamma^l_{ik}$ si trasformano con legge tensoriale.

EQUAZIONI DI PALATINI.
Faremo ora una breve digressione per dimostrare un risultato fondamentale dovuto al matematico italiano Attilio Palatini[59], consistente in una formula semplice per δR_{ik}.
Molte identità tensoriali si possono dedurre, come abbiamo visto a suo tempo, utilizzando in maniera opportuna *un sistema di coordinate geodetiche*, scegliendo un punto arbitrario $P \in \mathcal{V}^4$ in cui si annullano i simboli di Christoffel di 2^a specie ma non le loro derivate prime: $(\Gamma^l_{ik})_P \stackrel{*}{=} 0$, $(\partial_j\Gamma^l_{ik})_P \neq 0$. Allora, in particolare, le derivate covarianti nel punto P si riducono alle derivate ordinarie.
In particolare il tensore di Ricci (11.210) in un tale sistema di coordinate si riduce a:

$$R_{ik} \stackrel{*}{=} \partial_l\Gamma^l_{ik} - \partial_k\Gamma^l_{il} \stackrel{*}{=} \nabla_l\Gamma^l_{ik} - \nabla_k\Gamma^l_{il}. \tag{11.212}$$

Consideriamo ora una variazione della connessione Γ^l_{ik} che dia una nuova connessione $\overline{\Gamma}^l_{ik}$:

$$\Gamma^l_{ik} \to \overline{\Gamma}^l_{ik} = \Gamma^l_{ik} + \delta\Gamma^l_{ik}. \tag{11.213}$$

[59] Attilio Palatini fu allievo di Levi-Civita a Padova e successivamente professore ordinario di Meccanica Razionale e Fisica Matematica a Pavia, il metodo che lui seguì per trovare l'equazione (11.216), scoperta nel 1919, è oggi noto come **metodo di Palatini.**

Poiché $\delta\Gamma^l_{ik}$ è la differenza di due connessioni, essa è un tensore di tipo $(1,2)$. Questa variazione modifica a sua volta il tensore di Ricci

$$R_{ik} \to \overline{R}_{ik} = R_{ik} + \delta R_{ik}, \tag{11.214}$$

dove, sempre nel considerato sistema di coordinate geodetiche, si ha, per la (11.212):

$$\delta R_{ik} \stackrel{*}{=} \partial_l(\delta\Gamma^l_{ik}) - \partial_k(\delta\Gamma^l_{il}) \stackrel{*}{=} \nabla_l(\delta\Gamma^l_{ik}) - \nabla_k(\delta\Gamma^l_{il}), \tag{11.215}$$

perché la derivata parziale ∂ commuta con la variazione δ ed è equivalente alla derivata covariante ∇ in coordinate geodetiche.
Sia il 1° che il 3° membro della (11.215) sono manifestamente dei tensori essendo entrambi differenze di tensori, perciò, in base al **risultato fondamentale del calcolo tensoriale** - *se un'equazione tensoriale vale in un determinato sistema di coordinate, allora vale in tutti i sistemi* - possiamo dunque dire che, in un generico punto P di $\mathcal{V}^4$ ed in un arbitrario sistema di coordinate, è verificata la seguente **equazione di Palatini**:

$$\delta R_{ik} = \nabla_l(\delta\Gamma^l_{ik}) - \nabla_k(\delta\Gamma^l_{il}). \tag{11.216}$$

Poiché P è un punto arbitrario dello spaziotempo **l'equazione di Palatini** (11.216) **vale in generale in ogni punto della varietà $\mathcal{V}^4$ ed in ogni sistema di coordinate.**

Tornando alla (11.202), il 2° termine entro l'integrale, in virtù dell'equazione di Palatini, potrà ora scriversi così:

$$\sqrt{-g}g^{ik}\delta R_{ik} = \sqrt{-g}\left\{g^{ik}\nabla_l(\delta\Gamma^l_{ik}) - g^{il}\nabla_l(\delta\Gamma^k_{ik})\right\}, \tag{11.217}$$

avendo nel 2° termine, in parentesi graffe, cambiato l'indice muto k con l. Ricordando che la derivata covariante del tensore metrico, per un noto teorema di Ricci, è identicamente nulla, la (11.217) si può scrivere in quest'altro modo:

$$\sqrt{-g}g^{ik}\delta R_{ik} = \sqrt{-g}\,\nabla_l\left\{g^{ik}\delta\Gamma^l_{ik} - g^{il}\delta\Gamma^k_{ik}\right\} = \sqrt{-g}\,\nabla_l v^l, \tag{11.218}$$

avendo indicato con v^l il vettore che compare entro parentesi graffe.
La derivata covariante $\nabla_l v^l$, che figura nella (11.218), non è altro che la divergenza del campo vettoriale v^l, che abbiamo calcolato nel paragrafo 10.7.3 con la formula (10.55) che ora riscriviamo così:

$$\nabla_l\, v^l \equiv \frac{1}{\sqrt{-g}}\frac{\partial(\sqrt{-g}\,v^l)}{\partial x^l}. \tag{11.219}$$

La (11.218) in virtù della (11.219) diventa allora:

$$\sqrt{-g}g^{ik}\delta R_{ik} = \frac{\partial(\sqrt{-g}\,v^l)}{\partial x^l}. \tag{11.220}$$

Di conseguenza il 2° termine della (11.202) dà luogo all'integrale:

$$\int_\Omega g^{ik}\delta R_{ik}\sqrt{-g}\,d^4x = \int_\Omega \frac{\partial(\sqrt{-g}\,v^l)}{\partial x^l}d^4x. \tag{11.221}$$

Per un noto *teorema di Gauss* (generalizzato alle varietà riemanniane), tale integrale di volume può trasformarsi in un integrale esteso alla ipersuperficie $\Sigma \equiv \partial\Omega$ bordo o frontiera di Ω, cioè

$$\int_\Omega \frac{\partial(\sqrt{-g}\,v^l)}{\partial x^l}d^4x = \int_\Sigma \sqrt{-g}\,v^l\,d\Sigma. \tag{11.222}$$

Quest'ultimo integrale però è nullo, non appena si rammenti, per le (11.198) e (11.199), che sulla frontiera $\partial\Omega$ del dominio Ω, sono nulle le variazioni delle g^{ik} e delle loro derivate prime, e quindi anche le variazioni dei simboli di Christoffel che figurano nell'espressione di $v^l = g^{ik}\delta\Gamma^l_{ik} - g^{il}\delta\Gamma^k_{ik}$. Avremo perciò:

$$\int_\Sigma \sqrt{-g}\, v^l\, d\Sigma = 0.$$

Infine rimane da studiare il 5° termine della (11.202), cioè

$$\kappa\delta(g^{ik}T_{ik}\sqrt{-g}) = \kappa\delta g^{ik}T_{ik}\sqrt{-g} + \kappa g^{ik}\delta T_{ik}\sqrt{-g} + \kappa g^{ik}T_{ik}\delta\sqrt{-g}, \tag{11.223}$$

il 2° termine della somma a 2° membro non deve intervenire perché noi stiamo cercando le equazioni di campo, vale a dire le condizioni da imporre alle g^{ik} per una data distribuzione di materia-energia definita dal tensore energia-impulso T_{ik}, il quale si ritiene fissato, dunque non suscettibile di variazione, perciò $\delta T_{ik} = 0$. Avremo allora per la (11.223), e ricordando la (11.208), l'espressione

$$\kappa\,\delta(g^{ik}T_{ik}\sqrt{-g}) = \kappa\,\delta g^{ik}T_{ik}\sqrt{-g} - \frac{1}{2}\kappa\sqrt{-g}\,g_{ik}g^{ik}T_{ik}\,\delta g^{ik} = -\kappa\,T_{ik}\,\delta g^{ik}\sqrt{-g}. \tag{11.224}$$

Raccogliendo tutti i risultati precedentemente ottenuti, l'equazione variazionale (11.202) diventa infine:

$$\int_\Omega \left\{R_{ik} - \frac{1}{2}Rg_{ik} + \Lambda g_{ik} - \kappa T_{ik}\right\}\delta g^{ik}\sqrt{-g}\,d^4x = 0, \tag{11.225}$$

e siccome le δg^{ik} sono del tutto arbitrarie nel dominio 4-dimensionale Ω, per il Lemma fondamentale del Calcolo delle Variazioni, la precedente equazione variazionale implica le celeberrime:

Equazioni di Einstein del campo gravitazionale:

$$R_{ik} - \frac{1}{2}Rg_{ik} + \Lambda g_{ik} = \kappa T_{ik} \qquad (i,k = 0,1,2,3), \tag{11.226}$$

ovvero, ricordando la definizione del tensore di Einstein, più sinteticamente:

$$G_{ik} + \Lambda g_{ik} = \kappa T_{ik} \quad (i,k = 0,1,2,3). \tag{11.227}$$

Le equazioni gravitazionali (11.226) o (11.227), ricavate da un principio variazionale, coincidono perfettamente sia con la deduzione euristica che ne abbiamo fatto nel paragrafo 11.15.1, sia con la deduzione basata sul teorema di Cartan del paragrafo 11.15.2.

Esse soddisfano completamente il requisito di bellezza e semplicità sempre perseguito da Einstein e ribadito più volte da Paul A. M. Dirac! [60]

11.17 Tensore energia-impulso in RR ed in RG

I tensori energia-impulso più importanti sono sostanzialmente tre:

1) il tensore energia-impulso per **materia disgregata**;[61]

[60] A tal proposito si consiglia il lettore di consultare l'interessantissimo libro di Paul Adrien Maurice Dirac, *La bellezza come metodo, Saggi e riflessioni sulla fisica e matematica*, a cura di Vincenzo Barone, Indiana Editore srl, Milano, 2013.

[61] Come abbiamo già detto in una nota precedente, dicesi **materia disgregata** detta anche **materia incoerente non interagente, o polvere**, un sistema di particelle materiali, per il quale siano trascurabili, rispetto alle forze di volume, le forze di contatto, cioè in sostanza gli sforzi.

2) Il tensore energia-impulso per un **fluido perfetto**;

3) il tensore energia-impulso del **campo elettromagnetico**.

In vista degli sviluppi ai quali siamo interessati ci basterà occuparci soltanto dei primi due tensori.

1) **Tensore energia-impulso per materia disgregata**
Cominciamo col considerare, dapprima nell'ambito della Relatività Ristretta (RR), il tipo più semplice di campo materiale, cioè quello della **materia incoerente non interagente o materia disgregata**.
Tale campo è caratterizzato da due quantità fondamentali:
- il campo vettoriale delle 4-velocità di flusso delle singole particelle costituenti il campo di materia disgregata:

$$U^i = \frac{dx^i}{d\tau}, \tag{11.228}$$

dove τ è *il tempo proprio* lungo la linea di universo di una particella del flusso[62], e un
- campo scalare:

$$\mu_0 = \mu_0(x^i), \tag{11.229}$$

che descrive la **densità propria di massa propria** del flusso, cioè la densità che sarebbe misurata da un osservatore che si muovesse insieme al flusso (chiamato osservatore **comobile**, in inglese **comoving**).
Il più semplice tensore simmetrico di rango 2 che possiamo costruire a partire da queste due quantità è:

$$T^{ik} = \mu_0 U^i U^k, \quad (T^{ik} = T^{ki}) \tag{11.230}$$

che viene assunto, per definizione, come il *tensore energia-impulso del campo di materia disgregata.*[63]
Studiamo ora dettagliatamente questo tensore nell'ambito della Relatività Ristretta in coordinate di Minkowski (cioè in coordinate relative ad un sistema inerziale). Dalla

[62]Si tenga presente che molti autori (Per esempio: Landau-Lifschitz, Adler-Bazin-Schiffer, Carmeli, ecc.) definiscono la 4-velocità in maniera leggermente diversa da come l'abbiamo definita noi (seguendo Rindler, Cattaneo, Ray d'Inverno, ecc.), cioè tramite la (11.228). I primi autori citati pongono precisamente

$$u^i = \frac{dx^i}{ds}.$$

Osserviamo che con la nostra definizione (11.228) la 4-velocità U^i (indicata con una lettera maiuscola) ha le dimensioni di una effettiva velocità, mentre la u^i (indicata con lettera minuscola) è adimensionale, inoltre fra le due definizioni sussiste l'ovvia relazione $U^i = cu^i$. In unità relativistiche, cioè ponendo $c = 1$, le due definizioni, almeno formalmente, coincidono, avendosi $U^i = u^i$.

[63]Il tensore energia-impulso, dato dalla (11.230), è espresso in termini della 4-velocità definita in funzione del tempo proprio $d\tau$ e indicata con una lettera maiuscola, cioè $U^i = \frac{dx^i}{d\tau}$, ma è evidente che esso può esprimersi, seguendo Landau-Lifschitz, anche in termini della 4-velocità definita in funzione dell'intervallo spazio-temporale ds e indicata con una lettera minuscola, cioè $u^i = \frac{dx^i}{ds}$. Infatti, tenendo conto che $U^i = cu^i$, la (11.230) si potrà riscrivere così:

$$T^{ik} = \mu_0 c^2 u^i u^k.$$

È opportuno avvertire il lettore che alcuni autori, per esempio Adler-Bazin-Schiffer, definiscono il tensore energia-impulso, nel modo seguente:

$$\widetilde{T}^{ik} = \mu_0 u^i u^k.$$

Osserviamo che quest'ultimo, essendo le u^i, u^k adimensionali, ha le dimensioni di una *densità di massa*, mentre il tensore energia-impulso definito con la (11.230) ha le dimensioni di una *densità di energia*.

definizione (11.228) di 4-velocità di una particella avremo:

$$U^i = \frac{dx^i}{d\tau} = \left(\frac{dx^0}{d\tau}, \frac{dx^\alpha}{d\tau}\right) = \left(\frac{dx^0}{dt}\frac{dt}{d\tau}, \frac{dx^\alpha}{dt}\frac{dt}{d\tau}\right) = \left(c\gamma, v^\alpha\gamma\right) = \gamma(c, \mathbf{v}), \tag{11.231}$$

dove le $v^\alpha = \dfrac{dx^\alpha}{dt}$ sono le componenti della 3-velocità $\mathbf{v}$ della particella, $v = |\mathbf{v}|$ il suo modulo e $\gamma = (1 - \frac{v^2}{c^2})^{-1/2}$ il corrispondente **fattore di Lorentz**.
Il tempo proprio è definito da:

$$d\tau^2 = \frac{1}{c^2}ds^2 = \frac{1}{c^2}\left[c^2dt^2 - (dx^2 + dy^2 + dz^2)\right] =$$

$$= \frac{1}{c^2}\left[c^2dt^2 - dl^2\right] = \frac{1}{c^2}\left[c^2dt^2 - v^2dt^2\right] = dt^2 - \frac{v^2}{c^2}dt^2$$

cioè:

$$d\tau^2 = (1 - \beta^2)dt^2 = \gamma^{-2}dt^2, \tag{11.232}$$

con $\beta = \dfrac{v}{c}$ la cosiddetta velocità römeriana.
La componente T^{00} di T^{ik} vale allora:

$$T^{00} = \mu_0(U^0)^2 = \mu_0\left(\frac{dx^0}{d\tau}\right)^2 = \mu_0 c^2\gamma^2. \tag{11.233}$$

Questa quantità ha una interpretazione fisica semplice di cui ora ci occuperemo.
Prima di tutto, in RR, la massa di un corpo in moto è maggiore della sua massa a riposo di un fattore γ (essendo $\gamma > 1$), cioè:

$$m = m_0\gamma. \tag{11.234}$$

Inoltre, se consideriamo un elemento di volume cubico 3-dimensionale che a riposo abbia il lato uguale ad L_0 e perciò di volume $V_0 = L_0^3$, quando esso è in moto nella direzione di un suo spigolo questo subisce la contrazione di Lorentz, assumendo la lunghezza $\dfrac{L_0}{\gamma}$ (rispetto ad un osservatore fisso), mentre gli altri due lati, muovendosi ortogonalmente alla direzione di moto, non subiscono alcuna contrazione. Perciò il volume del cubo, rispetto ad un osservatore fisso, sarà

$$V = \frac{L_0}{\gamma} \cdot L_0 \cdot L_0 = \frac{L_0^3}{\gamma} = \frac{V_0}{\gamma}. \tag{11.235}$$

Per un'osservatore comobile col flusso di materia, la densità propria di massa propria è $\mu_0 = \dfrac{m_0}{V_0}$, mentre la densità misurata da un'osservatore fisso, tenuto conto delle (11.234) e (11.235), sarà:

$$\mu = \frac{m}{V} = \frac{m_0}{V_0}\gamma^2,$$

cioè:

$$\mu = \mu_0\gamma^2. \tag{11.236}$$

Se quindi un campo di materia di densità propria di massa propria μ_0 scorre con velocità $\mathbf{v}$ di fronte ad un osservatore fisso, quest'ultimo misurerà una densità alterata di un fattore

$\gamma^2(v)$ come dalla (11.236).
La componente T^{00} del tensore energia-impulso, espresso dalla (11.233), diventa:

$$T^{00} = \mu_0 c^2 \gamma^2 = \mu c^2 = \frac{mc^2}{V} = \frac{E}{V}, \tag{11.237}$$

e quindi può essere interpretata come *densità di energia relativistica* del campo di materia, perché il solo contributo all'energia del campo è il moto della materia.
Rappresentiamo le componenti del tensore T^{ik} con gli elementi della matrice

$$T^{ik} = \begin{pmatrix} T^{00} & T^{01} & T^{02} & T^{03} \\ T^{10} & T^{11} & T^{12} & T^{13} \\ T^{20} & T^{21} & T^{22} & T^{23} \\ T^{30} & T^{31} & T^{32} & T^{33} \end{pmatrix}, \tag{11.238}$$

se ora teniamo conto della (11.231) e delle (11.236), (11.237), gli elementi della precedente matrice saranno così espressi:

$$T^{00} = \mu_0 c^2 \gamma^2 = \mu c^2, \; T^{0\alpha} = T^{\alpha 0} = \mu_0 U^0 U^\alpha = \mu_0 c \gamma^2 v^\alpha = \mu c v^\alpha,$$
$$T^{\alpha\beta} = T^{\beta\alpha} = \mu_0 \gamma^2 v^\alpha v^\beta = \mu\, v^\alpha v^\beta, \tag{11.239}$$

dove $v^\alpha \equiv (v^1, v^2, v^3) \equiv (v_x, v_y, v_z)$ sono le componenti della 3-velocità $\mathbf{v}$ delle singole particelle del flusso.[64]
La (11.238), per le (11.239) si scriverà così:

$$T^{ik} = \begin{pmatrix} \mu c^2 & \mu c v_x & \mu c v_y & \mu c v_z \\ \mu c v_x & \mu v_x^2 & \mu v_x v_y & \mu v_x v_z \\ \mu c v_y & \mu v_x v_y & \mu v_y^2 & \mu v_y v_z \\ \mu c v_z & \mu v_x v_z & \mu v_y v_z & \mu v_z^2 \end{pmatrix} \tag{11.240}$$

cioè, raccogliendo a fattore comune il termine μc^2, possiamo riscrivere la precedente matrice:

$$T^{ik} = \mu c^2 \begin{pmatrix} 1 & v_x/c & v_y/c & v_z/c \\ v_x/c & v_x^2/c^2 & v_x v_y/c^2 & v_x v_z/c^2 \\ v_y/c & v_y v_x/c^2 & v_y^2/c^2 & v_y v_z/c^2 \\ v_z/c & v_z v_x/c^2 & v_z v_y/c^2 & v_z^2/c^2 \end{pmatrix}. \tag{11.241}$$

A questo punto osserviamo che, come la componente g_{00} del tensore metrico corrisponde al potenziale gravitazionale classico secondo la relazione approssimata $g_{00} \approx 1 + \frac{2\Phi}{c^2}$, così la componente T^{00} del tensore energia-impulso corrisponde alla densità di energia nello spazio secondo la relazione approssimata $T^{00} \approx \mu_0 c^2$. In entrambi i casi la validità

[64] È opportuno osservare che in RR, la 3-velocità di una particella l'abbiamo sempre indicata con la lettera minuscola $\mathbf{u} \equiv (u^\alpha)$ per non confonderla con la velocità di traslazione di un riferimento inerziale $\mathcal{R}$ rispetto ad un altro $\mathcal{R}'$ che abbiamo, generalmente, indicato con $\mathbf{v}$. In questo paragrafo, e nel seguito, ci discosteremo leggermente dalle precedenti notazioni in quanto la 3- velocità di una particella la indicheremo con $\mathbf{v} \equiv (v^\alpha)$, anziché con $\mathbf{u} \equiv (u^\alpha)$, e ciò per non fare confusione con la 4- velocità che, come abbiamo detto più sopra, viene indicata da alcuni autori, con la lettera minuscola u^i, $(i = 0, 1, 2, 3)$.

dell'approssimazione dipende dall'assumere *campi deboli e basse velocità* $v \ll c$, *cioè* $\gamma \simeq 1$.
Dimostriamo ora l'importantissimo:
Teorema di conservazione:
In assenza di forze, le equazioni che descrivono il moto di un campo di materia disgregata (precisamente le leggi di conservazione) si possono scrivere in modo molto conciso ed elegante, sotto la seguente forma tensoriale:

$$\partial_k T^{ik} = 0, \tag{11.242}$$

espressione che possiamo enunciare concisamente dicendo che: **è nulla la divergenza del tensore energia-impulso** T^{ik}.
Dimostrazione:
Iniziamo col porre $i = 0$ nella (11.242), avremo

$$\partial_k T^{0k} \equiv \partial_0 T^{00} + \partial_\alpha T^{0\alpha} = 0$$

cioè

$$\frac{\partial}{\partial x^0}(\mu c^2) + \frac{\partial}{\partial x^\alpha}(\mu c v^\alpha) = 0,$$

ed essendo $x^0 = ct$, con qualche semplificazione, otteniamo:

$$\frac{\partial \mu}{\partial t} + \frac{\partial(\mu v^\alpha)}{\partial x^\alpha} = 0,$$

che equivale all'equazione vettoriale 3-dimensionale:

$$\frac{\partial \mu}{\partial t} + \operatorname{div}(\mu \mathbf{v}) = 0, \tag{11.243}$$

si tratta dunque **dell'equazione classica di continuità**.
Nella dinamica classica dei fluidi, questa equazione esprime **la conservazione della massa di densità** μ **che si muove con velocità** $\mathbf{v}$.[65] Poiché in RR massa ed energia sono equivalenti, ne segue che **l'equazione di conservazione della massa-energia per materia disgregata (o incoerente) è**:

$$\partial_k T^{0k} = 0. \tag{11.244}$$

Analogamente si trova che le tre equazioni corrispondenti ad $i = \alpha$ $(\alpha = 1, 2, 3)$ sono:

$$\partial_k T^{\alpha k} \equiv \partial_0 T^{\alpha 0} + \partial_\beta T^{\alpha\beta} = \frac{\partial}{\partial x^0}(\mu c v^\alpha) + \frac{\partial}{\partial x^\beta}(\mu v^\alpha v^\beta) = 0$$

o più esplicitamente, essendo $x^0 = ct$, $(x^1, x^2, x^3) \equiv (x, y, z)$ e $(v^1, v^2, v^3) \equiv (v_x, v_y, v_z)$:

$$\frac{\partial}{\partial t}(\mu v^\alpha) + \frac{\partial}{\partial x}(\mu v^\alpha v_x) + \frac{\partial}{\partial y}(\mu v^\alpha v_y) + \frac{\partial}{\partial z}(\mu v^\alpha v_z) = 0, \quad (\alpha = 1, 2, 3).$$

Queste tre equazioni scalari possono essere compendiate in un'unica equazione vettoriale:

$$\frac{\partial(\mu \mathbf{v})}{\partial t} + \frac{\partial}{\partial x}(\mu v_x \mathbf{v}) + \frac{\partial}{\partial y}(\mu v_y \mathbf{v}) + \frac{\partial}{\partial z}(\mu v_z \mathbf{v}) = 0,$$

cioè

$$\frac{\partial \mu}{\partial t}\mathbf{v} + \mu\frac{\partial \mathbf{v}}{\partial t} + \mu\Big[\frac{\partial}{\partial x}(v_x \mathbf{v}) + \frac{\partial}{\partial y}(v_y \mathbf{v}) + \frac{\partial}{\partial z}(v_z \mathbf{v})\Big] + \Big[\frac{\partial \mu}{\partial x}v_x + \frac{\partial \mu}{\partial y}v_y + \frac{\partial \mu}{\partial z}v_z\Big]\mathbf{v} = 0,$$

[65] Per gli elementi di base della Meccanica dei sistemi continui, in particolare dei fluidi, si consiglia il lettore di consultare: **T. Levi Civita e U. Amaldi**, *Compendio di Meccanica Razionale*, vol.II°, Cap.IX, X, XI, XII, Zanichelli Editore, Bologna, 1951, ristampato nel 2012 da Edizioni Compomat nel III° volume delle *Lezioni di Meccanica razionale* degli stessi autori.

ed essendo $\frac{\partial \mu}{\partial x} v_x + \frac{\partial \mu}{\partial y} v_y + \frac{\partial \mu}{\partial z} v_z = \mathbf{v} \cdot \text{grad}\mu$, quest'ultima equazione può essere riscritta così:

$$\left[\frac{\partial \mu}{\partial t} + \mathbf{v} \cdot \text{grad}\mu\right]\mathbf{v} + \mu\left[\frac{\partial \mathbf{v}}{\partial t} + (\text{div}\,\mathbf{v})\mathbf{v} + v_x \frac{\partial \mathbf{v}}{\partial x} + v_y \frac{\partial \mathbf{v}}{\partial y} + v_z \frac{\partial \mathbf{v}}{\partial z}\right] = 0$$

o ancora:

$$\left[\frac{\partial \mu}{\partial t} + \mathbf{v} \cdot \text{grad}\mu + \mu\,\text{div}\mathbf{v}\right]\mathbf{v} + \mu\left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}\right] = 0.$$

Com'è noto dal calcolo vettoriale vale l'identità $\text{div}(\mu\,\mathbf{v}) = \mu\,\text{div}\,\mathbf{v} + \mathbf{v} \cdot \text{grad}\mu$, perciò l'espressione entro la prima parentesi quadra, nella precedente equazione, è $\frac{\partial \mu}{\partial t} + \text{div}(\mu\,\mathbf{v})$ ed allora l'equazione stessa si scrive:

$$\left[\frac{\partial \mu}{\partial t} + \text{div}(\mu\,\mathbf{v})\right]\mathbf{v} + \mu\left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}\right] = 0,$$

ma, per l'equazione di continuità (11.243), l'equazione si riduce alla:

$$\mu\left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}\right] \equiv \mu\mathbf{a} = 0, \tag{11.245}$$

essendo **a** l'espressione euleriana dell'accelerazione della generica particella.
Confrontando questa equazione con **l'equazione di moto di Eulero per un fluido perfetto**[66], in dinamica classica dei fluidi, cioè:

$$\mu\left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}\right] = \mu\,\mathbf{F} - \text{grad}\,p, \tag{11.246}$$

dove p è la pressione del fluido ed **F** è la forza per unità di massa, vediamo che la (11.245) non è altro che l'equazione di Eulero (11.246) in assenza di pressione e di forze di massa esterne (il che equivale a considerare materia disgregata), dunque, com'è noto, in Meccanica Classica dei fluidi, la (11.245) esprime **la conservazione della quantità di moto** del flusso di particelle di cui si sta discutendo.
Dalle considerazioni svolte più sopra, abbiamo stabilito che: **In RR la richiesta che il tensore energia-impulso abbia divergenza nulla equivale a richiedere la conservazione della massa-energia e dell'impulso (o quantità di moto) del campo di materia;**[67] da qui deriva il nome di **tensore energia-impulso**. La (11.242), cioè $\partial_k T^{ik} = 0$, è nota anche come **legge di conservazione dell'energia-impulso.**
Se in RR usiamo **coordinate generali (non galileiane)** allora la (11.242) viene sostituita dalla sua corrispondente in termini di derivata covariante:

$$\nabla_k T^{ik} = 0. \tag{11.247}$$

Eseguiamo ora la transizione alla RG, definendo ancora il tensore energia-impulso per la materia incoerente con la

$$T^{ik} = \mu_0 U^i U^k, \quad (T^{ik} = T^{ki}) \tag{11.248}$$

e usando *il Principio di Trascrizione* avremo proprio la (11.247) come espressione della **legge di conservazione di** T^{ik}.

[66]Nella Meccanica classica dei continui (consultare l'opera citata nella nota precedente, edizioni Compomat del 2012, pagg. 210-213) si dimostra che l'equazione fondamentale di moto, detta 1^a *equazione indefinita* è $\mu\mathbf{a} = \mu\,\mathbf{F} + \boldsymbol{\varnothing}$, dove $\mathbf{a} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}$ è l'accelerazione euleriana del generico elemento del continuo, **F** è la forza per unità di massa e $\boldsymbol{\varnothing}$ riassume l'azione delle forze molecolari. Per i **fluidi perfetti** si ha $\boldsymbol{\varnothing} = -\text{grad}\,p$, p essendo la pressione in seno al fluido, perciò *l'equazione fondamentale dell'idrodinamica pura* ossia *l'equazione di moto dei fluidi perfetti*, detta anche **equazione di Eulero** è: $\mu\mathbf{a} = \mu\,\mathbf{F} - \text{grad}\,p$. Si rammenti che: **fluidi ideali o perfetti** *si dicono quei sistemi continui, in cui gli sforzi si riducono a pressioni normali.*

[67]Non si dimentichi che nell'ambito della RR, abbiamo considerato un flusso di materia con basse velocità ($\gamma \simeq 1$), per cui $\mu_0 \simeq \mu = cost.$.

2) **Tensore energia-impulso per un fluido perfetto**
Un **fluido perfetto** è caratterizzato da 3 quantità:

a) *il campo vettoriale delle* **4-velocità**: $U^i = \dfrac{dx^i}{d\tau}$;

b) *il campo scalare della* **densità propria di massa propria**: $\mu_0 = \mu_0(x^i) \;\; (i = 0, 1, 2, 3)$;

c) *il campo scalare di* **pressione**: $p = p(x^i) \;\; (i = 0, 1, 2, 3)$.

Nel limite in cui $p \to 0$, il fluido perfetto si riduce a materia incoerente. Questo ci suggerisce di prendere il tensore energia-impulso per un fluido perfetto nella forma

$$T^{ik} = \mu_0 U^i U^k + p\, S^{ik} \tag{11.249}$$

dove S^{ik} è un tensore simmetrico da determinare. Gli unici tensori simmetrici di rango 2 associati al fluido sono $U^i U^k$ ed il tensore metrico g^{ik}, quindi l'ipotesi più semplice che possiamo avanzare per S^{ik} è:

$$S^{ik} = \alpha\, U^i U^k + \beta\, g^{ik} \tag{11.250}$$

dove α e β sono delle costanti da determinarsi in modo opportuno.
Analogamente al procedimento seguito nel punto **A)** per la materia incoerente, studiamo la legge di conservazione

$$\partial_k T^{ik} = 0 \tag{11.251}$$

in RR in coordinate di Minkowski e richiediamo che, in un limite appropriato, essa contenga sia l'equazione di continuità (11.243) sia l'equazione di Eulero (11.246) in assenza di forze di massa esterne, cioè l'equazione

$$\mu \left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = -\text{grad}\, p. \tag{11.252}$$

Si può dimostrare[68], ma per brevità ci limitiamo ad affermarlo, che questa richiesta porta ad imporre, nella (11.250), $\alpha = 1$ e $\beta = -1$. Le (11.249) e (11.250) danno allora la seguente espressione per il **tensore energia-impulso per un fluido perfetto**:

$$T^{ik} = (\mu_0 + p) U^i U^k - p\, g^{ik}. \tag{11.253}$$

Se in RR usiamo **coordinate generali (non galileiane)**, riteniamo che sia ancora valida la forma covariante della **legge di conservazione**, cioè:

$$\nabla_k T^{ik} = 0. \tag{11.254}$$

Anche nella RG consideriamo la (11.253) come definizione del tensore energia-impulso di un fluido perfetto e la (11.254) come **equazione di conservazione.**
La pressione p e la densità μ inoltre sono legate, com'è noto, dalla **equazione di stato** che descrive il particolare tipo di fluido perfetto che si sta considerando. In generale si tratta di un'equazione del tipo

$$p = p(\mu, T), \tag{11.255}$$

dove T è la temperatura assoluta; spesso si considerano fluidi in cui T è praticamente costante, per cui l'equazione di stato si riduce a

$$p = p(\mu). \tag{11.256}$$

Terminiamo qui questo breve cenno sui fluidi perfetti.

[68] Si può trovare la dimostrazione in Adler- Bazin- Schiffer, *Introduction to General Relativity*, Second Edition, pag. 334-338, Mc Graw- Hill, Kogakusha, LTD, Tokyo, 1975

11.18 Limite newtoniano delle equazioni gravitazionali

Sotto opportune ipotesi, che preciseremo fra un momento, studieremo ora il limite newtoniano delle equazioni di campo della Relatività Generale, senza il termine cosmologico (ponendo quindi $\Lambda = 0$), scritte covariantemente nella forma alternativa (11.176), cioè:

$$R_{ik} = \kappa \left(T_{ik} - \frac{1}{2} g_{ik} T \right), \qquad (i, k = 0, 1, 2, 3) \tag{11.257}$$

che si presta bene alle considerazioni che dovremo fare.
Nell'ordine di approssimazione che useremo, otterremo tre risultati di capitale importanza:

1) le equazioni gravitazionali di Einstein sono, come noi desideriamo, una generalizzazione della classica equazione di campo newtoniana, cioè l'equazione di Poisson

$$\Delta_2 \Phi(P) = 4\pi G \mu(P), \tag{11.258}$$

che si presenta dunque come caso limite delle (11.257);

2) risulterà stabilito il valore della costante einsteiniana κ di gravitazione ed il suo legame con l'analoga costante gravitazionale newtoniana G;

3) la legge newtoniana di gravitazione universale $F = G\frac{Mm}{r^2}$, essendo determinata dal potenziale $\Phi(P) = -\frac{GM}{r}$ soluzione dell'equazione di Poisson, è, nell'approssimazione considerata, contenuta nelle equazioni di campo di Einstein.

A tal fine, consideriamo un flusso di **materia disgregata che si muova a bassa velocità con densità propria di massa propria piccola.** Il tensore energia-impulso per questa situazione si può ottenere dalla sua espressione valida in RR, ed è espresso in forma controvariante, come abbiamo visto nel paragrafo precedente, dalla (11.241) che riscriviamo così:

$$T^{ik} = \mu_0 \gamma^2 c^2 \begin{pmatrix} 1 & v_x/c & v_y/c & v_z/c \\ v_x/c & v_x^2/c^2 & v_x v_y/c^2 & v_x v_z/c^2 \\ v_y/c & v_y v_x/c^2 & v_y^2/c^2 & v_y v_z/c^2 \\ v_z/c & v_z v_x/c^2 & v_z v_y/c^2 & v_z^2/c^2 \end{pmatrix}. \tag{11.259}$$

Ora però, dovendo usare le equazioni gravitazionali di Einstein (11.257) in forma covariante è necessario scrivere il tensore energia-impulso sotto forma covariante T_{ik}. Dobbiamo allora ricordarci che, per definizione, si ha $T_{ik} = \mu_0 U_i U_k$ con $U_i = \eta_{ik} U^k$, cioè $U_0 = U^0 = \gamma c, U_\alpha = -U^\alpha = -\gamma v^\alpha$, $(\alpha = 1, 2, 3)$, dalle quali è facile dedurre che $T_{00} = T^{00} = \mu_0 \gamma^2 c^2$, $T_{\alpha 0} = -T^{\alpha 0} = -\mu_0 U_\alpha U_0 = -\mu_0 \gamma^2 c v^\alpha = -\mu_0 \gamma^2 c^2 \frac{v^\alpha}{c}$, e analogamente $T_{\alpha\beta} = T^{\alpha\beta} = \mu_0 \gamma^2 c^2 \frac{v^\alpha}{c} \frac{v^\beta}{c}$.
Avremo allora che l'espressione analoga alla (11.259), ma in forma covariante, del tensore

energia-impulso sarà:

$$T_{ik} = \mu_0 \gamma^2 c^2 \begin{pmatrix} 1 & -v_x/c & -v_y/c & -v_z/c \\ -v_x/c & v_x^2/c^2 & v_x v_y/c^2 & v_x v_z/c^2 \\ -v_y/c & v_y v_x/c^2 & v_y^2/c^2 & v_y v_z/c^2 \\ -v_z/c & v_z v_x/c^2 & v_z v_y/c^2 & v_z^2/c^2 \end{pmatrix}. \tag{11.260}$$

L'ipotesi fatta di **bassa velocità**, implica intanto che $\beta^2 = \left(\frac{v}{c}\right)^2$ e termini di ordine superiore a β^2 **siano trascurabili** e perciò si abbia anche $\gamma = (1-\beta^2)^{-1/2} \simeq 1$, inoltre, per l'ipotesi di **densità propria di massa propria** μ_0 **piccola**, se nella (11.260) introduciamo tale densità entro la matrice moltiplicando per essa ciascun suo termine, avremo che tutti i termini ad eccezione del primo sono trascurabili perché prodotti di termini piccoli del tipo $-\mu_0\beta_x$, $-\mu_0\beta_y$, $-\mu_0\beta_z$, $\mu_0\beta_x^2, \mu_0\beta_x\beta_y, ecc.$, avendo indicato con $\beta_x = \frac{v_x}{c}, \beta_y = \frac{v_y}{c}, \beta_z = \frac{v_z}{c}$ le velocità römeriane relative alle componenti cartesiane v_x, v_y, v_z della 3-velocità **v** di ciascuna particella del flusso della data materia disgregata.
La matrice (11.260) si semplifica perciò nella

$$T_{ik} \approx c^2 \begin{pmatrix} \mu_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{11.261}$$

che esprime, nell'ordine delle approssimazioni operate, il tensore energia- impulso in forma covariante.
Assumeremo inoltre che **il flusso sia stazionario** e perciò il tensore metrico g_{ik} sia **indipendente dal tempo**. Usando le coordinate galileiane della RR, ct, x, y, z, consideriamo una metrica indipendente dal tempo che sia somma della metrica di Minkowski η_{ik} e di un piccolo termine perturbativo ϵh_{ik} anch'esso indipendente dal tempo, come abbiamo già fatto nel paragrafo 11.14 studiando il limite newtoniano per le equazioni delle geodetiche:

$$g_{ik}(x^\alpha) = \eta_{ik} + \epsilon h_{ik}(x^\alpha) \qquad (\alpha = 1, 2, 3). \tag{11.262}$$

Affrontiamo dapprima lo studio del 2° membro delle equazioni di campo (11.257).
Vediamo intanto come si calcola **l'invariante lineare** (detto anche **scalare di Von Laue**) $T = T^i_i = Tr(T^i_k)$ relativo al tensore (11.261). Scriviamo dunque il tensore energia-impulso nella sua forma mista generale (si tenga presente che il tensore metrico sotto forma controvariante è espresso, come dimostrato nella nota 49 del paragrafo 11.14, da $g^{im} = \eta^{im} - \epsilon h^{im}$):

$$T^i_k = g^{im} T_{mk} = (\eta^{im} - \epsilon h^{im}) T_{mk} = \eta^{im} T_{mk} - \epsilon h^{im} T_{mk},$$

perciò, operando una contrazione sugli indici i e k, avremo:

$$T = T^i_i = g^{im} T_{mi} = (\eta^{im} - \epsilon h^{im}) T_{mi} = \eta^{im} T_{mi} - \epsilon h^{im} T_{mi} =$$

$$= \eta^{00} T_{00} + 2\eta^{0\alpha} T_{0\alpha} + \eta^{\alpha\beta} T_{\alpha\beta} - \epsilon h^{00} T_{00} - 2\epsilon h^{0\alpha} T_{0\alpha} - \epsilon h^{\alpha\beta} T_{\alpha\beta}$$

dove

$$\eta^{00} = 1,\ \eta^{0\alpha} = \eta^{\alpha 0} = 0,\ \eta^{\alpha\alpha} = -1,\ T_{00} = \mu_0 c^2,\ T_{0\alpha} = T_{\alpha 0} = 0,\ T_{\alpha\beta} = T_{\beta\alpha} = 0,\ (\alpha, \beta = 1, 2, 3)$$

trascurando ancora termini di ordine $\epsilon\mu_0$, avremo

$$T = Tr(T^i_k) \approx Tr \begin{pmatrix} \mu_0 c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mu_0 c^2. \tag{11.263}$$

In quanto al termine $\frac{1}{2} g_{ik} T$ otteniamo

$$\frac{1}{2} g_{ik} T = \frac{1}{2}(\eta_{ik} + \epsilon h_{ik})\mu_0 c^2 = \frac{1}{2}(\eta_{ik}\mu_0 + \epsilon h_{ik}\mu_0)c^2 \approx \frac{1}{2}\eta_{ik}\mu_0 c^2,$$

la matrice associata a questo termine sarà dunque

$$\left(\frac{1}{2} g_{ik} T\right) \approx c^2 \begin{pmatrix} \mu_0/2 & 0 & 0 & 0 \\ 0 & -\mu_0/2 & 0 & 0 \\ 0 & 0 & -\mu_0/2 & 0 \\ 0 & 0 & 0 & -\mu_0/2 \end{pmatrix}.$$

Il 2° membro delle equazioni di campo (11.258) al primo ordine nelle piccole quantità μ_0, β, ϵh_{ik} sarà allora

$$\left(\kappa(T_{ik} - \frac{1}{2} g_{ik} T)\right) \approx \kappa \left\{ c^2 \begin{pmatrix} \mu_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} - c^2 \begin{pmatrix} \mu_0/2 & 0 & 0 & 0 \\ 0 & -\mu_0/2 & 0 & 0 \\ 0 & 0 & -\mu_0/2 & 0 \\ 0 & 0 & 0 & -\mu_0/2 \end{pmatrix} \right\} =$$

$$= \kappa c^2 \begin{pmatrix} \mu_0/2 & 0 & 0 & 0 \\ 0 & \mu_0/2 & 0 & 0 \\ 0 & 0 & \mu_0/2 & 0 \\ 0 & 0 & 0 & \mu_0/2 \end{pmatrix} = \frac{1}{2}\kappa\mu_0 c^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{2}\kappa\mu_0 c^2 \delta_{ik}. \tag{11.264}$$

Occupiamoci ora del 1° membro delle equazioni di campo (11.258); trascurando termini del 2° ordine in ϵh_{ik} si ottiene la seguente forma approssimata del tensore di Ricci[69]

$$R_{ik} \approx \partial_m \Gamma^m_{ik} - \frac{1}{2}\partial_{ik}\log(-g) \qquad (g = \det(g_{ik}) < 0). \tag{11.265}$$

Perciò, tenuto conto delle (11.264) e (11.265) le equazioni di campo (11.257) approssimate saranno

$$\partial_m \Gamma^m_{ik} - \frac{1}{2}\partial_{ik}\log(-g) \approx \frac{1}{2}\kappa\mu_0 c^2 \delta_{ik}. \tag{11.266}$$

Consideriamo dapprima l'equazione corrispondente agli indici $i = k = 0$, tenuto conto che $\delta_{00} = 1$, si ha

$$\partial_m \Gamma^m_{00} - \frac{1}{2}\partial_{00}\log(-g) = \frac{1}{2}\kappa\mu_0 c^2, \tag{11.267}$$

[69] Il tensore di Ricci sappiamo che si ottiene dal tensore di Riemann sotto forma mista del tipo $(1,3)$ contraendo 1° e 3° indice

$$R_{ik} = R^m_{.imk} = \partial_m \Gamma^m_{ik} - \partial_k \Gamma^m_{im} + \Gamma^m_{lm}\Gamma^l_{ik} - \Gamma^m_{lk}\Gamma^l_{im},$$

per le ipotesi fatte gli ultimi due termini sono trascurabili, il primo termine $\partial_m \Gamma^m_{ik}$ lo lasciamo inalterato, mentre il secondo termine diventa $\partial_k \Gamma^m_{im} = \frac{1}{2}\partial_{ik}\log(-g)$, tenuto conto che è $\Gamma^m_{im} = \frac{1}{2}\partial_i \log|g| = \frac{1}{2}\partial_i \log(-g)$, essendo, in RG, sempre $g = \det(g_{ik}) < 0$.
Osserviamo che il procedimento che stiamo seguendo in questo paragrafo, applicato alle equazioni gravitazionali (11.257) è fondato soprattutto sull'ipotesi (11.262) di campo debole e costituisce un primo esempio di **linearizzazione delle equazioni di campo.** Ci occuperemo di questo problema, in tutta la sua generalità, nell'APPENDICE alla fine del Cap.12.

poiché stiamo considerando una metrica indipendente dal tempo, il 2° termine del 1° membro della (11.267) è nullo, perciò avremo:

$$\partial_m(g^{mn}\Gamma_{00,n}) = \frac{1}{2}\kappa\mu_0 c^2. \tag{11.268}$$

Il simbolo di Christoffel di prima specie che figura nella (11.268) è così definito:

$$\Gamma_{00,n} = \frac{1}{2}(\partial_0 g_{0n} + \partial_0 g_{n0} - \partial_n g_{00}),$$

essendo il tensore metrico $g_{ik} = \eta_{ik} + \epsilon h_{ik}$ indipendente dal tempo ed il tensore metrico dello spazio di Minkowski costante sia nelle coordinate spaziali sia in quella temporale, la precedente espressione, per $n = 0$, si riduce alla $\Gamma_{00,0} = 0$, mentre per $n = \alpha$ si ha

$$\Gamma_{00,\alpha} = -\frac{\epsilon}{2}\partial_\alpha h_{00}.$$

Trascurando ancora termini del 2° ordine in ϵ, l'espressione entro parentesi nella (11.268), si approssima nel modo seguente:

$$g^{mn}\Gamma_{00,n} = g^{m0}\Gamma_{00,0} + g^{m\alpha}\Gamma_{00,\alpha} = (\eta^{m\alpha} - \epsilon h^{m\alpha})(-\frac{\epsilon}{2}\partial_\alpha h_{00}) =$$

$$= -\frac{\epsilon}{2}\eta^{m\alpha}\partial_\alpha h_{00} + \frac{\epsilon^2}{2}h^{m\alpha}\partial_\alpha h_{00} \simeq -\frac{\epsilon}{2}\eta^{m\alpha}\partial_\alpha h_{00}.$$

Sostituendo nel 1° membro dell'equazione di campo (11.268) otteniamo

$$\partial_m(g^{mn}\Gamma_{00,n}) = \partial_m(-\frac{\epsilon}{2}\eta^{m\alpha}\partial_\alpha h_{00}) =$$

$$= -\frac{\epsilon}{2}\Big\{\partial_0(\eta^{0\alpha}\partial_\alpha h_{00}) + \partial_1(\eta^{1\alpha}\partial_\alpha h_{00}) + \partial_2(\eta^{2\alpha}\partial_\alpha h_{00}) + \partial_3(\eta^{3\alpha}\partial_\alpha h_{00})\Big\} =$$

$$= -\frac{\epsilon}{2}\Big\{\partial_1[\eta^{11}\partial_1 h_{00} + \eta^{12}\partial_2 h_{00} + \eta^{13}\partial_3 h_{00}] + \partial_2[\eta^{21}\partial_1 h_{00} + \eta^{22}\partial_2 h_{00} + \eta^{23}\partial_3 h_{00}] +$$

$$+\partial_3[\eta^{31}\partial_1 h_{00} + \eta^{32}\partial_2 h_{00} + \eta^{33}\partial_3 h_{00}]\Big\} =$$

$$= -\frac{\epsilon}{2}\Big\{\partial_1[-\partial_1 h_{00}] + \partial_2[-\partial_2 h_{00}] + \partial_3[-\partial_3 h_{00}]\Big\} =$$

$$= \frac{\epsilon}{2}\Big\{\frac{\partial^2 h_{00}}{\partial(x^1)^2} + \frac{\partial^2 h_{00}}{\partial(x^2)^2} + \frac{\partial^2 h_{00}}{\partial(x^3)^2}\Big\} = \frac{\epsilon}{2}\Delta_2 h_{00},$$

e per la (11.267) l'equazione di campo corrispondente agli indici $i = k = 0$ può essere scritta nella forma

$$\Delta_2\Big(\frac{\epsilon h_{00}}{\kappa c^2}\Big) = \mu_0. \tag{11.269}$$

Se, in quest'ultima equazione, moltiplichiamo 1° e 2° membro per la costante $4\pi G$ otteniamo la classica equazione di Poisson

$$\Delta_2\Phi = 4\pi G\mu_0, \tag{11.270}$$

pur di porre

$$\Phi = 4\pi G\frac{\epsilon h_{00}}{\kappa c^2}. \tag{11.271}$$

Abbiamo così stabilito l'importantissimo risultato:
La teoria classica della gravitazione newtoniana è caso limite della teoria relativistica einsteiniana con metrica indipendente dal tempo e leggermente perturbata rispetto alla metrica

di Minkowski. Se combiniamo la (11.271) con il risultato già ottenuto nel paragrafo 11.14 con la formula (11.155), che fornisce il legame tra il potenziale classico ed il primo coefficiente del termine perturbativo della metrica, cioè

$$\Phi = \frac{c^2 \epsilon h_{00}}{2}, \tag{11.272}$$

troviamo un secondo notevole risultato: **il valore della costante di accoppiamento einsteiniana deve essere**[70]:

$$\kappa = \frac{8\pi G}{c^4}, \tag{11.273}$$

nel sistema C.G.S. il suo valore molto piccolo è approssimativamente

$$\kappa \approx 2.08 \times 10^{-48} \mathrm{cm}^{-1} \cdot \mathrm{g}^{-1} \cdot \mathrm{s}^2.$$

Perciò, con le precedenti considerazioni, risulta rispettato *il Principio di Corrispondenza*, il quale richiede che le nuove equazioni di campo debbano contenere, come caso limite, le equazioni classiche. Inoltre *siamo riusciti a determinare con precisione il legame tra la costante newtoniana di gravitazione G e la corrispondente costante einsteiniana* κ.
Per completezza il lettore può verificare che l'intero sistema delle equazioni di campo (11.266) *è soddisfatto dal tensore metrico*[71]

$$g_{ik} = \begin{pmatrix} 1+\epsilon h_{00} & 0 & 0 & 0 \\ 0 & -1+\epsilon h_{00} & 0 & 0 \\ 0 & 0 & -1+\epsilon h_{00} & 0 \\ 0 & 0 & 0 & -1+\epsilon h_{00} \end{pmatrix}$$

che ci fornisce una metrica approssimata della forma

$$ds^2 = \left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 - \left(1 - \frac{2\Phi}{c^2}\right) d\sigma^2,$$

dove $d\sigma^2 = dx^2 + dy^2 + dz^2$, la quale è dello stesso tipo del ds^2 relativo allo spazio vuoto attorno ad un corpo sferico omogeneo o formato da strati sferici omogenei che genera un campo a simmetria sferica come vedremo nel prossimo capitolo studiando la soluzione di Schwarzschild.
Il più importante risultato di questo paragrafo è l'aver determinato la costante einsteiniana di accoppiamento κ *ed il suo legame con l'analoga costante newtoniana G.*
Le equazioni di campo di Einstein nella loro forma covariante esplicita sono dunque:

$$R_{ik} - \frac{1}{2} g_{ik} R = \frac{8\pi G}{c^4} T_{ik}, \tag{11.274}$$

o equivalentemente in forma controvariante

$$R^{ik} - \frac{1}{2} g^{ik} R = \frac{8\pi G}{c^4} T^{ik}, \tag{11.275}$$

[70]Dobbiamo avvertire il lettore che in alcuni testi sulla Relatività Generale, per esempio su Adler-Bazin-Schiffer, si trova per la costante einsteiniana di gravitazione - che per non confonderla con la (11.273) indicheremo con $\widetilde{\kappa}$ - il valore $\widetilde{\kappa} = \frac{8\pi G}{c^2}$, ciò è dovuto al fatto che si assume per il tensore energia-impulso, anziché l'espressione $T^{ik} = \mu_0 U^i U^k$ che abbiamo utilizzato noi, equivalente alla $T^{ik} = \mu_0 c^2 u^i u^k$ - usata da Landau-Lifschitz e da molti altri autori - ed avente le dimensioni di una **densità di energia**, quest'altra $\widetilde{T}^{ik} = \mu_0 u^i u^k$, avente le dimensioni di una **densità di massa**, dove, si badi bene, le $u^i = \frac{dx^i}{ds}$ sono le componenti adimensionali della 4-velocità delle particelle e non le componenti della 3-velocità che abbiamo indicato con v^α.

[71]Si vedano i paragrafi 9.1 e 9.2 del Cap.9 del libro di Adler- Bazin-Schiffer, già citato nella nota 28 di pagina 169.

oppure ancora

$$R^i_k - \frac{1}{2}\delta^i_k R = \frac{8\pi G}{c^4} T^i_k, \tag{11.276}$$

nella forma mista.
Le equazioni di Einstein del campo gravitazionale possono infine scriversi anche nella 2^a forma convariante

$$R_{ik} = \frac{8\pi G}{c^4}(T_{ik} - \frac{1}{2}g_{ik}T) \tag{11.277}$$

con le analoghe in forma controvariante o mista.
Osservazione:
Aggiungiamo un'altra osservazione importante. Ricavato il teorema di Poisson (11.270) abbiamo stabilito che *la teoria classica newtoniana della gravitazione è caso limite della teoria relativistica con metrica quasi-minkowskiana indipendente dal tempo.* Una soluzione particolare dell'equazione di Poisson

$$\Delta_2\Phi(P) = 4\pi G\mu(P)$$

è data, com'è noto, da

$$\Phi(P) = -G\int_S \frac{\mu(Q)}{r(P,Q)}dS,$$

la soluzione generale si ottiene aggiungendo all'integrale precedente una arbitraria funzione armonica. Questa formula determina nell'approssimazione non relativistica, il potenziale del campo gravitazionale di qualsiasi distribuzione di massa in un volume S. In particolare, per il potenziale del campo gravitazionale generato da un corpo a simmetria sferica di massa M, nei punti P esterni alla sfera ed a distanza r dal suo centro, si ha:

$$\Phi(P) = -\frac{GM}{r}$$

e di conseguenza, la forza agente in questo campo su un altro corpo di massa m è in modulo $F = m\frac{\partial\Phi}{\partial r}$, cioè

$$F = G\frac{Mm}{r^2},$$

si ottiene dunque **la legge della gravitazione universale di Newton!**
Questo risultato è estremamente soddisfacente e può essere interpretato alla luce dei postulati della RG di Einstein dicendo che: **non esistono fra le masse dell'Universo azioni immediate a distanza e che perciò il moto dei corpi non è determinato da forze ma è causato dalla curvatura dello spaziotempo la quale è generata proprio dalla distribuzione e dal moto delle masse.**

11.19 Osservazioni sulla costante cosmologica Λ

Le originarie equazioni del campo gravitazionale scritte da Einstein furono, come già dicemmo, le (11.274), cioè:

$$R_{ik} - \frac{1}{2}g_{ik}R = \frac{8\pi G}{c^4}T_{ik} \qquad (i,k = 0,1,2,3). \tag{11.278}$$

Appena due mesi dopo la loro pubblicazione, Karl Schwarzschild fornì la prima soluzione esatta di tali equazioni, che si riferiva al campo gravitazionale del Sole, come vedremo dettagliatamente nel prossimo capitolo. Einstein stesso si impegnò presto nel compito di applicare la Relatività Generale all'universo nella sua totalità. A quel tempo si pensava che, salvo moti locali minori, l'universo fosse statico, perciò Einstein cercava naturalmente una soluzione statica delle sue equazioni di campo. Non ne trovò nessuna. Sembrava che la Relatività Generale

non si applicasse all'universo nella sua totalità. Einstein perciò modificò le equazioni (11.278) aggiungendo al primo membro un termine Λg_{ik}, oggi noto col nome di *termine cosmologico* in quanto la costante Λ che vi figura è chiamata *costante cosmologica* per un motivo che vedremo in seguito. La modifica doveva essere tale da non alterare l'accordo della teoria originale con i fatti astronomici del sistema solare. Questa condizione era soddisfatta dal termine cosmologico che, data l'estrema piccolezza della costante cosmologica ($|\Lambda| \leq 10^{-55}\,\mathrm{cm}^{-2}$), aveva un effetto trascurabile sul campo gravitazionale per piccole distanze (quali il diametro del sistema solare, o anche della Via Lattea) ma che per distanze grandi dal punto di vista cosmologico ha un'influenza notevole. La forma modificata delle equazioni di campo, comprendenti quindi il termine cosmologico, fornì ad Einstein quello che cercava, un modello di universo statico! Esso è detto *universo di Einstein* e fu il primo modello cosmologico.
Nello stesso anno 1917 l'astronomo olandese Willem de Sitter prese in considerazione le equazioni di campo con termine cosmologico e ne trovò un'importante soluzione. I particolari della soluzione di de Sitter saranno descritti in seguito nel capitolo sulla Cosmologia, ma possiamo anticipare che la soluzione si riferiva ad un universo vuoto, ma in espansione, nel senso che se vi viene introdotta una piccola particella, essa si allontana dall'osservatore proprio come fanno le galassie. Questa era la prima indicazione precisa che l'universo poteva essere in espansione. Va sottolineato che il modello apparve molti anni prima che le osservazioni di Slipher prima e di Hubble poi, nel 1929, mostrassero chiaramente che vi è una *recessione generale delle galassie*, come vedremo in seguito.
Si può affermare che la Relatività Generale ha fatto in questo caso una previsione teorica della più grande importanza. È appunto ciò che dovrebbe fare ogni buona teoria, dire agli osservatori che cosa cercare, invece di avere il senno di poi, e presentare una spiegazione ad osservazioni avvenute. Questo, in verità, fu il secondo successo della Relatività Generale. Il primo fu dovuto alla soluzione data da Schwarzschild alle equazioni di Einstein, che prevedeva la deviazione dei raggi di luce in un campo gravitazionale. È chiaro che per Einstein, la scoperta di de Sitter dev'essere stata un vero schock, e fu una delle ragioni che più tardi lo costrinsero ad abbandonare il termine cosmologico, ed a ritornare alle equazioni del campo originali. Sembra che Einstein abbia dapprima cercato la soluzione che poi de Sitter trovò, ma senza successo. Se Einstein l'avesse trovata, non avrebbe probabilmente proposto l'equazione di campo modificata, ma qualcun altro l'avrebbe certamente fatto in seguito.
La ripugnanza di Einstein per la soluzione di de Sitter sorse perché egli era profondamente contrario all'idea di un universo che, pur essendo vuoto, contenga tuttavia campi di forza (che ci siano tali campi nel modello di de Sitter si deduce dal moto delle particelle di prova). Einstein pensava che i campi di forza non possono esserci senza la materia e respinse pertanto fermamente, come non realistico, il modello di de Sitter e con esso il termine cosmologico che lo aveva originato.
Nel 1922 si ebbe un importante progresso quando il matematico e cosmologo russo Aleksandr Aleksandrovic Friedmann trovò tre modelli in espansione, che descriveremo nel capitolo dedicato alla Cosmologia. Come si è detto, questi soddisfacevano alle equazioni gravitazionali non modificate, e pertanto dimostravano che modelli in espansione possono ottenersi anche senza il termine cosmologico. Poco dopo la scoperta di Friedmann, gli astronomi cominciarono a capire che l'universo è in espansione, e ad Einstein e a molti altri parve non esserci più ragione per mantenere il termine Λg_{ik}, cioè il termine cosmologico, nelle equazioni di campo. Tale situazione portò Einstein ad esclamare: *"l'introduzione del termine cosmologico fu l'errore più grande della mia vita."* Non bisogna però dimenticare che il termine era stato introdotto per ottenere *un modello statico di universo*, ora eliminato dalle osservazioni.
Ciò nondimeno molti cosmologi hanno studiato modelli che si ottengono soltanto mantenendo il termine cosmologico. Alcuni di questi sono di grande interesse e due cosmologi importanti come Eddington e Lemaitre, hanno fermamente sostenuto che il termine cosmologico deve

essere conservato nelle equazioni di campo. I loro modelli saranno descritti nel capitolo sulla Cosmologia.

11.20 APPENDICE: Osservazioni su alcune convenzioni fondamentali

È opportuno far osservare al lettore che nell'immensa letteratura concernente la Relatività Generale, i vari autori purtroppo non usano le stesse convenzioni riguardanti le definizioni fondamentali del tensore di curvatura di Riemann-Christoffel $R^l_{.ijk}$ e del tensore di Ricci R_{ik} e conseguentemente dello scalare di curvatura R e del tensore di Einstein G_{ik}: ciò può condurre il lettore, che desidera consultare libri diversi, a dei fastidiosi malintesi.
Ci sono essenzialmente due modi per definire sia il tensore di Riemann-Christoffel che il tensore di Ricci, i quali differiscono fra di loro soltanto per il segno. Precisamente una prima definizione del tensore di Riemann-Christoffel (risalente a Tolman e Landau-Lifschitz) è:

$$R^l_{.ijk} = \partial_j \Gamma^l_{ik} - \partial_k \Gamma^l_{ij} + \Gamma^l_{mj}\Gamma^m_{ik} - \Gamma^l_{mk}\Gamma^m_{ij}, \qquad (1^a \text{ definizione}). \tag{11.279}$$

Questa definizione adotta la seguente convenzione: la prima derivata parziale, che compare a 2° membro col segno $+$, è fatta rispetto al <u>3° indice</u> di $R^l_{.ijk}$ indicato con j, mentre la seconda derivata, col segno $-$, è fatta rispetto al <u>4° indice</u> di $R^l_{.ijk}$ indicato con k; gli altri due termini $+\Gamma^l_{mj}\Gamma^m_{ik} - \Gamma^l_{mk}\Gamma^m_{ij}$, formati da somme di prodotti di simboli di Christoffel di seconda specie, sono costruiti come spiegato nel paragrafo 10.11.2 quando abbiamo definito il tensore di Riemann-Christoffel e la loro struttura rimane la stessa nelle due possibili definizioni. Noi abbiamo adottato questa prima definizione.
La seconda definizione è:

$$R^l_{.ijk} = \partial_k \Gamma^l_{ij} - \partial_j \Gamma^l_{ik} + \Gamma^l_{mk}\Gamma^m_{ij} - \Gamma^l_{mj}\Gamma^m_{ik}, \qquad (2^a \text{ definizione}) \tag{11.280}$$

questa, a differenza della prima, adotta la convenzione che: la prima derivata parziale, che compare a 2° membro col segno $+$, è fatta rispetto al <u>4° indice</u> di $R^l_{.ijk}$ indicato con k, mentre la seconda derivata, col segno $-$, è fatta rispetto al <u>3° indice</u> di $R^l_{.ijk}$ indicato con j.
Anche per il tensore di Ricci esistono due definizioni: la prima opera, sul tensore di Riemann-Christoffel $R^l_{.ijk}$, facendo una contrazione sul 1° e 3° indice, cioè su l e j:

$$R_{ik} \equiv R^l_{.ilk} = \partial_l \Gamma^l_{ik} - \partial_k \Gamma^l_{il} + \Gamma^l_{ml}\Gamma^m_{ik} - \Gamma^l_{mk}\Gamma^m_{il}, \qquad (1^a \text{ definizione}) \tag{11.281}$$

la seconda definizione opera, sul tensore di Riemann-Christoffel $R^l_{.ijk}$, facendo invece una contrazione sul 1° e 4° indice, cioè su l e k:

$$R_{ij} \equiv R^l_{.ijl} = \partial_j \Gamma^l_{il} - \partial_k \Gamma^l_{il} + \Gamma^l_{ml}\Gamma^m_{ik} - \Gamma^l_{mk}\Gamma^m_{il}, \qquad (2^a \text{ definizione}) \tag{11.282}$$

è perciò necessario, nella lettura di testi di vari autori, tenere nel debito conto queste osservazioni.

Capitolo 12

SOLUZIONE DI SCHWARZSCHILD E SUE CONSEGUENZE

Ho letto il suo lavoro con il massimo interesse.
Non mi sarei mai aspettato che si potesse determinare
la soluzione esatta del problema in modo così semplice.
La formulazione analitica mi sembra splendida.
(Lettera del 9 Gennaio 1916 di Albert Einstein a Karl Schwarzschild)

Appena due mesi dopo la presentazione del lavoro fondamentale di Einstein all'Accademia Prussiana delle Scienze (25 Novembre 1915) il matematico ed astronomo Karl Schwarzschild (direttore dell'Osservatorio astrofisico di Potsdam vicino a Berlino), trovò una soluzione esatta delle equazioni gravitazionali di Einstein.[1]
A tal proposito, ecco come si esprime il fisico e biografo di Einstein, Abraham Pais.
"Due mesi più tardi, il 16 Gennaio 1916, Einstein lesse una comunicazione all'Accademia Prussiana per conto di Karl Schwarzschild, che a quell'epoca prestava servizio nell'esercito tedesco sul fronte russo. La comunicazione conteneva la soluzione esatta per il campo gravitazionale isotropo statico di una massa puntiforme, il primo esempio di soluzione rigorosa delle equazioni di campo complete di Einstein. Il 24 Febbraio 1916, Einstein lesse un'altra comunicazione di Schwarzschild, che forniva questa volta la soluzione per una massa puntiforme nel campo gravitazionale di una sfera di fluido incomprimibile: è qui che il raggio di Schwarzschild viene introdotto per la prima volta. Il 29 Giugno 1916 Einstein tenne un discorso all'Accademia Prussiana in memoria del fisico tedesco, morto l'11 Maggio dopo una breve malattia contratta sul fronte russo. In quella occasione ricordò il grande talento di Schwarzschild e i suoi importanti contributi sia sperimentali che teorici, come pure i risultati da lui raggiunti (a partire dal 1909) come direttore dell'Osservatorio astrofisico di Potsdam. Il contributo di Schwarzschild, dichiarò a conclusione del suo intervento, avrebbe continuato a rappresentare uno stimolo per la scienza."[2]

[1]Il fondamentale lavoro di K. Schwarzschild è *Ueber das Gravitationsfeld einer Massenpunktes nach der Einsteinchen Theorie*, Sitzungsberichte Akad. Berlin, 1916, pagg. 189-196; il famoso $\mathbf{ds}^2$ è dato a pag.194. Per l'esattezza storica (anche se la quasi totalità degli autori non ne fa menzione) è doveroso rilevare che la soluzione trovata da Schwarzschild fu trovata, qualche mese dopo ed indipendentemente, anche dal matematico olandese Johannes Droste nel lavoro *Het veld van een enkel centrum in Einstein's teorie der zwaartekracht en de beweging van een stoffelijk punt in dat veld*, Versl. gewone Vergad. Akad. Amsterdam, t.25, mai 1916, pagg. 163-180. Veramente Droste aveva già affrontato un anno prima il problema della ricerca di soluzioni delle equazioni gravitazionali di Einstein nel lavoro *On the field of a single center in Einstein's theory of gravitation*, Proc. Akad. Sc. Amsterdam, t.17, juin 1915, pagg. 998-1026, utilizzando però una loro versione non definitiva e non corretta pubblicata precedentemente da Einstein e M. Grossman, che portava evidentemente ad un risultato errato per il $\mathbf{ds}^2$.

[2]Cfr. Abraham Pais, *Sottile è il Signore....*, Boringhieri, Torino 1987, pag. 277 (sesta riga a partire dal basso).

12.1 Metrica indipendente dal tempo, a simmetria radiale

Studieremo ora la prima applicazione concreta e rigorosa della teoria einsteiniana della gravitazione a cui abbiamo accennato alla fine del capitolo precedente: il cosiddetto **problema esterno di Schwarzschild**, *che permette di spiegare in maniera piuttosto semplice, anche quantitativamente, il fenomeno di* **precessione del perielio dei pianeti**, *in particolare di Mercurio, e la* **deflessione** *che i raggi luminosi provenienti da una stella lontana subiscono per effetto del campo gravitazionale del Sole*. Per l'esattezza storica dobbiamo dire che questi fenomeni vennero spiegati da Einstein, attraverso calcoli laboriosi basati su complessi procedimenti di approssimazione, senza conoscere la soluzione di Schwarzschild la quale venne scoperta appena qualche mese dopo la pubblicazione del suo lavoro fondamentale sulla Relatività Generale.
Ammettiamo che il campo gravitazionale esistente nella varietà spaziotemporale $\mathcal{V}^4$ sia generato da un'unica massa M, **a simmetria sferica o radiale** e proponiamoci di determinare, a partire dall'equazione di campo di Einstein scritta in forma covariante:

$$R_{ik} - \frac{1}{2} R g_{ik} = \kappa T_{ik} \qquad (i, k = 0, 1, 2, 3) \tag{12.1}$$

e priva del termine cosmologico $(\Lambda = 0)$, l'espressione completa della metrica di $\mathcal{V}^4$ all'esterno di tale massa. Ma all'esterno di M lo spazio è vuoto, cioè privo di materia e perciò il tensore energia-impulso è nullo $T_{ik} = 0$, ed allora la precedente equazione gravitazionale si riduce nel vuoto alla forma semplice [3]

$$R_{ik} = 0, \tag{12.2}$$

che scriveremo esplicitamente così:

$$R_{ik} \equiv R^l_{ilk} = \partial_l \Gamma^l_{ik} - \partial_k \Gamma^l_{il} + \Gamma^l_{ml} \Gamma^m_{ik} - \Gamma^l_{mk} \Gamma^m_{il} = 0. \tag{12.3}$$

L'equazione di campo nel vuoto (12.3), nonostante l'apparente semplicità è un'equazione alle derivate parziali non lineare le cui funzioni incognite sono i coefficienti g_{ik} della metrica dello spaziotempo $\mathcal{V}^4$, ed essendo non lineare è difficile da risolvere. Tuttavia, imponendo delle *condizioni di simmetria* dettate da considerazioni fisiche sul tensore metrico g_{ik} e quindi sulla metrica

$$\mathbf{ds}^2 = g_{00}(x^i)(dx^0)^2 + 2g_{0\alpha}(x^i)dx^0 dx^\alpha + g_{\alpha\beta}(x^i)dx^\alpha dx^\beta \qquad (i = 0, 1, 2, 3;\ \alpha, \beta = 1, 2, 3),$$

possiamo semplificare fortemente tale equazione ed addirittura determinare una soluzione esatta: che è appunto la **soluzione di Schwarzschild.**
Data la presenza in $\mathcal{V}^4$ di un'unica massa dotata di **simmetria sferica** si è naturalmente portati ad ammettere che di analoga simmetria spaziale sia dotata la varietà $\mathcal{V}^4$. Ma che cosa significa che $\mathcal{V}^4$ abbia simmetria spaziale sferica?
Def.1ª : *Diremo che la $\mathcal{V}^4$ è dotata di* **simmetria spaziale sferica** *se esiste un sistema di coordinate (fisicamente ammissibili) $x^0 = ct, x^1 = x, x^2 = y, x^3 = z$, tale che di fronte ad un cambiamento di coordinate puramente spaziale del tipo*

$$x^{\alpha'} = A^{\alpha'}_{\alpha} x^\alpha, \qquad (\alpha, \alpha', = 1, 2, 3) \tag{12.4}$$

dove la matrice $A^{\alpha'}_{\alpha}$ è una matrice ortogonale[4]*, soddisfacente perciò alla $\sum_{\alpha'} A^{\alpha'}_{\alpha} A^{\alpha'}_{\beta} = \delta_{\alpha\beta}$, il suo $\mathbf{ds}^2$ risulti formalmente invariante.*

[3] Se nell'equazione di campo (12.1) poniamo $T_{ik} = 0$ otteniamo che sarà nullo il tensore di Einstein, cioè $R_{ik} - \frac{1}{2} R g_{ik} = 0$. Se questa equazione la riscriviamo sotto forma mista, avremo $R^k_i - \frac{1}{2} R \delta^k_i = 0$, contraendo ora sugli indici i, k si ottiene $R = 0$, dunque sarà nullo anche il tensore di Ricci: $R_{ik} = 0$ cioè la (12.2).

[4] Si ricordi che una matrice si dice **ortogonale** se la sua trasposta coincide con la sua inversa: $A^t = A^{-1}$.

È quasi evidente che il più generale $\mathbf{ds}^2$ a simmetria spaziale sferica potrà costruirsi mediante gli invarianti spaziali finiti e differenziali del gruppo delle rotazioni dello spazio euclideo E^3, cioè, introducendo in E^3, accanto alle coordinate cartesiane x, y, z, anche le coordinate polari sferiche r(raggio vettore), θ (colatitudine), φ (longitudine) con polo nel centro O della massa sferica, tali invarianti sono[5]:

$$x^2 + y^2 + z^2 = r^2, \quad dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2, \quad xdx + ydy + zdz = rdr. \tag{12.5}$$

In luogo dei precedenti invarianti si potranno assumere più semplicemente questi altri fra loro indipendenti

$$r, \quad dr, \quad r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{12.6}$$

Inoltre ammettiamo le seguenti definizioni:
Def.2ª : *Diremo che un* **campo gravitazionale è stazionario** *se nella varietà spaziotemporale* $\mathcal{V}^4$ *esiste un sistema di coordinate fisicamente ammissibili per le quali i potenziali gravitazionali verificano la condizione*

$$\partial_0 g_{ik} = 0, \qquad (i, k = 0, 1, 2, 3) \tag{12.7}$$

cioè **la metrica è indipendente dal tempo.**
Def.3ª : *Diremo invece che un* **campo gravitazionale è statico** *se sono verificate le due condizioni*

$$\partial_0 g_{ik} = 0, \quad (i, k = 0, 1, 2, 3), \qquad g_{0\alpha} = g_{\alpha 0} = 0, \quad (\alpha = 1, 2, 3). \tag{12.8}$$

Cioè un campo statico è un campo stazionario per il quale si annullano i coefficienti misti $g_{0\alpha} = g_{\alpha 0}$ della metrica di $\mathcal{V}^4$. Quest'ultima condizione significa **invarianza rispetto ad una inversione temporale**, cioè il $\mathbf{ds}^2$ deve essere invariante rispetto allo scambio di dx^0 con $-dx^0$, il che comporta che devono essere nulli i termini del tipo $2g_{0\alpha}dx^0dx^\alpha$, cioè appunto $g_{0\alpha} = g_{\alpha 0} = 0$. La metrica dello spaziotempo $\mathcal{V}^4$, all'esterno della sfera di massa M e centro O, sarà perciò, nelle coordinate x^0, x^1, x^2, x^3:

$$\mathbf{ds}^2 = g_{00}(x^\delta)(dx^0)^2 + g_{\alpha\beta}(x^\delta)dx^\alpha dx^\beta, \qquad (\alpha, \beta, \delta = 1, 2, 3). \tag{12.9}$$

Se lavoriamo in coordinate polari sferiche possiamo porre $x^0 = ct, x^1 = r$(raggio vettore), $x^2 = \theta$ (colatitudine), $x^3 = \varphi$ (longitudine).
La simmetria spaziale sferica o radiale implica ovviamente che i coefficienti della metrica (12.9), oltre ad essere indipendenti dal tempo in virtù delle (12.8), dipendano soltanto dalla coordinata radiale $x^1 = r$, cioè:

$$g_{00} = g_{00}(r), \quad g_{\alpha\beta} = g_{\alpha\beta}(r), \qquad (\alpha, \beta = 1, 2, 3). \tag{12.10}$$

La stessa simmetria radiale implica anche **isotropia spaziale**, cioè **mancanza di direzioni privilegiate** e perciò, in coordinate polari, ciò significa invarianza del $\mathbf{ds}^2$ di fronte allo scambio di $d\theta$ in $-d\theta$ e, separatamente, di $d\varphi$ in $-d\varphi$, il che comporta che la metrica non può contenere termini del tipo $dr\,d\theta$, $dr\,d\varphi$, $d\theta\,d\varphi$, cioè si avrà $g_{\alpha\beta} = 0$ con $\alpha \neq \beta = 1, 2, 3$; la matrice delle g_{ik} è dunque diagonale e la metrica si riduce alla forma

$$\mathbf{ds}^2 = g_{00}(r)(dx^0)^2 + g_{\alpha\alpha}(r)(dx^\alpha)^2. \tag{12.11}$$

[5]Le coordinate polari sferiche sono legate alle coordinate cartesiane dalle relazioni:

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta, \quad (r \geq 0, \ -\pi/2 \leq \theta \leq +\pi/2, \ 0 \leq \varphi < 2\pi).$$

Inoltre, a grande distanza dalla massa sferica M che genera il campo gravitazionale, ci aspettiamo che la metrica di $\mathcal{V}^4$ si riduca ad una metrica di Minkowski, che in coordinate polari si esprime così[6]:

$$\mathbf{ds}^2 = c^2dt^2 - d\sigma^2, \qquad con \quad d\sigma^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{12.12}$$

Se ora, per semplicità di scrittura, poniamo $g_{00}(r) = A(r)$ e $g_{\alpha\alpha}(r) = B(r)$, dal complesso delle condizioni, di natura euristica, fatte più sopra possiamo perciò presumere che la metrica che stiamo cercando, in coordinate polari, abbia la forma[7]:

$$\mathbf{ds}^2 = A(r)c^2dt^2 - B(r)dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\varphi^2, \tag{12.13}$$

con

$$\lim_{r\to+\infty} A(r) = 1, \qquad \lim_{r\to+\infty} B(r) = 1, \tag{12.14}$$

in modo tale che a distanze molto grandi da M essa si riduca alla metrica di Minkowski (12.12). La metrica (12.13) contiene soltanto due funzioni incognite puramente radiali $A(r), B(r)$ soggette alle sole condizioni all'infinito (12.14).
Osserviamo che il determinante della metrica (12.13) è

$$g = \det(g_{ik}) = \det\begin{pmatrix} A(r) & 0 & 0 & 0 \\ 0 & -B(r) & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix} = -A(r)B(r)r^4\sin^2\theta. \tag{12.15}$$

Affinché si possa dire chiaramente quale sia la segnatura della metrica (12.13), cioè quali siano i segni dei suoi coefficienti g_{ik}, basta stabilire quali siano i segni delle due funzioni incognite $A(r)$ e $B(r)$ in quanto i segni dei rimanenti due termini sono manifestamente negativi.
A tal fine, osserviamo che il segno di g, come già sappiamo, deve essere negativo $g < 0$ e perciò, per la (12.15), $A(r)$ e $B(r)$ devono essere o entrambi positivi o entrambi negativi. Ma la funzione $A(r)$ deve essere necessariamente positiva in quanto il primo termine della metrica (12.9), cioè $g_{00}(dx^0)^2 = g_{00}c^2dt^2$ deve coincidere col primo termine della stessa metrica scritta nella forma (12.13), ciò implica $g_{00} = A(r)$. Ma già sappiamo che $g_{00} > 0$,[8] perciò deve essere necessariamente $A(r) > 0$. Ne segue che anche $B(r)$ deve essere positivo.
Dunque tali funzioni si possono rappresentare entrambe con funzioni esponenziali:

$$A(r) = e^{\nu(r)} > 0, \qquad B(r) = e^{\lambda(r)} > 0, \tag{12.16}$$

ed allora la metrica (12.13) assume la forma definitiva

$$\mathbf{ds}^2 = e^{\nu(r)}c^2dt^2 - e^{\lambda(r)}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\varphi^2, \tag{12.17}$$

della quale dobbiamo determinare soltanto le due funzioni incognite $e^{\nu(r)}$ e $e^{\lambda(r)}$ assoggettate alle condizioni (12.14) che ora possono essere scritte così:

$$\lim_{r\to+\infty} e^{\nu(r)} = 1, \qquad \lim_{r\to+\infty} e^{\lambda(r)} = 1. \tag{12.18}$$

La (12.17) rappresenta la forma finale della metrica che useremo per ottenere la soluzione esatta di Schwarzschild delle equazioni di campo di Einstein. Essa, dal modo come l'abbiamo

[6]Il $d\sigma^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$ non è altro che la metrica dello spazio euclideo ordinario E^3 in coordinate polari sferiche con polo nel centro O della data massa sferica.

[7]Il ragionamento fatto sopra è piuttosto approssimato; per considerazioni più precise e dettagliate si veda il libro citato più volte di Adler- Bazin-Schiffer, *Introduction to General Relativity*, Second edition, pag. 186-187, McGraw Hill, Tokyo, Kogagusha, LTD, 1975.

[8]Cfr. la formula (11.38) del Cap.11, paragrafo 11.8.

costruita, soddisfa pienamente le richieste di **indipendenza dal tempo**, di **simmetria radiale**, e di **isotropia spaziale**.
In ogni caso la forma (12.17) della metrica potrebbe essere assunta a priori come un **ansatz**[9] per la ricerca della soluzione di Schwarzschild.
La coordinata radiale r usata nella (12.17) ha un chiaro significato fisico. Consideriamo una superficie sferica definita da un valore costante di r (il che implica $dr = 0$), i cui punti sono individuati dalle due coordinate angolari θ e φ. L'elemento di linea su questa superficie ad un dato istante t, in cui è $dt = 0$, sarà dato per la (12.17) da

$$\mathbf{ds}^2 = -r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{12.19}$$

La lunghezza fisica dell'equatore, definito come la linea $\theta = \dfrac{\pi}{2}$, si ottiene integrando $\sqrt{|\mathbf{ds}^2|} = \sqrt{-\mathbf{ds}^2}$ da $\varphi = 0$ a $\varphi = 2\pi$, avremo allora

$$\ell = \int_0^{2\pi} \sqrt{-\mathbf{ds}^2} = \int_0^{2\pi} r d\varphi = 2\pi r. \tag{12.20}$$

Questo risultato è identico a quello dello spazio piatto per una superficie sferica. Perciò per mezzo della misura della lunghezza fisica di un cerchio massimo noi possiamo determinare il valore della coordinata r per la sfera considerata. Similmente è facile mostrare che l'area fisica di tale sfera è $4\pi r^2$, come in uno spazio piatto, la quale ci permette ancora di determinare il valore di r. *È perciò chiaro che r è geometricamente definito, ed, inoltre, le tre coordinate spaziali r, θ, φ corrispondono da vicino alle variabili usate dagli astronomi nelle osservazioni reali.*

12.2 Simboli di Christoffel dedotti dalle equazioni delle geodetiche

Anche con la forma semplificata (12.17) della metrica, il lavoro per calcolare i 40 simboli di Christoffel che appaiono nelle equazioni di campo è piuttosto lungo e noioso. Esiste tuttavia un semplice e conveniente artificio che possiamo usare per ottenere tutti i simboli di Christoffel non nulli.
Le equazioni di Eulero-Lagrange per le geodetiche di una varietà riemanniana, com'è lo spaziotempo $\mathcal{V}^4$, nella forma (10.106) stabilita nel paragrafo 10.9, cioè:

$$\ddot{x}^i + \Gamma^i_{hk}\dot{x}^h\dot{x}^k = 0, \qquad \left(\dot{x}^i \equiv \frac{dx^i}{ds},\ \ddot{x}^i \equiv \frac{d^2x^i}{ds^2}, \quad i,h,k = 0,1,2,3\right) \tag{12.21}$$

contengono tutti i simboli di Christoffel di 2^a specie che ci occorrono.
Inversamente, se noi conoscessimo le equazioni di Eulero-Lagrange per le geodetiche, potremmo identificare facilmente tutti i simboli di Christoffel nulli e non nulli.
Le equazioni di Eulero-Lagrange, com'è noto, sono equivalenti alla seguente equazione variazionale

$$\delta \int ds = 0,$$

nella quale ds è l'elemento lineare definito dalla $ds \stackrel{\text{def}}{=} \sqrt{|\mathbf{ds}^2|} > 0$, essendo $\mathbf{ds}^2$ la metrica dello spaziotempo, che sappiamo essere indefinita cioè $\mathbf{ds}^2 \gtrless 0$ (il caso $\mathbf{ds}^2 = 0$ va trattato a parte).

[9] In fisica e in matematica il termine tedesco **ansatz** viene utilizzato per indicare una ipotesi o una assunzione, che pare ragionevole assumere tenendo conto dei dati a disposizione e **la cui validità deve essere verificata successivamente dopo aver risolto il problema**. Il termine tedesco **ansatz** significa appunto "approccio" o "tentativo".

Se ora introduciamo un piccolo trucco ponendo $\dfrac{\mathbf{ds}^2}{\mathbf{ds}^2} \equiv \dfrac{ds^2}{ds^2} = 1$, la precedente equazione variazionale può essere scritta così:

$$\delta \int \frac{ds^2}{ds^2} ds = 0. \tag{12.22}$$

Questo trucco si rivelerà utile nelle considerazioni che seguono.
Infatti, per la (12.17) si ha

$$ds = \left[e^{\nu(r)} (dx^0)^2 - e^{\lambda(r)} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\varphi^2 \right]^{1/2}, \tag{12.23}$$

ovvero

$$\frac{ds^2}{ds^2} = 1 = e^{\nu(r)} (\dot{x}^0)^2 - e^{\lambda(r)} \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 \sin^2\theta \dot{\varphi}^2, \tag{12.24}$$

e perciò il principio variazionale (12.22) si può scrivere più semplicemente così:

$$\delta \int \left[e^{\nu(r)} (\dot{x}^0)^2 - e^{\lambda(r)} \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 \sin^2\theta \dot{\varphi}^2 \right] ds = 0. \tag{12.25}$$

L'integrando della (12.25) può essere evidentemente interpretato come una lagrangiana

$$\mathcal{L}(r, \theta, \dot{x}^0, \dot{r}, \dot{\theta}, \dot{\varphi}) = e^{\nu(r)} (\dot{x}^0)^2 - e^{\lambda(r)} \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 sin^2\theta \dot{\varphi}^2, \tag{12.26}$$

e perciò possiamo scrivere le equazioni di Eulero-Lagrange, per il nostro problema variazionale, nella forma

$$\frac{d}{ds}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i} \right) - \frac{\partial \mathcal{L}}{\partial x^i} = 0. \qquad (i = 0, 1, 2, 3). \tag{12.27}$$

Osserviamo che nell'equazione (12.27) abbiamo indicato le coordinate che figurano nella lagrangiana (12.26) così: $x^0 = ct, x^1 = r, x^2 = \theta, x^3 = \varphi$, inoltre x^0 e φ non figurano esplicitamente in $\mathcal{L}$ e perciò sono coordinate **cicliche o ignorabili** e, com'è noto, forniscono due integrali primi. Un terzo integrale primo è costituito dall'espressione (12.24) che riscriviamo sotto la forma

$$e^{\nu(r)} (\dot{x}^0)^2 - e^{\lambda(r)} \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 \sin^2\theta \dot{\varphi}^2 - 1 = 0. \tag{12.28}$$

Scriveremo le equazioni (12.27) in forma esplicita e le confronteremo con le (12.21), in tal modo è facile determinare i simboli di Christoffel non nulli, *ed è proprio in tale confronto che consiste l'artificio più sopra menzionato.*
L'equazione di Eulero-Lagrange per la coordinata $x^i = x^0$ si ottiene immediatamente dalla (12.27) perchè x^0 è *una coordinata ciclica* per $\mathcal{L}$ e perciò $\dfrac{\partial \mathcal{L}}{\partial x^0} = 0$, inoltre si ha $\dfrac{\partial \mathcal{L}}{\partial \dot{x}^0} = 2e^{\nu(r)} \dot{x}^0$, per cui la corrispondente equazione di Eulero-Lagrange è $\dfrac{d}{ds}(2e^{\nu(r)} \dot{x}^0) = 0$ che fornisce l'integrale primo

$$e^{\nu(r)} \dot{x}^0 = cost. = k_2. \tag{12.29}$$

Indicando la derivazione rispetto ad r con un apice $(')$, l'equazione di Lagrange diventa:

$$e^{\nu(r)} \left[\ddot{x}^0 + \nu' \dot{r} \dot{x}^0 \right] = 0,$$

ed essendo $e^{\nu(r)} \neq 0$, si ha

$$\ddot{x}^0 + \nu' \dot{r} \dot{x}^0 = 0. \tag{12.30}$$

Questa è un'equazione differenziale ordinaria del 2° ordine particolarmente semplice perché la metrica è indipendente dal tempo. Confrontandola con la (12.21) per $i = 0$, avremo:

$$\ddot{x}^0 + 2\Gamma^0_{10} \dot{x}^1 \dot{x}^0 + ... = 0 \qquad \Longleftrightarrow \qquad \ddot{x}^0 + \nu' \dot{x}^1 \dot{x}^0 = 0, \tag{12.31}$$

che implica $2\Gamma^0_{10} = \nu'$, da cui

$$\Gamma^0_{10} = \Gamma^0_{01} = \frac{1}{2}\nu', \tag{12.32}$$

tutti gli altri simboli di Christoffel che intervengono nella (12.21), per $i = 0$, sono evidentemente nulli.
Essendo la lagrangiana del nostro problema

$$\mathcal{L} = e^{\nu(r)}(\dot{x}^0)^2 - e^{\lambda(r)}\dot{r}^2 - r^2\dot{\theta}^2 - r^2\sin^2\theta\dot{\varphi}^2,$$

in maniera analoga a quanto fatto per x^0, per la variabile $x^1 = r$, si deducono le seguenti espressioni

$$\frac{\partial\mathcal{L}}{\partial\dot{r}} = -2e^{\lambda(r)}\dot{r}; \quad \frac{d}{ds}\Big(\frac{\partial\mathcal{L}}{\partial\dot{r}}\Big) = \frac{d}{ds}\Big(-2e^{\lambda(r)}\dot{r}\Big) = -2e^{\lambda(r)}\ddot{r} - 2e^{\lambda(r)}\lambda'\dot{r}^2;$$

$$\frac{\partial\mathcal{L}}{\partial r} = e^{\nu(r)}\nu'(\dot{x}^0)^2 - e^{\lambda(r)}\lambda'\dot{r}^2 - 2r\dot{\theta}^2 - 2r\,\sin^2\theta\dot{\varphi}^2.$$

Dopo facili semplificazioni si ottiene l'equazione di Lagrange per $i = 1$:

$$\ddot{r} + \frac{1}{2}\lambda'\dot{r}^2 + \frac{1}{2}\nu' e^{\nu(r)-\lambda(r)}(\dot{x}^0)^2 - re^{-\lambda}\dot{\theta}^2 - re^{-\lambda}\sin^2\theta\dot{\varphi}^2 = 0. \tag{12.33}$$

I soli simboli di Christoffel non nulli con indice superiore uguale ad 1 sono perciò:

$$\Gamma^1_{11} = \frac{1}{2}\lambda', \quad \Gamma^1_{00} = \frac{1}{2}\nu' e^{\nu-\lambda}, \quad \Gamma^1_{22} = -re^{-\lambda}, \quad \Gamma^1_{33} = -re^{-\lambda}\sin^2\theta. \tag{12.34}$$

Continuando in maniera analoga, avremo per la variabile $x^2 \equiv \theta$:

$$\frac{\partial\mathcal{L}}{\partial\dot{\theta}} = -2r^2\dot{\theta}; \quad \frac{d}{ds}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\theta}}\Big) = \frac{d}{ds}\Big(-2r^2\dot{\theta}\Big) = -2r^2\ddot{\theta} - 4r\dot{r}\dot{\theta};$$

$$\frac{\partial\mathcal{L}}{\partial\theta} = -2r^2\sin\theta\cos\theta\dot{\varphi}^2;$$

e perciò la corrispondente equazione di Eulero-Lagrange, per $i = 2$, è:

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2 = 0. \tag{12.35}$$

Sicché i corrispondenti simboli di Christoffel non nulli sono:

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}, \quad \Gamma^2_{33} = -\sin\theta\cos\theta. \tag{12.36}$$

L'ultima equazione di Eulero-Lagrange per la variabile $x^3 \equiv \varphi$ si ottiene con analogo procedimento seguito in precedenza:

$$\frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = -2r^2\sin^2\theta\,\dot{\varphi}; \quad \frac{d}{ds}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\Big) = \frac{d}{ds}\Big(-2r^2\sin^2\theta\,\dot{\varphi}\Big) =$$

$$-2r^2\sin^2\theta\,\ddot{\varphi} - 4r\dot{r}\sin^2\theta\,\dot{\varphi} - 4r^2\,\sin\theta\,\cos\theta\,\dot{\theta}\,\dot{\varphi},$$

$$\frac{\partial\mathcal{L}}{\partial\varphi} = 0; \quad (\varphi \text{ coordinata} \quad \text{ciclica})$$

e perciò la corrispondente equazione di Eulero-Lagrange è:

$$\ddot{\varphi} + \frac{2}{r}\dot{r}\dot{\varphi} + 2\cot\theta\,\dot{\theta}\,\dot{\varphi} = 0, \tag{12.37}$$

ed i corrispondenti coefficienti di Christoffel sono:

$$\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r}, \quad \Gamma^3_{23} = \Gamma^3_{32} = \cot\theta. \tag{12.38}$$

Essendo φ una coordinata ciclica ad essa corrisponde un integrale primo che, tenuto conto dei calcoli precedenti, è ovviamente

$$r^2 \sin^2\theta\, \dot{\varphi} = cost. = k_1. \tag{12.39}$$

12.3 Soluzione esatta delle equazioni di campo: metrica di Schwarzschild

Siamo ora in grado di scrivere esplicitamente le equazioni di campo (12.3); esse contengono due simboli di Christoffel contratti del tipo Γ^l_{il}, Γ^l_{ml}, questi com'è noto possono essere scritti nella forma $\Gamma^l_{il} = \partial_i(\log\sqrt{-g})$, $\Gamma^l_{ml} = \partial_m(\log\sqrt{-g})$, perciò le equazioni di campo diventano

$$R_{ik} \equiv \partial_l \Gamma^l_{ik} - \partial_{ki}(\log\sqrt{-g}) + \Gamma^m_{ik}\partial_m(\log\sqrt{-g}) - \Gamma^l_{mk}\Gamma^m_{il} = 0. \tag{12.40}$$

Usando la metrica (12.17), possiamo scrivere $\log\sqrt{-g}$ esplicitamente in termini delle coordinate r, θ, φ e t. Intanto scriviamo il tensore metrico in forma esplicita

$$(g_{ik}) = \begin{pmatrix} e^{\nu(r)} & 0 & 0 & 0 \\ 0 & -e^{\lambda(r)} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}, \tag{12.41}$$

il cui determinante è

$$g = \det(g_{ik}) = -e^{\nu(r)+\lambda(r)} r^4 \sin^2\theta < 0. \tag{12.42}$$

Perciò:

$$\log\sqrt{-g} = \log(e^{\nu(r)+\lambda(r)} r^4 \sin^2\theta)^{1/2} = \frac{1}{2}\log(e^{\nu(r)+\lambda(r)} r^4 \sin^2\theta),$$

cioè:

$$\log\sqrt{-g} = \frac{\nu+\lambda}{2} + 2\log r + \log|\sin\theta|. \tag{12.43}$$

Siamo ora pronti per scrivere le equazioni di campo (12.40) in termini delle coordinate r, θ, φ e t.
Consideriamo dapprima $i = k = 0$, cioè

$$R_{00} \equiv \partial_l \Gamma^l_{00} - \partial_{00}(\log\sqrt{-g}) + \Gamma^m_{00}\partial_m(\log\sqrt{-g}) - \Gamma^l_{m0}\Gamma^m_{0l} = 0. \tag{12.44}$$

Molti dei termini che compaiono in questa equazione sono identicamente nulli: i termini non nulli sono quelli che contengono i simboli di Christoffel espressi nelle (12.32), (12.34), (12.36) e (12.38), che qui riscriviamo:

$$\Gamma^0_{10} = \Gamma^0_{01} = \frac{1}{2}\nu',$$

$$\Gamma^1_{11} = \frac{1}{2}\lambda', \quad \Gamma^1_{00} = \frac{1}{2}\nu' e^{\nu-\lambda}, \quad \Gamma^1_{22} = -re^{-\lambda}, \quad \Gamma^1_{33} = -re^{-\lambda}\sin^2\theta.$$

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}, \quad \Gamma^2_{33} = -\sin\theta\cos\theta.$$

$$\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r}, \quad \Gamma^3_{23} = \Gamma^3_{32} = \cot\theta.$$

Avremo:

$$R_{00} = \partial_1 \Gamma^1_{00} + \Gamma^1_{00} \partial_1 \log \sqrt{-g} - (\Gamma^l_{00}\Gamma^0_{0l} + \Gamma^l_{10}\Gamma^1_{0l} + \Gamma^l_{20}\Gamma^2_{0l} + \Gamma^l_{30}\Gamma^3_{0l})$$

cioè

$$R_{00} = \partial_1 \Gamma^1_{00} + \Gamma^1_{00} \partial_1 \log \sqrt{-g} - 2\Gamma^1_{00}\Gamma^0_{01} =$$

$$= \left(\frac{1}{2}\nu' e^{\nu-\lambda}\right)' + \left(\frac{1}{2}\nu' e^{\nu-\lambda}\right)\left(\frac{\nu'+\lambda'}{2} + \frac{2}{r}\right) - \frac{1}{2}\nu'^2 e^{\nu-\lambda}.$$

Questa si riduce alla

$$\nu'' + \frac{1}{2}\nu'^2 - \frac{1}{2}\lambda'\nu' + 2\frac{\nu'}{r} = 0. \tag{12.45}$$

Procediamo allo stesso modo per $i = k = 1$, avremo:

$$R_{11} \equiv \partial_l \Gamma^l_{11} - \partial_{11}(\log \sqrt{-g}) + \Gamma^m_{11}\partial_m(\log \sqrt{-g}) - \Gamma^l_{m1}\Gamma^m_{1l} = 0. \tag{12.46}$$

Eliminando i termini nulli si ottiene:

$$R_{11} \equiv \partial_l \Gamma^l_{11} - \partial_{11}(\log \sqrt{-g}) + \Gamma^1_{11}\partial_1(\log \sqrt{-g}) - (\Gamma^0_{10}\Gamma^0_{10} + \Gamma^1_{11}\Gamma^1_{11} + \Gamma^2_{21}\Gamma^2_{21} + \Gamma^3_{31}\Gamma^3_{31}) = 0. \tag{12.47}$$

cioè

$$R_{11} \equiv -\left[\left(\frac{\nu''+\lambda''}{2} - \frac{2}{r^2}\right) - \frac{1}{2}\lambda'' + \frac{1}{4}\nu'^2 + \frac{1}{4}\lambda'^2 + \frac{2}{r^2} - \frac{1}{2}\lambda'\left(\frac{\nu'+\lambda'}{2} + \frac{2}{r}\right)\right] = 0, \tag{12.48}$$

e questa si riduce alla:

$$\nu'' + \frac{1}{2}\nu'^2 - \frac{1}{2}\lambda'\nu' - 2\frac{\lambda'}{r} = 0. \tag{12.49}$$

Le equazioni (12.45), (12.49) costituiscono un sistema di due equazioni differenziali ordinarie del secondo ordine nelle due funzioni incognite $\nu(r)$, $\lambda(r)$: sottraendo la seconda dalla prima otteniamo:

$$(\nu + \lambda)' = 0,$$

cioè:

$$\nu + \lambda = cost. = k. \tag{12.50}$$

Possiamo scegliere il valore della costante k uguale a zero per mezzo di un semplice stratagemma. Sostituiamo il tempo coordinato t con un'altra coordinata $te^{k/2}$; dalla (12.17) è chiaro che ciò equivale a sostituire ν con $\nu + k$, sicché la (12.50) diventa

$$\nu = -\lambda. \tag{12.51}$$

Come vedremo questa scelta della coordinata temporale ha la caratteristica desiderabile di rendere l'elemento di linea asintoticamente uguale a quello di uno spazio piatto di Minkowski (12.12). La coordinata temporale, selezionata in questo modo, si vedrà che corrisponde al **tempo fisico** com'è misurato da un osservatore all'infinito.
Sostituendo $-\lambda$ al posto di ν nella (12.49), otteniamo un'equazione differenziale del 2° ordine nella sola funzione incognita $\lambda(r)$:

$$\lambda'' - \lambda'^2 + \frac{2}{r}\lambda' = 0. \tag{12.52}$$

Questa può essere scritta più convenientemente così:

$$(re^{-\lambda})'' = 0. \tag{12.53}$$

L'integrazione di quest'ultima è banale e ci da:

$$(re^{-\lambda})' = cost. \tag{12.54}$$

Conviene lasciare così la (12.54) (con una costante indeterminata) e procedere alla determinazione di R_{22} nelle equazioni di campo (12.40): la ragione di ciò apparirà dalle considerazioni che seguono e precisamente dalla (12.56).
Procedendo allo stesso modo col quale si sono determinate R_{00} ed R_{11}, abbiamo:

$$R_{22} \equiv \partial_l \Gamma^l_{22} - \partial_{22}(\log\sqrt{-g}) + \Gamma^m_{22}\partial_m(\log\sqrt{-g}) - \Gamma^l_{m2}\Gamma^m_{2l} = 0.$$

La sostituzione dei termini contenenti i simboli di Christoffel non nulli, come in precedenza, tenuto conto della (12.43) comporta

$$R_{22} \equiv \partial_1 \Gamma^1_{22} - \partial_{22}(\log\sqrt{-g}) + \Gamma^1_{22}\partial_1(\log\sqrt{-g}) - \Gamma^1_{22}\Gamma^2_{21} - \Gamma^2_{21}\Gamma^1_{22} - \Gamma^3_{23}\ \Gamma^3_{23} = 0,$$

cioè:

$$R_{22} \equiv e^{-\lambda}\left(1 + \frac{r}{2}\nu' - \frac{r}{2}\lambda'\right) - 1 = 0. \tag{12.55}$$

In virtù della (12.51) e della sua derivata rispetto ad r, questa si può scrivere più semplicemente

$$(re^{-\lambda})' = 1. \tag{12.56}$$

Questa è precisamente la stessa della (12.54) con la sola differenza che la costante incognita che appare nella (12.54) è ora identificata in 1. L'integrazione della (12.56) fornisce immediatamente

$$re^{-\lambda} = r + C, \tag{12.57}$$

dove C è una costante arbitraria di integrazione generalmente indicata con $C = -2m$.
Perciò dalle tre equazioni $R_{00} = 0, R_{11} = 0, R_{22} = 0$, siamo riusciti a determinare le due funzioni incognite $\nu(r)$ e $\lambda(r)$ che appaiono nella metrica di partenza (12.17), il che equivale ad aver determinato le due funzioni esponenziali:

$$e^{\nu} = 1 + \frac{C}{r} = 1 - \frac{2m}{r}, \qquad e^{\lambda} = \frac{1}{1 + \frac{C}{r}} = \frac{1}{1 - \frac{2m}{r}}. \tag{12.58}$$

Consideriamo per un momento il risultato (12.58). Come possiamo notare sono state necessarie soltanto tre equazioni delle dieci che formano le equazioni di campo (12.40) per ottenere ciò che appare una soluzione completa del nostro problema. Manifestamente abbiamo un problema di compatibilità: le altre sette equazioni del sistema (12.40) sono compatibili con la soluzione trovata?
Mostreremo che il rimanente elemento diagonale R_{33} del tensore di Ricci è invero identicamente nullo in virtù della soluzione (12.58), sicché quest'ultima e l'equazione $R_{33} = 0$ sono compatibili.
Come abbiamo fatto prima, scriviamo l'equazione di campo relativa ad R_{33}, avremo:

$$R_{33} \equiv \partial_l \Gamma^l_{33} - \partial_{33}(\log\sqrt{-g}) + \Gamma^m_{33}\partial_m(\log\sqrt{-g}) - \Gamma^l_{m3}\Gamma^m_{3l} = 0.$$

Essendo $\partial_{33}(\log\sqrt{-g}) = 0$ ed eliminando gli altri termini nulli si ottiene:

$$R_{33} \equiv \partial_1\Gamma^1_{33} + \partial_2\Gamma^2_{33} + \Gamma^1_{33}\partial_1(\log\sqrt{-g}) + \Gamma^2_{33}\partial_2(\log\sqrt{-g}) + 2\Gamma^3_{13}\Gamma^1_{33} + +2\Gamma^3_{32}\Gamma^2_{33} = 0,$$

cioè

$$R_{33} = R_{22}\sin^2\theta = \left[e^{-\lambda}\left(1 + \frac{r}{2}\nu' - \frac{r}{2}\lambda'\right) - 1\right]\sin^2\theta = \left[(re^{-\lambda})' - 1\right]\sin^2\theta = 0. \tag{12.59}$$

In virtù della (12.56) questa è identicamente nulla, sicché l'equazione $R_{33} = 0$ è consistente con la soluzione (12.58).
Si può verificare che tutti gli elementi fuori della diagonale principale del tensore di Ricci sono identicamente nulli quando sono scritti in termini delle coordinate che abbiamo usato r, θ, φ, t e che perciò la (12.58) è la soluzione completa delle equazioni di campo (12.40).
Riassumendo i risultati di questo paragrafo e dei due precedenti possiamo dire che: **lo spaziotempo al di fuori di una massa M a simmetria sferica ammette la seguente metrica di Schwarzschild**:

$$\mathbf{ds}^2 = \left(1 + \frac{C}{r}\right)c^2dt^2 - \frac{1}{1 + \frac{C}{r}}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\varphi^2, \tag{12.60}$$

dove C è, come abbiamo visto, una costante arbitraria di integrazione; se questa costante, com'è d'uso, si scrive $C = -2m$ allora la metrica di Schwarzschild si presenta nella forma diagonale:

$$\mathbf{ds}^2 = \left(1 - \frac{2m}{r}\right)c^2dt^2 - \frac{1}{1 - \frac{2m}{r}}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\varphi^2. \tag{12.61}$$

I coefficienti della metrica sotto forma covariante sono evidentemente:

$$g_{00} = 1 - \frac{2m}{r}, \quad g_{11} = -\left(1 - \frac{2m}{r}\right)^{-1}, \quad g_{22} = -r^2, \quad g_{33} = -r^2\sin^2\theta, \tag{12.62}$$

i corrispondenti espressi in forma controvariante sono invece:

$$g^{00} = \left(1 - \frac{2m}{r}\right)^{-1}, \quad g^{11} = -\left(1 - \frac{2m}{r}\right), \quad g^{22} = -\frac{1}{r^2}, \quad g^{33} = -\frac{1}{r^2\sin^2\theta}, \tag{12.63}$$

ed il determinante della metrica, per le (12.42) e (12.58), è

$$g = \det g_{ik} = -r^4\sin^2\theta < 0. \tag{12.64}$$

Il risultato espresso dalla metrica di Schwarzschild** (12.61) **deve essere considerato come la maggiore conquista della teoria della Relatività Generale nel campo della Meccanica Celeste; esso è una soluzione esatta che corrisponde storicamente alla trattazione newtoniana della Legge di Attrazione della teoria classica della gravitazione.

12.4 Alcune considerazioni sulla metrica di Schwarzschild

È evidente che la metrica di Schwarzschild approssima la metrica piatta di Minkowski (12.12) a grandi distanze, cioè per $r \to +\infty$. *Possiamo perciò identificare il tempo t con il tempo misurato da un osservatore a grande distanza dall'origine.* Perciò **il tempo coordinato** è, in questo senso, una coordinata di riguardo. È importante ricordare che il significato fisico delle coordinate è intimamente legato alla metrica, come dimostrato da questo esempio.
La costante d'integrazione incognita $C = -2m$ che appare nella metrica di Schwarzschild può essere determinata per mezzo della corrispondenza con la teoria newtoniana. Ricordiamo che la teoria geometrica della gravitazione (Relatività Generale) si riduce nel limite classico di campi deboli e basse velocità dei corpi, alla teoria newtoniana se $g_{00} \simeq 1 + \frac{2\phi}{c^2}$, dove ϕ è il potenziale classico del campo gravitazionale. Nel caso qui considerato di una massa puntiforme (equivalente ad una massa a simmetria sferica), il potenziale gravitazionale, come sappiamo, è dato da $\phi = -\frac{GM}{r}$, dove M è la massa del corpo e $G = 6,6726 \times 10^{-11}\mathrm{N} \cdot \mathrm{m}^2 \cdot \mathrm{kg}^{-2}$ è la costante

newtoniana di gravitazione nel SI. Allora, nel limite classico, $g_{00} \simeq 1 - \frac{2GM}{c^2 r}$. Confrontando g_{00}, espresso da quest'ultima relazione, con il primo termine della metrica di Schwarzschild avremo:

$$m = \frac{GM}{c^2}. \tag{12.65}$$

La costante m ha perciò le dimensioni di una lunghezza e talvolta viene chiamata impropriamente **massa geometrica** del corpo centrale.
Abbiamo ottenuto la soluzione di Schwarzschild (12.61) imponendo le condizioni di **simmetria sferica ed indipendenza dal tempo**. Tuttavia, il grande matematico statunitense George David Birkhoff nel 1923 dimostrò che il requisito di indipendenza dal tempo è superfluo. In effetti il risultato a cui pervenne consiste nel seguente notevole:
Teorema di Birkhoff: *Ogni distribuzione di materia a simmetria sferica, anche in moto radiale, conduce necessariamente alla metrica statica di Schwarzschild all'esterno della materia.*
La dimostrazione è analoga a quella da noi fatta ma più involuta poiché, in tal caso, λ e ν sono trattate come funzioni di r, t e non solamente come funzioni di r.[10] Una conseguenza importante del teorema di Birkhoff è che: *Una distribuzione a simmetria sferica di materia pulsante radialmente* **non può emettere onde gravitazionali** *perché la metrica al di fuori della materia è statica. Tali onde possono perciò essere emesse soltanto attraverso una deformazione più complicata del corpo massivo.* Per una trattazione succinta dell'importantissimo argomento delle onde gravitazionali rimandiamo all'APPENDICE alla fine di questo capitolo.
Prima di continuare con altre considerazioni, che faremo nel prossimo paragrafo, è intanto opportuno notare che, sulla sfera di raggio $r = 2m$, la metrica di Schwarzschild (12.61) diventa degenere in quanto il coefficiente di $(dx^0)^2 = c^2 dt^2$ si annulla ed il coefficiente di dr^2 diventa infinito, perciò $2m$ ha la caratteristica di *un raggio critico*. Tenuto conto della (12.65) tale raggio critico può scriversi così:

$$R_S = 2m = \frac{2GM}{c^2}, \tag{12.66}$$

e viene **è chiamato raggio di Schwarzschild o raggio gravitazionale**. Esso assume un'importanza fondamentale nello studio del cosiddetto **collasso gravitazionale** delle stelle come vedremo meglio un po' più avanti. Per stelle ordinarie, considerate come sfere massive a simmetria sferica, tale raggio è molto piccolo; per esempio per il nostro Sole il raggio di Schwarzschild è di circa 3 km, e perciò ben interno al Sole, dove le equazioni di campo (12.40) non sono più valide e la metrica esterna di Schwarzschild (12.61) non è più appropriata alla descrizione della geometria dello spaziotempo. *L'esistenza di questa singolarità è perciò di nessuna conseguenza per la descrizione dei moti planetari.*
È chiaro che una stella avente approssimativamente la massa del Sole dovrebbe essere fortemente compressa per essere tutta contenuta nel raggio di Schwarzschild. Lo studio di questa situazione è diventato di grande interesse negli ultimi decenni, nell'ambito dell'Astrofisica, poiché adesso appare chiaro da studi teorici che una frazione significativa delle stelle dell'universo possono realmente raggiungere e superare grandi densità durante un **collasso gravitazionale**, che costituisce la fase finale della loro esistenza come stelle normali dando origine ai cosiddetti **buchi neri**. Torneremo con più dettagli, nei prossimi paragrafi, su questo affascinante argomento.
A conclusione di questo paragrafo, raccogliamo ora in un quadro riassuntivo le equazioni differenziali delle geodetiche 4-dimensionali, ricavate nel paragrafo precedente, relative alla metrica esterna di Schwarzschild (12.61) delle quali faremo una diretta applicazione nel capitolo successivo per studiare l'avanzamento del perielio dei pianeti, in particolare di Mercurio, e la

[10] Lo studioso può trovare la dimostrazione del teorema di Birkhoff, in Hans C. Ohanian e Remo Ruffini, *Gravitazione e Spazio-Tempo*, Zanichelli, Bologna, 1997, Cap.7, paragrafo 7.5, pp. 337-340. Oppure in Ray d'Inverno, *Introduzione alla Relatività di Einstein*, CLUEB, 2001, Cap. 14, paragrafo 14.5, pp. 233-237.

deflessione dei raggi luminosi in un campo gravitazionale.
Si tratta delle equazioni di Eulero-Lagrange relative alle coordinate, $x^1 = r$, $x^2 = \theta$, $x^3 = \varphi$, $x^0 = ct$. Con più precisione, per la coordinata $x^1 = r$ utilizzeremo anziché l'equazione del 2° ordine (12.33), l'integrale primo (12.28) dedotto direttamente dal $\mathbf{ds}^2$; per la coordinata $x^2 = \theta$, l'equazione del 2° ordine (12.35); per la terza coordinata spaziale $x^3 = \varphi$, l'integrale primo (12.39); ed infine per la coordinata temporale $x^0 = ct$ l'integrale primo (12.2).
Le equazioni differenziali delle geodetiche 4-dimensionali del problema di Schwarzschild, scritte nell'ordine in cui le abbiamo appena elencate, sono dunque:

$$\begin{cases} \left(1-\dfrac{2m}{r}\right)c^2\dot{t}^2 - \left(1-\dfrac{2m}{r}\right)^{-1}\dot{r}^2 - r^2\left(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2\right) - 1 = 0, \\ \ddot{\theta} + \dfrac{2}{r}\dot{r}\dot{\theta} - \sin\theta\,\cos\theta\,\dot{\varphi}^2 = 0, \\ r^2\sin^2\theta\,\dot{\varphi} = cost. = k_1 \quad (\varphi \text{ coordinata ciclica}), \\ \left(1-\dfrac{2m}{r}\right)\dot{t} = cost. = k_2, \end{cases} \tag{12.67}$$

dove, è bene ricordarlo, il puntino sopra le lettere indica derivazione rispetto al parametro s che figura nell'elemento lineare ds.

12.5 L'idea di "buco nero" nella teoria newtoniana

Prima di discutere le singolarità della metrica di Schwarzschild che conducono al concetto di "buco nero" nell'ambito della Teoria della Relatività Generale, desideriamo occuparci brevemente di una curiosità interessante.
Ad un esame superficiale sembrerebbe che *l'idea di* **buco nero** *fosse già insita nella teoria gravitazionale di Newton.* Infatti John Michell, astronomo di Cambridge è stato il primo, in un saggio del 1783, ad aver concepito come *una stella con una massa sufficientemente elevata possa essere in grado di trattenere la sua stessa luce e dunque apparire* **"scura e nera" (dark star)**.
Alla stessa conclusione pervenne il grande matematico ed astronomo francese Pierre Simon de Laplace nel 1796. La spiegazione di questo fatto è piuttosto semplice e conduce ad un **"raggio di non ritorno"** che stranamente coincide col raggio di Schwarzschild.
Infatti, consideriamo una particella di massa m che si allontana radialmente da un corpo sferico S o avente simmetria sferica, di raggio R, e massa M. Se la particella ha una velocità v alla distanza r dal centro O della sfera, per il teorema di conservazione dell'energia si ha

$$E = T + V = \frac{1}{2}mv^2 - \frac{GMm}{r} = cost., \tag{12.68}$$

dove $T = \dfrac{1}{2}mv^2$ e $V = -\dfrac{GMm}{r}$ sono rispettivamente l'energia cinetica e l'energia potenziale della particella ed E la sua energia totale. Se per $t = 0$, cioè all'istante iniziale, è $r = R$, cioè la particella si trova sulla superficie della sfera, e $v = v_0$ sia la sua velocità radiale iniziale, avremo

$$E = \frac{1}{2}mv_0^2 - \frac{GMm}{R}. \tag{12.69}$$

Sottraendo la (12.69) dalla (12.68) e semplificando per m avremo

$$v^2 = v_0^2 - 2GM\left(\frac{1}{R} - \frac{1}{r}\right); \tag{12.70}$$

è ovvio che se per un certo valore di r il 2° membro della (12.70) si annulla, dunque si annulla anche il 1° membro v ($v = 0$), la particella cade sul corpo sferico S per effetto della sua gravitazione. Affinché la particella sfugga all'azione gravitazionale di S il secondo membro della (12.70) dev'essere positivo o al più nullo per qualsiasi valore di r, quindi anche per $r \to +\infty$, ossia deve essere

$$v_0 \geq \sqrt{\frac{2GM}{R}}. \tag{12.71}$$

L'espressione al secondo membro è la cosiddetta **velocità di fuga** v_f della particella:

$$v_f = \sqrt{\frac{2GM}{R}}, \tag{12.72}$$

essa è definita come **la minima velocità radiale** *che deve avere la particella sulla superficie del corpo sferico S affinché possa scappare all'infinito dove la sua velocità si annulla.*
Se la particella, parte dalla superficie del dato corpo sferico, con velocità radiale iniziale minore di v_f, cioè $v_0 < v_f$, per le considerazioni fatte più sopra, non può sfuggire al campo gravitazionale del corpo, perciò essa si allontanerà, con velocità sempre decrescente, fino ad una determinata distanza r_* in cui la sua velocità si annulla $v_* = 0$. La particella essendo però soggetta al campo gravitazionale del corpo ricadrà su di esso.
Poiché l'espressione (12.72) della velocità di fuga è indipendente dalla massa della particella, possiamo supporre che si applichi anche ai fotoni dei quali è composta la luce ed avendo questa velocità uguale a c, essa potrà scappare all'infinito solo se $v_f = c$, cioè solo se la massa del corpo ed il suo raggio verificano la relazione

$$c^2 = \frac{2GM}{R}. \tag{12.73}$$

Concludendo, osserviamo che se dalla condizione limite (12.73) ricaviamo il raggio R, si ha:

$$R = \frac{2GM}{c^2}, \tag{12.74}$$

questo valore, per una strana coincidenza, è uguale al raggio R_S di Schwarzschild. Perciò, se aumentassimo la massa M (tenendo il raggio costante) o, equivalentemente diminuissimo il raggio (tenendo la massa costante), ne seguirebbe che la luce non potrebbe più scappare. Tutto ciò è proprio quello che è stato osservato da Laplace nel 1798, il quale mostrò che un corpo con circa la stessa massa del Sole, ma con un raggio 250 volte maggiore non permetterebbe alla luce di scappare dal suo campo gravitazionale.
Il risultato newtoniano appena ottenuto è però fuorviante perché fornisce un'immagine errata di ciò che succede in un **buco nero di Schwarzschild**. I segnali luminosi emessi ad un raggio $r \leq R_S$ non si allontanano prima, poi si fermano, e quindi tornano indietro, essi piuttosto cadono immediatamente e non iniziano mai a muoversi nella direzione esterna. Lo stesso vale per qualsiasi altro corpo anche se munito di propulsione come le astronavi. Se la luce non può sfuggire, neanche altri corpi possono farlo perché la loro velocità deve rimanere sempre inferiore alla velocità della luce.

12.6 Singolarità della metrica di Schwarzschild: buchi neri

In questo paragrafo e nei successivi vogliamo dare, almeno un cenno, di quegli oggetti bizzarri del nostro universo che vanno sotto il nome di **buchi neri (black holes)**. Prima di darne una definizione esplicita è opportuno riassumere brevemente alcune conclusioni a cui si perviene

in Astrofisica nello studio dell'evoluzione stellare.[11]
Le stelle aventi massa inferiore a 1,3 masse solari terminano pacificamente la loro esistenza sotto forma di **nane bianche.** *Se la massa è troppo grande perché si abbia questo caso, la stella subisce un'ulteriore evoluzione, che termina con il collasso del suo nucleo e (almeno in taluni casi) con un'esplosione del tipo* **supernova**, *nel corso della quale vengono allontanati gli strati più esterni. Se la massa del nucleo così rimasto è inferiore a 2 masse solari, il collasso si arresta allorché si raggiungono densità dell'ordine di quella dei nuclei atomici, e si ha una* **stella di neutroni.**
Che cosa succede, invece, quando la massa del nucleo supera il doppio della massa solare, *oppure* quando la velocità d'implosione è troppo grande per potersi fermare, anche se la massa è inferiore a 2 masse solari, *oppure* quando altra massa cade su una stella di neutroni - o su una **nana bianca** - e porta la massa totale a superare il valore critico che permette la stabilità? *In questi casi, tutti i possibili metodi, sia quelli provvisori, sia quelli permanenti, per stabilizzare la stella ed impedire il collasso sono già stati sfruttati fino in fondo, si ha allora* **il totale collasso gravitazionale!**
Che cosa succede quando una stella subisce questo tipo di catastrofico collasso gravitazionale? Che cosa succede alla stella e al suo campo gravitazionale dopo il collasso? Gli studi teorici hanno condotto a concludere che *il completo collasso gravitazionale di un corpo dà sempre come risultato un oggetto chiamato* **buco nero**, *e che la natura del buco nero così prodotto, generalmente, dipende soltanto dalla sua* **massa, dal suo momento angolare e dalla sua carica elettrica.**
Purtroppo, nella Relatività Generale, l'analisi teorica di ciò che avviene in situazioni generiche durante il collasso gravitazionale è assai difficile: è assai difficile trovare soluzioni esatte delle equazioni di campo di Einstein che descrivono il collasso gravitazionale. Fortunatamente la situazione migliora notevolmente quando si pone l'ipotesi semplificativa *dell'esatta simmetria sferica* come abbiamo fatto noi nel trattare il problema esterno di Schwarzschild. Nel caso della simmetria sferica, le equazioni si possono risolvere, e si può analizzare nei particolari ciò che avviene durante il collasso. Noi non possiamo percorrere un così lungo cammino che ci porterebbe fuori dagli scopi che ci siamo prefissi nel redigere questo libro. Ci accontenteremo, come abbiamo detto, di dare qualche cenno facendo delle semplici considerazioni teoriche sulla metrica di Schwarzschild ottenuta alla fine del paragrafo 12.3.
Per discutere le caratteristiche generali della soluzione di Schwarzschild, faremo l'ipotesi che il corpo sia collassato completamente, in modo tale che la densità della materia sia nulla in ogni punto (eccetto forse in $r = 0$). Questo significa che consideriamo la soluzione di Schwarzschild come una soluzione esatta nel vuoto delle equazioni gravitazionali di Einstein.
Come abbiamo accennato più sopra, nel caso del collasso gravitazionale totale di una stella, la sua massa ed il suo raggio subiscono una fortissima contrazione generando quella che si dice una **singolarità.** Perciò è interessante, discutendo della metrica di Schwarzschild, assumere a priori che essa sia valida per valori arbitrari di r $(0 \leq r < +\infty)$, cosa che constateremo non essere vera. A tal proposito è opportuno ricordarsi che un sistema di coordinate descrive, in generale, solo una porzione della varietà sulla quale sono definite. Nel caso della metrica di Schwarzschild le coordinate usate non descrivono i punti per i quali è $\theta = 0$ e $\theta = \pi$ perché l'elemento di linea ivi diventa degenere in quanto la metrica cessa di essere di rango 4, cioè esprimibile in 4 coordinate. Questa degenerazione potrebbe essere rimossa introducendo le coordinate cartesiane (x, y, z) dove come al solito si ha: $x = r \sin\theta \cos\varphi, \quad y = r \sin\theta \sin\varphi, \quad z = r \cos\theta$. Tali punti vengono

[11]Sui buchi neri ci sono moltissimi libri ad essi completamente dedicati. Una lettura divulgativa semplice ed introduttiva a questo argomento è: Robert M. Wald, *Teoria del Big Bang e Buchi Neri*, Universale Scientifica Boringhieri,Torino, 1983. Altra ottima lettura ad un livello più avanzato, ma comunque per principianti, è: Derek Raine and Edwin Thomas, *Black Holes, An Introduction*, World Scientific Publishing Co. Pte. Ltd., 2005. Una buona trattazione di livello universitario, si trova nel più volte citato, Ray d'Inverno, *Introduzione alla Relatività di Einstein*, Clueb, Bologna, 2001, cap.16. Un'ottima trattazione si trova in: Wolfgang Rindler, *Relativity, Special, General, and Cosmological*, Oxford University Press, Second Edition, 2006, Reprinted 2008, Cap.11, Cap. 12. Buona è anche la trattazione di Hans C. Ohanian e Remo Ruffini, *Gravitazione e spazio-tempo*, Zanichelli, Bologna, 1997, Cap. 8.
Due libri fondamentali ma molto avanzati sono S. Chandrasekhar, *The Mathematical Theory of Black Holes*, Oxford University Press, 1983, e V. P. Frolov and I. D. Novikov, *Balck Holes Physics*, Kluwer Dordrecht, 1998.

chiamati **singolarità di coordinate** e sono generalmente **rimovibili**.
Riprendiamo dunque quella che chiameremo **forma standard della metrica di Schwarzschild** (12.61), nelle coordinate (t, r, θ, φ), scrivendola in maniera leggermente diversa, tenendo conto di ciò che abbiamo detto nel paragrafo 12.4:

$$\mathbf{ds}^2 = \left(1 - \frac{R_S}{r}\right)c^2dt^2 - \frac{1}{1 - \frac{R_S}{r}}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\varphi^2, \tag{12.75}$$

nella quale

$$R_S = 2m = \frac{2GM}{c^2}, \tag{12.76}$$

è il raggio di Schwarzschild.
Osserviamo subito che tale metrica presenta un comportamento anomalo sia nei punti della **sfera di Schwarzschild** di raggio $R_S = \frac{2GM}{c^2}$, sia nell'origine delle coordinate in cui è $r = 0$. Infatti la prima anomalia consiste nel fatto che se la coordinata radiale r tende al raggio di Schwarzschild, cioè $r \to R_S$, la metrica standard (12.75) diventa degenere perché il suo primo coefficiente tende a zero mentre il secondo coefficiente tende all'infinito:

$$g_{00} = 1 - \frac{R_S}{r} \to 0, \tag{12.77}$$

$$g_{11} = -\frac{1}{1 - \frac{R_S}{r}} \to -\infty, \tag{12.78}$$

Inoltre, quando $r \to 0$, $g_{00} \to +\infty$, e $g_{11} \to 0$ cioè il comportamento di questi due primi coefficienti della metrica si inverte, infatti si ha

$$g_{00} > 0, \quad g_{11} < 0, \quad \text{per} \quad r > R_S,$$

e

$$g_{00} < 0, \quad g_{11} > 0, \quad \text{per} \quad r < R_S,$$

il che significa che r diventa una coordinata temporale, mentre t diventa una coordinata spaziale.[12] Di conseguenza la ipersuperficie $r = R_S$ divide la varietà spaziotempo in due regioni sconnesse:

$$I. \quad R_S < r < +\infty$$
$$II. \quad 0 < r < R_S,$$

all'interno delle quali le coordinate t ed r hanno invertito il loro carattere. In entrambi i casi diciamo che si ha una **singolarità** della metrica di Schwarzschild, ma di differente natura. Per accertarsi se tale singolarità è una **genuina singolarità di curvatura** è necessario calcolare l'invariante scalare costruito col tensore di Riemann (scalare di Ricci), si ottiene:

$$R_{ijhk}R^{ijhk} = \frac{12R_S^2}{r^6}, \tag{12.79}$$

e si vede chiaramente che esso non è divergente per $r = R_s$. Poiché è un invariante, il suo valore rimane lo stesso in tutti i sistemi di coordinate. Per lo stesso motivo, questo invariante è divergente nell'origine $r = 0$. Perciò la singolarità per $r = 0$ non è rimovibile ed è nota come singolarità **fisica, intrinseca, essenziale o di curvatura**.
D'altra parte, per la (12.64) anche il determinante della metrica $g = \det g_{ik} = -r^4\sin^2\theta < 0$ rimane finito per $r = R_S$, quindi non si tratta di una singolarità "fisica" ma di una singolarità dovuta alla scelta delle coordinate, cioè, come anche si dice, di una **pseudosingolarità** che può essere rimossa con un conveniente cambio di variabili come accenneremo in seguito.

[12] Si ricordi che la segnatura della metrica che abbiamo sempre utilizzato è $(+---)$ dove il segno $+$ è quello che compete al termine contenete la coordinata temporale, mentre i tre segni $-$ sono relativi alle coordinate spaziali.

12.7 Diagrammi spazio-temporali in coordinate di Schwarzschild

La tecnica principale che si usa per l'interpretazione della soluzione di Schwarzschild è lo studio della struttura locale della falda superiore del cono di luce, cioè quella rivolta verso il futuro, che chiameremo, per brevità, *cono luce futuro*. Siccome, sia in RR che in RG, i raggi di luce sono caratterizzati sempre dal $\mathbf{ds}^2 \equiv ds^2 = 0$, possiamo porre la seguente:
Definizione: Dicesi **cono di luce locale**, di vertice un punto prefissato x_0^i, il luogo geometrico dei punti $x_0^i + dx^i$ dello spaziotempo soddisfacente alla seguente equazione:

$$ds^2 \equiv g_{ik}(x)dx^i dx^k = 0, \qquad (i,k = 0,1,2,3). \tag{12.80}$$

Il cono di luce locale pone delle restrizioni ai possibili percorsi spazio-temporali di un osservatore, dato che esso si muove su di una linea oraria del genere tempo, la cui direzione in qualsiasi punto deve giacere all'interno del cono luce futuro avente il vertice in quel punto. Spesso, per semplicità usiamo diagrammi spaziotemporali con due dimensioni soppresse, dunque su di un piano, come abbiamo sempre fatto per i diagrammi nello spaziotempo di Minkowski. In questo spazio, essendo la metrica a coefficienti costanti, in due dimensioni l'elemento di linea (con $c = 1$) lo si scrive così:

$$ds^2 \equiv dt^2 - dx^2 = 0. \tag{12.81}$$

Questa equazione equivale a scrivere:

$$\frac{dt}{dx} = \pm 1, \tag{12.82}$$

dove $\frac{dt}{dx} = \tan\psi$, essendo ψ l'angolo formato dalla linea oraria di un raggio di luce con l'asse t, il che significa, come sappiamo, che i coni luce nello spazio di Minkowski hanno tutti la stessa apertura $\psi = \frac{\pi}{4}$, cioè i raggi di luce sono tutti inclinati di $\pm\frac{\pi}{4}$ sull'asse delle ordinate che funge da asse dei tempi t.
In uno spaziotempo **curvo**, la curvatura si manifesta, nei diagrammi spaziotemporali, attraverso la distorsione dei coni di luce che vengono schiacciati, stirati, inclinati o rovesciati in vari modi, come vedremo fra poco.
Un'anomalia riguardante la metrica, molto più seria di quelle discusse nel paragrafo precedente, è la seguente. Si può verificare immediatamente usando la 1ª delle equazioni differenziali (12.67) delle geodetiche 4-dimensionali (con $c = 1$), che le **geodetiche radiali**, cioè quelle aventi le coordinate t, θ, φ costanti, hanno la seguente equazione:

$$-\left(1 - \frac{R_S}{r}\right)^{-1} \dot{r}^2 - 1 = 0, \tag{12.83}$$

dalla quale, tenendo conto che $\dot{r} = \frac{dr}{ds}$, si ottiene la condizione:[13]

$$\mathbf{ds}^2 = -\frac{1}{1 - \frac{R_S}{r}} dr^2. \tag{12.84}$$

Dalla precedente relazione si deduce che per $r > R_S$, cioè all'esterno della sfera di Schwarzschild, si ha $\mathbf{ds^2} < 0$, dunque le relative geodetiche sono **del genere spazio**, mentre per $0 < r < R_S$, cioè all'interno di tale sfera, essendo $\mathbf{ds^2} > 0$ le geodetiche sono **del genere tempo**. Ma nel Cap.10, paragrafo 10.9.1, abbiamo visto che le geodetiche sono curve autoparallele, cioè

[13] Osserviamo, attenendoci alla notazione di Rindler da noi adottata, che al primo membro della formula (12.84) è necessario scrivere $\mathbf{ds}^2 \gtreqless 0$ e non semplicemente $ds^2 > 0$ affinché il primo e secondo membro di tale formula abbiano lo stesso segno sia per $r > R_S$, sia per $r < R_S$.

un vettore tangente alla geodetica in un suo punto P, com'è il vettore infinitesimo **ds**, viene trasportato per parallelismo quando il punto P si sposta lungo la geodetica stessa. Perciò esso non può cambiare da un vettore del genere spazio in uno del genere tempo. Ne segue che le due regioni definite da $r > R_S$ e $0 < r < R_S$ non si raccordano in modo liscio sulla superficie di Schwarzschild $r = R_S$.
Ciò può essere visto in maniera più stringente se consideriamo una **geodetica radiale nulla** per la quale si ha $d\theta = d\varphi = 0$ con $ds^2 = 0$. Ponendo $c = 1$ (come abbiamo fatto più volte per semplificare un po' i calcoli), dalla metrica standard di Schwarzschild otteniamo allora:

$$ds^2 = \left(1 - \frac{R_S}{r}\right)dt^2 - \left(1 - \frac{R_S}{r}\right)^{-1} dr^2 = 0. \tag{12.85}$$

Conseguentemente **le direzioni radiali nulle**, cioè le direzioni tangenti alle geodetiche radiali nulle, soddisfano la relazione:

$$\frac{dt}{dr} = \pm \frac{1}{1 - \frac{R_S}{r}}. \tag{12.86}$$

Questa formula (generalizzazione della (12.82)) ci dà l'inclinazione, sull'asse t, delle linee di universo della luce, cioè dei coni di luce in un diagramma spaziotemporale del piano (r, t), con r ascissa e t ordinata. Riscriviamo la (12.86) nella forma

$$dt = \pm \frac{r}{r - R_S} dr. \tag{12.87}$$

Essa è un'equazione differenziale a variabili separate che si integra facilmente, ottenendo:

$$t = \pm \left[r + R_S \log |r - R_S| + cost.\right]; \tag{12.88}$$

l'equazione precedente rappresenta una doppia congruenza di geodetiche radiali nulle, una caratterizzata dal segno $+$, l'altra dal segno $-$ (Fig.12.1)

Figura 12.1: Soluzione 2-dimensionale di Schwarzschild nelle coordinate omonime.

Le curve della prima congruenza, diciamola Γ_1, aventi equazione:

$$t = r + R_S \log |r - R_S| + cost., \tag{12.89}$$

si dicono *geodetiche radiali nulle* **uscenti** e se le rappresentiamo in un diagramma spaziotemporale bidimensionale (r, t), esse sono distribuite in due insiemi di curve, l'uno appartenente alla I^a regione (quella situata a destra della retta $r = R_S$), in cui è $r > R_S$, e per la (12.86)

(considerata col segno $+$), essendo $\frac{dr}{dt} > 0$, ne segue che r è una funzione crescente di t. L'altro insieme di curve della stessa congruenza uscente, giace invece nella II^a regione (quella situata a sinistra della retta $r = R_S$), caratterizzata da $0 < r < R_S$ e per esso si ha, sempre in virtù della (12.86), $\frac{dr}{dt} < 0$ dunque con r funzione decrescente di t. La congruenza delle geodetiche radiali nulle uscenti è rappresentata nella (Fig.12.1) con linee blu.
Allo stesso modo le curve della seconda congruenza Γ_2 aventi equazione:

$$t = -\big(r + R_S \log|r - R_S| + cost.\big), \tag{12.90}$$

si dicono *geodetiche radiali nulle* **entranti**. Se le rappresentiamo nello stesso diagramma spazio-temporale bidimensionale (r, t) esse sono distribuite in due insiemi di curve, l'uno appartenente alla I^a regione, in cui è $r > R_S$, e per la (12.86) (considerata col segno $-$), essendo $\frac{dr}{dt} < 0$, ne segue che r è una funzione decrescente di t. L'altro insieme di curve della stessa Γ_2 entrante, giace invece nella II^a regione caratterizzata da $0 < r < R_S$ e per esso si ha, sempre in virtù della (12.86), $\frac{dr}{dt} > 0$ dunque con r funzione crescente di t. La congruenza Γ_2 delle geodetiche radiali nulle entranti è rappresentata nella (Fig.12.1) con linee rosse, mentre la superficie di Schwarzschild $r = R_S$, è rappresentata da una retta parallela all'asse delle t. Se teniamo conto del fatto che le direzioni del genere tempo sono contenute nel cono di luce, troviamo che nella I^a regione con $r > R_S$ i coni di luce hanno, nel piano (r, t), l'orientamento mostrato in Fig.12.1 con asse verticale (parallelo all'asse t). L'apertura del cono di luce per $r \gg R_S$ ($r \to +\infty$), in cui la metrica è quasi minkowskiana, assume la massima apertura prossima a $\pi/4$, decresce al decrescere di r e tende a zero quando $r \to R_S$. Le linee di universo della luce, tangenti in ogni loro punto ai coni luce, tendono asintoticamente alla retta $r = R_S$ perché il loro coefficiente angolare $\frac{dt}{dr} \to +\infty$.
Inoltre, stante al diagramma spaziotemporale di Fig.12.1, *un raggio di luce, o un corpo qualsiasi, che parte dall'infinito, cioè nella regione $r \gg R_S$ e va verso la superficie di Schwarzschild $r = R_S$, in direzione radiale,* **non la raggiungerà mai** *in quanto, come si vede bene dal diagramma, si avvicina ad essa asintoticamente.* Questo comportamento però è poco credibile dal punto di vista fisico ed è dovuto alla cattiva scelta della coordinata temporale t; basta cambiare tale coordinata, passando per esempio al tempo proprio τ della particella, perché la particella, attraversando la sfera di Schwarzschild, vada a cadere, in un tempo finito, nella singolarità $r = 0$, come vedremo nel prossimo paragrafo.
Al contrario, nella regione II^a con $0 < r < R_S$, come abbiamo visto si invertono i ruoli delle due coordinate r e t, in quanto r diventa una coordinata temporale mentre t diventa una coordinata spaziale; i coni di luce, nel passare dalla regione $r > R_S$ all'altra $0 < r < R_S$, invertono bruscamente il proprio asse che da verticale diventa orizzontale Fig.12.1. Essi hanno una grande apertura, di circa $\frac{\pi}{4}$, per $0 < r < R_S$ ma molto prossimi ad R_S (per intenderci sulla sinistra della retta $r = R_S$ ma molto vicino ad essa). Poi per $r \to 0$ essi restringono sempre di più la loro apertura diventando sottilissimi in prossimità di $r = 0$. Un raggio di luce che parte dalla singolarità $r = 0$ e va verso la superficie di Schwarzschild $r = R_S$ non la raggiungerà mai in quanto, come si vede bene dal diagramma, si avvicina ad essa asintoticamente: ciò significa che **un raggio di luce, o peggio ancora, un qualunque corpo, il quale si muove certamente con velocità inferiore a quella della luce, non può mai uscire dalla regione definita dalla disuguaglianza** $0 < r \leq R_S$.
È proprio per quest'ultima proprietà: che la luce non può sfuggire dalla regione di spazio contenuta nella sfera di Schwarzschild e che tale regione appare, ad un osservatore lontano, buia e nera tanto da giustificare l'appellativo datogli, nel 1967, dal grande relativista statunitense John Archibald Wheeler, di **buco nero (black hole)**.

A questo punto possiamo dare una semplice definizione di buco nero di Schwarzschild:
Definizione: *Chiamiamo* **buco nero** *la regione dello spaziotempo, attorno ad una stella a simmetria sferica, in cui sia* $0 < r \leq R_S$*, cioè la regione interna alla sfera avente il raggio di Schwarzschild, inclusa la sfera stessa, caratterizzata dalla proprietà che il campo gravitazionale è talmente intenso che qualsiasi cosa che entri nella regione non può più uscirne.*
Il confine di un buco nero è chiamato **orizzonte degli eventi.** Nel caso di un corpo sferico, l'orizzonte degli eventi coincide con la sfera di Schwarzschild di raggio $r = R_S$. Per ciò che abbiamo detto più sopra, una volta attraversato l'orizzonte degli eventi ed entrati nel buco nero, non si può più ritornare nella lontana regione dello spaziotempo in cui il campo gravitazionale è debole. Alla stessa stregua, un osservatore che resti all'esterno di un buco nero non può vedere nulla di ciò che avviene al suo interno.

12.8 Caduta libera radiale di una particella in un buco nero

Nel paragrafo precedente abbiamo osservato che usando le coordinate di Schwarzschild (t, r, θ, φ) *un raggio di luce, o un corpo qualsiasi, che parte dall'infinito, cioè nella regione* $r \gg R_S$ *e va verso la superficie di Schwarzschild* $r = R_S$*, in direzione radiale,* **non la raggiungerà mai**. Abbiamo anche detto però che ciò non è attendibile dal punto di vista fisico ed è dovuto alla cattiva scelta della coordinata temporale t, infatti basta sostituire al tempo coordinato t il tempo proprio τ, perché la particella vada a cadere, in un tempo finito, nella singolarità $r = 0$.
Consideriamo dunque il moto di una particella che sia in caduta libera radialmente verso un buco nero, essa si muoverà lungo una geodetica radiale del genere tempo ($\theta = cost.$, $\varphi = cost.$), data dalle equazioni (cfr. le (12.67)):

$$\left(1 - \frac{R_S}{r}\right)\dot{t} = k_2, \tag{12.91}$$

$$\left(1 - \frac{R_S}{r}\right)\dot{t}^2 - \frac{1}{1 - \frac{R_S}{r}}\dot{r}^2 = 1, \tag{12.92}$$

nelle quali si deve ricordare che $\dot{t} \equiv \frac{dt}{ds}$ ed $\dot{r} \equiv \frac{dr}{ds}$ e che k_2 è una costante d'integrazione. Differenti scelte di questa costante d'integrazione corrispondono a diverse condizioni iniziali. La scelta più naturale delle condizioni iniziali per un corpo in caduta libera radiale consiste nell'assumere che esso parta dall'infinito ($r = +\infty$) con velocità iniziale nulla $\left(\frac{dr}{dt} = 0\right)$. Questa scelta comporta, come il lettore può verificare utilizzando le (12.91), (12.92) che $k_2 = 1$.
Indichiamo ora, come di consueto, con τ il tempo proprio lungo la linea di universo della particella ricordandoci che $d\tau = \frac{ds}{c}$, ma avendo supposto di usare unità relativistiche, si ha $c = 1$, dunque $d\tau = ds$. Perciò nelle (12.91) e (12.92) il puntino sopra t ed r può essere interpretato come derivazione rispetto a τ. Combinando queste due equazioni avremo:

$$\frac{d\tau}{dr} = \pm\frac{\sqrt{r}}{\sqrt{R_S}}. \tag{12.93}$$

Tenendo conto del fatto che al decrescere di r da $r = +\infty$ fino ad $r = 0$, centro del buco nero, il tempo proprio cresce, cioè $\frac{d\tau}{dr}$ è una funzione decrescente, per cui bisogna prendere il segno meno nella precedente equazione, ed integrando otteniamo:

$$\tau - \tau_0 = \frac{2}{3\sqrt{R_S}}\left(r_0^{\frac{3}{2}} - r^{\frac{3}{2}}\right), \tag{12.94}$$

Figura 12.2: Particella in caduta libera radiale nelle coordinate (r, t) ed (r, τ).

dove la particella occupa la posizione r_0 al tempo proprio τ_0 (Fig.12.2).

È piuttosto sorprendente che questo risultato è precisamente quello che si ottiene classicamente nell'ambito newtoniano (esercizio).

Come si vede non compare nessun tipo di comportamento singolare al raggio di Schwarzschild ed il corpo cade tranquillamente in maniera continua fino ad $r = 0$ in un tempo proprio finito dato da $\tau - \tau_0 = \dfrac{2r_0^{3/2}}{3\sqrt{R_S}}$.

Se invece, descriviamo il moto della particella utilizzando il tempo coordinato t di Schwarzschild, dalle (12.91) e (12.92) avremo:

$$\frac{dt}{dr} = \frac{\dot{t}}{\dot{r}} = -\frac{1}{\sqrt{R_S}} \frac{r\sqrt{r}}{r - R_S}. \tag{12.95}$$

Integrando, con calcoli un pò tediosi[14], otteniamo:

$$t - t_0 = -\frac{2}{3\sqrt{R_S}} \left(r^{3/2} - r_0^{3/2} + 3R_S\sqrt{r} - 3R_S\sqrt{r_0} \right) + R_S \log \frac{(\sqrt{r} + \sqrt{R_S})(\sqrt{r_0} - \sqrt{R_S})}{(\sqrt{r} - \sqrt{R_S})(\sqrt{r_0} + \sqrt{R_S})}. \tag{12.96}$$

Nei casi in cui r_0 ed r sono molto maggiori di R_S, i risultati (12.94) e (12.96) sono approssimativamente uguali, come ci si aspetta dato che molto lontano dalla superficie di Schwarzschild $r = R_S$ lo spaziotempo è quasi minkowskiano. Se, invece r è molto prossimo ad R_S, dalla (12.96) approssimandola avremo:

$$t - t_0 \approx R_S \log \frac{r_0 - R_S}{r - R_S}, \tag{12.97}$$

da cui risulta chiaro che se $t \to +\infty$ ciò implica $r \to R_S$, ma questo valore non lo raggiunge mai in un tempo finito Fig.12.2.

Possiamo dunque concludere che la coordinata temporale t, che figura nella metrica di Schwarzschild, ha un utile significato fisico perché corrisponde al tempo proprio misurato da un osservatore a riposo posto molto lontano dall'origine (all'infinito), come già supponemmo fin dall'inizio di questo capitolo. Dal punto di vista di tale osservatore, una particella che cade verso il buco nero impiega un tempo infinito per raggiungere la sfera di Schwarzschild $r = R_S$. Dal punto di vista di un osservatore che viaggia con la particella, il tempo τ a lei necessario per raggiungere sia $r = R_S$ sia $r = 0$ è finito. Perciò la coordinata temporale di Schwarzschild t non è appropriata per descrivere la caduta della particella in un buco nero. Inoltre, come abbiamo

[14]Si suggerisce di effettuare sulla (12.95), riscritta con le variabili separate, il cambio di variabile $x = \sqrt{r}$. Ciò conduce al calcolo di $2\int \frac{x^4}{x^2 - R_S} dx$.

visto, il sistema delle coordinate di Schwarzschild t, r, θ, φ rende singolare la metrica per $r = R_S$. Si sente dunque la necessità di operare una trasformazione delle coordinate di Schwarzschild, soprattutto di cambiare la variabile temporale t, in modo tale che nella caduta libera radiale si possa rimuovere tale singolarità: è ciò che faremo nel prossimo paragrafo.

12.9 Coordinate e metrica di Eddington-Finkelstein

Ci sono diversi modi di introdurre nuove coordinate; ci occupiamo ora di uno di questi, definendo una nuova coordinata temporale t', e lasciando invariate le altre coordinate, in maniera tale che le geodetiche radiali nulle diventino linee rette. Per far ciò teniamo conto dell'equazione (12.88) che definisce la doppia congruenza delle geodetiche radiali nulle che nelle nuove coordinate vogliamo che si trasformi in una doppia famiglia di rette. È immediato verificare che la trasformazione cercata è quella che definisce le cosiddette **coordinate di Eddington-Finkelstein**[15]:

$$\begin{cases} t = t' \mp R_S \log\left|\dfrac{r}{R_S} - 1\right|, \qquad \text{per} \quad r \gtrless R_S, \\ \\ r = r', \quad \theta = \theta', \quad \varphi = \varphi'. \end{cases} \tag{12.98}$$

Introducendo queste formule di trasformazione nella metrica di Schwarzschild (12.75) (con $c = 1$) otteniamo:

$$\mathbf{ds}^2 = \left(1 - \frac{R_S}{r}\right)dt'^2 - \left(1 + \frac{R_S}{r}\right)dr^2 \mp \frac{2R_S}{r}dt'\,dr - r^2 d\theta^2 - r^2 \sin^2\theta d\varphi^2. \tag{12.99}$$

Questa è la **forma di Eddington-Finkelstein della metrica di Schwarzschild** che ora discuteremo con qualche dettaglio. Notiamo che la (12.99) ci fornisce in realtà due differenti forme della metrica, corrispondenti al doppio segno del terzo termine. Scriviamo esplicitamente queste due metriche:

$$\mathbf{ds}^2 = \left(1 - \frac{R_S}{r}\right)dt'^2 - \left(1 + \frac{R_S}{r}\right)dr^2 - \frac{2R_S}{r}dt'\,dr - r^2 d\theta^2 - r^2 \sin^2\theta d\varphi^2, \tag{12.100}$$

$$\mathbf{ds}^2 = \left(1 - \frac{R_S}{r}\right)dt'^2 - \left(1 + \frac{R_S}{r}\right)dr^2 + \frac{2R_S}{r}dt'\,dr - r^2 d\theta^2 - r^2 \sin^2\theta d\varphi^2. \tag{12.101}$$

La (12.100) è la cosiddeta soluzione in **coordinate di Eddington-Finkelstein avanzate**, mentre la (12.101) è l'analoga soluzione in **coordinate di Eddington-Finkelstein ritardate**.
Per prima cosa osserviamo che le due metriche (12.100), (12.101) di Eddington-Finkelstein sono ancora *stazionarie*, cioè a coefficienti indipendenti dal tempo, ma ora esse hanno l'inconveniente di non essere più diagonali per la presenza del termine misto in $dt'\,dr$. In compenso *esse hanno il grande pregio di* **non essere più singolari** *sulla superficie $r = R_S$ di Schwarzschild, ma regolari.* Inoltre, come vedremo fra poco, le due regioni spaziali I^a e II^a poste da entrambi i lati della superficie $r = R_S$ ora si raccordano in modo liscio sulla superficie stessa.
Una comprensione più dettagliata della situazione si ottiene esaminando la struttura generale dei coni di luce, delle due metriche ora introdotte, nel piano (t', r) Fig.12.3.

Cominciamo col discutere il caso della prima metrica, la (12.100) in coordinate avanzate. Le direzioni radiali nulle - caratterizzate da $\mathbf{ds}^2 = 0$, θ e φ costanti - sono determinate dall'equazione:

$$\left(1 - \frac{R_S}{r}\right)dt'^2 - \left(1 + \frac{R_S}{r}\right)dr^2 - \frac{2R_S}{r}dt'\,dr = 0. \tag{12.102}$$

[15]Queste coordinate furono scoperte dapprima da Sir Arthur Stanley Eddington nel 1924, e pubblicate su *Nature*, **113** (2832), 192. Riscoperte da David Finkelstein nel 1958 e pubblicate in *Physical Review*, **110** (3), 965-967. Riprese infine da Roger Penrose nel 1965 in un lavoro pubblicato su *Physical Review Letters*, **14** (3), 57-59.

Figura 12.3: Soluzione di Schwarzschild nelle coordinate di Eddington- Finkelstein avanzate $(r.t')$.

Questa può essere così trasformata:

$$dt'^2 - \frac{R_S}{r}dt'^2 - dr^2 - \frac{R_S}{r}dr^2 - \frac{R_S}{r}dt'\,dr - \frac{R_S}{r}dt'\,dr = 0,$$

nella quale mettendo in evidenza, in maniera opportuna, dei fattori comuni avremo:

$$-\frac{R_S}{r}dt'(dt'+dr) + (dt'+dr)(dt'-dr) - \frac{R_S}{r}dr(dt'+dr) = 0,$$

da cui:

$$(dt'+dr)\Big[-\frac{R_S}{r}dt' + dt' - dr - \frac{R_S}{r}dr\Big] = 0,$$

infine otteniamo:

$$(dt'+dr)\Big\{\Big(1-\frac{R_S}{r}\Big)dt' - \Big(1+\frac{R_S}{r}\Big)dr\Big\} = 0. \tag{12.103}$$

Ne segue che le direzioni radiali nulle sono determinate dalle due equazioni differenziali:

$$\frac{dr}{dt'} = -1, \qquad \text{e} \qquad \frac{dr}{dt'} = \frac{r-R_S}{r+R_S}. \tag{12.104}$$

Le linee di universo le cui tangenti hanno l'una o l'altra di queste direzioni (12.104) sono **geodetiche radiali nulle**. Si vede immediatamente che la prima famiglia di tali geodetiche, ha la semplicissima equazione:

$$t' + r = cost. \tag{12.105}$$

Esse costituiscono dunque, come volevamo, un fascio di rette parallele del piano (t', r) inclinate di 45° sull'asse t', ovvero di 135° sull'asse r delle ascisse, come mostra la Fig.12.3 nella quale tali rette sono disegnate in rosso.
Integrando la seconda equazione che figura nella (12.104) otteniamo l'altra congruenza di **geodetiche radiali nulle**:

$$t' = r + 2R_S \log|r - R_S| + cost. \tag{12.106}$$

Sia da questa equazione che dalla seconda delle (12.104) si può vedere facilmente che le tangenti alle geodetiche di questa seconda famiglia hanno le seguenti proprietà:

$$\frac{dr}{dt'} \to 1 \quad \text{per} \quad r \to +\infty; \qquad \frac{dr}{dt'} \to 0 \quad \text{per} \quad r \to R_S;$$

$$\frac{dr}{dt'} \to -1 \quad \text{per} \quad r \to 0.$$

Ne segue, dalla seconda di queste proprietà, che le geodetiche non attraversano la superficie $r = R_S$ ma tendono asintoticamente ad essa per $t' \to -\infty$: esse hanno perciò la forma mostrata in Fig.12.3 nella quale esse sono disegnate in blu.
Si può dimostrare, a partire dalla (12.99), che le direzioni non-radiali nulle come le direzioni del genere tempo hanno valori di $\frac{dr}{dt'}$ compresi fra i due valori (12.104). Conseguentemente i coni di luce, nel piano (t', r), hanno la falda, contenente gli eventi futuri, tangente alle geodetiche radiali nulle, come mostra la Fig.12.3.
Ricordiamo che le particelle materiali si muovono su linee di universo del genere tempo o su linee nulle (cioè del genere luce nel caso dei fotoni), cioè su linee che giacciono all'interno o sulla superficie dei coni di luce. Ne segue allora, come mostra la Fig.12.3 , che nessuna particella può attraversare la superficie $r = R_S$ muovendosi verso l'esterno. Inoltre, ogni particella che si trova, ad un dato momento, dentro la superficie $r = R_S$ (cioè dentro il buco nero) deve muoversi necessariamente verso la linea $r = 0$ (asse t') del piano (t', r) raggiungendola in un tempo finito sia coordinato t' sia proprio τ. Il fatto che nessuna particella può attraversare la superficie $r = R_S$ verso l'esterno, significa che ogni osservatore P situato nella regione $r > R_S$ non può ricevere alcuna informazione circa eventi che accadono dentro la superficie $r = R_S$, essa si comporta quindi come una **membrana a senso unico** a cui abbiamo già dato il nome di **orizzonte degli eventi**.
Il caso della metrica (12.101) può essere discusso ovviamente in maniera analoga del caso della metrica (12.100) per mezzo di una inversione del tempo. Ne segue che in questo caso al posto della Fig.12.3 avremo la nuova Fig.12.4, che si ottiene dalla precedente mediante una riflessione rispetto all'asse r.

Figura 12.4: Soluzione di Schwarzschild nelle coordinate di Eddington- Finkelstein ritardate (r, t').

Ripetendo sulla (12.101) calcoli analoghi a quelli visti in precedenza per la (12.100) otteniamo:

$$(dt' - dr)\left\{\left(1 - \frac{R_S}{r}\right)dt' + \left(1 + \frac{R_S}{r}\right)dr\right\} = 0. \tag{12.107}$$

Ne segue che le direzioni radiali nulle sono determinate dalle due equazioni differenziali:

$$\frac{dr}{dt'} = 1, \qquad \text{e} \qquad \frac{dr}{dt'} = -\frac{r - R_S}{r + R_S}. \tag{12.108}$$

Le linee di universo le cui tangenti hanno l'una o l'altra di queste direzioni (12.108) sono ancora **geodetiche radiali nulle**. Si vede immediatamente che la prima famiglia di tali geodetiche, ha

la semplicissima equazione:

$$t' - r = cost. \tag{12.109}$$

Esse costituiscono un fascio di rette parallele del piano (t', r) inclinate di 45° sia sull'asse t' sia sull'asse r delle ascisse, come mostra la Fig.12.4 nella quale esse sono disegnate in rosso. Integrando la seconda equazione (12.108) otteniamo l'altra congruenza di **geodetiche radiali nulle**:

$$t' = -\Big(r + 2R_S \log|r - R_S| + cost.\Big). \tag{12.110}$$

Sia da questa equazione che dalla seconda delle (12.108) si può vedere facilmente che le tangenti alle geodetiche di questa seconda famiglia hanno le seguenti proprietà:

$$\frac{dr}{dt'} \to -1 \quad \text{per} \quad r \to +\infty; \qquad \frac{dr}{dt'} \to 0 \quad \text{per} \quad r \to R_S;$$

$$\frac{dr}{dt'} \to 1 \quad \text{per} \quad r \to 0.$$

Ne segue, dalla seconda di queste proprietà, che le geodetiche non attraversano la superficie $r = R_S$ ma tendono asintoticamente ad essa per $t' \to +\infty$: esse hanno perciò la forma mostrata in Fig.12.4 nella quale sono disegnate in blu.
Si può dimostrare, a partire dalla (12.99), che le direzioni non-radiali nulle come le direzioni del genere tempo hanno valori di $\frac{dr}{dt'}$ compresi fra i due valori (12.108). Conseguentemente i coni di luce, nel piano (t', r), hanno la falda del futuro tangente alle geodetiche radiali nulle, come mostra la Fig.12.4. Possiamo vedere dalla Fig.12.4 che nessuna particella può attraversare l'orizzonte dall'esterno verso l'interno mentre quelle particelle che ad un certo momento si trovano nella regione $0 < r < R_S$ necessariamente si muovono verso l'esterno e raggiungono l'orizzonte in un tempo proprio finito. L'orizzonte anche qui agisce come **una membrana a senso unico** ma in senso contrario al precedente, possono uscire segnali verso l'esterno ma dall'esterno non entra niente, si tratta di un cosiddetto **buco bianco**. Si osservi che sia i buchi neri che i buchi bianchi sono soluzioni matematiche delle equazioni di Einstein nel vuoto, ciò non vuol dire che abbiano entrambe un significato fisico.[16] Ci sono argomenti che supportano l'esistenza dei buchi neri ma, per il momento, non ci sono evidenze per l'esistenza dei buchi bianchi.

12.10 Coordinate di Eddington-Finkelstein avanzate e ritardate

a) Coordinate avanzate e buchi neri.
Le due metriche (12.100) e (12.101) vengono talvolta trasformate in maniera opportuna introducendo una nuova coordinata temporale.
Cominciamo con l'introdurre questa coordinata per la metrica (12.100), chiamata **coordinata temporale avanzata o coordinata entrante nulla**[17] **di Eddington-Finkelstein** che indicheremo

[16]Un comportamento analogo si ha nello studio della propagazione ondosa riguardo ai potenziali ritardati ed ai potenziali anticipati: sono tutti e due soluzioni analitiche dell'equazione delle onde, ma si ha evidenza fisica soltanto dei primi.

[17]Soffermiamoci un momento sul concetto di **coordinate nulle** che talvolta viene utilizzato già nello spazio di Minkowski la cui metrica, in coordinate polari è $\mathbf{ds}^2 = dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2)$(supposto ovviamente $c = 1$). Come sappiamo i raggi di luce sono caratterizzati dalla $\mathbf{ds}^2 = 0$ e se li rappresentiamo nel piano (r, t) (dove $d\theta = d\varphi = 0$) avranno equazione $(dt - dr)(dt + dr) = 0$, cioè $t \pm r = cost$.
Si dicono **raggi di luce uscenti** quelli per i quali r cresce al crescere di t, essi sono rappresentati nel piano (r, t) dal fascio di rette di equazione $t - r = u = cost.$ inclinate di 45° sull'asse r. Diversi valori di u numerano (sono delle etichette) per i differenti raggi di luce uscenti, a partire dal valore di r corrispondente a $t = 0$. Similmente si dicono **raggi di luce entranti** quelli per i quali r decresce al crescere di t, essi sono rappresentati nel piano (r, t) dal fascio di

con v, ponendo:

$$v = t' + r, \qquad -\infty < v < +\infty. \tag{12.111}$$

Da questa semplice relazione, differenziando si ottiene $dt' = dv - dr$, ed elevando primo e secondo membro al quadrato si ha $dt'^2 = dv^2 - 2dv\,dr + dr^2$ che sostituita nella metrica (12.100), con facili calcoli si ottiene la *metrica di Schwarzschild nelle coordinate* (v, r, θ, φ) *di Eddington-Finkelstein avanzate*:

$$\mathbf{ds}^2 = \left(1 - \frac{R_S}{r}\right) dv^2 - 2dv\,dr - r^2 d\Omega^2, \tag{12.112}$$

avendo posto, per brevità, $d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2$ che è la metrica di una 2-sfera di raggio unitario in coordinate polari sferiche.
In questa forma la metrica, pur essendo non diagonale per la presenza del termine misto $-2dv\,dr = -dv\,dr - dr\,dv$, è ancora non singolare per $r = R_S$, perché il suo determinante è diverso da zero: $g = \det(g_{ik}) = -R_S^4 \sin^2\theta \neq 0$. Consideriamo il diagramma spazio-temporale nel piano (t, r), illustrato nella Fig.12.5,

Figura 12.5: Diagramma di un osservatore in caduta libera in un buco nero nelle coordinate di Eddington-Finkelstein $(r.t')$.

esteso da $r = 0$ a $r = +\infty$. In particolare, possiamo vedere che cosa capita ai coni di luce quando ci avviciniamo alla superficie $r = R_S$ che, in queste coordinate, non è una superficie singolare. Dall'espressione (12.112) della metrica si può vedere che la linea di universo caratterizzata da $r = R_S$, $\theta = cost.$, $\varphi = cost.$ è **un raggio nullo** in quanto si ha $\mathbf{ds}^2 = 0$ in virtù del fatto che si annullano tutti e tre i termini del secondo membro: il 1° è nullo in quanto è $r = R_S$, il 2° è nullo perché essendo $r = R_S = cost.$ si ha $dr = 0$, il 3° è nullo perché essendo costanti sia θ che φ si ha $d\Omega^2 = 0$. Osserviamo infine che la superficie $r = R_S$ è generata da questi raggi nulli (per diversi valori di θ e φ).
Siamo ora in grado, utilizzando la nuova metrica (12.112), di esaminare le proprietà di **un buco nero** all'interno dell'orizzonte degli eventi, riottenendo i risultati già visti utilizzando la metrica (12.100) ma in una nuova luce.
Riscriviamo la metrica (12.112) nella forma:

$$2dv\,dr = -\left[\mathbf{ds}^2 + \left(\frac{R_S}{r} - 1\right) dv^2 + r^2 d\Omega^2\right]. \tag{12.113}$$

rette di equazione $t + r = v = cost.$ inclinate di 135° sull'asse r. Diversi valori di v numerano (sono delle etichette) per i differenti raggi di luce entranti a partire dal valore di r corrispondente a $t = 0$. Lungo un raggio entrante u varia da punto a punto *sicché u numera i punti su di un dato raggio entrante*, viceversa *v numera i punti su di un dato raggio uscente. Per questa loro caratteristica di numerare i raggi di luce (caratterizzati da $\mathbf{ds}^2 = 0$), le coordinate u e v sono spesso chiamate* **coordinate nulle rispettivamente uscenti ed entranti**. Questa terminologia viene qui estesa al caso delle coordinate di Eddington-Finkelstein.

Per uno spostamento del genere tempo o nullo, essendo $\mathbf{ds}^2 \geq 0$, se $0 < r < R_S$ la somma fra parentesi quadre della (12.113) è positiva e perciò il secondo membro è negativo, dunque deve essere negativo anche il primo membro. Ricordiamo ora che su di un raggio radiale nullo è $dv = d(t' + r) = 0$, e quindi uno spostamento dv giace entro il cono di luce futuro se $dv > 0$. Perciò per rendere negativo il primo membro della (12.113) per uno spostamento che punta verso il futuro si deve avere $dr < 0$. Dunque uno spostamento che punta verso il futuro è necessariamente diretto verso l'interno.
L'orizzonte degli eventi taglia fuori dal mondo esterno tutti gli eventi che accadono all'interno della superficie $r = R_S$, sicché un esploratore, proveniente da lontano, che oltrepassi l'orizzonte degli eventi di un buco nero non può mai più tornare indietro né inviare segnali verso l'esterno. Ovviamente tale osservatore non può evitare la sua distruzione nella singolarità non eliminabile $r = 0$ che è nel suo futuro.
a) Coordinate ritardate e buchi bianchi.
Analoghe considerazioni si possono fare per la metrica (12.101), nel senso che anche per essa è possibile introdurre una nuova coordinata temporale chiamata **coordinata temporale ritardata o coordinata uscente nulla di Eddington-Finkelstein** che indicheremo con w, ponendo:

$$w = t' - r, \qquad -\infty < w < +\infty. \tag{12.114}$$

Con analogo procedimento di quello visto per la coordinata temporale avanzata otteniamo la *metrica di Schwarzschild nelle coordinate* (w, r, θ, φ) *di Eddington-Finkelstein ritardate*:

$$\mathbf{ds}^2 = \Big(1 - \frac{R_S}{r}\Big) dw^2 + 2dw\, dr - r^2 d\Omega^2 \tag{12.115}$$

Anche qui la metrica, pur essendo non diagonale per la presenza del termine misto $2dw\, dr = dw\, dr + dr\, dw$, è ancora non singolare per $r = R_S$, perché il suo determinante è diverso da zero: $g = \det(g_{ik}) = -R_S^4 \sin^2\theta \neq 0$ (verificare per esercizio).
Riconsideriamo il diagramma spazio-temporale nel piano (t, r) della Fig.12.5 già esaminata, nella quale si mostra il piano esteso da $r = 0$ a $r = +\infty$. In particolare, possiamo vedere che cosa capita ai coni di luce quando ci avviciniamo alla superficie $r = R_S$ che, in queste coordinate, non è una superficie singolare.
Riscriviamo la metrica (12.115) nella forma:

$$2dw\, dr = \mathbf{ds}^2 + \Big(\frac{R_S}{r} - 1\Big) dw^2 + r^2 d\Omega^2. \tag{12.116}$$

Nella regione $r \leq R_S$ risulta:

$$dw\, dr \geq 0, \quad \text{per} \quad \mathbf{ds}^2 \geq 0. \tag{12.117}$$

Dato che deve essere $dw \geq 0$ per linee di universo dirette verso il futuro, deve anche essere $dr \geq 0$. Il segno d'uguaglianza vale solo nel caso in cui $r = R_S$, $d\Omega^2 = 0$, e $\mathbf{ds}^2 = 0$. In questo caso una stella con una superficie di raggio $r < R_S$ deve per forza espandersi oltre la superficie $r = R_S$. A questa soluzione che è l'inverso temporale del buco nero, come abbiamo già detto, è stato dato il nome di **buco bianco**.
Vi è però un limite all'uso di queste coordinate di Eddington-Finkelstein: quelle avanzate (v, r, θ, φ) descrivono bene la caduta verso l'interno della sfera di Schwarzschild e sono quindi utilizzate per lo studio del collasso gravitazionale e dei buchi neri, anche se ci può essere qualche piccola complicazione per il fatto che la metrica non è diagonale. Esse però non sono adeguate allo studio di particelle che dall'interno dell'orizzonte degli eventi si muovono verso l'esterno. Le coordinate ritardate (w, r, θ, φ) invece descrivono bene l'uscita di particelle da $r = 0$ attraverso $r = R_S$, ma hanno lo stesso problema delle coordinate di Schwarzschild nel descrivere il moto opposto (ovvero qualsiasi particella entrante raggiungerebbe $r = R_S$ in un tempo infinito).

Inoltre vi è un comportamento insoddisfacente tra i due sistemi per le geodetiche in $r = R_S$: in coordinate avanzate le geodetiche descrivono fotoni entranti, mentre in quelle ritardate si prevedono solo fotoni uscenti. Questa discrepanza si risolve utilizzando un sistema di coordinate più regolare, come quello scoperto nel 1960 dal grande matematico americano Martin David Kruskal, del quale ci occuperemo dettagliatamente nei seguenti due paragrafi.

12.11 Relazioni fra le coordinate di Schwarzschild e quelle di Kruskal

Abbiamo visto nel paragrafo precedente che la soluzione di Schwarzschild per $R_S < r < +\infty$ può essere estesa sia alla soluzione avanzata (12.112) che alla soluzione ritardata (12.115) di Eddington-Finkelstein, nelle quali $0 < r < +\infty$. Che ciò sia possibile è indicato dal fatto che una geodetica radiale nulla può essere estesa attraverso $r = R_S$ ed arrivare fino alla singolarità essenziale $r = 0$. Sorge allora spontanea la domanda: è possibile estendere ulteriormente queste soluzioni? O meglio, è possibile trovare un sistema di coordinate in modo tale che la metrica (la soluzione) di Schwarzschild espressa per mezzo di queste coordinate copra in *modo massimale* lo spaziotempo eliminando qualsiasi singolarità che non sia essenziale? Per rispondere a queste domande è necessario introdurre due semplici definizioni riguardanti le varietà.
Definizione 1:
Una varietà dotata di una geometria metrica o affine è detta **geodeticamente massima** *se qualsiasi geodetica che parte da un punto arbitrario della varietà, o può essere estesa per valori infiniti del parametro affine lungo la geodetica un entrambe le direzioni, o termina in una singolarità intrinseca.*
Definizione 2:
Se, in particolare, **tutte** *le geodetiche che partono da un* **qualsiasi punto** *possono essere estese per valori infiniti del parametro affine in entrambe le direzioni, la varietà è detta* **geodeticamente completa**.
È chiaro che una varietà geodeticamente completa è anche massima, l'inverso, in generale, è falso. Per esempio lo spazio di Minkowski è ovviamente una varietà geodeticamente completa. Invece, né la varietà spaziotemporale descritta dalla metrica di Schwarzschild, né le due metriche, avanzata e ritardata, di Eddington- Finkelstein sono geodeticamente massime. Perciò è un fatto notevole che Kruskal abbia trovato un sistema di coordinate in cui la soluzione di Schwarzschild, espressa in queste coordinate, ci fornisca **la massima estensione analitica** ed inoltre è **unica**. Tuttavia anche la soluzione di Kruskal, anche se massima, non è completa per l'esistenza di una singolarità intrinseca. L'idea geniale di Kruskal consiste nell'effettuare una trasformazione di coordinate in maniera tale che le geodetiche radiali nulle, sia uscenti che entranti, siano rette. Ma il prezzo che bisogna pagare per queste coordinate, per così dire, "buone" è che la metrica non risulta più statica, neanche nella regione esterna alla sfera di Schwarzschild.
Svilupperemo ora dettagliatamente l'idea di Kruskal.[18] Prima di far ciò, osserviamo un aspetto particolare che presenta la propagazione della luce in termini delle coordinate di Schwarzschild. Le linee di universo dei raggi di luce sono caratterizzati dall'annullarsi dell'elemento di linea, cioè da $\mathbf{ds^2} = 0$. Perciò per un raggio di luce che si propaga radialmente (θ e φ costanti), la metrica di Schwarzschild (12.75) ci dà:

$$\left(1 - \frac{R_S}{r}\right)c^2dt^2 - \left(1 - \frac{R_S}{r}\right)^{-1}dr^2 = 0. \tag{12.118}$$

Ciò significa che il quadrato della velocità radiale della luce è dato da:

$$\left(\frac{dr}{dt}\right)^2 = c^2\left(1 - \frac{R_S}{r}\right)^2 = c^2 g_{00}^2, \tag{12.119}$$

[18] Il lavoro fondamentale di M. D. Kruskal è: *Maximal Extension of Schwarzschild metric*, Phys. Rev., **119**, 1743, (1960).

sicché per $r = R_S$ la velocità radiale della luce si annulla. Questo fatto ottenuto nelle coordinate di Schwarzschild (t, r, θ, φ) è assolutamente indesiderabile: esso può essere eliminato come segue. Cerchiamo una trasformazione che lasciando invariate θ, φ trasformi le coordinate t, r in due nuove coordinate v, u ($t \to v$, $r \to u$; v coordinata temporale e u coordinata spaziale) in modo tale che tutte le geodetiche radiali nulle si trasformino in rette.[19] L'idea geniale di Kruskal consistette nello scrivere la metrica nelle nuove coordinate (v, u, θ, φ) nel modo seguente:

$$ds^2 = f^2(u, v)(dv^2 - du^2) - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{12.120}$$

con f^2 funzione da determinarsi purché sia priva di zeri. Infatti le geodetiche radiali nulle di questa metrica (rappresentanti raggi di luce) hanno ovviamente equazioni date da $(dv - du)(dv + du) = 0$, cioè $v - u = cost.$ e $v + u = cost.$ che nel piano (u, v) rappresentano due fasci di rette rispettivamente parallele alle bisettrici del 1° e 3° quadrante e del 2° e 4° quadrante.
Utilizzando lo stesso procedimento di prima possiamo constatare che la velocità radiale della luce è ovunque unitaria :

$$\left(\frac{du}{dv}\right)^2 = 1. \tag{12.121}$$

Perciò nelle nuove coordinate u, v non esistono limiti imposti alla propagazione della luce come quelli che invece impone la metrica di Schwarzschild, e ciò è soddisfacente.
A questo punto abbiamo bisogno dell'equazione tensoriale che ci fa passare dal tensore metrico $g_{ik}(x^j)$, espresso per mezzo delle coordinate $(x^j) \equiv (x^0, x^\alpha) \equiv (x^0, x^1, x^2, x^3)$, al tensore $g_{i'k'}(x'^j)$, espresso per mezzo delle nuove coordinate $(x^{j'}) \equiv (x'^0, x'^\alpha) \equiv (x'^0, x'^1, x'^2, x'^3)$:

$$g_{ik} = \frac{\partial x^{j'}}{\partial x^i}\frac{\partial x^{h'}}{\partial x^k} g_{j'h'}. \tag{12.122}$$

Identifichiamo le g_{ik} con i coefficienti della **metrica di Schwarzschild**, espressa nelle coordinate $(x^0, r, \theta, \varphi)$, che qui riscriviamo, ricordando che $x^0 = ct$:

$$\mathbf{ds^2} = \left(1 - \frac{R_S}{r}\right)(dx^0)^2 - \left(1 - \frac{R_S}{r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{12.123}$$

inoltre identifichiamo le $g_{i'k'}(x'^j)$ con i coefficienti della **metrica di Kruskal** (12.120) tenendo conto che le nuove coordinate sono $(x'^0, x'^1, x'^2, x'^3) \equiv (v, u, \theta, \varphi)$, perciò, si badi bene, che v ed u sono funzioni delle sole variabili x^0 ed r, cioè: $v = v(x^0, r)$ e $u = u(x^0, r)$.
Se ora applichiamo la formula di trasformazione (12.122) non è difficile, anche se un po' laborioso, ottenere le seguenti equazioni differenziali che è necessario risolvere per ottenere l'espressione della funzione incognita f^2:

$$\begin{cases} 1 - \dfrac{R_S}{r} = f^2\left[\left(\dfrac{\partial v}{\partial x^0}\right)^2 - \left(\dfrac{\partial u}{\partial x^0}\right)^2\right], \\[2ex] -\left(1 - \dfrac{R_S}{r}\right)^{-1} = f^2\left[\left(\dfrac{\partial v}{\partial r}\right)^2 - \left(\dfrac{\partial u}{\partial r}\right)^2\right], \\[2ex] 0 = \dfrac{\partial u}{\partial x^0}\dfrac{\partial u}{\partial r} - \dfrac{\partial v}{\partial x^0}\dfrac{\partial v}{\partial r}. \end{cases} \tag{12.124}$$

Si osservi che i segni di u e v non sono determinati da queste equazioni. Al fine di semplificare i calcoli, a questo punto, è utile introdurre un nuovo parametro radiale ξ (chiamato da qualche

[19] Si badi bene a non confondere le coordinate di Kruskal u, v, che abbiamo denotato con le notazioni usate nel lavoro originale citato nella nota precedente, con le analoghe di Eddington-Finkelstein generalmente indicate con le stesse lettere. Le coordinate di Kruskal son talvolta indicate con lettere maiuscole U, V, oppure con T, R, o ancora con T, X.

autore **coordinata radiale di Regge-Wheeler** o anche **"coordinata tartaruga"**)[20] ed una nuova funzione incognita $F(\xi)$, così definiti:

$$\begin{cases} \xi = r + R_S \log\left|\dfrac{r}{R_S} - 1\right|, \\ \\ F(\xi) = \dfrac{1 - R_S/r}{f^2(r)}. \end{cases} \tag{12.125}$$

Qui si introduce un'ipotesi fondamentale, che la funzione f da determinarsi sia funzione della sola r; **questo è un punto critico** *perché esistono infinite trasformazioni che conducono ad una metrica della forma* (12.120)*, ma soltanto tale assunzione conduce alla metrica di Kruskal rimuovendo la singolarità di coordinate per $r = R_S$.*
Le relazioni (12.124) ora si semplificano nelle:

$$\left(\frac{\partial v}{\partial x^0}\right)^2 - \left(\frac{\partial u}{\partial x^0}\right)^2 = F(\xi), \tag{12.126}$$

$$\left(\frac{\partial v}{\partial \xi}\right)^2 - \left(\frac{\partial u}{\partial \xi}\right)^2 = -F(\xi), \tag{12.127}$$

$$\frac{\partial u}{\partial x^0}\frac{\partial u}{\partial \xi} = \frac{\partial v}{\partial x^0}\frac{\partial v}{\partial \xi}. \tag{12.128}$$

Se sommiamo l'equazione (12.126) con la (12.127) si ottiene

$$\left(\frac{\partial v}{\partial x^0}\right)^2 + \left(\frac{\partial v}{\partial \xi}\right)^2 - \left[\left(\frac{\partial u}{\partial x^0}\right)^2 + \left(\frac{\partial u}{\partial \xi}\right)^2\right] = 0. \tag{12.129}$$

Utilizzando le equazioni (12.128) e (12.129) è immediato verificare che:

$$\left(\frac{\partial v}{\partial x^0} + \frac{\partial v}{\partial \xi}\right)^2 = \left(\frac{\partial u}{\partial x^0} + \frac{\partial u}{\partial \xi}\right)^2, \tag{12.130}$$

$$\left(\frac{\partial v}{\partial x^0} - \frac{\partial v}{\partial \xi}\right)^2 = \left(\frac{\partial u}{\partial x^0} - \frac{\partial u}{\partial \xi}\right)^2. \tag{12.131}$$

Estraendo la radice quadrata di ambo i membri delle (12.130), (12.131) avremo:

$$\frac{\partial v}{\partial x^0} + \frac{\partial v}{\partial \xi} = \pm\left(\frac{\partial u}{\partial x^0} + \frac{\partial u}{\partial \xi}\right), \tag{12.132}$$

$$\frac{\partial v}{\partial x^0} - \frac{\partial v}{\partial \xi} = \pm\left(\frac{\partial u}{\partial x^0} - \frac{\partial u}{\partial \xi}\right). \tag{12.133}$$

Se consideriamo la prima equazione col segno $+$ al secondo membro, mentre la seconda la consideriamo col segno $-$ e poi sommiamo e sottraiamo membro a membro otteniamo le due relazioni[21]:

$$\frac{\partial v}{\partial x^0} = \frac{\partial u}{\partial \xi}, \qquad \frac{\partial v}{\partial \xi} = \frac{\partial u}{\partial x^0} \tag{12.134}$$

dalle quali è facile dedurre:

$$\frac{\partial^2 u}{\partial (x^0)^2} - \frac{\partial^2 u}{\partial \xi^2} = 0, \qquad \frac{\partial^2 v}{\partial (x^0)^2} - \frac{\partial^2 v}{\partial \xi^2} = 0. \tag{12.135}$$

[20] Il parametro ξ, definito dalla prima delle (12.125), viene chiamato **"coordinata tartaruga"** in quanto rallenta, in qualche modo, la variazione della coordinata radiale perché, al crescere o al decrescere di r, e precisamente per $r \to R_S$, si ha $\log\left|\dfrac{r}{R_S} - 1\right| \to -\infty$.

[21] Se dovessimo prendere gli stessi segni si annullerebbe lo Jacobiano della trasformazione.

Dunque entrambe le nuove variabili u e v soddisfano l'equazione alle derivate parziali delle onde in x^0 e ξ (se avessimo scelto nelle (12.132), (12.133) i segni opposti a quelli scelti in precedenza si sarebbero trovate le stesse due equazioni (12.135)).
Com'è noto, la soluzione generale dell'equazione delle onde (12.135) è la somma algebrica di due funzioni arbitrarie, che siano almeno di classe C^2, l'una funzione dell'argomento $\xi + x^0$, l'altra funzione di $\xi - x^0$. Perciò le soluzioni sia delle equazioni (12.134) sia delle (12.135), si può facilmente verificare che sono:

$$\begin{cases} v = h(\xi + x^0) + g(\xi - x^0), \\ \\ u = h(\xi + x^0) - g(\xi - x^0), \end{cases} \tag{12.136}$$

dove h e g sono funzioni da determinarsi. Adesso sostituiamo u e v espresse dalle (12.136) nelle equazioni (12.126), (12.127), (12.128). Si trova allora che la (12.128) è automaticamente soddisfatta perché si riduce ad una identità, mentre le (12.126), (12.127) sono equivalenti e forniscono una sola equazione:

$$-4h'(\xi + x^0)g'(\xi - x^0) = F(\xi), \tag{12.137}$$

dove l'apice posto sopra le funzioni h e g indica derivazione rispetto al loro argomento. Questa è un'equazione rimarchevole che ci condurrà alla determinazione univoca, a meno di una costante inessenziale, delle funzioni h, g ed F.
Finora non abbiamo posto nessuna limitazione alla variabilità della coordinata radiale r. Adesso dobbiamo specificare se $r > R_S$ ovvero se $r < R_S$, poiché nelle due regioni avremo differenti funzioni di trasformazione che devono essere raccordate assieme sul confine delle due regioni, cioè in $r = R_S$. Consideriamo dapprima $r > R_S$, nel qual caso F è positiva per la seconda della (12.125). Per risolvere la (12.137) deriviamola rispetto a ξ ed a x^0, ottenendo:

$$\frac{F'(\xi)}{F(\xi)} = \frac{h''(\xi + x^0)}{h'(\xi + x^0)} + \frac{g''(\xi - x^0)}{g'(\xi - x^0)}, \tag{12.138}$$

$$0 = \frac{h''(\xi + x^0)}{h'(\xi + x^0)} - \frac{g''(\xi - x^0)}{g'(\xi - x^0)}. \tag{12.139}$$

Allora, tenendo conto di queste ultime due relazioni, possiamo scrivere:

$$\Big[\log F(\xi)\Big]' = 2\big[\log h'(\xi + x^0)\big]'. \tag{12.140}$$

Possiamo trattare ξ ed $y = \xi + x^0$ come variabili indipendenti e ciò significa che il primo membro della (12.140) è funzione di una variabile che è indipendente dalla variabile da cui dipende il secondo membro, ciò implica che entrambi i membri devono essere uguali ad una stessa costante arbitraria che indichiamo, per convenienza di calcolo, con 2λ. Possiamo dunque riscrivere la (12.140) con più precisione nel modo seguente:

$$\frac{d}{d\xi}\Big[\log F(\xi)\Big] = 2\frac{d}{dy}\big[\log h'(y)\big] = 2\lambda. \tag{12.141}$$

Quest'ultima equivale a due equazioni; cominciamo con la prima:

$$\frac{d}{d\xi}\Big[\log F(\xi)\Big] = 2\lambda, \tag{12.142}$$

essa ammette la soluzione

$$F(\xi) = Ae^{2\lambda\xi}, \tag{12.143}$$

dove A è una costante positiva (il lettore provi quest'ultima affermazione). L'altra equazione contenuta nella (12.141) si riduce alla:

$$\frac{d}{dy}\big[\log h'(y)\big] = \lambda$$

la cui soluzione è

$$h'(y) = Be^{\lambda y}\,, \tag{12.144}$$

dove B è una costante positiva (anche in questo caso il lettore è invitato a verificare questo fatto). Sostituendo le formule (12.143) e (12.144) nella (12.137), si trova

$$g'(\xi - x^0) = -De^{\lambda(\xi - x^0)}\,, \tag{12.145}$$

dove $D = \frac{A}{4B} > 0$.
È immediato risolvere le equazioni (12.144) e (12.145) e si trova:

$$h(\xi + x^0) = \frac{B}{\lambda}e^{\lambda(\xi + x^0)} + H_1, \qquad g(\xi - x^0) = \frac{-D}{\lambda}e^{\lambda(\xi - x^0)} + H_2,$$

dove H_1 e H_2 sono due costanti additive arbitrarie che scegliamo entrambe uguali a zero: $H_1 = H_2 = 0$. Le costanti moltiplicative le scegliamo invece entrambe di modulo $\frac{1}{2}$, cioè $\frac{B}{\lambda} = \frac{1}{2}$ e $\frac{-D}{\lambda} = -\frac{1}{2}$. Queste scelte delle costanti conducono alle seguenti espressioni, particolarmente semplici, per le funzioni $h(\xi + x^0)$ e $g(\xi - x^0)$:

$$h(\xi + x^0) = \frac{1}{2}e^{\lambda(\xi + x^0)}, \qquad g(\xi - x^0) = -\frac{1}{2}e^{\lambda(\xi - x^0)}.$$

Poiché $h'(\xi + x^0) = \frac{\lambda}{2}e^{\lambda(\xi + x^0)}$ e $g'(\xi - x^0) = -\frac{\lambda}{2}e^{\lambda(\xi - x^0)}$, è facile dalla (12.137) determinare l'espressione per la funzione $F(\xi)$. In definitiva, per le tre funzioni h, g ed F avremo le seguenti espressioni particolarmente semplici:

$$h(\xi + x^0) = \frac{1}{2}e^{\lambda(\xi + x^0)}, \qquad g(\xi - x^0) = -\frac{1}{2}e^{\lambda(\xi - x^0)}, \qquad F(\xi) = \lambda^2 e^{2\lambda\xi}. \tag{12.146}$$

Ora, tenendo conto delle (12.136) e delle prime due delle (12.146) otteniamo:[22]

$$\begin{cases} u = \dfrac{1}{2}\Big[e^{\lambda(\xi + x^0)} + e^{\lambda(\xi - x^0)}\Big] = e^{\lambda\,\xi}\Big[\dfrac{e^{\lambda x^0} + e^{-\lambda x^0}}{2}\Big] = e^{\lambda\,\xi}\,\cosh(\lambda x^0), \\[2ex] v = \dfrac{1}{2}\Big[e^{\lambda(\xi + x^0)} - e^{\lambda(\xi - x^0)}\Big] = e^{\lambda\,\xi}\Big[\dfrac{e^{\lambda x^0} - e^{-\lambda x^0}}{2}\Big] = e^{\lambda\,\xi}\,\sinh(\lambda x^0). \end{cases} \tag{12.147}$$

Ci rimane da valutare $e^{\lambda\,\xi}$: ciò si fa agevolmente ricordando la definizione della variabile ξ data dalla prima delle (12.125), con semplici passaggi avremo:

$$e^{\lambda\,\xi} = \Big(\frac{r}{R_S} - 1\Big)^{\lambda R_S} e^{\lambda r}.$$

[22] Ricordiamo al lettore che le funzioni iperboliche sono così definite:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2}, \quad \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^{2x} - 1}{e^{2x} + 1} = \frac{1 - e^{-2x}}{1 + e^{-2x}}.$$

Perciò le (12.147) ci forniscono la trasformazione:

$$\begin{cases} u = \left(\dfrac{r}{R_S} - 1\right)^{\lambda R_S} e^{\lambda r} \cosh(\lambda x^0), \\ \\ v = \left(\dfrac{r}{R_S} - 1\right)^{\lambda R_S} e^{\lambda r} \sinh(\lambda x^0), \end{cases} \tag{12.148}$$

inoltre, per la seconda delle (12.125), per la terza delle (12.146), tenuto conto dell'espressione data più sopra per $e^{\lambda \xi}$, avremo anche:

$$f^2 = \frac{R_S}{\lambda^2 r}\left(\frac{r}{R_S} - 1\right)^{1-2\lambda R_S} e^{-2\lambda r}. \tag{12.149}$$

Rimane infine da stabilire il valore del parametro arbitrario λ; per far ciò richiediamo che la funzione f^2 non abbia zeri né singolarità per $r = R_S$, ciò implica che sia $\lambda = \dfrac{1}{2R_S}$ (sarà allora $\mathbf{ds^2} = 0$ soltanto sul cono di luce). Se usiamo le unità in cui $c = 1$, dato che $x^0 = ct$, possiamo, nelle formule precedenti sostituire x^0 con t, perciò **la trasformazione di coordinate appropriata per la regione esterna al raggio di Schwarzschild**, cioè **all'esterno di un buco nero**, è **la celebre trasformazione di Kruskal**:

$$\begin{cases} u = \sqrt{\dfrac{r}{R_S} - 1}\, e^{r/2R_S} \cosh\left(\dfrac{t}{2R_S}\right), \\ \\ v = \sqrt{\dfrac{r}{R_S} - 1}\, e^{r/2R_S} \sinh\left(\dfrac{t}{2R_S}\right), \end{cases} \quad \text{per } r > R_S, \tag{12.150}$$

È inoltre stabilito in maniera univoca il valore di:

$$f^2 = \frac{4R_S^3}{r}\, e^{-r/R_S}. \tag{12.151}$$

Nelle relazioni (12.146), abbiamo scelto arbitrariamente la funzione h positiva, sicché g doveva essere negativa per la condizione $F > 0$. Naturalmente noi avremmo potuto scegliere ugualmente bene i segni opposti per le funzioni h e g ottenendo così segni opposti per u e v nelle (12.150).
Per $r < R_S$ è necessario ripetere parte delle nostre precedenti considerazioni. Dalla seconda delle (12.125) segue che $F(\xi)$ deve essere negativa, ciò implica che le funzioni h e g devono essere entrambe positive. Procedendo come prima, otteniamo **la trasformazione di Kruskal appropriata per la regione interna al raggio di Schwarzschild, cioè nel buco nero**:

$$\begin{cases} u = \sqrt{1 - \dfrac{r}{R_S}}\, e^{r/2R_S} \sinh\left(\dfrac{t}{2R_S}\right), \\ \\ v = \sqrt{1 - \dfrac{r}{R_S}}\, e^{r/2R_S} \cosh\left(\dfrac{t}{2R_S}\right), \end{cases} \quad \text{per } r < R_S, \tag{12.152}$$

con f^2 dato dalla stessa espressione (12.151).
Le coordinate (u, v) utilizzate in queste trasformazioni sono note anche come **coordinate di Kruskal-Szekeres**, in quanto sembra che il matematico ungherese George Szekeres sia arrivato più o meno agli stessi risultati di Martin D. Kruskal.
Per renderci conto di tutti gli aspetti salienti delle trasformazioni di Kruskal (che esamineremo

nel prossimo paragrafo) ricaviamoci le trasformazioni inverse, facilmente deducibili sia dalle (12.150), sia dalle (12.152), avremo rispettivamente:[23]

$$u^2 - v^2 = \left(\frac{r}{R_S} - 1\right) e^{r/R_S}, \qquad \frac{v}{u} = \tanh\left(\frac{t}{2R_S}\right) \qquad \text{per } r > R_S, \tag{12.153}$$

e

$$v^2 - u^2 = \left(1 - \frac{r}{R_S}\right) e^{r/R_S}, \qquad \frac{u}{v} = \tanh\left(\frac{t}{2R_S}\right) \qquad \text{per } r < R_S. \tag{12.154}$$

Si noti che entrambe le coppie di equazioni (12.153), (12.154) definiscono implicitamente le coordinate r e t di Schwarzschild come funzioni delle coordinate di Kruskal: $r = r(u, v)$ e $t = t(u, v)$.
Infine, val la pena osservare che le due possibili differenze dei quadrati delle coordinate di Kruskal, $u^2 - v^2$, o $v^2 - u^2$, che compaiono nelle (12.153), (12.154), forniscono in effetti, una sola equazione

$$u^2 - v^2 = \left(\frac{r}{R_S} - 1\right) e^{r/R_S}. \tag{12.155}$$

12.12 Diagramma spazio-temporale e metrica di Kruskal

Per studiare in maniera adeguata le trasformazioni di Kruskal (12.150), (12.152) è opportuno disegnare un diagramma spazio-temporale 2-dimensionale riportando su di un piano, come coordinate cartesiane ortogonali, le coordinate di Kruskal (u, v): u sull'asse delle ascisse e v sull'asse delle ordinate (Fig.12.6).

Figura 12.6: Le quattro regioni del diagramma di Kruskal.

Questo diagramma rappresenta *l'estensione massimale della geometria di Schwarzschild*; le coordinate coprono ciò che noi pensiamo sia la varietà descritta da questa soluzione.
Le coordinate originali di Schwarzschild (t, r), erano buone per $r > R_S$, che è solo una parte della varietà disegnata nel diagramma di Kruskal (regione I).
Da entrambe le seconde formule che figurano nelle (12.153), (12.154) si deduce che, per $t \to +\infty$ si ha $u = v$ (bisettrice del 1° e 3° quadrante) e per $t \to -\infty$ si ha invece $u = -v$ (bisettrice del 2° e 4° quadrante)[24]. Da ciò segue che le trasformazioni (12.150) e (12.152) si raccordano sulle

[23] Si ricordi che $\cosh^2 x - \sinh^2 x = 1$.

[24] Per vedere ciò, basta tener presente che $\tanh x = \dfrac{e^{2x} - 1}{e^{2x} + 1} = \dfrac{1 - e^{-2x}}{1 + e^{-2x}}$ per cui è

$$\lim_{x \to +\infty} \tanh x = 1, \qquad \lim_{x \to -\infty} \tanh x = -1.$$

due rette limite (vedi le (12.153), (12.154)): $u = v\,(r = R_S,\ t = +\infty)$, $u = -v\,(r = R_S,\ t = -\infty)$. Queste trasformazioni sono perciò compatibili.
Le due bisettrici, $u = \pm v$, dividono il piano (u, v) in quattro regioni che indicheremo con I, II, III, IV, la prima e la terza di queste regioni contengono l'asse Ou, la seconda e la quarta contengono l'asse Ov, come indicato nella Fig.12.6.
La regione del piano (u, v) definita dalle (12.150) per $r > R_S$ si identifica con $u > |v|$ $(-v < u < v)$: essa è indicata con I nella Fig.12.6. Invece la regione definita dalle (12.152) per $r < R_S$ si identifica con la $v > |u|$ $(-u < v < u)$ ed è indicata con II nella Fig.12.6.

Figura 12.7: Diagramma spazio-temporale della soluzione di Kruskal.

Alcune linee speciali del diagramma di Kruskal sono di notevole interesse: per esempio per ogni valore fissato di x^0 e quindi di t, la linea $r = R_S$ del piano (r, t), di separazione fra la regione esterna e quella interna all'orizzonte degli eventi (cfr. le Fig.12.1, Fig.12.3, Fig.12.4), corrisponde nel piano (u, v) all'origine delle coordinate, cioè al punto $u = v = 0$ come si vede chiaramente dalle formule di trasformazione (12.150), (12.152).
Considerevole comprensione delle trasformazioni (12.150) e (12.152) si ottiene se si disegnano, sul diagramma di Kruskal, le linee $r = cost.$ e $t = cost.$
A) Cominciamo col considerare le $r = cost. \neq R_S$. Dalla (12.155), si vede chiaramente che tali linee hanno equazione

$$u^2 - v^2 = cost., \tag{12.156}$$

e perciò si tratta di due fasci di **iperbole equilatere** del piano (u, v), l'uno caratterizzato da un valore positivo della costante (per $r > R_S$), giace nelle regioni I, III, l'altro caratterizzato da un valore negativo della costante (per $r < R_S$), giace nelle regioni II, IV (vedi Fig.12.7).
Se $r = R_S$ la costante al secondo membro dell'equazione (12.156), si annulla riducendosi dunque alla $u^2 - v^2 = 0$ che rappresenta le due rette $v = \pm u$: cioè le iperboli equilatere degenerano nelle bisettrici dei quattro quadranti in cui il piano (u, v) è diviso dagli assi coordinati Ou (asse delle ascisse) e Ov (asse delle ordinate), come abbiamo già visto.
Al valore singolare $r = 0$ (compatibile soltanto con $r < R_S$), per la (12.155), corrisponde invece l'iperbole:

$$u^2 - v^2 = -1, \qquad \text{per } r < R_S. \tag{12.157}$$

Come sappiamo, questa iperbole è formata da due rami: il ramo superiore giacente nella regione II (detto talvolta *singolarità futura*), ed il ramo inferiore giacente nella regione IV (detto talvolta

singolarità passata)(disegnate in blu nella Fig. 12.7). Di questi due rami di iperbole va considerato soltanto il ramo superiore (il solo che abbia significato fisico) di equazione:

$$v = +\sqrt{u^2 + 1}, \tag{12.158}$$

come si vede dalla (12.152) essendo il cosh $\left(\frac{t}{2R_S}\right)$ positivo per qualsiasi valore di t $(-\infty < t < +\infty)$. *La regione del diagramma di Kruskal compresa fra questo ramo superiore dell'iperbole* (12.157), *rappresentante la singolarità, e le due bisettrici $v = \pm u$ è la regione che si identifica col* **buco nero** (Fig.12.7).
Si osservi infine che le regioni del piano (u, v), nel diagramma di Kruskal, situate *al di sopra della singolarità futura e al disotto della singolarità passata non fanno parte dello spaziotempo.* La regione compresa tra il ramo inferiore dell'iperbole (12.157) e le rette $v = \pm u$ è ciò che talvolta viene chiamato **"buco bianco"**, esso è l'opposto temporale del **buco nero**.
B) Consideriamo ora, nel diagramma di Kruskal, le linee $t = cost.$ (Fig.12.7). Dalle seconde relazioni riportate nelle (12.153), (12.154) si deduce che le linee $t = cost.$ sono linee di equazione $\frac{v}{u} = cost.$, ovvero, il che è lo stesso, $\frac{u}{v} = cost.$, si tratta cioè del fascio di rette passanti per l'origine O del piano (u, v) privato delle due bisettrici $\frac{v}{u} = \pm 1$, come abbiamo già stabilito. Il valore $t = +\infty$ corrisponde alla retta $v = u$, mentre $t = -\infty$ corrisponde alla retta $v = -u$.
Il valore $t = 0$, invece, corrisponde alla retta $v = 0$ (asse u delle ascisse) per $r > R_S$, mentre per $r < R_S$ corrisponde alla retta $u = 0$ (asse v delle ordinate), come si vede dalle (12.153), (12.154) rispettivamente.
Nelle coordinate di Kruskal la geometria dello spaziotempo è rappresentata dalla metrica (12.120) con f^2 dato dalla (12.151). La metrica di Kruskal è dunque:

$$\mathbf{ds}^2 = \frac{4R_S^3}{r} e^{-r/R_S}(dv^2 - du^2) - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{12.159}$$

essa è soluzione delle equazioni gravitazionali di Einstein ed è **non singolare quasi ovunque**. *Soltanto lungo l'iperbole $u^2 - v^2 = -1$, corrispondente ad $r = 0$ si ha una* **singolarità non eliminabile** *sia nella metrica che nel tensore di curvatura di Riemann, come si può verificare.* Inoltre, come avevamo richiesto come primo requisito, i raggi di luce viaggiano sempre lungo linee rette, con la stessa velocità $\left(\frac{du}{dv}\right)^2 = 1$, come nella Relatività Ristretta. Perciò u funge da **coordinata radiale globale**, e v da **coordinata temporale globale.** Esse tuttavia non corrispondono asintoticamente all'infinito a coordinate sferiche per uno spazio piatto, come invece succede per le coordinate di Schwarzschild.
Osservazione 1:
Abbiamo già osservato che è necessario richiedere che il 2° membro della (12.155), cioè

$$\left(\frac{r}{R_S} - 1\right) e^{r/R_S},$$

definisca univocamente r come funzione implicita delle due variabili u e v. Ciò si verifica se tale espressione è una funzione monotòna di r. Poiché la sua derivata è $\frac{r}{R_S^2} e^{r/R_S}$, essa è certamente monotòna per $r > 0$. Per $r = 0$ la derivata si annulla e la conclusione precedente non è più valida. Poiché questo corrisponde a $u^2 - v^2 = -1$, vediamo che le due regioni ombreggiate con linee orizzontali della Fig.12.7 devono essere escluse dalle regioni ammissibili del diagramma di Kruskal.
Osservazione 2:
La metrica di Schwarzschild è una soluzione non-singolare delle equazioni di Einstein nelle

regioni I, II, III e IV ma ha una singolarità sul contorno corrispondente ad $r = R_S$, cioè sulla retta $u = v$. La metrica di Kruskal è anch'essa una soluzione non-singolare delle equazioni di Einstein in queste regioni ed è equivalente alla soluzione di Schwarzschild, ma non ha singolarità sul contorno $u = v$. Questo risultato è analogo alla *continuazione analitica* nella teoria delle funzioni complesse ed è appropriatamente riferito come una **estensione analitica** della varietà spaziotempo.
Usiamo ora le coordinate di Kruskal per studiare i raggi di luce che viaggiano radialmente verso l'interno della sfera di raggio $r = R_S$ (che nel piano (u, v) è rappresentata dalla retta $u = v$ ($t = +\infty$) (**orizzonte degli eventi**), problema che abbiamo già trattato usando le coordinate di Schwarzschild. Questi raggi si muovono lungo linee per le quali è $d\theta = d\varphi = 0$ (linee radiali) e per le quali $\mathbf{ds^2} = 0$ (linee nulle). Dall'espressione (12.159) della metrica di Kruskal queste linee sono definite da:

$$v = \pm u + cost. \tag{12.160}$$

I raggi di luce radiali si muovono dunque lungo rette inclinate di 45° sugli assi del diagramma di Kruskal, ed i coni di luce hanno perciò un angolo di apertura di 45°. Uno di questi raggi è rappresentato dalla linea L della Fig.12.8, esso ha pendenza -1, cioè è inclinato di $-45°$ sull'asse u.

Figura 12.8: Linea di universo di un raggio di luce radiale che si muove verso la singolarità.

Le linee di universo di particelle materiali essendo del genere tempo devono giacere all'interno del cono di luce relativo a ciascun punto in cui passano, perciò, in ogni loro punto devono avere una pendenza maggiore di 1. Infatti per tali particelle che si muovano radialmente o meno, si può dimostrare che $|dv/du| > 1$.
Infine consideriamo due osservatori uno dei quali si trova in quiete ($r = cost.$, $r \gg R_S$) molto lontano dall'orizzonte degli eventi, l'altro osservatore è in caduta libera verso la singolarità $r = 0$. La Fig.12.9 mostra le linee di universo di questi due osservatori: la linea di universo del primo è un ramo di iperbole che giace nella regione $r > R_S$ (cioè nella regione I) di equazione $r = cost.$ ($u^2 - v^2 = cost. > 0$). La linea di universo del secondo osservatore che cade radialmente verso la singolarità è una linea retta che va ad intersecare l'iperbole $r = 0$. Nel cadere questo osservatore emette dei lampi di luce (disegnati nella Fig.12.9 come segmenti tratteggiati) ad intervalli regolari verso l'osservatore lontano il quale li riceve ad intervalli, misurati secondo il suo tempo proprio, che vanno man mano crescendo col progredire della caduta del primo osservatore. L'ultimo raggio di luce che riceve l'osservatore lontano è inclinato di 45° sugli assi del diagramma di Kruskal appena sotto all'orizzonte degli eventi che è la retta $v = u$. Dopodiché, l'osservatore in caduta libera, attraversa l'orizzonte cadendo nel buco nero, e raggiunta la

Figura 12.9: Linee di universo di due osservatori: uno in quiete all'infinito e l'altro in caduta libera radiale verso il buco nero.

singolarità viene distrutto!

Conclusione[25]

Le coordinate originali di Schwarzschild (t, r), si sono dimostrate buone per $r > R_S$, cioè nella regione I, che è solo una parte della varietà rappresentata nel diagramma di Kruskal.
Partendo dalla regione I e seguendo raggi nulli diretti verso il futuro si raggiunge la regione II, e seguendo raggi nulli diretti verso il passato è possibile raggiungere la regione IV.
Se avessimo considerato geodetiche del genere spazio, saremmo stati condotti dalla regione I alla regione III.
La regione II è quella che noi associamo al **buco nero**. Una volta che un qualunque oggetto viaggia dalla regione I alla regione II, esso non può più tornare indietro. Infatti ogni percorso diretto verso il futuro nella regione II termina colpendo la singolarità in $r = 0$; *una volta superato l'orizzonte degli eventi esso è completamente destinato al non ritorno.* Questo fatto va fortemente rimarcato: non solo non torna indietro nella regione I, ma non smetterà mai di muoversi nella direzione delle r decrescenti, poiché questa è semplicemente la direzione del genere tempo. Questo potrebbe essere visto anche nel nostro originario sistema di coordinate: per $r < R_S$ sappiamo che *t diventa del genere spazio ed r del genere tempo.* Perciò il movimento verso la singolarità non si può interrompere.
Le regioni III e IV potrebbero sembrare in qualche modo inattese ma, va rimarcato, che anch'esse rappresentano soluzioni delle equazioni di Einstein. La regione IV è semplicemente la regione con il tempo invertito della regione II, una parte dello spaziotempo dalla quale gli oggetti possono fuggire verso di noi, mentre nessun oggetto che parta dalla regione I può arrivare nella regione IV. Questa regione, in contrapposizione con il nome dato alla regione II, si chiama **buco bianco**. C'è *"una singolarità nel passato"* (ramo inferiore dell'iperbole), che potrebbe essere interpretata come la singolarità da cui proviene l'universo. Il bordo sinistro della regione IV è **l'orizzonte degli eventi passato**, mentre il bordo destro della regione II è **l'orizzonte degli eventi futuro**. La regione III non può essere raggiunta dalla nostra regione I sia muovendosi in avanti che indietro nel tempo e niente e nessuno può raggiungerci provenendo da lì. È un'altra regione asintoticamente piatta dello spazio tempo, un'immagine speculare al nostro spaziotempo. La regione III può essere pensata come se fosse connessa alla regione I attraverso un **condotto spaziotemporale** detto **wormhole o ponte di Einstein-Rosen**, un tipo di configurazione a forma di collo che unisce due regioni distinte. Queste ultime questioni

[25]Per questa conclusione ci siamo valsi di Sean M. Carroll, *Spacetime and Geometry*, Addison Wesley, 2004, pagg. 226-227.

appena accennate sono di grande interesse speculativo, come lo sono una gran quantità di altri argomenti come: i buchi neri rotanti per lo studio dei quali è necessario utilizzare la cosiddetta metrica di Kerr, i buchi neri elettricamente carichi, i diagrammi di Penrose, ecc., ma tutto ciò esorbita dagli scopi del libro.

12.13 APPENDICE: Cenno sulle onde gravitazionali

12.13.1 Introduzione

In questa APPENDICE ci occuperemo, molto brevemente, di un argomento importantissimo che ha rivoluzionato l'Astronomia e l'Astrofisica moderne, per la trattazione adeguata del quale sarebbero necessari alcuni capitoli. Noi ci limiteremo ad alcune nozioni essenziali dalle quali non si può prescindere: una trattazione più approfondita uscirebbe fuori dagli scopi del libro.[26]
Molto brevemente possiamo dire che un'onda gravitazionale è una deformazione della curvatura dello spazio tempo che si propaga come un'onda in un liquido.
La loro esistenza fu prevista teoricamente nel 1916 da Einstein come conseguenza della sua teoria della Relatività Generale,[27] ed è stata confermata sperimentalmente soltanto 100 anni dopo la loro previsione teorica, nel 2015. La teoria delle onde gravitazionali è un campo di studio estremamente ricco che fa confluire insieme differenti domini come Relatività Generale, Teoria dei campi, Astrofisica e Cosmologia. D'altra parte l'apparato sperimentale ha acquisito tecniche straordinarie che permettono oggigiorno agli strumenti di avere delle sensibilità impensabili fino a pochi anni fa. Torneremo brevemente in seguito sulla scoperta sperimentale delle onde gravitazionali, che oggi è considerata un grande successo dell'Astronomia e del notevole sviluppo tecnologico dell'interferometria.

12.13.2 Alcune analogie fra campo elettromagnetico e campo gravitazionale

Il concetto di **radiazione o propagazione per onde** si è presentato, storicamente, nello stesso tempo in cui è sorto il concetto di **campo elettromagnetico** scaturito dallo studio quantitativo dei fenomeni elettromagnetici come son descritti dalla teoria di Maxwell della quale abbiamo discusso molto sinteticamente, dal punto di vista classico, nel paragrafo 1.4 del Cap.1 e diffusamente, dal punto di vista della Relatività Ristretta, nel Cap. 7.
Come abbiamo visto, le equazioni di Maxwell legano il campo elettromagnetico, dunque **E** ed **H**, con le sue sorgenti: le cariche elettriche ρ e le correnti elettriche **j**.
Il problema della risoluzione delle equazioni di Maxwell è stato ricondotto, sotto l'ipotesi della validità del cosiddetto *gauge di Lorentz*:

$$\eta^{ik}\partial_k A_i = 0, \quad \text{con} \quad \eta^{ik} \text{ tensore metrico dello spazio di Minkowski,} \tag{12.161}$$

[26]Per una trattazione approfondita e moderna delle onde gravitazionali il testo più importante di riferimento è l'opera in 2 volumi di Michele Maggiore, *Gravitational waves*, volume 1: Theory and Experiments, Oxford University Press, 2008 - 554 pages; *Gravitational waves*, volume 2: Astrophysics and Cosmology, Oxford University Press, 2018 - 848 pages.
Trattazioni più elementari e molto meno estese dei volumi di M. Maggiore sono: James B. Hartle, *Gravity: An introduction to Einstein's General Relativity*, Capp. 16, 21, 23, Pearson Educational Limited, 2014. Una classica trattazione, sempre valida, è quella di Steven Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, John Wiley and Sons, New York, 1972, Cap.10. U'altra trattazione, più elementare, si può trovare, nel più volte citato, Ray d'Inverno, *Introduzione alla Relatività di Einstein*, Capp. 20, 21, CLUEB, Bologna, 2004; infine molto buona, ma come al solito piuttosto stringata è la trattazione di L. D. Landau e E. M. Lifschitz, *Teoria dei Campi*, Editori Riuniti, Edizione 2010, Cap. XIII, pag.447.

[27]La previsione teorica dell'esistenza delle onde gravitazionali è stata conseguita da Einstein, utilizzando le equazioni di campo linearizzate, nell'importante lavoro: *Näerungsweise Integration der Feldgleichungen der Gravitation*, Sitzber. Preuss. Akad. Wiss. Berlin, 1916, pp.688-696; la questione venne ripresa dallo stesso Einstein, due anni dopo, nell'altro lavoro: *Über Gravitationswellen*, Sitzber. Preuss. Akad. Wiss. Berlin, 1918, pp. 154-167.

a quello della ricerca delle soluzioni dell'equazione ondulatoria per i potenziali A^i:

$$\Box A^i = \frac{4\pi}{c} J^i, \tag{12.162}$$

nella quale $\Box$ è il cosiddetto *operatore di d'Alembert o d'Alembertiano*

$$\Box A^i \stackrel{\text{def}}{=} \eta^{ik} \frac{\partial^2 A^i}{\partial x^i \partial x^k} \equiv \eta^{ik} \partial_{ik} A^i = \frac{1}{c^2} \frac{\partial^2 A^i}{\partial t^2} - \Delta_2 A^i. \tag{12.163}$$

In regioni prive di sorgenti essendo $J^i = 0$ la (12.162) diventa:

$$\Box A^i = 0, \tag{12.164}$$

da cui segue che il 4-potenziale elettromagnetico A_i e quindi anche il tensore del campo elettromagnetico $F_{ik} = \partial_i A_k - \partial_k A_i$, soddisfano entrambi l'equazione d'onda.
Perciò la caratteristica matematica fondamentale della teoria elettromagnetica consiste nel fatto che le (12.162) *costituiscono un sistema di 4 equazioni differenziali (disaccoppiate) alle derivate parziali del 2° ordine di tipo iperbolico, le cui funzioni incognite sono i* **potenziali elettromagnetici**: *il potenziale scalare* $A^0 = \varphi$ e il potenziale vettore $\mathbf{A} = (A^\alpha)$ $(\alpha = 1, 2, 3)$ compendiati nell'unico 4-potenziale elettromagnetico $A^i = (A^0, A^1, A^2 A^3)$.
Ciò ha come conseguenza che variazioni delle forze elettromagnetiche causate da perturbazioni nella distribuzione delle sorgenti elettromagnetiche, come succede, per esempio, nel moto accelerato di corpi carichi, si propagano nello spazio con una *velocità finita* uguale alla *velocità c della luce nel vuoto*. Conseguentemente l'azione di un corpo carico su un altro non è simultanea ma *ritardata*. Veramente le equazioni di campo iperboliche non ammettono soltanto soluzioni ritardate ma anche soluzioni *anticipate*. La richiesta della validità del *Principio di Causalità* ha dimostrato che possiamo usare soltanto le soluzioni ritardate. Nel caso delle equazioni di Maxwell, si hanno formule esplicite che determinano le soluzioni ritardate, per assegnate distribuzioni delle sorgenti. Utilizzando queste formule si trova che sorgenti in moto accelerato creano un campo contenente una distribuzione di energia e quantità di moto elettromagnetiche che si irradia dalle sorgenti. Questa è la **radiazione elettromagnetica** o, se si vuole, si tratta di **onde elettromagnetiche** emesse da tali sorgenti.
In maniera analoga, le equazioni di Einstein del campo gravitazionale nella prima forma:

$$R_{ik} - \frac{1}{2} g_{ik} R = \kappa T_{ik}, \qquad \left(\kappa = \frac{8\pi G}{c^4}\right) \quad (i, k = 0, 1, 2, 3). \tag{12.165}$$

legano la curvatura dello spaziotempo, compendiata nel primo membro delle (12.165) nel tensore R_{ik} di Ricci e nello scalare di curvatura R, alla sua sorgente: la massa-energia della materia, la cui distribuzione è compendiata nel tensore energia-impulso T_{ik}. Tali equazioni, come sappiamo, possono essere espresse anche in una forma alternativa che abbiamo chiamato seconda forma:

$$R_{ik} = \frac{8\pi G}{c^4} \left(T_{ik} - \frac{1}{2} g_{ik} T\right) \qquad (i, k = 0, 1, 2, 3), \tag{12.166}$$

dove $T = T^i_i$ è il cosiddetto *scalare materiale o scalare di Laue.*
Sia le (12.165) che le (12.166) costituiscono un sistema di 10 equazioni differenziali alle derivate parziali del 2° ordine di tipo iperbolico le cui funzioni incognite sono i coefficienti $g_{ik}(x)$ della metrica dello spaziotempo, cioè i cosiddetti **potenziali gravitazionali**. Un'importante differenza fra il sistema differenziale delle equazioni di Maxwell del campo elettromagnetico e le equazioni di Einstein del campo gravitazionale è che *le prime sono* **lineari** *mentre le seconde sono* **non lineari**. Perciò, nella Relatività Generale, abbiamo una situazione simile a quella della teoria di Maxwell ma matematicamente più complicata. In particolare, è ragionevole

supporre che, come le equazioni di Maxwell hanno come conseguenza necessaria l'esistenza delle **onde elettromagnetiche**, così le equazioni di Einstein della Relatività Generale implicano l'esistenza delle **onde gravitazionali** come effettivamente fu previsto da Einstein nel 1916. A causa della non linearità delle equazioni gravitazionali di Einstein lo studio dettagliato delle onde gravitazionali è un problema estremamente difficile, molto più difficile dello studio del campo elettromagnetico in cui si ha a che fare con equazioni lineari.
In questa APPENDICE discuteremo alcuni aspetti della teoria einsteiniana della gravitazione linearizzata. Abbiamo visto, all'inizio del capitolo, che la soluzione più semplice delle equazioni gravitazionali (12.165), cioè la soluzione esatta di Schwarzschild, si ottiene considerando tali equazioni in una regione dello spazio vuota di materia, all'esterno di una massa a simmetria sferica. In tal caso essendo $T_{ik} = 0$, l'equazione (12.165) si riduce alla forma più semplice:

$$R_{ik} = 0 \qquad (i, k = 0, 1, 2, 3), \tag{12.167}$$

nella quale il tensore di Ricci R_{ik} è espresso dalla:

$$R_{ik} = g^{ml} R_{milk} = R^l_{ilk} = \partial_l \Gamma^l_{ik} - \partial_k \Gamma^l_{il} + \Gamma^l_{ml} \Gamma^m_{ik} - \Gamma^l_{mk} \Gamma^m_{il} = 0. \tag{12.168}$$

Anche in questa forma notevolmente semplificata, la (12.167) è ancora costituita da un sistema di 10 equazioni alle derivate parziali non lineari nelle 10 funzioni incognite $g_{ik}(x)$. La non linearità comporta, com'è noto, che non vale *il principio di sovrapposizione degli effetti*, cioè, se g_{ik} e g^*_{ik} sono due soluzioni distinte dell'equazione di campo (12.167), una loro combinazione lineare non è necessariamente una loro soluzione. La ragione fisica di ciò è facilmente comprensibile: il campo gravitazionale di un corpo deve contenere energia, e poiché possiede energia deve necessariamente possedere una massa e perciò creare anch'esso un ulteriore campo gravitazionale che va ad aggiungersi al campo creato dalle masse sorgenti. Cioè, il campo stesso contribuisce al campo creato dalla propria sorgente. A causa di questo feedback il campo gravitazionale di due corpi non è semplicemente uguale alla somma dei due campi considerati separatamente, ma coinvolge la struttura dettagliata dei campi interagenti.
Nel paragrafo 12.13.4 otterremo un insieme di equazioni gravitazionali lineari, approssimate, ignorando perciò l'effetto di feedback appena discusso. In tale teoria linearizzata della gravitazione gli effetti saranno considerati semplicemente additivi, come succede nella teoria classica newtoniana della gravitazione.

12.13.3 Condizioni sulle coordinate: invarianza di gauge delle equazioni di Einstein

A questo punto - per una chiara trattazione teorica delle onde gravitazionali - torna utile una discussione estremamente importante, che finora non abbiamo mai fatto, sul significato delle soluzioni delle equazioni di campo della Relatività Generale. Utilizzando il tensore di Einstein $G_{ik} \equiv R_{ik} - \frac{1}{2} g_{ik} R$, tali equazioni possono essere scritte sinteticamente così:

$$G_{ik} = \kappa T_{ik}, \qquad \left(\kappa = \frac{8\pi G}{c^4}\right) \quad (i, k = 0, 1, 2, 3). \tag{12.169}$$

Il tensore simmetrico G_{ik} ha 10 componenti indipendenti, perciò le equazioni di campo (12.169) costituiscono un sistema di 10 equazioni alle derivate parziali algebricamente indipendenti. L'incognito tensore metrico g_{ik} ha anch'esso 10 componenti algebricamente indipendenti, e perciò, a prima vista, si potrebbe pensare che le equazioni di Einstein (con appropriate condizioni al contorno) siano sufficienti a determinare le g_{ik} univocamente. Tuttavia, non è così. Benché algebricamente indipendenti le 10 G_{ik}, sono legate dalle 4 identità differenziali conseguenza delle identità di Bianchi:

$$\nabla_k G^{ik} = 0. \tag{12.170}$$

Dunque non abbiamo 10 equazioni funzionalmente indipendenti ma soltanto $10 - 4 = 6$, che ci lasciano 4 gradi di libertà per le 10 incognite g_{ik}. Perché abbiamo questi 4 gradi di libertà? Supponiamo che g_{ik} sia una soluzione delle equazioni di Einstein, se noi operiamo una arbitraria trasformazione di coordinate $x^i \to x^{i'} = x^{i'}(x)$, il tensore trasformato $g'_{ik} = g_{i'k'}$ è ancora una soluzione, come stabilisce il Principio di covarianza generale. Ciò significa anche che g_{ik} e g'_{ik} rappresentano la stessa soluzione fisica (la stessa geometria) vista da differenti sistemi di riferimento.
La trasformazione di coordinate coinvolge 4 funzioni arbitrarie $x^{i'}$, perciò i 4 gradi di libertà derivano dalla libertà di scelta del sistema di coordinate e spariscono allorquando si fa una specifica scelta del sistema di coordinate. Per esempio, possiamo scegliere un sistema di di coordinate nel quale 4 dei 10 potenziali g_{ik} sono nulli. *Le equazioni di Einstein non determinano le soluzioni g_{ik} in modo unico ma a meno di un'arbitraria trasformazione di coordinate.*
L'insuccesso delle equazioni di Einstein nel determinare le g_{ik} univocamente è l'analogo dell'incapacità delle equazioni di Maxwell, in RR, nel determinare in maniera univoca i 4- potenziali A_i.
Nello stesso modo in cui si procede nella teoria di Maxwell, nella teoria einsteiniana della gravitazione, possiamo eliminare l'ambiguità nel tensore metrico adottando qualche particolare sistema di coordinate. La scelta di un sistema di coordinate può essere espressa da *quattro condizioni sulle coordinate*, che aggiunte al sistema delle sei equazioni indipendenti di Einstein, determina la soluzione senza ambiguità.
Un esempio notevole di un particolare e conveniente sistema di coordinate è rappresentata dal cosiddetto *sistema di coordinate armoniche: il cosiddetto gauge armonico.*
Coordinate armoniche: il gauge armonico.
Le **coordinate armoniche** (o **gauge armonico**, detto anche gauge di Hilbert, o gauge di De Donder) sono definite dalla condizione di dar luogo a potenziali gravitazionali g^{ik} che siano soluzioni delle 4 equazioni:

$$\Gamma^i \equiv g^{mn}\Gamma^i_{mn} = 0, \qquad (i, m, n = 0, 1, 2, 3) \tag{12.171}$$

Secondo il fisico russo Vladimir Fock, esse hanno addirittura, per i sistemi isolati, lo stesso ruolo delle coordinate galileiane in fisica classica.
Alcuni autori utilizzano il gauge armonico nello studio delle onde gravitazionali per semplificare le equazioni linearizzate di Einstein in maniera analoga all'uso del gauge di Lorentz per semplificare le equazioni di Maxwell.[28]
Per vedere com'è sempre possibile scegliere un sistema di coordinate per il quale vale la (12.171) cioè un sistema di coordinate armoniche (gauge armonico), basta ricordarsi come si trasformano, in un cambiamento arbitrario di coordinate, i simboli di connessione Γ^i_{mn} (si vedano le (10.213) e (10.214) del Cap.10):

$$\Gamma^{i'}_{m'n'} = \frac{\partial x^{i'}}{\partial x^r}\frac{\partial x^t}{\partial x^{m'}}\frac{\partial x^s}{\partial x^{n'}}\Gamma^r_{ts} - \frac{\partial x^r}{\partial x^{n'}}\frac{\partial x^s}{\partial x^{m'}}\frac{\partial^2 x^{i'}}{\partial x^r \partial x^s} \tag{12.172}$$

Se operiamo una moltiplicazione contratta con $g^{m'n'}$, otteniamo:

$$\Gamma^{i'} = \frac{\partial x^{i'}}{\partial x^r}\Gamma^r - g^{rs}\frac{\partial^2 x^{i'}}{\partial x^r \partial x^s}, \tag{12.173}$$

[28]La denominazione di *"gauge di Lorentz"* deriva il suo nome dall'analogia con l'omonimo gauge dell'elettromagnetismo. È curioso osservare che si tratta di fatto di un errore di nome. In elettromagnetismo, questo gauge è stato usato per la prima volta dal fisico danese Ludvig Valentin Lorenz (senza la "t") (1829-1891), lo scienziato che è stato anche l'inventore, nel 1867, dei potenziali ritardati, quando il più noto fisico olandese Hendrik Antoon Lorentz (1853-1928) aveva appena 14 anni! Tuttavia, questo errore è ora universalmente entrato nell'uso e perciò non ci si deve preoccupare eccessivamente.

dove si è fatto uso della relazione:

$$g^{rs} = g^{m'n'} \frac{\partial x^r}{\partial x^{m'}} \frac{\partial x^s}{\partial x^{n'}}.$$

Perciò se nella (12.173) è $\Gamma^r \neq 0$, possiamo sempre trovare un nuovo sistema di coordinate $x^{i'}$, in cui sia $\Gamma^{i'} = 0$ e ridursi quindi al gauge armonico, basta risolvere le equazioni differenziali alle derivate parziali del secondo ordine:

$$g^{rs} \frac{\partial^2 x^{i'}}{\partial x^r \partial x^s} = \frac{\partial x^{i'}}{\partial x^r} \Gamma^r. \tag{12.174}$$

Siamo ora in grado di spiegare il perché della denominazione del termine *"coordinate armoniche"*. Una funzione Φ si dice armonica se soddisfa l'equazione di Laplace quadridimensionale:

$$\Box \Phi = 0, \tag{12.175}$$

dove $\Box$ è l'operatore d'Alembertiano covariante definito dalla:

$$\Box \Phi \equiv g^{ik} \nabla_i (\nabla_k \Phi), \tag{12.176}$$

essendo ∇_i la derivata covariante. Poiché

$$g^{ik} \nabla_i (\nabla_k \Phi) = g^{ik} \left(\frac{\partial}{\partial x^i} (\nabla_k \Phi) - \Gamma^j_{ik} (\nabla_j \Phi) \right) = \tag{12.177}$$

$$= g^{ik} \left[\frac{\partial^2 \Phi}{\partial x^i \partial x^k} - \Gamma^j_{ik} \frac{\partial \Phi}{\partial x^j} \right] = g^{ik} \frac{\partial^2 \Phi}{\partial x^k \partial x^i} - \Gamma^j \frac{\partial \Phi}{\partial x^j}. \tag{12.178}$$

Se $\Gamma^j = 0$ la condizione di gauge armonico diventa

$$\Box \Phi = g^{ik} \frac{\partial^2 \Phi}{\partial x^i \partial x^k} = 0, \tag{12.179}$$

in tal caso **le coordinate stesse sono funzioni armoniche**, infatti ponendo $\Phi = x^m$ nella (12.179) si trova:

$$\Box x^m = g^{ik} \frac{\partial^2 x^m}{\partial x^i \partial x^k} = g^{ik} \frac{\partial}{\partial x^i} \delta^m_k = 0, \tag{12.180}$$

ciò giustifica l'appellativo di "coordinate armoniche" dato a questo sistema di coordinate.
In assenza di campi gravitazionali, l'ovvio sistema di coordinate armoniche è quello di Minkowski, nel quale $g^{ik} = \eta^{ik}$ sicché, si può verificare che la condizione di armonicità (12.171) è banalmente soddisfatta. In presenza di campi gravitazionali deboli il sistema delle coordinate armoniche può essere considerato come quasi minkowskiano.

12.13.4 Soluzioni di campo debole

Affinché si possa ignorare l'effetto di feedback di un campo gravitazionale come parte della propria sorgente, è chiaro che dobbiamo assumere che si tratti di un **campo debole**. Questa ipotesi è analoga a quella già fatta nel paragrafo 11.18 del Cap.11 quando abbiamo studiato il limite newtoniano delle equazioni gravitazionali, l'unica differenza essenziale è che ora il campo gravitazionale, oltre che essere debole, non è più stazionario.
D'altra parte il campo gravitazionale all'esterno del Sole è descritto da una metrica che differisce molto poco dalla metrica di Minkowski: la differenza, come abbiamo visto, è espressa dal

termine $\frac{2m}{r} = \frac{R_S}{r}$ della metrica di Schwarzschild il cui massimo valore è dell'ordine di 10^{-5} alla superficie del Sole. All'interno del Sole questa differenza sarà più grande, e dipende dalla dettagliata distribuzione della massa del Sole, ma non ci si aspetta che ecceda l'ordine 10^{-4}.
Un campo gravitazionale descritto da una metrica che differisce molto poco dalla metrica di Minkowski è chiamato **campo gravitazionale debole**. Tenuto conto delle considerazioni precedenti il campo gravitazionale del Sole è debole ovunque all'esterno del Sole stesso. Veramente ciò si può affermare per lo spazio intorno ad una stella qualunque. Le *nane bianche* possono avere un raggio 100 volte minore del raggio del Sole; la differenza dalla metrica di Minkowski si stima essere dell'ordine di 10^{-2} al centro della stella. La sola eccezione, in cui il campo gravitazionale non può essere considerato debole, è il caso delle *stelle di neutroni (pulsar)*, il raggio delle quali è dell'ordine del raggio di Schwarzschild, cioè $r \approx R_S$, per le quali il rapporto $\frac{2m}{r} = \frac{R_S}{r}$ è piuttosto grande risultando $\frac{R_S}{r} \approx 1$.
Dalle precedenti considerazioni si deduce che i campi gravitazionali deboli meritano una speciale attenzione. In ciò che segue svilupperemo un metodo approssimato applicabile a tali campi. Il problema si semplifica se usiamo le coordinate galileiane della RR, $(x^0 = ct, x^1 = x, x^2 = y, x^3 = z)$, nelle quali la metrica di Minkowski è diagonale ed ha la semplice forma:

$$\eta_{ik} = \eta^{ik} = (1, -1. - 1, -1). \tag{12.181}$$

La definizione di **campo gravitazionale debole** è ora immediata: *il campo gravitazionale descritto dalla metrica g_{ik} è debole se la differenza $g_{ik} - \eta_{ik}$ è molto piccola rispetto all'unità*, cioè:

$$|g_{ik} - \eta_{ik}| \ll 1. \tag{12.182}$$

Il che significa che si considera una metrica che differisce leggermente dalla metrica piatta di Minkowski, sia perciò somma della metrica di Minkowski η_{ik} e di un piccolo termine perturbativo h_{ik}, sia cioè:

$$g_{ik}(x^j) = \eta_{ik} + h_{ik}(x^j) \quad \text{con} \quad |h_{ik}| \ll 1, \quad (i, j, k = 0, 1, 2, 3), \tag{12.183}$$

dove, è bene sottolinearlo, il termine perturbativo $h_{ik}(x^j) = h_{ki}(x^j)$, simmetrico rispetto ai suoi indici, è funzione di tutte le coordinate $(x^j) = (x^0, x^1, x^2, x^3)$ compreso il tempo.
Assumeremo che la condizione (12.183) sia soddisfatta soltanto per una certa classe di coordinate che corrispondono approssimativamente ai sistemi inerziali dello spazio di Minkowski. Inoltre supporremo che lo spaziotempo sia asintoticamente piatto, cioè, se r è un parametro radiale, allora:

$$\lim_{r \to \infty} h_{ik}(x) = 0. \tag{12.184}$$

Il metodo generale di approssimazione per campi deboli è basato sull'uso appropriato di questo piccolo parametro $h_{ik}(x)$, generalmente dipendente da tutte le coordinate, e che in seguito specificheremo meglio. Inoltre, il metodo si fonda essenzialmente sul fatto che trascureremo sistematicamente termini dal 2° ordine in poi nelle h_{ik}, quindi, nelle equazioni di campo, ci limiteremo a considerare soltanto termini lineari in h_{ik} e loro derivate. È per questo motivo che il procedimento che seguiremo è chiamato *approssimazione lineare* e le corrispondenti equazioni di campo, che ora ricaveremo, **equazioni di campo linearizzate**.

Il cuore del problema consiste nel fatto che *la Relatività Generale è invariante rispetto ad un enorme gruppo di simmetria, il gruppo di ogni possibile trasformazione di coordinate*,

$$x^i \to x'^{\,i}(x), \tag{12.185}$$

dove $x'^{\,i}(x)$ è un'arbitraria funzione regolare di x^i. Più precisamente, richiediamo che le $x'^{\,i}(x)$ siano invertibili, differenziabili, con inverse differenziabili, supponiamo cioè che le $x'^{\,i}(x)$ definiscano un arbitrario diffeomorfismo. Applicando la trasformazione (12.185) la metrica dello spaziotempo si trasforma così:

$$g_{ik}(x) \to g'_{ik}(x') = \frac{\partial x^r}{\partial x'^{\,i}} \frac{\partial x^s}{\partial x'^{\,k}} g_{rs}(x). \tag{12.186}$$

Ci riferiremo a questa simmetria come alla **simmetria di gauge della RG**.
Poiché i valori numerici delle componenti di un tensore dipendono dal sistema di coordinate, quello che veramente vogliamo intendere con la (12.183), nelle situazioni fisiche alle quali siamo interessati, è che esista un sistema di coordinate dove essa sia valida in una regione sufficientemente vasta dello spazio.
Operando sul tensore h_{ik} conveniamo che l'innalzamento e l'abbassamento degli indici verrà effettuato con la metrica di Minkowski $\eta^{ik} = \eta_{ik}$ "imperturbata": $h_i^k = \eta^{kl} h_{il}$, $h^{ik} = \eta^{il}\eta^{km} h_{lm}$, ecc. Con questa convenzione, è subito visto che il tensore metrico in forma controvariante, a meno di termini di ordine superiore a 2 in h, è dato da:[29]

$$g^{ik} = \eta^{ik} - h^{ik}, \qquad (g^{ik} = \eta^{ik} - h^{ik} + O(h^2)). \tag{12.187}$$

Sottolineiamo subito che la condizione per cui le h_{ik} debbono essere grandezze infinitesime non determina una scelta univoca del sistema di coordinate. Se questa condizione è soddisfatta in qualche sistema di coordinate, essa sarà soddisfatta anche per qualsiasi trasformazione di coordinate del tipo

$$x^m \to x'^{\,m} = x^m + \xi^m(x) \tag{12.188}$$

(**trasformazione di "gauge"**), dove le ξ^m sono quattro grandezze arbitrarie limitate, dello stesso ordine delle h_{ik}, cioè tali che $|\xi^m| \ll 1$ e dove anche le derivate $|\partial_i x^m|$ sono tutt'al più dell'ordine di grandezza delle $|h_{ik}|$, sicché, nell'approssimazione di campo debole, si dovranno trascurare non solo i termini dal 2° ordine in poi in h_{ik} ma anche i loro prodotti ed i prodotti con le loro derivate. Per una tale trasformazione, il tensore metrico si trasforma nel seguente modo:

$$g'_{mn} = \frac{\partial x^a}{\partial x'^{\,m}} \frac{\partial x^b}{\partial x'^{\,n}} g_{ab} \approx (\delta^a_m - \partial_m \xi^a)(\delta^b_n - \partial_n \xi^b)(\eta_{ab} + h_{ab}) \approx \eta_{mn} + (h_{mn} - \partial_n \xi_m - \partial_m \xi_n), \tag{12.189}$$

dove si è posto $\xi_m = \eta_{mn}\xi^n$. Di conseguenza,

$$g'_{mn} \approx \eta_{mn} + h'_{mn} \tag{12.190}$$

è della stessa forma della (12.183) se si fa l'identificazione

$$h'_{mn} = h_{mn} - \partial_n \xi_m - \partial_m \xi_n. \tag{12.191}$$

Questa libertà di "gauge" verrà usata per scrivere le equazioni di campo linearizzate.

12.13.5 Linearizzazione delle equazioni di Einstein. Onde gravitazionali

Il nostro scopo è dunque quello di linearizzare le equazioni di campo della RG, le (12.165), dobbiamo perciò calcolarci il tensore di Ricci R_{ik}, lo scalare di curvatura R ed il tensore di Einstein.
Cominciamo col calcolare i simboli di Christoffel di seconda specie:

$$\Gamma^m_{ab} = g^{ml}\Gamma_{l,ab} = \frac{1}{2} g^{ml}\Big(\partial_a g_{lb} + \partial_b g_{la} - \partial_l g_{ab}\Big), \tag{12.192}$$

[29] Infatti, abbiamo $g_{ik}g^{ij} = (\eta_{ik} + h_{ik})(\eta^{ij} - h^{ij}) = \eta_{ik}\eta^{ij} - \eta_{ik}h^{ij} + \eta^{ij}h_{ik} - h_{ik}h^{ij} = \delta^j_k - h^j_k + h^j_k - h_{ik}h^{ij} = \delta^j_k + O(h^2)$.

dove il tensore controvariante g^{ml} è espresso dalla (12.187).
Se ora introduciamo le (12.183) e (12.187) nella (12.192) troviamo i simboli di Christoffel nell'approsimazione lineare:

$$\Gamma^m_{ab} \approx \frac{1}{2}\eta^{ml}\Big(\partial_a h_{lb} + \partial_b h_{la} - \partial_l h_{ab}\Big). \tag{12.193}$$

Calcoliamo ora il tensore di Riemann nella forma completamente covariante (10.148), che qui riportiamo per comodità del lettore:

$$R_{imkn} = \frac{1}{2}\big(\partial_{in}g_{mk} + \partial_{mk}g_{in} - \partial_{ik}g_{mn} - \partial_{mn}g_{ik}\big) + g_{rs}\big(\Gamma^r_{in}\Gamma^s_{mk} - \Gamma^r_{ik}\Gamma^s_{mn}\big). \tag{12.194}$$

Essendo le h_{ik} molto piccole, i simboli di Christoffel di seconda specie Γ^i_{kl}, espressi in funzione delle derivate delle h_{ik}, sono anch'essi molto piccoli, trascurando le potenze superiori alla prima, nel tensore di Riemann possiamo lasciare soltanto i termini entro la prima parentesi ai quali dobbiamo apportare ancora l'approssimazione di campo debole, avremo dunque:

$$R^{(1)}_{imkn} \approx \frac{1}{2}\big(\partial_{in}h_{mk} + \partial_{mk}h_{in} - \partial_{ik}h_{mn} - \partial_{mn}h_{ik}\big) \tag{12.195}$$

Utilizziamo la (12.195) per calcolarci il tensore di Ricci $R_{mn} = g^{ik}R_{imkn}$ apportandovi l'approssimazione di campo debole, avremo:

$$R^{(1)}_{mn} \approx \eta^{ik}R^{(1)}_{imkn}, \tag{12.196}$$

cioè:

$$R^{(1)}_{mn} \approx \frac{1}{2}\eta^{ik}\big(\partial_{in}h_{mk} + \partial_{mk}h_{in} - \partial_{ik}h_{mn} - \partial_{mn}h_{ik}\big),$$

o meglio:

$$R^{(1)}_{mn} \approx \frac{1}{2}\Big(\eta^{il}\partial_{in}h_{ml} + \eta^{il}\partial_{mk}h_{in} - \Box h_{mn} - \partial_{mn}h\Big). \tag{12.197}$$

È possibile ottenere un'ulteriore semplificazione se si introducono le nuove variabili:

$$\overline{h}_{ml} \stackrel{\text{def}}{=} h_{ml} - \frac{1}{2}\eta_{ml}h, \qquad h = \eta^{rs}h_{rs} = h^r_r, \tag{12.198}$$

dalle quali discende che[30]

$$h_{ml} = \overline{h}_{ml} - \frac{1}{2}\eta_{ml}\overline{h}, \qquad \overline{h} = \eta^{rs}\overline{h}_{rs} = h - 2h = -h. \tag{12.199}$$

In queste nuove variabili la (12.197) si riscriverà così:

$$R^{(1)}_{mn} \approx \frac{1}{2}\Big[\eta^{ik}\partial_{in}\Big(\overline{h}_{mk} - \frac{1}{2}\eta_{mk}\overline{h}\Big) + \eta^{ik}\partial_{mk}\Big(\overline{h}_{in} - \frac{1}{2}\eta_{in}\overline{h}\Big) - \Box\Big(\overline{h}_{mn} - \frac{1}{2}\eta_{mn}\overline{h}\Big) + \partial_{mn}\overline{h}\Big]. \tag{12.200}$$

Svolgendo i calcoli, la (12.200) ci fornisce:

$$R^{(1)}_{mn} \approx \frac{1}{2}\Big[\eta^{ik}\partial_{in}\overline{h}_{mk} - \frac{1}{2}\eta^{ik}\eta_{mk}\partial_{in}\overline{h} + \eta^{ik}\partial_{mk}\overline{h}_{in} - \frac{1}{2}\eta^{ik}\eta_{in}\partial_{mk}\overline{h} - \Box\overline{h}_{mn} + \frac{1}{2}\eta_{mn}\Box\overline{h} + \partial_{mn}\overline{h}\Big].$$

[30] Infatti dalla prima delle (12.198) otteniamo $h_{ml} = \overline{h}_{ml} + \frac{1}{2}\eta_{ml}h$ con $h = \eta^{rs}h_{rs}$, d'altra parte, in modo analogo, possiamo scrivere $\overline{h} = \eta^{rs}\overline{h}_{rs} = \eta^{rs}(h_{rs} - \frac{1}{2}\eta_{rs}h) = h - \frac{1}{2}\eta^{rs}\eta_{rs}h$ ma $\eta^{rs}\eta_{rs} = 4$, dunque $\overline{h} = h - 2h$, cioè $\overline{h} = -h$.

Si vede subito che entro la parentesi quadra della precedente espressione la somma algebrica del secondo, quarto ed ultimo termine è nulla perciò tale espressione si riduce alla:

$$R^{(1)}_{mn} \approx \frac{1}{2}\Big[\eta^{ik}\partial_{in}\overline{h}_{mk} + \eta^{ik}\partial_{mk}\overline{h}_{in} - \Box h_{mn}\Big]. \tag{12.201}$$

Ci rimane da calcolare lo scalare di curvatura linearizzato, cioè:

$$R^{(1)} \approx \eta^{rs}R^{(1)}_{rs} = \frac{1}{2}\eta^{rs}\Big[\eta^{ik}\partial_{is}\overline{h}_{rk} + \eta^{ik}\partial_{rk}\overline{h}_{is} - \Box h_{rs}\Big] =$$

$$= \frac{1}{2}\eta^{rs}\eta^{ik}\partial_{is}\overline{h}_{rk} + \frac{1}{2}\eta^{rs}\eta^{ik}\partial_{rk}\overline{h}_{is} - \frac{1}{2}\Box h.$$

Osserviamo che i primi due termini della precedente espressione sono perfettamente uguali, per accertarsene basta cambiare in maniera opportuna il nome degli indici muti nelle sommatorie che vi compaiono, sicché, in definitiva, per lo scalare di Ricci linearizzato avremo la seguente espressione:

$$R^{(1)} \approx= -\frac{1}{2}\Box h + \eta^{ik}\eta^{rs}\partial_{is}\overline{h}_{rl}. \tag{12.202}$$

Con le due formule (12.201) e (12.202) siamo infine in grado di scrivere il tensore di Einstein linearizzato $G^{(1)}_{mn} \equiv R^{(1)}_{mn} - \frac{1}{2}\eta_{mn}R^{(1)}$, cioè:

$$G^{(1)}_{mn} \equiv \frac{1}{2}\Big[\eta^{ik}\partial_{in}\overline{h}_{mk} + \eta^{ik}\partial_{mk}\overline{h}_{in} - \Box h_{mn} - \eta_{mn}\Big(-\frac{1}{2}\Box h + \eta^{ik}\eta^{rs}\partial_{is}\overline{h}_{rl}\Big)\Big]$$

ovvero:

$$G^{(1)}_{mn} \equiv \frac{1}{2}\Big[\eta^{ik}\partial_{in}\overline{h}_{mk} + \eta^{ik}\partial_{mk}\overline{h}_{in} - \Box\Big(h_{mn} - \frac{1}{2}\eta_{mn}h\Big) - \eta_{mn}\eta^{ik}\eta^{rs}\partial_{is}\overline{h}_{rl}\Big],$$

infine:

$$G^{(1)}_{mn} \equiv \frac{1}{2}\Big[\eta^{ik}\partial_{in}\overline{h}_{mk} + \eta^{ik}\partial_{mk}\overline{h}_{in} - \Box\overline{h}_{mn} - \eta_{mn}\eta^{ik}\eta^{rs}\partial_{is}\overline{h}_{rl}\Big]. \tag{12.203}$$

Ricordando che le equazioni gravitazionali di Einstein sono $G_{mn} = \kappa T_{mn}$, $\Big(\kappa = \frac{8\pi G}{c^4}\Big)$, previo cambio di alcuni indici muti, le cercate **equazioni linearizzate di Einstein del campo gravitazionale** sono perciò:

$$\Box\overline{h}_{mn} - \eta^{lr}\big(\partial_{nr}\overline{h}_{lm} + \partial_{mr}\overline{h}_{ln}\big) + \eta_{mn}\eta^{lr}\eta^{ks}\partial_{rs}\overline{h}_{lk} = -2\kappa T_{mn}. \tag{12.204}$$

Si può verificare direttamente che, se applichiamo al primo membro della (12.204) l'operatore $\eta^{ln}\frac{\partial}{\partial x^l}$, il risultato è identicamente nullo. Perciò il tensore energia-impulso T_{mn} deve soddisfare la condizione:

$$\eta^{ln}\partial_l T_{mn} = 0 \quad \Longleftrightarrow \quad \partial_l T^l_m = 0. \tag{12.205}$$

L'equazione (12.205) è manifestamente **l'equazione di conservazione dell'energia e quantità di moto** della RR. Questa osservazione chiarisce il significato fisico dell'equazione di campo linearizzata (12.204): noi partiamo col tensore energia-impulso T_{mn} che descrive il sistema fisico in considerazione nella Relatività Ristretta, cioè non considerando il campo gravitazionale. Questo tensore energia-impulso è allora la sorgente del campo gravitazionale nell'approssimazione lineare.
Consideriamo ancora la trasformazione di coordinate (12.188), cioè:

$$x'^{\,m} = x^m + \xi^m(x^n), \tag{12.206}$$

ξ^m essendo dell'ordine di h_{mn}. Avremo allora:

$$g_{mn} = \eta_{mn} + h_{mn} = \frac{\partial x'^{\,r}}{\partial x^m}\frac{\partial x'^{\,s}}{\partial x^n} g'_{rs} = (\delta^r_m + \partial_m \xi^r)(\delta^s_n + \partial_n \xi^s) g'_{rs}. \tag{12.207}$$

Ponendo

$$g'_{rs} = \eta_{rs} + h'_{rs}, \tag{12.208}$$

e ritenendo soltanto termini del primo ordine, otteniamo:

$$h_{mn} = h'_{mn} + (\eta_{mr}\partial_n \xi^r + \eta_{rn}\partial_m \xi^r) = h'_{mn} + (\partial_n \xi_m + \partial_m \xi_n). \tag{12.209}$$

Questa relazione mostra che h'_{mn} e h_{mn} sono dello stesso ordine di grandezza. Conseguentemente, possiamo usare le equazioni di campo linearizzate (12.204) tanto nel sistema di coordinate x^m quanto nell'altro $x'^{\,m}$.
Definiamo ora, nel nuovo sistema di coordinate $x'^{\,m}$, le quantità $\overline{h}'_{mn}$ in analogia con le (12.198), avremo:

$$\overline{h}'_{mn} = h'_{mn} - \frac{1}{2}\eta_{mn} h', \qquad h' = \eta^{rs} h'_{rs}. \tag{12.210}$$

Dalla (12.209) possiamo dedurre un'utile relazione fra le $\overline{h}_{mn}$ e le corrispondenti quantità $\overline{h}'_{mn}$. Questa relazione è:[31]

$$\overline{h}_{mn} = \overline{h}'_{mn} + \eta_{mr}\partial_n \xi^r + \eta_{nr}\partial_m \xi^r - \eta_{mn}\partial_r \xi^r. \tag{12.211}$$

Moltiplicando questa relazione per $\eta^{ln}\dfrac{\partial}{\partial x^l}$ troviamo:

$$\eta^{ln}\partial_l \overline{h}_{mn} = \eta^{ln}\partial_l \overline{h}'_{mn} + \eta_{mr}\Box \xi^r. \tag{12.212}$$

L'equazione (12.212) permette un'essenziale semplificazione nelle equazioni di campo (12.204). Infatti, supponendo assegnate le $\overline{h}_{mn}$, possiamo applicare una trasformazione del tipo (12.206) con le ξ^m soddisfacenti l'equazione:

$$\eta^{ln}\partial_l \overline{h}_{mn} = \eta_{mr}\Box \xi^r. \tag{12.213}$$

Poiché questa è **l'equazione delle onde** non omogenea nello spazio di Minkowski, possiamo sempre determinare soluzioni ξ^r per ogni assegnato $\overline{h}_{mn}$. Ma allora dalla (12.212) avremo l'importantissima condizione che (in analogia con il *gauge di Lorentz* dell'elettrodinamica) potremmo appropriatamente chiamare *gauge di Einstein*:[32]

$$\eta^{ln}\partial_l \overline{h}'_{mn} = 0 \quad \Longleftrightarrow \quad \partial_l \overline{h}'^{\,l}_{m} = 0. \tag{12.214}$$

Nel sistema di coordinate x' le equazioni di campo linearizzate sono le (12.204) con le $\overline{h}_{mn}$ sostituite dalle $\overline{h}'_{mn}$, cioè:

$$\Box \overline{h}'_{mn} - \eta^{lr}\partial_{nr}\overline{h}'_{lm} - \eta^{lr}\partial_{mr}\overline{h}'_{ln} + \eta_{mn}\eta^{lr}\eta^{ab}\partial_{rb}\overline{h}'_{la} = -2\kappa T_{mn}. \tag{12.215}$$

[31] A beneficio dello studente dimostriamo la (12.211) partendo dalla (12.209), cioè: $h_{mn} = h'_{mn} + \partial_n \xi_m + \partial_m \xi_n$. Utilizziamo ora sia le (12.198) sia le (12.210), avremo $\overline{h}_{mn} + \frac{1}{2}\eta_{mn} h = \overline{h}'_{mn} + \frac{1}{2}\eta_{mn} h' + \partial_n \xi_m + \partial_m \xi_n$ da cui la formula $\overline{h}_{mn} = \overline{h}'_{mn} + \frac{1}{2}\eta_{mn}(h' - h) + \partial_n \xi_m + \partial_m \xi_n$ che indichiamo con **(a)**. Moltiplichiamo primo e secondo membro della (12.209) per η^{mn}, otteniamo $h = h' + \partial_n \xi^n + \partial_m \xi^m$, ed essendo $\partial_n \xi^n = \partial_m \xi^m = \partial_r \xi^r$ avremo $h' - h = -2\partial_r \xi^r$ che sostituita nella precedente formula **(a)** ci dà infine $\overline{h}_{mn} = \overline{h}'_{mn} + \eta_{mr}\partial_n \xi^r + \eta_{nr}\partial_m \xi^r - \eta_{mn}\partial_r \xi^r$. c.d.d.

[32] È interessante osservare che si può ottenere un'analoga semplificazione delle equazioni linearizzate di Einstein anziché usare la (12.214) utilizzare la condizione di **gauge armonico**, cioè scegliere un sistema di coordinate definito dalla relazione $\Gamma^i \equiv g^{mn}\Gamma^i_{mn} = 0$.

Introducendo dunque la (12.214) nella (12.215) si annullano ciascuno degli ultimi tre termini del primo membro e perciò si ha una drastica semplificazione delle equazioni di Einstein linearizzate che non solo risultano molto più semplici:

$$\Box \bar{h}'_{mn} = -2\kappa T_{mn}, \tag{12.216}$$

ma indicano palesemente che **la perturbazione $\bar{h}'_{mn}$ si propaga come un'onda alla velocità della luce.**
Per maggiore chiarezza, e tenuto conto che $\kappa = \dfrac{8\pi G}{c^4}$, sintetizziamo il risultato precedente scrivendo **le equazioni di campo di Einstein linearizzate assieme alla condizione necessaria per la loro validità**:

$$\begin{cases} \Box \bar{h}'_{mn} = -\dfrac{16\pi G}{c^4} T_{mn}, & \text{(Equazioni di Einstein linearizzate)} \\ \\ \eta^{ln}\partial_l \bar{h}'_{mn} = 0, & \text{(Gauge di Einstein).} \end{cases} \tag{12.217}$$

Abbiamo dunque raggiunto l'importantissimo risultato:
In un campo gravitazionale debole causato da una piccola perturbazione della metrica piatta di Minkowski, le equazioni di campo linearizzate di Einstein, sotto opportune condizioni per le coordinate ("gauge di Einstein"), implicano l'esistenza di "onde gravitazionali" che si propagano nello spaziotempo alla velocità della luce.
Si noti l'analogia delle equazioni di campo $(12.217)_1$, valide sotto la *condizione del gauge di Einstein* $(12.217)_2$, con le equazioni di Maxwell alle quali si sia associata la condizione del *gauge di Lorenz*:

$$\begin{cases} \Box A'^{\,i} = -\dfrac{4\pi}{c} J^i, & \text{(Equazioni di Maxwell)} \\ \\ \partial_i A'^{\,i} = 0 & \text{(Gauge di Lorenz).} \end{cases} \tag{12.218}$$

Il risultato sintetizzato nelle (12.217) costituisce la previsione teorica fatta da Einstein, nel 1916, dell'esistenza delle onde gravitazionali (OG), la cui scoperta sperimentale, dopo anni ed anni di duro lavoro, è avvenuta soltanto 100 anni dopo! Della scoperta sperimentale delle OG parleremo nell'ultimo paragrafo, nei prossimi due ci limitiamo ad illustrare brevemente due soluzioni semplici: i potenziali ritardati e le onde gravitazionali piane.

12.13.6 Soluzione in termini di potenziali ritardati delle equazioni linearizzate di Einstein.

Come in elettrodinamica, la soluzione delle equazioni $(12.217)_1$ può essere scritta in termini di *potenziali ritardati* quando è nota la sorgente T_{mn}, avremo:

$$\bar{h}'_{mn}(t, x^\alpha) = -\frac{\kappa}{2\pi} \int \frac{1}{R} T_{mn}(x'^{\,\alpha}, t - R/c) d^3x', \quad (\alpha = 1, 2, 3) \tag{12.219}$$

dove

$$R^2 = \sum_{\alpha=1}^{3} (x^\alpha - x'^{\,\alpha})^2. \tag{12.220}$$

L'integrazione nella (12.219) è estesa alla falda inferiore (passato) del cono di luce minkowskiano di vertice il punto P le cui coordinate sono (t, x^α).
Per una distribuzione di materia in equilibrio (caso statico) il tensore energia impulso T_{mn} nella più bassa approssimazione è dato da:

$$T_{00} = c^2\rho, \qquad T_{mn} = 0, \quad (m, n \neq 0). \tag{12.221}$$

Per questo caso statico l'equazione $(12.217)_1$ si riduce all'equazione di Poisson:

$$\Delta_2 \overline{h}'_{mn} = \frac{16\pi G\rho}{c^2}, \tag{12.222}$$

per $m = n = 0$, e all'equazione di Laplace:

$$\Delta_2 \overline{h}'_{mn} = 0, \tag{12.223}$$

per $m, n \neq 0$. Le soluzioni delle (12.222) e (12.223) sono allora date rispettivamente da

$$\overline{h}'_{00} = -\frac{4G}{c^2} \int \frac{\rho}{R} d^3x' \tag{12.224}$$

$$\overline{h}'_{mn} = 0, \qquad (m, n \neq 0). \tag{12.225}$$

12.13.7 Cenno sulle onde gravitazionali piane

In questo paragrafo consideriamo l'equazione delle onde $(12.217)_1$ nel vuoto, dunque con $T_{mn} = 0$, che scriveremo, più semplicemente, sopprimendo l'apice su $\overline{h}'_{mn}$, così:

$$\Box \overline{h}_{mn} = 0. \tag{12.226}$$

La più semplice soluzione dell'equazione delle onde $(12.217)_1$ nel vuoto è un'onda piana monocromatica:

$$\overline{h}_{mn} = \Re\{A_{mn} e^{ik_r x^r}\}, \tag{12.227}$$

dove A_{mn} è *il tensore di polarizzazione* ovvero *il tensore ampiezza dell'onda* e k_r è *il quadrivettore d'onda*. Se sostituiamo la (12.227) nella (12.226), con un calcolo diretto, otteniamo:

$$\Box \overline{h}_{mn} = \eta^{rs} \frac{\partial}{\partial x^r} \frac{\partial}{\partial x^s} (e^{ik_l x^l}) = \eta^{rs} \frac{\partial}{\partial x^r} \left[ik_l \frac{\partial x^l}{\partial x^s} e^{ik_l x^l} \right] = \eta^{rs} \frac{\partial}{\partial x^r} \left[ik_l \delta^l_s e^{ik_l x^l} \right] =$$

$$= \eta^{rs} \frac{\partial}{\partial x^r} \left[ik_s e^{ik_l x^l} \right] = -\eta^{rs} k_r k_s e^{ik_l x^l} = 0, \tag{12.228}$$

se

$$\eta^{rs} k_r k_s = 0, \tag{12.229}$$

cioè la (12.227) è soluzione dell'equazione delle onde (12.226) se k_r è un quadrivettore nullo. Inoltre la condizione di *gauge armonico* richiede che

$$\frac{\partial}{\partial x^m} \overline{h}^m_n = 0, \tag{12.230}$$

che può essere riscritta come

$$\eta^{mr} \frac{\partial}{\partial x^m} \overline{h}_{rn} = 0. \tag{12.231}$$

Usando la (12.227), la precedente relazione ci dà:

$$\eta^{mr} \frac{\partial}{\partial x^m} A_{rn} e^{ik_l x^l} = 0 \quad \rightarrow \quad \eta^{mr} A_{rn} k_m = 0 \quad \rightarrow \quad k_m A^m_n = 0. \tag{12.232}$$

Quest'ulteriore condizione esprime l'ortogonalità fra il vettore d'onda ed il tensore di polarizzazione. Poiché $\overline{h}_{mn}$ è costante su quelle superficie d'onda sulle quali

$$k_r x^r = cost., \tag{12.233}$$

queste sono le equazioni dei **fronti d'onda**. È convenzionale riferirsi a k^0 come a $\frac{\omega}{c}$ dove ω è la *frequenza delle onde*. Conseguentemente

$$\overrightarrow{k} = (\frac{\omega}{c}, \mathbf{k}), \tag{12.234}$$

avendo indicato con $\overrightarrow{k}$ il 4-vettore d'onda e con $\mathbf{k}$ il corrispondente 3-vettore unitario. Poiché $\overrightarrow{k}$ è un vettore nullo, avremo

$$k_0^2 - k_1^2 - k_2^2 - k_3^2 = 0, \tag{12.235}$$

per cui sarà

$$\omega = ck_0 = \sqrt{k_1^2 + k_2^2 + k_3^2}, \tag{12.236}$$

dove (k_1, k_2, k_3) sono le componenti del 3-vettore unitario $\mathbf{k}$.

12.13.8 Scoperta sperimentale delle onde gravitazionali

Gli studi teorici sulle onde gravitazionali, a partire dal 1916 in cui Einstein ne predisse l'esistenza, hanno stimolato costantemente la ricerca della loro rivelazione sperimentale. Ma fu soltanto nel 1960 che il fisico americano Joseph Weber, dell'Università del Maryland, riuscì, per primo, a costruire un'apparecchiatura idonea - a suo dire - a rivelare le onde gravitazionali: le cosiddette "sbarre di Weber" consistenti di cilindri multipli di alluminio di 2 m di lunghezza ed 1 m di diametro quali antenne di rilevazione. Weber fu convinto, per parecchi anni, di aver rivelato le onde gravitazionali, ma i segnali ricevuti erano estremamente deboli e molto probabilmente si trattava di rumori spuri ed inoltre l'apparato non aveva la necessaria sensibilità. Inoltre le grosse difficoltà di costruire apparecchiature per esperimenti analoghi, e le controversie che ne nacquero raffreddarono gli entusiasmi e tutto venne poi abbandonato.
In quest'ultimo paragrafo diamo un breve resoconto dell'effettiva recente scoperta sperimentale delle onde gravitazionali.
Appena cinque anni fa, alle 11.51 (ora italiana) del 14 settembre 2015, venivano rilevate le onde gravitazionali prodotte dalla collisione di due buchi neri. Due eventi, questi, che fino ad allora non erano mai stati "osservati" o meglio "rivelati". I primi italiani ad accorgersi di ciò che stava accadendo furono Gabriele Vedovato, fisico dell'INFN di Padova, e Marco Drago, addottorato nell'ateneo padovano, in quel momento all'Albert Einstein Institute di Hannover, che immediatamente dettero la notizia a centinaia di ricercatori in tutto il mondo.
Con una massa rispettivamente di 29 e 36 volte quella del Sole, i due buchi neri hanno spiraleggiato, scontrandosi a una velocità di circa 150.000 chilometri al secondo e culminando nella formazione di un singolo buco nero rotante di massa pari a 62 volte quella del Sole, rilasciando quindi una energia pari a 3 masse solari. L'evento avvenne alla distanza di 1,3 miliardi di anni luce ed ebbe una durata di soli 0,2 secondi.
A permettere la scoperta furono i due *interferometri* (con bracci di circa 4 km di lunghezza) del progetto LIGO (**Laser Interferometer Gravitational-wave Observatory**), situati rispettivamente a Livingstone in Louisiana e a Hanford, Washington. Congiuntamente con il gruppo VIRGO (che comprendeva fisici e ingegneri di cinque Paesi europei tra cui l'Italia), operante a Cascina vicino a Pisa, furono poi studiati i dati prodotti dai due rivelatori.[33]
Quella delle onde gravitazionali è stata definita la scoperta del secolo. Nel 2017, proprio per aver ideato e promosso la costruzione degli interferometri LIGO, il premio Nobel per la fisica

[33] È doveroso ricordare che l'esperimento VIRGO di Cascina (Pisa) fu promosso, fin dal 1970, su iniziativa del fisico italiano Prof. Guido Pizzella dell'Università d Roma "La Sapienza" stimolato e incoraggiato dal suo maestro, il grande allievo, della scuola di Enrico Fermi, il Prof. Edoardo Amaldi.
Un'ottima descrizione di questi esperimenti e della loro storia si trova su Wikipedia alle voci LIGO e VIRGO.

venne assegnato a Rainer Weiss, professore emerito di fisica al Massachussets Institute of Technology, a Barry Clark Barish, professore emerito di fisica al California Institute of Technology (Caltech) ed a Kip Stephen Thorne, professore di fisica teorica al Caltech fino al 2009.
Un'altra tappa importante è stata raggiunta il 14 agosto 2017, infatti, i due rivelatori LIGO e VIRGO *hanno captato un nuovo segnale di onde gravitazionali che non proveniva dalla collisione di due buchi neri, ma dalla fusione di due stelle di neutroni.* Per la prima volta, oltre alle onde gravitazionali vennero rilevate anche onde elettromagnetiche, emesse in tutto lo spettro di frequenza: **è la nascita dell'Astronomia Multimessaggera** (radiazione elettromagnetica + onde gravitazionali + neutrini), in cui confluisce l'Astrofisica, la Fisica delle particelle elementari, la Relatività Generale e la Cosmologia.
E ancora, sempre la collaborazione LIGO e VIRGO, in uno studio pubblicato nel mese di Giugno del 2020, annuncia la scoperta di un oggetto di 2,6 masse solari, rivelato il 14 agosto 2019 mentre si fondeva con un buco nero di 23 masse solari, generando onde gravitazionali. Non si sa se si sia trattato di un buco nero o di una stella di neutroni, ma ha comunque rappresentato un record: sarebbe infatti il buco nero più leggero ad oggi conosciuto (quelli finora noti non sono inferiori a 5 masse solari) o la stella di neutroni più pesante (quelle conosciute raggiungono le 2,5 masse solari).
La scoperta delle onde gravitazionali ha un altro importante aspetto: apre un nuovo tipo di scienza. Sarà possibile innanzitutto testare in modo dettagliato la teoria della Relatività Generale di Einstein, confrontandola con certe proposte alternative avanzate da altri scienziati. Queste implicano differenze tecniche che si traducono in differenti osservazioni, incluse teorie vettoriali, teorie scalari e la cosiddetta costante cosmologica. La loro verifica comporta misurazioni dettagliate e il confronto con i dati o forme d'onda delle onde gravitazionali. Un obiettivo più lontano è la speranza di sviluppare una teoria quantistica della gravità. Ciò potrebbe unificare la fisica quantistica e la relatività generale. La scoperta delle onde gravitazionali, inoltre, introduce un modo totalmente nuovo ed eccitante di guardare al nostro universo, attraverso i fenomeni gravitazionali. Quello che sappiamo sull'Universo proviene da molti tipi di telescopi, ma tutti osservano segnali di origine elettromagnetica, inclusi la luce visibile, gli infrarossi, gli ultravioletti, microonde, onde radio. Molti fenomeni non emettono radiazioni elettromagnetiche, ma possono avere effetti gravitazionali che danno onde gravitazionali. Un buon esempio sono i buchi neri, che danno segnali gravitazionali, ma non segnali elettromagnetici. Alcuni fenomeni producono entrambi i tipi di onde come si è rilevato con LIGO. Essere in grado di vedere i segnali delle onde gravitazionali, oltre alle onde elettromagnetiche, ci permette di osservare in modo più ampio il nostro universo. Abbiamo imparato che i buchi neri grandi decine di volte la massa del nostro sole esistono, si trovano in sistemi binari, collidono e si fondono. Inoltre, abbiamo visto anche stelle di neutroni collidere, fornendo una nuova spiegazione sul modo in cui elementi pesanti, come l'oro, abbiano avuto origine sulla Terra quando si è formata.
L'esistenza delle onde gravitazionali fornisce la conferma sperimentale che l'interazione gravitazionale non consiste in un'azione a distanza fra corpi massivi (come era supposto nel modello newtoniano della gravità), ma implica l'esistenza dello spaziotempo curvo che deformandosi lievissimamente consente la propagazione dell'onda.

Capitolo 13

TEST CLASSICI DELLA RELATIVITÀ GENERALE

In un certo senso una teoria basata sulla deduzione è nemica della fisica sperimentale. Quest'ultima lotta continuamente per stabilire mediante prove cruciali la natura delle cose fondamentali; la prima, lotta per minimizzare i successi ottenuti dimostrando quanto vasta sia la natura delle cose che può accordarsi con tutti i risultati sperimentali.
(A. S. Eddington, *The Mathematical Theory of Relativity*, Cambridge University Press (1923))

Lo sviluppo efficace della scienza richiede il giusto equilibrio tra due metodi: da un lato, la costruzione a partire dalle osservazioni, dall'altro, la deduzione - mediante il puro ragionamento - da ipotesi congetturali.
(Paul A. M. Dirac, *Physical Science and Philosophy,* Nature, vol.139, 1937, pag. 1001.)

13.1 Introduzione

In questo capitolo ci occuperemo dettagliatamente delle prove sperimentali che vennero indicate da Einstein, già nel lavoro fondamentale del 1916, a sostegno della validità della sua Teoria della Relatività Generale[1]. Si tratta dei cosiddetti tre *test classici*: ***la precessione del perielio di Mercurio, la deflessione dei raggi luminosi e lo spostamento gravitazionale verso il rosso delle linee spettrali*** (*red shift*). Vedremo che la validità di questi pochi test è molto diversa dai moltissimi che supportano la Teoria della Relatività Speciale, le cui fondamenta sperimentali sono le meglio verificate di tutta la fisica. Il motivo principale di questi pochi test consiste nel fatto che i campi gravitazionali vicini a noi, sui quali è facile sperimentare, sono molto deboli ed i loro effetti relativistici non sono molto diversi dagli analoghi newtoniani. A questi test se ne è aggiunto, più recentemente, un quarto: ***il ritardo di un segnale luminoso in un campo gravitazionale***. Come per tutte le teorie fisiche, la validità della Relatività Generale, deve concordare con i risultati sperimentali e le osservazioni. *Per quanto possa essere profonda, bella ed*

[1]Per approfondire, sia dal punto di vista concettuale sia dal punto di vista storico, le classiche prove sperimentali e talune più moderne, si consiglia vivamente allo studente la lettura dei seguenti volumi: Arthur S. Eddington, *Spazio, Tempo e Gravitazione*, Universale scientifica Boringhieri, Torino, 1971; Clifford M. Will, *Einstein aveva ragione? Le prove sperimentali della Relatività Generale*, Bollati Boringhieri, Torino, 1989.

elegante, essa dovrebbe essere subito abbandonata se risultasse in disaccordo con i dati osservativi.
D'altra parte Einstein nel 1930 scrisse: *"Non ritengo che il significato principale della teoria della relatività generale consista nella previsione di qualche piccolo effetto sperimentalmente osservabile, ma piuttosto nella semplicità dei suoi fondamenti e nella sua coerenza interna."*

13.2 Avanzamento del perielio dei pianeti

Come già dicemmo all'inizio del Cap.11, tra la fine del XIX° secolo e l'inizio del XX°, poche teorie fisiche potevano essere paragonate alla teoria della gravitazione di Newton, in quanto a precisione delle loro previsioni, che qui, per comodità del lettore, ripetiamo. Il moto e la posizione dei pianeti e dei satelliti, entro il sistema solare, le eclissi del sole e della luna, il moto delle comete, ecc. sono previsti dalla meccanica celeste classica con precisione a dir poco sbalorditiva, inoltre tali previsioni sono in accordo con i dati osservativi entro pochi secondi d'arco su intervalli temporali di parecchi anni. Tuttavia la teoria di Newton si rivelò non del tutto esatta: il moto di precessione del perielio dei pianeti più vicini al sole, previsti dalla teoria newtoniana, si discostano leggermente dai valori osservati. Nel caso del pianeta Mercurio, tale scostamento è maggiore che per gli altri pianeti ed ammonta a circa $43''$ per secolo e non è assolutamente spiegato dalla meccanica newtoniana. Questa piccola deviazione fu scoperta attraverso calcoli molto laboriosi sia dal matematico ed astronomo francese Francois Le Verrier nel 1845, sia dall'astronomo e matematico statunitense Simon Newcomb nel 1882, che ricalcolò tale deviazione.
La spiegazione della precessione del perielio di Mercurio fu uno dei primi successi della teoria relativistica della gravitazione di Einstein.
La spiegazione di tale avanzamento fu data da Einstein nel lavoro originale del 1916 prima che Karl Schwarzschild scoprisse la metrica che porta ora il suo nome. Tuttavia la formulazione della soluzione esatta del problema esterno di Schwarzschild, che abbiamo studiato nel capitolo precedente, permette agevolmente di confermare le osservazioni astronomiche sull'avanzamento del perielio del pianeta Mercurio, che prima della formulazione della Relatività Generale da parte di Einstein non aveva avuto alcuna conferma teorica. E noi per dimostrarlo, nei paragrafi seguenti, forniremo una trattazione basata appunto sulla metrica di Schwarzschild.

13.2.1 Moto ellittico: perielio ed afelio

È ben noto, dalla teoria della gravitazione di Newton, che *se un punto materiale P, di massa m, è attirato in ragione inversa del quadrato della distanza $r = OP$ da un punto fisso O di massa M (**corpo centrale**), e se la velocità iniziale $\mathbf{v}_0$ è esterna al raggio vettore iniziale OP_0, la traiettoria del punto P è o una ellisse, o un arco di parabola o un arco di iperbole, aventi un fuoco in $F \equiv O$.* La traiettoria è un'ellisse, e quindi chiusa, quando la velocità iniziale $\mathbf{v}_0$ è in valore assoluto minore della cosiddetta **velocità parabolica** $v_p = \sqrt{\dfrac{2\mu}{r_0}}$, cioè quando:

$$v_0 < \sqrt{\frac{2\mu}{r_0}} = \sqrt{\frac{2G(M+m)}{r_0}}, \tag{13.1}$$

dove la costante $\mu = G(M+m)$ è il **coefficiente di attrazione del problema dei due corpi** e G la costante di gravitazione newtoniana. Il moto del punto P è allora quello che in Astronomia si chiama **moto ellittico o moto kepleriano** (Fig.13.1).
Se la massa m di P è trascurabile di fronte alla massa M del corpo centrale, cioè se $m \ll M$, la costante di attrazione si riduce a $\mu = GM$. Questo moto è periodico, e, a ciascuna rivoluzione del punto P sull'ellisse, purché questa non sia una circonferenza, il raggio vettore OP passa

Figura 13.1: Moto ellittico o kepleriano.

per un **minimo** e per un **massimo**, corrispondenti ai due vertici dell'ellisse situati sull'asse maggiore, o, usando il linguaggio Astronomico, al **perielio** e **all'afelio** dell'orbita[2]. In tutte le rivoluzioni del moto ellittico, questi due punti, **perielio** ed **afelio**, sono fissi: in particolare l'angolo dei raggi vettori di due perieli consecutivi è uguale a 2π.

Figura 13.2: Avanzamento o retrocessione del perielio di un pianeta.

Supponiamo ora che l'azione del punto fisso O sul punto mobile P non sia più l'attrazione newtoniana, diretta da P verso O e di valore assoluto $\frac{\mu}{r^2}$, ma sia una forza vicina a questa attrazione. Supponiamo in particolare che il moto di P abbia ancora luogo su un piano fisso passante per O, nel senso positivo (antiorario) attorno a questo punto, e che OP abbia un minimo ed un massimo successivi, corrispondenti a dei punti della traiettoria che possiamo chiamare ancora **perielio** ed **afelio**. L'angolo dei raggi vettori di due perieli consecutivi non è più in generale uguale a 2π, ma è vicino a questo valore.[3] Durante la rivoluzione limitata al passaggio

[2]Secondo l'etimologia il **perielio** di un pianeta è il punto dell'orbita in cui il pianeta è più vicino al Sole, e **l'afelio** è il punto dell'orbita del pianeta più lontano dal Sole. Analoga terminologia si usa per il moto della Luna attorno alla Terra, si dice allora **perigeo** ed **apogeo**. Per un astro qualsiasi si suol dire: **periastro** e **apoastro**.

[3]Per esempio, Newton calcolò che se nel valore dell'attrazione $\frac{\mu}{r^2}$, si sostituisce l'esponente 2 con l'esponente N, N essendo un numero vicino a 2, l'angolo dei raggi vettori di due perieli consecutivi è in prima approssimazione $\frac{2\pi}{\sqrt{3-N}}$ anziché 2π: si ha **avanzamento** o **retrocessione** del perielio a seconda che $N > 2$ o $N < 2$.
Per uno studio approfondito dell'avanzamento del perielio dei pianeti, in particolare di Mercurio, si consiglia di consultare l'opera, in 2 volumi, di Jean Chazy, *La Théorie de la Relativité et la Mécanique Céleste*, Gauthier-Villars et Cie. Editeurs, Paris (1928),Tome I, Chap. III, pagg. 56-85 e Chap.V, pagg.190 -193. Si tratta di un'opera notevole, oggi un

di P per due perieli consecutivi Π e Π' (Fig.13.2), si ha **avanzamento** o **retrocessione** del perielio secondo che l'angolo $\Pi O \Pi'$, descritto dal raggio vettore, sia superiore o inferiore a 2π; **il valore algebrico dell'avanzamento del perielio** nella rivoluzione considerata è la differenza fra l'angolo $\Pi O \Pi'$ e 2π.

13.2.2 Equazioni differenziali delle geodetiche del $\mathbf{ds}^2$ esterno di Schwarzschild

Nel capitolo precedente sono state determinate le equazioni differenziali delle geodetiche 4-dimensionali relative al $\mathbf{ds}^2$ esterno di Schwarzschild, le quali, interpretate nello spazio fisico 3-dimensionale ci forniranno le equazioni delle orbite dei pianeti in moto di rivoluzione attorno al Sole. Lo studio di queste orbite ci potrà dire, in particolare, se i loro perieli, durante più rivoluzioni successive attorno al Sole, subiscono o meno degli spostamenti e valutare quali sono le loro entità.
Tali equazioni, espresse in funzione delle variabili di Schwarzschild θ, ϕ, t e r, considerate come funzioni di s, sono le (12.67) del capitolo precedente, che qui riscriviamo in un ordine leggermente diverso:

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2 = 0, \tag{13.2}$$

$$r^2 \sin^2\theta\dot{\varphi} = k_1, \tag{13.3}$$

$$\left(1 - \frac{\alpha}{r}\right)\dot{t} = k_2, \tag{13.4}$$

$$c^2\left(1 - \frac{\alpha}{r}\right)\dot{t}^2 - \left(1 - \frac{\alpha}{r}\right)^{-1}\dot{r}^2 - r^2\left(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2\right) - 1 = 0, \tag{13.5}$$

dove $\alpha = \dfrac{2GM}{c^2} = \dfrac{2\mu}{c^2}$, $\mu = GM$, M massa centrale che crea il campo gravitazionale e k_1, k_2, costanti arbitrarie d'integrazione.
Per semplificare queste equazioni scegliamo i seguenti valori iniziali

$$\theta = \frac{\pi}{2} \iff \cos\theta = 0,\ \dot{\theta} = \frac{d\theta}{ds} = 0,$$

i quali, sostituiti nell'equazione (13.2) forniscono il valore $\ddot{\theta} = 0$, che implica $\dot{\theta}$ =cost. Quindi se $\dot{\theta} = 0$ all'istante iniziale, esso si mantiene tale durante il moto del pianeta. Allo stesso modo, poichè θ =cost, se $\theta = \dfrac{\pi}{2}$ all'istante iniziale, si mantiene tale durante il moto. Allora la soluzione $\theta = \dfrac{\pi}{2}$ è l'unica soluzione ammissibile per l'equazione (13.2) con queste condizioni iniziali. Ciò significa che *la traiettoria spaziale è piana come nel caso newtoniano e giace nel piano equatoriale definito da* $\theta = \dfrac{\pi}{2}$.
Per $\theta = \dfrac{\pi}{2}$ si ha $\sin\theta = 1$ e le rimanenti tre equazioni differenziali (13.3) (13.4) e (13.5) diventano

$$r^2\frac{d\varphi}{ds} = k_1 \iff \frac{d\varphi}{ds} = \frac{k_1}{r^2} \tag{13.6}$$

$$\left(1 - \frac{\alpha}{r}\right)\frac{dt}{ds} = k_2 \iff \left(1 - \frac{\alpha}{r}\right)^{-1} k_2 = \frac{dt}{ds} \tag{13.7}$$

$$c^2\left(1 - \frac{\alpha}{r}\right)\left(\frac{dt}{ds}\right)^2 - \left(1 - \frac{\alpha}{r}\right)^{-1}\left(\frac{dr}{ds}\right)^2 - r^2\left(\frac{d\varphi}{ds}\right)^2 - 1 = 0. \tag{13.8}$$

po' dimenticata, che, nonostante alcune parti siano un po' invecchiate, conserva ancora la propria freschezza ed, a nostro avviso, rimane insuperata in molti punti importanti.

Dall'equazione (13.8), cambiando di segno a tutti i termini, moltiplicando per $\dfrac{1}{1-\dfrac{\alpha}{r}}$ ed utilizzando l'equazione (13.7), si ottiene:

$$\left(\frac{dr}{ds}\right)^2 + r^2\left(1-\frac{\alpha}{r}\right)\left(\frac{d\varphi}{ds}\right)^2 - c^2k_2^2 + \left(1-\frac{\alpha}{r}\right) = 0.$$

Moltiplicando per $\left(\dfrac{ds}{d\varphi}\right)^2$, avremo:

$$\left(\frac{dr}{d\varphi}\right)^2 + r^2\left(1-\frac{\alpha}{r}\right) + \left(1-c^2k_2^2\right)\left(\frac{ds}{d\varphi}\right)^2 - \frac{\alpha}{r}\left(\frac{ds}{d\varphi}\right)^2 = 0. \tag{13.9}$$

Tenendo conto della (13.6) la precedente equazione si trasforma nella:

$$\left(\frac{dr}{d\varphi}\right)^2 + r^2 - \alpha r + \left(1-c^2k_2^2\right)\frac{r^4}{k_1^2} - \alpha\frac{r^3}{k_1^2} = 0. \tag{13.10}$$

Facendo un cambio di variabile, $r = \dfrac{1}{\rho}$, si ha[4]

$$\frac{dr}{d\varphi} = \frac{d\dfrac{1}{\rho}}{d\varphi} = -\frac{1}{\rho^2}\frac{d\rho}{d\varphi}.$$

L'equazione (13.10) allora si trasforma in

$$\frac{1}{\rho^4}\left(\frac{d\rho}{d\varphi}\right)^2 + \frac{1}{\rho^2} - \frac{\alpha}{\rho} + \frac{1-c^2k_2^2}{k_1^2}\frac{1}{\rho^4} - \frac{\alpha}{k_1^2}\frac{1}{\rho^3} = 0.$$

Ed infine moltiplicando 1° e 2° membro per ρ^4 si ottiene

$$\left(\frac{d\rho}{d\varphi}\right)^2 - \alpha\rho^3 + \rho^2 - \frac{\alpha}{k_1^2}\rho - \frac{c^2k_2^2-1}{k_1^2} = 0 \tag{13.11}$$

o equivalentemente

$$\left(\frac{d\rho}{d\varphi}\right)^2 = \alpha\rho^3 - \rho^2 + \frac{\alpha}{k_1^2}\rho + \frac{c^2k_2^2-1}{k_1^2}. \tag{13.12}$$

Questa è l'equazione differenziale del 1° ordine delle orbite.
Se si vuole l'equazione differenziale del 2° ordine, equivalente alla (13.12), basta derivare 1° e 2° membro di quest'ultima, rispetto a φ, si otterrà:

$$2\frac{d\rho}{d\varphi}\left[\frac{d^2\rho}{d\varphi^2} + \rho - \frac{3}{2}\alpha\rho^2 + \frac{\alpha}{2k_1^2}\right] = 0.$$

Se $\dfrac{d\rho}{d\varphi} \neq 0$, cioè *se escludiamo le orbite circolari*, avremo l'equazione del 2° ordine

$$\frac{d^2\rho}{d\varphi^2} + \rho = \frac{\alpha}{2k_1^2}\left(3k_1^2\rho^2 + 1\right) = \frac{1}{p} + \frac{3GM}{c^2}\rho^2, \tag{13.13}$$

[4]Molti autori, in teoria delle orbite, pongono $u = \dfrac{1}{r}$.

con $p = \dfrac{c^2k_1^2}{GM}$ e $\dfrac{1}{p} = \dfrac{GM}{c^2k_1^2} = \dfrac{\alpha}{2k_1^2}$; utilizzeremo quest'ultima equazione nel paragrafo 13.3.1 per studiare la deflessione dei raggi luminosi.

Osservazione sulla legge delle aree:

Il prodotto $r^2\dfrac{d\varphi}{dt}$, costante nella teoria newtoniana, qui dipende dalla posizione del pianeta sulla sua orbita (salvo il caso in cui essa sia circolare).

In effetti dai due integrali primi:

$$\frac{d\varphi}{ds} = \frac{k_1}{r^2}, \qquad \frac{dt}{ds} = \frac{k_2}{1-\dfrac{\alpha}{r}}$$

avremo:

$$r^2\frac{d\varphi}{dt} = r^2\frac{d\varphi}{ds}\frac{ds}{dt} = r^2\frac{\dfrac{d\varphi}{ds}}{\dfrac{dt}{ds}} = \frac{k_1}{k_2}\left(1-\frac{\alpha}{r}\right) = \frac{k_1}{k_2}\left(1-\frac{2GM}{c^2r}\right).$$

Perciò chi rimane costante è il prodotto $r^2\dfrac{d\varphi}{ds} = k_1$, e siccome ds è proporzionale al tempo proprio, $ds = cd\tau$, si può dire che *l'espressione usuale della legge delle aree resta valida, non rispetto al tempo coordinato t, ma rispetto al tempo proprio τ.*

13.2.3 Equazione differenziale dell'orbita ed avanzamento del perielio

Nel paragrafo precedente abbiamo stabilito l'equazione differenziale del 1° ordine dell'orbita (13.12):

$$\left(\frac{d\rho}{d\varphi}\right)^2 = P(\rho) \equiv \alpha\rho^3 - \rho^2 + B\rho + D, \tag{13.14}$$

dove $\rho = \dfrac{1}{r}$, $\alpha = \dfrac{2GM}{c^2} = \dfrac{2\mu}{c^2}$, $B = \dfrac{\alpha}{k_1^2}$, $D = \dfrac{c^2k_2^2 - 1}{k_1^2}$, k_1 e k_2 essendo due costanti d'integrazione delle quali k_1 ha il significato di costante delle aree.

Abbiamo dunque che l'equazione differenziale del 1° ordine delle orbite dedotta dal $\mathbf{ds}^2$ di Schwarzschild è caratterizzata, al secondo membro, da un polinomio di 3° grado nella variabile $\rho = \dfrac{1}{r}$ a coefficienti costanti. Dunque: *il quadrato della derivata $\dfrac{d\frac{1}{r}}{d\varphi}$, che sulle traiettorie newtoniane è un polinomio di 2° grado in $\dfrac{1}{r}$, sulle traiettorie dedotte dal $\mathbf{ds}^2$ di Schwarzschild è un polinomio di 3° grado in $\dfrac{1}{r}$, dove il termine di 2° grado è $-\dfrac{1}{r^2}$ come sulle traiettorie newtoniane, e dove il termine di 3° grado è indipendente dalle costanti d'integrazione.*[5]

L'equazione differenziale (13.14) dell'orbita contiene il differenziale $d\varphi$, ma non la variabile φ, e perciò si integra immediatamente con una quadratura:

$$\varphi = \int \frac{d\rho}{\sqrt{P(\rho)}} + cost., \tag{13.15}$$

avendo scelto il segno + davanti all'integrale.

Nell'approssimazione newtoniana il polinomio $P(\rho)$ è di 2° grado, cioè la (13.14) si riduce alla:

$$\left(\frac{d\rho}{d\varphi}\right)^2 = -\rho^2 + B\rho + D, \tag{13.16}$$

[5]Per la trattazione dello stesso problema in teoria newtoniana si può consultare: T. Levi Civita e U. Amaldi, *Lezioni di Meccanica Razionale*, vol 2, parte prima, Cap.III, pagg. 212-213 (si veda in particolare la formula (11)), Zanichelli Editore, Bologna, 1974.

ed avendo posto $\rho = \dfrac{1}{r}$ la precedente equazione diventa

$$\left(\frac{d\frac{1}{r}}{d\varphi}\right)^2 = -\frac{1}{r^2} + B\frac{1}{r} + D,$$

e perciò la quadratura si effettua per mezzo di funzioni circolari; la soluzione, com'è noto, è un'ellisse di equazione polare:

$$r = \frac{p}{1 + e\cos(\varphi - \varphi_0)}, \tag{13.17}$$

con $0 \leq e < 1$ l'eccentricità, p il parametro dell'orbita e φ_0 l'anomalia iniziale del raggio vettore supposto che l'asse polare abbia origine nel Sole e sia orientato verso il perielio.
Se supponiamo $\varphi_0 = 0$ e indichiamo con r_1 e r_2 i valori *minimo* (**perielio**)(per $\varphi = 0$) e *massimo* (**afelio**)(per $\varphi = \pi$) del raggio vettore, avremo:

$$r_1 = r_{min} = \frac{p}{1+e}, \qquad r_2 = r_{max} = \frac{p}{1-e}, \quad (r_1 < r_2) \tag{13.18}$$

e perciò ρ_1 sarà il valore di ρ corrispondente *al perielio* e ρ_2 *all'afelio*:

$$\rho_1 = \rho_{max} = \frac{1+e}{p}, \qquad \rho_2 = \rho_{min} = \frac{1-e}{p}, \quad (\rho_1 > \rho_2), \tag{13.19}$$

ed essendo i valori dei raggi vettori r_1 e r_2 abbastanza grandi ne segue che ρ_1 e ρ_2 sono piuttosto piccoli. Questi valori non sono altro che le radici dell'equazione di 2° grado

$$-\rho^2 + B\rho + D = 0 \tag{13.20}$$

che si ottengono dalla condizione di estremo $\dfrac{d\rho}{d\varphi} = 0$ che determina i massimi e minimi del polinomio di 2° grado che figura al 1° membro della (13.20), avremo perciò:

$$-\rho^2 + B\rho + D = -(\rho - \rho_1)(\rho - \rho_2) = (\rho - \rho_1)(\rho_2 - \rho) = -\rho^2 + (\rho_1 + \rho_2)\rho - \rho_1\rho_2, \tag{13.21}$$

nella quale $B = \rho_1 + \rho_2$ è la somma delle radici e $D = -\rho_1\rho_2$ è l'opposto del prodotto delle radici.
La (13.16) si può scrivere allora così:

$$\frac{d\rho}{d\varphi} = \sqrt{(\rho - \rho_1)(\rho_2 - \rho)}. \tag{13.22}$$

Allorquando ρ passa da ρ_1 a ρ_2, φ varia della quantità[6]:

$$\delta\varphi \equiv \varphi_2 - \varphi_1 = \int_{\rho_1}^{\rho_2} \frac{d\rho}{\sqrt{(\rho - \rho_1)(\rho_2 - \rho)}} = \left[\arcsin \frac{\rho - \frac{\rho_1+\rho_2}{2}}{\frac{\rho_2-\rho_1}{2}}\right]_{\rho_1}^{\rho_2} = \pi, \tag{13.23}$$

[6]L'integrale che figura nella (13.23), tenuto conto della (13.21), si calcola ricordando il risultato di Analisi Matematica:

$$\int \frac{Ax + C}{\sqrt{-x^2 + px + q}}dx = -A\sqrt{-x^2 + px + q} + \left(C + \frac{Ap}{2}\right)\arcsin\frac{x - \frac{p}{2}}{\sqrt{q + \frac{p^2}{4}}}, \quad \left(q + \frac{p^2}{4} > 0\right)$$

con $x = \rho$, $A = 0$, $C = 1$, $p = \rho_1 + \rho_2$, $q = -\rho_1\rho_2$. Per il calcolo di questo integrale si può consultare: Aldo Ghizzetti, *Complementi ed Esercizi di Analisi Matematica*, vol.II, Terza edizione 1961, Libreria Eredi Virgiglio Veschi, Roma, pag. 91-92.

che rappresenta l'angolo formato dai due raggi vettori che dal Sole vanno rispettivamente al perielio e all'afelio, risultato facilmente prevedibile dall'esame della (13.17). *Ovviamente fra due perieli successivi l'angolo formato dai rispettivi raggi vettori sarà il doppio del precedente, cioè* $\Delta\varphi = 2\delta\varphi = 2\pi$.
Nel caso generale relativistico in cui si deve ragionare sull'equazione (13.15) nella quale $P(\rho) \equiv \alpha\rho^3 - \rho^2 + B\rho + D$ è un polinomio di 3° grado, *l'integrazione conduce a funzioni ellittiche.* Ma si può vedere che in pratica, per studiare moti analoghi a quelli planetari, ed in particolare per calcolare in tali moti, l'avanzamento del perielio, è sufficiente effettuare nell'equazione (13.15) una quadratura approssimata.
La ricerca dei *valori estremi (minimi e massimi)* di $\rho(\varphi)$ per determinare i quali è necessario che sia soddisfatta la condizione $\dfrac{d\rho}{d\varphi} = 0$, conduce, per la (13.14), all'equazione di 3° grado:

$$\alpha\rho^3 - \rho^2 + B\rho + D = 0. \tag{13.24}$$

Si hanno dunque, per la (13.24), tre radici ρ_1, ρ_2, ρ_3 legate dalla relazione[7]

$$\rho_1 + \rho_2 + \rho_3 = \frac{1}{\alpha}, \tag{13.25}$$

che esprime la somma delle radici di un'equazione algebrica di 3° grado, come il rapporto cambiato di segno tra il 2° ed il 1° coefficiente, cioè $-\left(\dfrac{-1}{\alpha}\right) = \dfrac{1}{\alpha}$, come insegna l'Algebra.
Il termine in ρ^3 essendo correttivo rispetto al caso newtoniano, ed essendo il suo coefficiente $\alpha = \dfrac{2GM}{c^2}$ molto piccolo, implica che il polinomio di 3° grado $\alpha\rho^3 - \rho^2 + B\rho + D$ differisce di pochissimo dal polinomio di 2° grado $-\rho^2 + B\rho + D$ del caso newtoniano. Ne consegue che le radici ρ_1 e ρ_2 di quest'ultimo sono praticamente anche radici dell'equazione di 3° grado (13.24). Ma le radici newtoniane ρ_1 e ρ_2 rappresentano, rispettivamente, i valori di ρ corrispondenti al perielio e all'afelio. Inoltre, dal momento che α, ρ_1, ρ_2 sono piccole, ρ_3 per la (13.25) deve essere necessariamente grande rispetto a ρ_1 e ρ_2.
L'equazione (13.14) si scrive allora così[8]:

$$\begin{aligned}\frac{d\rho}{d\varphi} &= \sqrt{\alpha(\rho-\rho_1)(\rho-\rho_2)(\rho-\rho_3)} = \sqrt{(\rho-\rho_1)(\rho_2-\rho)\alpha(\rho_3-\rho)} = \\ &= \sqrt{(\rho-\rho_1)(\rho_2-\rho)[1-\alpha(\rho_1+\rho_2)-\alpha\rho]} = \\ &= \sqrt{(\rho-\rho_1)(\rho_2-\rho)\Big[1-\alpha(\rho_1+\rho_2)\Big]\Big[1-\frac{\alpha\rho}{1-\alpha(\rho_1+\rho_2)}\Big]}.\end{aligned}$$

[7]In ciò che segue si ritiene che l'equazione di 3° grado (13.24), a coefficienti reali, abbia tutte e tre le radici ρ_1, ρ_2, ρ_3 reali e positive ciò che, com'è noto dalla teoria delle equazioni algebriche, non sempre è vero. Per accertarsene è necessario trasformare la (13.24) nella cosiddetta *equazione ridotta* $x^3 + px + q = 0$ con il cambio di variabile $\rho = x + \dfrac{1}{3\alpha}$ e verificare, tenendo conto della **formula di Cardano** che **il discriminante** dell'equazione ridotta **sia negativo**, cioè: $\Delta = \dfrac{q^2}{4} + \dfrac{p^3}{27} < 0$. Il verificarsi di questa condizione (il cosiddetto *caso irriducibile*) dipende dai valori che assumono i coefficienti α, B e D, dei quali il primo è indipendente dalle costanti d'integrazione, mentre vi dipendono gli altri due. Nel nostro caso si può comunque verificare dall'esame dei coefficienti della (13.24) che, per opportuni valori delle costanti d'integrazione k_1 e k_2, che figurano soltanto in B e D, si può avere $\Delta < 0$ e perciò, in tal caso, la (13.24) possiede tre radici reali e distinte.

[8]Si tenga presente che $\rho_1 + \rho_2 + \rho_3 = \dfrac{1}{\alpha}$, dunque $\rho_3 = \dfrac{1}{\alpha} - (\rho_1 + \rho_2)$ e perciò $\alpha(\rho_3 - \rho) = \alpha\rho_3 - \alpha\rho =$ $= \alpha\Big[\dfrac{1}{\alpha} - (\rho_1+\rho_2)\Big] - \alpha\rho = 1 - \alpha(\rho_1+\rho_2) - \alpha\rho$. Inoltre $1 - \alpha(\rho_1+\rho_2) - \alpha\rho = 1 - \alpha(\rho_1+\rho_2) - \dfrac{\alpha\rho[1-\alpha(\rho_1+\rho_2)]}{1-\alpha(\rho_1+\rho_2)} =$ $[1-\alpha(\rho_1+\rho_2)]\Big[1 - \dfrac{\alpha\rho}{1-\alpha(\rho_1+\rho_2)}\Big]$.

Possiamo scrivere la precedente equazione nel modo seguente:

$$d\varphi = \frac{1}{\sqrt{(\rho-\rho_1)(\rho_2-\rho)}}\frac{1}{\sqrt{1-\alpha(\rho_1+\rho_2)}}\frac{1}{\sqrt{1-\dfrac{\alpha\rho}{1-\alpha(\rho_1+\rho_2)}}}d\rho,$$

ed essendo $\alpha(\rho_1+\rho_2)$ molto piccolo, possiamo trascurarlo entro la radice quadrata che sta al denominatore del 3° fattore, del secondo membro, della precedente equazione, che diventerà

$$d\varphi \approx \frac{1}{\sqrt{(\rho-\rho_1)(\rho_2-\rho)}}\frac{1}{\sqrt{1-\alpha(\rho_1+\rho_2)}}\frac{1}{\sqrt{1-\alpha\rho}}d\rho. \tag{13.26}$$

Ora, data la piccolezza sia di $\alpha(\rho_1+\rho_2)$ sia di $\alpha\rho$, il 2° e 3° fattore del secondo membro della (13.26) possiamo svilupparli in serie di potenze trascurando in tali sviluppi i termini di ordine superiore a 2, avremo allora[9]:

$$d\varphi \approx \frac{\left[1+\frac{\alpha}{2}(\rho_1+\rho_2)\right]\left(1+\frac{\alpha\rho}{2}\right)}{\sqrt{(\rho-\rho_1)(\rho_2-\rho)}}d\rho. \tag{13.27}$$

Quindi la variazione di φ tra ρ_1 e ρ_2 è

$$\varphi_2-\varphi_1 \approx \left[1+\frac{\alpha}{2}(\rho_1+\rho_2)\right]\int_{\rho_1}^{\rho_2}\frac{1+\frac{\alpha\rho}{2}}{\sqrt{-\rho^2+(\rho_1+\rho_2)\rho-\rho_1\rho_2}}d\rho, \tag{13.28}$$

e per la formula d'integrazione data nella precedente nota a piè dipagina, nella quale si ponga $x=\rho$, $A=\frac{\alpha}{2}$, $C=1$, $p=\rho_1+\rho_2$, $q=-\rho_1\rho_2$, $\sqrt{q+\frac{p^2}{4}}$, si trova:

$$\begin{aligned}\varphi_2-\varphi_1 &\approx \left[1+\frac{\alpha}{2}(\rho_1+\rho_2)\right]\left[\left(-\frac{\alpha}{2}\right)\sqrt{(\rho-\rho_1)(\rho_2-\rho)}+\left(1+\frac{\alpha}{4}(\rho_1+\rho_2)\right)\arcsin\frac{\rho-\frac{\rho_1+\rho_2}{2}}{\frac{\rho_2-\rho_1}{2}}\right]_{\rho_1}^{\rho_2}= \\ &= \left[1+\frac{\alpha}{2}(\rho_1+\rho_2)\right]\left[1+\frac{\alpha}{4}(\rho_1+\rho_2)\right]\pi = \\ &= \left[1+\frac{\alpha}{4}(\rho_1+\rho_2)+\frac{\alpha}{2}(\rho_1+\rho_2)+\frac{\alpha^2}{8}(\rho_1+\rho_2)^2\right]\pi.\end{aligned}$$

Trascurando il termine molto piccolo $\frac{\alpha^2}{8}(\rho_1+\rho_2)^2$, si ha:

$$\varphi_2-\varphi_1 \approx \left[1+\frac{3}{4}\alpha(\rho_1+\rho_2)\right]\pi, \tag{13.29}$$

quindi per due perieli successivi avremo il doppio della precedente grandezza, cioè

$$2\pi+\frac{3}{2}\pi\alpha\left(\rho_1+\rho_2\right).$$

[9] Lo sviluppo in serie, cui si allude sopra, è: $\frac{1}{\sqrt{1-x}}=1+\frac{1}{2}x-\frac{1\cdot 3}{2\cdot 4}x^2+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3-$ .....con $|x|<1$, che al 1° ordine, cioè trascurando i termini dal secondo ordine in poi, ci dà: $\frac{1}{\sqrt{1-x}}\approx 1+\frac{1}{2}x$.

Come si può notare, questo valore è superiore del valore newtoniano 2π, della quantità[10]

$$\Delta\varphi = \frac{3\pi}{2}\alpha\left(\rho_1 + \rho_2\right). \tag{13.30}$$

Sostituiamo ora ad α, ρ_1, ρ_2 i loro valori derivanti dal significato fisico e geometrico del sistema studiato

$$\alpha = \frac{2GM}{c^2}, \quad \rho_1 + \rho_2 = \frac{1+e}{p} + \frac{1-e}{p} = \frac{2}{p},$$

dove e è l'eccentricità e p il parametro dell'ellisse orbitale. Per quest'ultimo scegliamo il valore newtoniano in quanto si discosta di pochissimo da quello einsteiniano, quindi $p = a(1-e^2)$, dove a è il semiasse maggiore dell'orbita. La (13.30) si trasforma perciò nella:

$$\Delta\varphi = \frac{6\pi GM}{e^2 a\left(1-e^2\right)}. \tag{13.31}$$

Questa è la formula di avanzamento del perielio di un qualsiasi pianeta durante una rivoluzione attorno al Sole. (Fig.13.3).

Figura 13.3: Avanzamento del perielio di un pianeta in particolare di Mercurio durante due rivoluzioni successive.

Al fine di utilizzare unicamente dati provenienti dalle osservazioni astronomiche, sostituiamo $GM = \mu$ con la sua espressione in funzione del semiasse maggiore a e del periodo T di una rivoluzione, a partire dalla terza legge di Keplero:

$$\frac{4\pi^2 a^3}{T^2} = \mu = GM.$$

Finalmente, **l'avanzamento del perielio in un giro del pianeta attorno al Sole**, è dato dalla formula, immediatamente calcolabile, dovuta ad Einstein (1916):

$$\Delta\varphi = \frac{24\pi^3 a^2}{\left(1-e^2\right)c^2 T^2} \quad \text{rad/giro}, \tag{13.32}$$

nella quale $\Delta\varphi$ è espresso in radianti/giro, a in cm, c in cm/s ($c \approx 3 \times 10^{10}$ cm/s) e T in secondi.

[10]Non si confonda la differenza $\delta\varphi = \varphi_2 - \varphi_1$ che compare nella (13.29) e che individua l'angolo formato dai raggi vettori r_1 ed r_2 che vanno dal Sole rispettivamente al perielio ed all'afelio in una prima rivoluzione del pianeta, con il $\Delta\varphi$ che figura invece nella (13.30) in quanto rappresenta l'angolo formato dai due raggi vettori che vanno dal Sole a due perieli consecutivi raggiunti dal pianeta durante il suo moto.

13.2.4 Calcolo numerico dell'avanzamento del perielio di Mercurio

La formula (13.32) del paragrafo precedente, fornisce l'avanzamento in *radianti/giro* che è una quantità estremamente piccola e perciò di difficile misura ed anche poco significativa. È molto più significativo determinare l'avanzamento del perielio in secondi d'arco per secolo in quanto gli astronomi sono in grado di misurarlo.
A tal fine bisogna calcolare dapprima *l'avanzamento in secondi d'arco per rivoluzione* $(''/giro)$, e per ottenere ciò bisogna moltiplicare il valore precedente (13.32) per

$$\frac{180 \times 3600}{\pi}.$$

Infine per ottenere *l'avanzamento in secondi d'arco per secolo* $(''/secolo)$, occorre *moltiplicare per il numero di rivoluzioni per secolo*, cioè per

$$\frac{T_{Terra}}{T_{Pianeta}} \times 100 = \frac{\text{Numero di rivoluzioni}}{\text{secolo}},$$

dove

$$T_{Terra} = 365,25 \text{ (Numero dei giorni solari medi contenuti in un anno)},$$

e

$$T_{Pianeta} = \text{(Numero dei giorni solari medi corrispondenti ad una rivoluzione del Pianeta)}.$$

Nel caso di Mercurio si ha

$$T_{Mercurio} = 87,97 \times 3600 \times 24 \ s.,$$

cioè in questo caso si ha:

$$\frac{T_{Terra}}{T_{Mercurio}} \times 100 = \frac{365,25}{87,97} \times 100 = 415,20 \text{ (Numero delle rivoluzioni di Mercurio per secolo)}.$$

Nella tabella 13.1 abbiamo riportato dati orbitali provenienti da osservazioni astronomiche [11]. I semiassi maggiori a delle orbite sono espressi in funzione della distanza media a_0 della Terra

Pianeta	$\frac{\mathbf{a}}{\mathbf{a_0}}$	**e**	**T**	$\boldsymbol{\Delta\varphi}$
Mercurio	0,3871	0,2056	87,97	42″,9
Venere	0,7233	0,0068	224,70	8″,67
Terra	1	0,0167	365,25	3″,83
Marte	1,5237	0,0933	686,98	1″,35

Tabella 13.1:

dal Sole, presa per unità, il cui valore è di circa:

$$a_0 = 14,95 \times 10^{12} \text{ cm}.$$

In base alla precedente tabella, per il pianeta Mercurio si ha il rapporto $a/a_0 = 0,3871$ e perciò

$$a_{Mercurio} = 0,3871 \times 14,95 \times 10^{12} \text{ cm} = 5,78145 \times 10^{12} \text{ cm},$$

[11] Per i dati numerici riportati qui di seguito il lettore può consultare Henri Arzeliès, *Relativité Gnéralisée e Gravitation*, Fascicule II, Chap.V, paragrafo 2.[56] b., pag.86, Gauthier-Villars, Paris, 1963; oppure Jacques Lévy, *Astronomie de position et relativité générale*, Cahiers de physique, t.12, n° 99, novembre 1958, pagg.437-446; oppure un qualunque trattato moderno di Astronomia.

inoltre, per lo stesso Mercurio si ha $e_{Merc.} = 0,2056$ e $T_{Merc.} = 87,97 \times 24 \times 3600\,\mathrm{s/rivoluz.}$ Possiamo ora calcolare esplicitamente il valore dell'avanzamento del perielio del pianeta Mercurio applicando la formula (13.32) in maniera opportuna, tenuto conto delle considerazioni precedenti ed impostando il calcolo così:

$$\Delta\varphi = \frac{24\pi^3 a^2}{(1-e^2)\,c^2T^2} \times \frac{180 \times 3600}{\pi} \times 415,20\,''/\mathrm{secolo}. \tag{13.33}$$

Il calcolo esplicito è opportuno effettuarlo nel modo seguente:

$$\Delta\varphi = \frac{24}{1-e^2} \times \left(\frac{\pi a}{cT}\right)^2 \times 269.049.600 =$$

$$= \frac{24}{1-(0,2056)^2} \times \left(\frac{\pi \times 5,781 \times 10^{12}\,\mathrm{cm}}{3 \times 10^{10}\,\mathrm{cm/s} \times 7.600.608\,\mathrm{s}}\right)^2 \times 269.049.600\,''/\mathrm{secolo} =$$

$$= 25,0592902807 \times \left(\frac{1818,03966863}{22.801.824}\right)^2 \times 269.049.600 \approx 42'',9\,''/\mathrm{secolo},$$

cioè

$$\Delta\varphi \approx 43\,''/\mathrm{secolo}.$$

Questo è **l'avanzamento del perielio del pianeta Mercurio** in secondi d'arco per secolo: si tratta della conferma teorica dei dati osservati sperimentalmente dagli Astronomi su un arco temporale di qualche secolo, e non spiegato in nessun modo dalla teoria newtoniana della gravitazione. *In effetti è stato spiegato per la prima volta da Einstein (1916) in base alla sua teoria relativistica della gravitazione e costituisce perciò, senza ombra di dubbio, uno dei maggiori successi della Relatività Generale!*

13.3 Deflessione dei raggi luminosi

A soli tre anni dalla pubblicazione dell'articolo fondamentale di Einstein (1916) sulla Relatività Generale, nel 1919, fu possibile effettuare le prime verifiche sperimentali sulla deflessione dei raggi luminosi di certe stelle, da parte del nostro Sole, grazie all'opera dell'astronomo inglese Sir Arthur Stanley Eddington[12](1882-1944). Egli, all'inizio della prima guerra mondiale, occupava una posizione di primo piano nel panorama scientifico britannico ed era stato da poco eletto membro della Royal Society ed era anche direttore dell'Osservatorio Reale di Greenwich. Nonostante l'interruzione, durante il periodo bellico, degli scambi scientifici diretti tra scienziati britannici e tedeschi, il cosmologo olandese Willem de Sitter riuscì a far pervenire ad Eddington gli ultimi articoli pubblicati da Einstein. Eddington si rese subito conto delle profonde implicazioni della relatività generale e ne divenne uno dei primi e più profondi studiosi. Nel 1917 preparò un dettagliato rapporto per la Physical Society di Londra e questo contribuì a diffondere la conoscenza di questa nuova teoria. Eddington e l'astronomo reale Frank Dyson nel frattempo progettarono una duplice spedizione per misurare la deviazione della luce in occasione di una eclisse solare prevista per il 29 Maggio 1919. Essi si resero conto che tale eclisse sarebbe stata una splendida opportunità in quanto ci si attendeva che un gran numero di stelle molto luminose costituisse il campo di sfondo in prossimità del Sole. Delle due spedizioni, progettate dai due scienziati britannici, una fu guidata da Eddington in persona ed ebbe come meta l'isola di Principe, al largo della costa della Guinea Spagnola; l'altra, guidata da Andrew Crommelin, era

[12]Si consiglia vivamente allo studente la lettura del libro, già citato all'inizio del capitolo, divenuto ormai un classico: Arthur S. Eddington, *Spazio, Tempo e Gravitazione*, con Appendice di Tullio Regge, Universale Scientifica Boringhieri, 1971. Una descrizione approfondita e dettagliata della deflessione di un raggio luminoso nel campo gravitazionale del Sole si trova nel Cap. 7.

diretta alla città di Sobral nel nord del Brasile. Il principio sul quale era basato l'esperimento è piuttosto semplice. Durante un'eclisse totale di Sole, la Luna nasconde completamente il disco solare e consente l'osservazione delle stelle che vengono a trovarsi in vicinanza del bordo del Sole. Con l'uso di telescopi e lastre fotografiche gli astronomi raccolgono le immagini del disco oscurato e delle stelle che vi si trovano intorno. Queste immagini vengono poi confrontate con quelle fatte nella stessa regione del cielo molti mesi prima o dopo l'eclisse, di notte, quando il Sole non si trova in quella regione del cielo e perciò le immagini delle stelle sono nella loro posizione effettiva sulla volta celeste. In questa maniera è possibile misurare le eventuali deviazioni. L'accordo, entro il limite degli errori sperimentali, decretò il successo della teoria einsteiniana che detronizzava dopo quasi due secoli e mezzo la teoria newtoniana della gravitazione.
I risultati delle due spedizioni, riportati da Eddington a pag.153 del suo libro, citato nella nota a piè di pagina, sono:

$$Sobral \quad 1,98 \pm 0,12 \ secondi\ d'arco,$$

$$Principe \quad 1,61 \pm 0,30 \ secondi\ d'arco.$$

Questi risultati erano in ottimo accordo con la previsione einsteiniana che è di $1'',75$ come vedremo nel prossimo paragrafo.Tale avvenimento fece di Einstein una celebrità mondiale.

13.3.1 Raggi luminosi in quanto geodetiche di lunghezza nulla

In questo paragrafo ci occuperemo del calcolo effettivo della deflessione dei raggi luminosi nel campo gravitazionale del Sole prevista da Einstein e ne constateremo la concordanza coi dati sperimentali forniti da Eddington.
La traiettoria di un raggio luminoso è una geodetica di lunghezza nulla, perciò basterebbe introdurre nell'espressione della metrica di Schwarzschild $\mathbf{ds}^2 = 0$ e procedere nella determinazione dell'equazione differenziale del raggio luminoso, tuttavia esiste un percorso più semplice che è quello che ora seguiremo. A tal fine ricordiamo che nel paragrafo 13.2.2 abbiamo già trovato le equazioni differenziali delle geodetiche (orbite) di una generica particella mobile nel campo gravitazionale generato dal corpo di massa M. Se questa particella è un fotone tali equazioni ci forniranno le equazioni differenziali di un raggio di luce. Prendiamo allora in considerazione l'equazione differenziale del 2° ordine delle geodetiche (13.13), che, come abbiamo visto nel paragrafo 13.2.2 è:

$$\frac{d^2\rho}{d\varphi^2} + \rho = \frac{3GM}{c^2}\rho^2 + \frac{GM}{c^2k_1^2}, \tag{13.34}$$

ed imponiamo la condizione $\mathbf{ds}^2 = 0$ equivalente a $ds = 0$. Ciò non è direttamente praticabile sulla (13.34) perché essa non contiene il ds ma può essere fatto sull'integrale primo (13.6) del paragrafo 13.2.2, che ora riscriviamo così:

$$\frac{ds}{d\varphi} = \frac{r^2}{k_1}. \tag{13.35}$$

Ma, porre nella (13.35) $ds = 0$ implica che la costante d'integrazione k_1 assuma valore infinito ($k_1 = \infty$). In tal caso, l'equazione differenziale (13.34) diventa:

$$\frac{d^2\rho}{d\varphi^2} + \rho = \frac{3GM}{c^2}\rho^2, \tag{13.36}$$

è questa l'equazione differenziale del 2° ordine dei raggi luminosi che si propagano in un campo gravitazionale a simmetria sferica generato da una massa sferica omogenea (o a strati concentrici omogenei). Come si vede si tratta di un'equazione non lineare per la presenza, al 2° membro, del termine di 2° grado ρ^2.

Nella più bassa approssimazione, cioè quando il campo gravitazionale generato dal corpo centrale è completamente trascurato, il 2° membro della (13.36) può essere considerato nullo, e perciò ρ soddisfa l'equazione differenziale lineare del 2° ordine a coefficienti costanti (dello stesso tipo del moto armonico):

$$\frac{d^2\rho}{d\varphi^2} + \rho = 0. \tag{13.37}$$

L'integrale generale di questa equazione, com'è noto, può essere scritto nel modo seguente:

$$\rho = A\sin(\varphi - \varphi_0),$$

con A e φ_0 due costanti arbitrarie. Se le variabili $r = \dfrac{1}{\rho}$ e φ si interpretano come coordinate polari su di un piano, la precedente equazione fornisce l'insieme delle curve integrali dell'equazione differenziale (13.37); passando a coordinate cartesiane ponendo $x = r\cos\varphi$, $y = r\sin\varphi$ si può vedere che si tratta di ∞^2 rette del piano Oxy in cui l'origine O coincide col centro del Sole.
Con un'appropriata scelta degli assi x ed y si può fare in modo che sia $\varphi_0 = 0$. La soluzione della (13.37) sarà perciò

$$\rho = \frac{1}{R}\sin\varphi, \tag{13.38}$$

avendo indicato con $\dfrac{1}{R}$ la costante A d'integrazione.

Figura 13.4: Deflessione dei raggi luminosi in un campo gravitazionale.

Questa equazione, tenuto conto che $y = r\sin\varphi$ equivale alla

$$y = r\sin\varphi = R = cost. \tag{13.39}$$

e rappresenta perciò, al variare di R, un fascio di rette parallele all'asse x, se invece il valore di $R = cost.$ è fissato, rappresenta una ben determinata retta parallela all'asse x (vedi Fig.13.4); se questa costante si identifica col raggio del Sole, $R = R_\odot$, la retta $y = R$ rappresenta un raggio di luce, proveniente da una stella lontana, che lambisce il bordo del Sole e che non è incurvato perché si è trascurato completamente la sua azione gravitazionale. Inoltre la (13.38) mostra che $r = \dfrac{1}{\rho}$ ha un minimo valore $r_{min.} = R$ per l'angolo $\varphi = \dfrac{\pi}{2}$ (Fig.13.4).
Se ora cerchiamo una soluzione della (13.36) in 2^a approssimazione, sostituendo in essa ρ con

$$\rho = \rho_0 + \rho_1, \tag{13.40}$$

essendo ρ_1 una piccola perturbazione e $\rho_0 = \dfrac{1}{R}\sin\varphi$ la soluzione di 1^a approssimazione, avremo

$$\frac{d^2(\rho_0 + \rho_1)}{d\varphi^2} + (\rho_0 + \rho_1) = \frac{3GM}{c^2}(\rho_0 + \rho_1)^2,$$

cioè

$$\frac{d^2\rho_0}{d\varphi^2} + \frac{d^2\rho_1}{d\varphi^2} + \rho_0 + \rho_1 = \frac{3GM}{c^2}\rho_0^2 + \frac{6GM}{c^2}\rho_0\rho_1 + \frac{3GM}{c^2}\rho_1^2.$$

Se in quest'ultima equazione trascuriamo gli ultimi due termini del 2° membro e teniamo presente che $\frac{d^2\rho_0}{d\varphi^2} + \rho_0 = 0$, avremo:

$$\frac{d^2\rho_1}{d\varphi^2} + \rho_1 = \frac{3GM}{c^2}\frac{1}{R^2}\sin^2\varphi. \tag{13.41}$$

Si può verificare che quest'equazione differenziale ha per *soluzione particolare*:

$$\rho_1 = \frac{GM}{c^2R^2}(1 + \cos^2\varphi). \tag{13.42}$$

Perciò la soluzione generale della (13.41) si ottiene aggiungendo alla soluzione particolare (13.42) la soluzione generale dell'omogenea associata che è la (13.38):

$$\rho = \frac{1}{R}sin\varphi + \frac{GM}{c^2R^2}(1 + \cos^2\varphi). \tag{13.43}$$

Introducendo le coordinate cartesiane $x = r\cos\varphi, y = r\sin\varphi$, la soluzione precedente si scriverà così:

$$y = R - \frac{GM}{c^2R}\left(r + \frac{x^2}{r}\right),$$

essendo però $r = \sqrt{x^2 + y^2}$ possiamo scriverla, più convenientemente, nella forma:

$$y = R - \frac{GM}{c^2R}\frac{2 + \left(\frac{y}{x}\right)^2}{\sqrt{1 + \left(\frac{y}{x}\right)^2}}|x|. \tag{13.44}$$

Si vede così che se $|x| \ggg |y|$, al numeratore ed al denominatore della frazione che figura nella formula precedente, si potranno trascurare gli addendi $\left(\frac{y}{x}\right)^2$, e la soluzione (13.44) si approssimerà asintoticamente all'espressione

$$y = R - \frac{2GM}{c^2R}|x|, \tag{13.45}$$

che rappresenta i due asintoti della curva (13.44). Con più precisione i due asintoti hanno equazioni date da:

$$y = R + \frac{2GM}{c^2R}x, \quad \text{per } x < 0, \tag{13.46}$$

$$y = R - \frac{2GM}{c^2R}x, \quad \text{per } x > 0. \tag{13.47}$$

Le due rette precedenti hanno per coefficienti angolari l'espressione

$$\tan\varphi = \pm\frac{2GM}{c^2R}, \tag{13.48}$$

perciò esse formano angoli supplementari con l'asse x, ed essendo $\frac{2GM}{c^2R}$ una quantità molto piccola, avremo $\tan\varphi \approx \varphi$. Dunque (Fig.13.4) l'asintoto 1, orientato concordemente col senso

di percorrenza della luce, forma con l'asse x l'angolo $\varphi_1 \approx \dfrac{2GM}{c^2R}$, mentre l'asintoto 2, anch'esso concordemente orientato col senso di percorrenza della luce, forma con lo stesso asse x un angolo φ_2 uguale al precedente ($\varphi_1 = \varphi_2$). È allora ovvio che l'angolo δ di deflessione del raggio luminoso è il doppio di quello che gli asintoti formano con la retta $y = R$, parallela all'asse x, avremo perciò:

$$\delta = \frac{4GM}{c^2R} \text{ rad.} \tag{13.49}$$

Nella (13.49), δ è espresso in radianti, M in grammi, R in centimetri e c in cm/s. Per avere δ in secondi d'arco è necessario moltiplicare la precedente espressione per $\dfrac{180 \times 3600}{\pi}$, si avrà perciò:

$$\delta = \frac{4GM}{c^2R} \times \frac{180 \times 3600}{\pi} \text{ secondi d'arco}(''). \tag{13.50}$$

In pratica, per un raggio che lambisce il bordo del Sole, si dovrà porre: $M = 1,99 \times 10^{33}$ g, $G = 6,6726 \times 10^{-8}\,\mathrm{cm}^3 \cdot \mathrm{g}^{-1} \cdot \mathrm{s}^{-2}$, $R = 6,95 \times 10^{10}$ cm, $c = 3 \times 10^{10}$ cm/s, avremo allora:

$$\delta = \frac{4 \times 6,67 \times 10^{-8} \times 1,99 \times 10^{33}}{9 \times 10^{20} \times 6,95 \times 10^{10}} \times \frac{180 \times 3600}{\pi} = 175.080,07 \times 10^{-5}, \tag{13.51}$$

cioè il classico risultato ottenuto da Einstein, che d'ora in poi indicheremo con δ_E:

$$\delta_E \approx 1'',75. \tag{13.52}$$

Questo risultato teorico concorda perfettamente con i dati sperimentali ottenuti da Eddington durante l'eclisse del 1919. Oggi però si dubita della precisione delle misure effettuate da Eddington e collaboratori che avevano incertezze variabili dal 10% al 30%. Questo margine d'incertezza piuttosto elevato, dovuto agli errori sperimentali, rese particolarmente importante ripetere le misurazioni. In occasione delle eclissi successive, numerose spedizioni tentarono di migliorare la precisione dei dati raccolti (vi furono tre spedizioni nel 1922, una nel 1929, due nel 1936, ed ancora una nel 1947, nel 1952 e, più di recente nel 1973), ma con scarso successo. Verifiche più precise sono state compiute di recente misurando la deflessione, in vicinanza del Sole, delle onde radio emesse dalle cosiddette *quasar*, che confermano, senza ombra di dubbio, la previsione di Einstein.
Per l'esattezza storica, è opportuno mettere in rilievo che già Newton predisse l'incurvamento dei raggi luminosi provenienti da stelle lontane e che passano vicino al bordo del Sole anche se, sembra, che lui non fece mai il calcolo per valutarne l'entità[13]. Tale calcolo venne effettuato, per la prima volta nel 1801, da un astronomo tedesco di nome Johann Georg von Soldner (1776-1833) che pubblicò il risultato nella rivista *Berliner Astr. Jahrb.* pag. 161-172; questo lavoro fu ristampato parzialmente in *Annalen der Physik*, **(4)**, Bd. 65, 1921, pag. 593-604. Secondo la teoria newtoniana della gravitazione è ben noto che l'orbita descritta da un corpo che si muova intorno al Sole è una conica: essa sarà una ellisse (in particolare un cerchio) o una parabola o una iperbole a seconda che la sua energia totale sia: $E < 0$, oppure, $E = 0$, oppure $E > 0$. Nel caso del raggio luminoso tangente al disco del Sole, poiché una particella di luce si muove con una velocità elevatissima, $c \simeq 300.000$ km/s, avrà una notevole energia, e l'orbita sarà certamente una iperbole col fuoco nel centro del Sole, dunque con un'eccentricità $e \gg 1$, talché tale iperbole si discosta poco dai propri asintoti. Tenendo conto di tutto ciò, Soldner riuscì a calcolare l'entità dell'angolo di deflessione, trovando:

$$\delta_N \simeq 0'',875,$$

[13]Newton espresse molto bene la sua convinzione che la luce avesse un peso, nella sua *Ottica* con le seguenti parole; Quesito 1. *Non agiscono i corpi sulla luce a distanza e non incurvano i suoi raggi con la loro azione, e non è tale azione (caeteris paribus) più forte alla distanza minima?* (Questa citazione è tratta da Arthur S. Eddington, *Spazio, Tempo e Gravitazione*, Boringhieri, Torino, 1971, Cap. 7, pag. 143).

cioè esattamente metà del valore previsto da Einstein, in base alla Teoria della Relatività Generale.
Il lavoro di Soldner fu però subito dimenticato, in parte perché l'effetto calcolato era al di là dei limiti di precisione dei telescopi dell'epoca, ed in parte a causa dell'affermarsi, nel corso del secolo XIX°, della teoria ondulatoria della luce, secondo la quale essa si muove come u'onda nell'ipotetico "etere", e quindi non subirebbe deflessioni.
Come Soldner circa un secolo prima, anche Einstein, che certamente non era a conoscenza del lavoro dell'astronomo tedesco, nel 1907 era interessato all'effetto della gravità sulla luce: se il *Principio di equivalenza* comportava un effetto sulla frequenza della luce (*il red shift gravitazionale*) - che studieremo nel prossimo paragrafo - esso doveva comportarne uno anche sulla sua traiettoria. Nel 1911 egli determinò l'entità della deflessione di un raggio luminoso tangente al Sole tenendo conto del *Principio di equivalenza*, ma trascurando l'effetto dovuto alla curvatura dello spazio-tempo, ottenendo stranamente lo stesso risultato di Soldner $0'',875$. Einstein suggerì di controllare tale effetto durante un'eclisse di Sole, ma l'effetto era troppo piccolo per poter essere misurato e la sua proposta, come nel caso di Soldner, venne presto dimenticata. Soltanto nel 1915 quando utilizzò pienamente le equazioni della Relatività Generale, che tengono conto della non euclideità dello spazio-tempo, ottenne il risultato corretto $\delta_E \simeq 1'',75$.
Ci occuperemo brevemente della deduzione newtoniana nel seguente paragrafo.

13.3.2 Deflessione dei raggi luminosi previsto dalla teoria newtoniana

Occupiamoci ora della deflessione dei raggi luminosi in senso newtoniano. Per quanto detto nel paragrafo precedente, in vicinanza del Sole, tali raggi sarebbero dei rami di iperbole aventi come fuoco il centro S del Sole. Proponiamoci allora di calcolare, sotto questa ipotesi, l'angolo di deviazione di un raggio luminoso che sia tangente al bordo del Sole e tale che quando la luce è infinitamente lontana da ogni massa, la sua velocità tenda al valore costante c della velocità della luce nel vuoto.
L'angolo di *deviazione newtoniana* δ_N si deduce immediatamente dalle formule classiche di geometria delle iperboli.
Supponiamo di considerare un raggio di luce che coincida con un ramo di iperbole il cui fuoco F, sia il centro S del Sole ($F \equiv S$) e che sia situato sul semiasse positivo delle x dunque di coordinate $F \equiv (0, f)$, con $f > 0$ *semidistanza focale*, talché si tratta della classica iperbole di equazione cartesiana $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$, con $a^2 + b^2 = f^2$.
Rispetto agli assi di simmetria dell'iperbole, scelti come assi cartesiani x e y, gli asintoti hanno per equazione (Fig.13.5):

$$y = \pm \frac{b}{a} x = \pm \sqrt{e^2 - 1}\, x, \tag{13.53}$$

essendo a e b rispettivamente *il semiasse trasverso* e *non trasverso* ed $e = \dfrac{f}{a}$ *l'eccentricità.*

È allora ovvio che δ_N coincide con l'angolo formato dagli asintoti e contenente l'asse non trasverso y, per cui è facile constatare che esso vale:

$$\delta_N = 2 \arctan \frac{1}{\sqrt{e^2 - 1}}, \tag{13.54}$$

con il relativo ramo d'iperbole molto teso (quasi rettilineo) e di conseguenza con una eccentricità e molto grande. D'altra parte, la distanza del fuoco S (coincidente con il centro del Sole) dal vertice V dell'iperbole, ad esso più vicino, è uguale al raggio R del Sole, cioè $VS = R$, inoltre essendo $OS = f$ la distanza focale e $OV = a$ il semiasse trasverso, sarà $VS = OS - OV = f - a$, cioè $R = f - a$. Ricordando infine che l'eccentricità è definita come $e = \dfrac{f}{a}$, avremo la semplice

Figura 13.5: Deflessione dei raggi luminosi previsto dalla teoria newtoniana.

relazione

$$R = a(e-1). \tag{13.55}$$

Se ora applichiamo ad una particella di luce, considerata di massa unitaria ed avente velocità v, il teorema dell'energia (equazione delle forze vive) $T - U = E$ e ricordiamo che il potenziale gravitazionale è $U = \dfrac{\mu}{r}$ e che inoltre, su di un'orbita iperbolica l'energia totale è $E = \dfrac{\mu}{2a}$, avremo:

$$\frac{1}{2}v^2 - \frac{\mu}{r} = \frac{\mu}{2a}. \tag{13.56}$$

D'altra parte v deve tendere alla velocità c della luce nel vuoto quando $r \to \infty$, sicché avremo $c^2 = \dfrac{\mu}{a}$. Dalle (13.55), (13.56) si ottiene allora

$$e = 1 + \frac{R}{a} = 1 + \frac{c^2 R}{\mu}. \tag{13.57}$$

Come si vede dalla precedente espressione, l'eccentricità dell'iperbole è molto grande, $e \gg 1$, per cui possiamo presumere che $e^2 - 1 \approx e^2$. Con tale approssimazione la (13.54) diventa[14]

$$\delta_N = 2 \arctan \frac{1}{e} \approx 2\,\frac{1}{e}, \tag{13.58}$$

e tenendo conto della (13.57) e che $\mu = GM$ (supponendo che la massa della particella sia trascurabile di fronte alla massa M del corpo centrale), avremo infine:

$$\delta_N = \frac{2GM}{c^2 R}. \tag{13.59}$$

[14]Si ricordi che lo sviluppo in serie di arctan x con $|x| < 1$ è la cosiddetta *serie ciclometrica*:

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots\ldots,$$

che ha raggio di convergenza uguale ad 1. Per valori molto piccoli di x, i termini dal secondo in poi sono trascurabili e perciò si ha

$$\arctan x \approx x.$$

Confrontando la (13.59) con l'analoga einsteiniana (13.49), cioè $\delta_E = \dfrac{4GM}{c^2R}$, possiamo concludere dicendo che: *la deviazione δ_N, calcolata nell'ipotesi che i raggi luminosi siano dei rami di iperbole del moto newtoniano, è la metà dell'analoga deviazione einsteiniana δ_E dedotta dalla metrica* $\mathbf{ds}^2$ *di Schwarzschild.* Per i raggi luminosi tangenti al disco solare essendo la deviazione einsteiniana $\delta_E = 1'',75$, la deviazione newtoniana sarà $\delta_N = 0'',875$, in accordo con i calcoli effettuati da Soldner.
Osservazione: Lenti Gravitazionali.
L'effetto della deflessione dei raggi luminosi applicato a corpi molto massivi, come può essere una galassia, un buco nero o un **quasar (quasi stellar radio source)**, può produrre un effetto molto spettacolare chiamato **lente gravitazionale**. Fu nel 1980 che gli astronomi riportarono per la prima volta l'identificazione di due oggetti, che precedentemente erano stati considerati come due quasar distinti, e sono invece risultati essere due immagini dello stesso oggetto (una reale ed una apparente) (Fig.13.6). L'interpretazione del fenomeno fu che l'effetto fosse dovuto alla presenza di una galassia piuttosto massiccia, distante circa un quarto della distanza che ci separa dal quasar che è la principale componente della lente gravitazionale.

Figura 13.6: Lente gravitazionale: una sorgente ed una sua immagine.

Analoga situazione si produce quando i raggi di luce provenienti da una sorgente lontana sono deviati dal campo gravitazionale generato da un oggetto sferico molto massivo: i raggi che passano alle estremità diametralmente opposte della massa si intersecano a grande distanza al di là di essa. Un osservatore posto nel punto d'intersezione di tali raggi vedrà contemporaneamente la sorgente in due punti distinti del cielo, vedrà cioè due immagini virtuali della sorgente (Fig.13.7).
Lo studio approfondito delle lenti gravitazionali occupa ormai interi volumi ed ovviamente esce dagli scopi di questo libro; ci basta avvertire il lettore, che nonostante il nome di lenti gravitazionali, esse non si comportano affatto allo stesso modo delle lenti ottiche, per cui necessitano di uno studio particolare adatto alla loro natura.[15]

13.4 Redshift delle righe spettrali in un campo gravitazionale

Lo spostamento verso il rosso (redshift) è lo spostamento in frequenza di un'onda elettromagnetica dovuto al campo gravitazionale creato da un corpo massivo. La luce ed ogni altra forma di radiazione elettromagnetica dotata di una determinata lunghezza d'onda, e quindi di una determinata frequenza, prodotta da una sorgente posta in un intenso campo gravitazionale, non appena entra in una regione dello spazio dove l'intensità del campo gravitazionale diminuisce, la sua lunghezza d'onda aumenta e perciò la frequenza diminuisce. Tale "stiramento della

[15]Ci sono ormai parecchie pubblicazioni sulle lenti gravitazionali, un libro molto completo fatto da noti specialisti di Relatività Generale, è : P. Schneider, J. Ehlers ed E. E. Falco, *Gravitational Lenses*, Springer-Verlag, Berlino, 1992.

Figura 13.7: Lente gravitazionale: una sorgente sdoppiata in due immagini.

lunghezza d'onda" appare, nelle regioni dello spettro visibile, come uno spostamento della radiazione verso la parte rossa dello spettro elettromagnetico.
Questo effetto fu previsto teoricamente da Einstein, nel 1916, sulla sola base del Principio di Equivalenza, senza l'intervento delle equazioni gravitazionali.
Esso consiste sostanzialmente nel *rallentamento del tempo nel campo gravitazionale - man mano che ci si allontana dal corpo che crea il campo - ed il conseguente* **redshift (spostamento verso il rosso)** *delle linee spettrali emesse da atomi di corpi massivi.*
Si tratta di un effetto molto difficile da verificare sperimentalmente, infatti ciò è stato fatto, per la prima volta, con un esperimento molto delicato da Pound e Rebka, soltanto nel 1960, alla Harvard University e ripetuto, con tecniche più raffinate, nel 1964 da Pound e Snider, per cui oggi si considera come un esperimento estremamente ben verificato.
La spiegazione teorica può essere facilmente conseguita applicando ciò che abbiamo stabilito nel paragrafo 11.14 del Cap.11 quando abbiamo studiato la gravità come fenomeno metrico utilizzando le considerazioni del precedente paragrafo 11.13 che sono fondate unicamente sul **Principio di equivalenza**.
Consideriamo, per esempio, un'onda luminosa emessa dal Sole e ricevuta sulla Terra. Il potenziale gravitazionale sulla superficie del Sole sia Φ_S. Ricordando la (11.36), abbiamo che l'intervallo di **tempo proprio** $d\tau$ è legato all'intervallo di **tempo coordinato** dt, dalla relazione

$$d\tau = \sqrt{g_{00}(x^i)}dt \tag{13.60}$$

ed utilizzando l'espressione di g_{00} dato dalla formula approssimata (11.156), sulla superficie del Sole, avremo:

$$d\tau_S = \sqrt{g_{00}(x^i_S)}dt = \sqrt{1 + \frac{2\Phi_S}{c^2}}dt. \tag{13.61}$$

Similmente, sulla Terra, l'intervallo di tempo proprio è legato all'intervallo di tempo coordinato dalla relazione:

$$d\tau_T = \sqrt{g_{00}(x^i_T)}dt = \sqrt{1 + \frac{2\Phi_T}{c^2}}dt, \tag{13.62}$$

dove Φ_T è il valore del potenziale gravitazionale sulla Terra. Supponiamo ora che n onde di frequenza ν_S siano emesse nel tempo proprio $\Delta\tau_S$ da un atomo del Sole. Allora

$$n = \nu_S \Delta\tau_S. \tag{13.63}$$

Sulla Terra certamente si riceveranno n onde, ma la frequenza e la durata del treno d'onde è cambiata. Usando per la Terra una analoga relazione frequenza-durata, come sul Sole, avremo:

$$n = \nu_T \Delta\tau_T, \tag{13.64}$$

e siccome il numero n delle onde è lo stesso, per le (13.63) e (13.64), avremo

$$\nu_S \Delta\tau_S = \nu_T \Delta\tau_T,$$

perciò:

$$\nu_T = \nu_S \frac{\Delta\tau_S}{\Delta\tau_T}. \tag{13.65}$$

Per la (13.60), l'intervallo di tempo coordinato dell'onda corrispondente a $\Delta\tau_S$ è:

$$\Delta t = \frac{\Delta\tau_S}{\sqrt{g_{00}(x_S^i)}} = \frac{\Delta\tau_S}{\sqrt{1 + \frac{2\Phi_S}{c^2}}}. \tag{13.66}$$

Noi supponiamo che l'intervallo di tempo coordinato Δt sia lo stesso sul Sole e sulla Terra.[16] Perciò la (13.62) fornisce:

$$\Delta t = \frac{\Delta\tau_T}{\sqrt{g_{00}(x_T^i)}} = \frac{\Delta\tau_T}{\sqrt{1 + \frac{2\Phi_T}{c^2}}}. \tag{13.67}$$

Per le (13.66) e (13.67) avremo:

$$\frac{\Delta\tau_S}{\Delta\tau_T} = \frac{\sqrt{g_{00}(x_S^i)}}{\sqrt{g_{00}(x_T^i)}} = \sqrt{\frac{1 + \frac{2\Phi_S}{c^2}}{1 + \frac{2\Phi_T}{c^2}}}, \tag{13.68}$$

sostituendo la (13.68) nella (13.65) si ottiene:

$$\nu_T = \nu_S \frac{\sqrt{g_{00}(x_S^i)}}{\sqrt{g_{00}(x_T^i)}} = \nu_S \sqrt{\frac{1 + \frac{2\Phi_S}{c^2}}{1 + \frac{2\Phi_T}{c^2}}}. \tag{13.69}$$

Se ora sviluppiamo sia il numeratore che il denominatore dell'ultima espressione nella (13.69) in serie di potenze delle piccole quantità $\frac{\Phi_S}{c^2}$ e $\frac{\Phi_T}{c^2}$, trascurando i termini d'ordine superiore al secondo, otteniamo[17]:

$$\nu_T = \nu_S \sqrt{1 + \frac{2\Phi_S}{c^2}} \frac{1}{\sqrt{1 + \frac{2\Phi_T}{c^2}}} \simeq \nu_S \left(1 + \frac{1}{2}\frac{2\Phi_S}{c^2}\right)\left(1 - \frac{1}{2}\frac{2\Phi_T}{c^2}\right) =$$

$$= \nu_S \left(1 - \frac{\Phi_T}{c^2} + \frac{\Phi_S}{c^2} - \frac{\Phi_S}{c^2}\frac{\Phi_T}{c^2}\right) \simeq \nu_S \left(1 + \frac{\Phi_S - \Phi_T}{c^2}\right).$$

dove si è trascurato il termine piccolissimo $\frac{\Phi_S}{c^2}\frac{\Phi_T}{c^2}$. Avremo allora:

$$\nu_T \simeq \nu_S + \nu_S \frac{\Phi_S - \Phi_T}{c^2}$$

[16] Sul *tempo proprio* ed il *tempo coordinato* si veda Adler-Bazin e Schiffer, pag. 120-126.

[17] Gli sviluppi in serie da utilizzare per il calcolo sopra riportato sono:

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8}x^4 + ...(|x| < 1)$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}x^4 - .....(|x| < 1)$$

ovvero

$$\frac{\nu_T - \nu_S}{\nu_S} \simeq \frac{\Phi_S - \Phi_T}{c^2}, \tag{13.70}$$

o con notazione più breve

$$\frac{\Delta\nu}{\nu_S} \simeq -\frac{\Delta\Phi}{c^2}, \tag{13.71}$$

avendo posto $\Delta\nu = \nu_T - \nu_S$ e $\Delta\Phi = \Phi_T - \Phi_S$.
Considerando la Terra immersa nel campo gravitazionale del Sole avremo

$$\Delta\Phi = \Phi_T - \Phi_S = -\frac{GM}{R_1} + \frac{GM}{R_2} = GM\Big(\frac{1}{R_2} - \frac{1}{R_1}\Big)$$

con M massa del Sole e con R_1 distanza fra il centro del Sole e la superficie della Terra ed R_2 raggio del Sole. Essendo $R_1 > R_2$ si avrà $\frac{1}{R_1} < \frac{1}{R_2}$, perciò $\Delta\Phi > 0$, ciò implica, per la (13.71) $\Delta\nu = \nu_T - \nu_S < 0$ ovvero $\nu_T < \nu_S$.
Perciò la frequenza della luce decresce da quando lascia il Sole a quando viene ricevuta sulla Terra, noi vediamo dunque uno **spostamento verso il rosso delle righe spettrali (redshift)**. Naturalmente, non vi è nulla di speciale usando punti della superficie dellaTerra e della superficie del Sole come punti in un campo gravitazionale, si potrebbero prendere, per esempio, due punti a differenti altezze rispetto alla superficie della Terra. Se le misure sperimentali fossero abbastanza precise da poter misurare il corrispondente piccolo spostamento spettrale si otterrebbe che la frequenza diminuisce passando da una regione in cui il campo gravitazionale è più intenso ad una regione in cui esso è meno intenso.

Capitolo 14

INTRODUZIONE ALLA COSMOLOGIA

Poiché lo spazio, al di là dei confini del mondo, si estende illimitatamente, l'animo razionale si chiede che cosa ci sia al di là (di quei confini), ove la mente anela di spingere lo sguardo; e l'animo stesso nella sua ansia di libertà vuole travalicare.
(Lucrezio, *De Rerum Natura,* libro 2, vv. 1044-1047)

La Cosmologia nata come un capitolo dell'Astronomia e dell'Astrofisica è diventata ormai una scienza a sé stante caratterizzata da uno sviluppo rapidissimo e tumultuoso. Moltissimi sono i libri ad essa completamente dedicati, purtroppo questi invecchiano molto rapidamente da un anno all'altro a causa delle molteplici scoperte osservative che i moderni e potentissimi telescopi ci offrono in continuazione. Questo capitolo è dedicato ad una brevissima introduzione alla **Cosmologia teorica o matematica** considerata, quasi esclusivamente, dal punto di vista del fisico matematico soprattutto per quel che concerne la formulazione dei cosiddetti modelli di Friedmann visti come applicazione delle equazioni di campo della Relatività Generale.[1]

14.1 Introduzione

"*Nessuna branca della scienza può sostenere di avere una maggiore area d'interesse della* **Cosmologia**, *in quanto la Cosmologia è lo studio dell'Universo, e l'Universo per definizione contiene* **ogni cosa.**"[2]
Più precisamente diciamo che: *la* **Cosmologia** *è quella parte dell'Astronomia che si propone lo studio dell'origine, della struttura e dell'evoluzione dell'Universo nella sua totalità: è perciò naturale che in essa confluiscono tutte le teorie fisiche classiche e moderne quali la Termodinamica, la Relatività Generale, la Meccanica Quantistica, la Teoria delle particelle elementari ecc. Essa può essere divisa, grosso modo, in due parti:* **la Cosmologia teorica o matematica e la Cosmologia osservativa o sperimentale** che si integrano e si compenetrano a vicenda.

[1]Il lettore desideroso di saperne di più su questo affascinante ramo della scienza, può consultare i seguenti volumi: Jayant Vishnu Narlikar: *An Introduction to Cosmology*, third edition, Cambridge University Press, 2002; oppure Jamal Nazrul Islam, *An Introduction to Mathematical Cosmology*, Second Edition, Cambridge University Press, 2002; o ancora Mark H. Jones, Robert J. A. Lambourne et al. *An Introduction to Galaxies and Cosmology*, Cambridge University Press, 2004. Un'altra lettura molto utile ed istruttiva, anche se un pò datata, è l'opera in 2 volumi di due eminenti astrofisici russi: Jakov B. Zeldovic e Igor D. Novikov, *Struttura ed evoluzione dell'Universo*, Vol.1: *Cosmologia Classica*, Vol.2: *Nascita delle Galassie e teorie cosmologiche alternative*, Editori Riuniti - Edizioni Mir - Mosca, 1982.

[2]L'affermazione riportata più sopra si trova, come frase introduttiva, a pag.1 dell'ottimo libro del celebre cosmologo indiano J. V. Narlikar citato nella nota precedente.

In questo capitolo noi ci occuperemo soprattutto della Cosmologia teorica, a carattere essenzialmente matematico, il cui compito principale è la formulazione dei cosiddetti **modelli cosmologici (relativistici).**
Nei due paragrafi che seguono daremo, per prima cosa, uno sguardo storico all'evoluzione delle nostre idee sull'universo fino alla scoperta, nel 1929, della sua espansione da parte dell'astronomo americano Edwin Powell Hubble.
Tutte le epoche, tutte le filosofie e tutte le religioni hanno avuto una loro ingenua "Cosmologia" ma essa come scienza moderna a sé stante è nata solo con l'avvento della Relatività Generale. È oggi possibile considerare una Cosmologia Newtoniana (come fanno certi autori) ma ciò è stato fatto in modo serio soltanto dopo la nascita della Cosmologia relativistica stimolata dalla Relatività Generale; ad ogni modo essa ha poca importanza nell'ambito della Cosmologia moderna che è essenzialmente relativistica e perciò noi non ce ne occuperemo.
La metodologia della Cosmologia non è diversa da quella degli altri capitoli della scienza: sulla base di alcuni dati di osservazione essa propone un *"modello teorico di Universo"*; questo modello e le sue conseguenze logiche vengono poi confrontati con nuovi risultati di osservazione; eventuali disaccordi conducono ad una modificazione, parziale, o radicale, del modello, e così via.
Le informazioni, inizialmente assai scarse, sul nostro Universo fisico hanno reso difficile la costruzione dei primi modelli, per i quali ci si è valsi di *"ipotesi di lavoro"* suggerite essenzialmente da motivi di semplicità. Ciò nonostante anche questi modelli semplicistici hanno costituito un buon innesco ed hanno dato luogo successivamente a modelli meno rozzi, dimostratisi utili per l'interpretazione dei dati che, in misura sempre maggiore, vengono forniti dall'osservazione astronomica e astrofisica.
A conclusione di queste osservazioni introduttive si potrebbe dire, con Livio Gratton[3], che: "La Cosmologia è veramente una *Historia Universalis*, da cui la storia umana viene esclusa semplicemente perché i suoi metodi sono di natura diversa. Questo è il motivo per cui la Cosmologia ha sempre avuto un'importanza enorme in tutte le religioni del mondo. Ogni religione ha avuto la sua Cosmologia, e molti filosofi e scienziati, da Anassagora a Giordano Bruno a Galileo hanno pagato a caro prezzo la loro opposizione alla Cosmologia ufficiale della religione dominante: Giordano Bruno fu messo al rogo per le sue idee cosmologiche. Galileo non fu processato per le sue scoperte scientifiche, ma per le sue dottrine cosmologiche!"

14.2 L'evoluzione delle nostre idee sull'Universo

Conformemente alle vedute di Tolomeo (2° sec. d.C.), tutto il Medioevo e parte dell'era moderna hanno ritenuto che tutte le stelle fossero fissate alla "*sfera cristallina (la sphaera fixarum degli antichi), avente la Terra al suo centro.*" Ancora con Niccolò Copernico (1473-1543) la concezione non è molto cambiata, benché al centro della sfera cristallina egli ponga il Sole anziché la Terra; la preminenza del Sole sulle altre stelle è ancora indiscussa.
La successiva evoluzione delle nostre idee sull'Universo ha avuto tre periodi principali prima che si arrivasse alla scoperta della sua espansione.

- **I° Periodo** (1540-1750).
  Il primo di questi periodi fu la persuasione, acquisita fra i primi da Giordano Bruno (1548-1600), che *le stelle occupano l'intero spazio e che il Sole non è che una stella tra tante.* Isaac Newton (1642-1727) completò quantitativamente questa ipotesi, basandosi sull'osservazione e sul calcolo. Ammettendo che le stelle fossero, mediamente, simili al Sole, e sapendo che la luminosità apparente decade col quadrato della distanza, dal confronto tra la luminosità

[3]Il Prof. Livio Gratton è stato un celebre scienziato italiano, professore di Astrofisica all'Università di Roma "La Sapienza" nel periodo in cui uno di noi (F. B.) era studente ed ha avuto il privilegio di averlo come docente di Astrofisica nell'Anno Accademico 1969-70. Il passo sopra riportato è preso dal suo bel libro: *Relatività, Cosmologia, Astrofisica*, Universale scientifica Boringhieri, Torino, 1968.

apparente di molte stelle e la luminosità apparente di alcune porzioni di Sole, egli valutò la distanza r delle stelle a noi più vicine ad almeno 100.000 volte quella r^* del Sole. Non si sbagliò di molto: oggi si ritiene che il rapporto sia $r/r^* \approx 270.000$. Anch'egli ebbe la convinzione che le stelle dovevano occupare in modo sensibilmente *uniforme* (a larghissima scala) e in maniera *statica* tutto lo spazio.

- **II° Periodo** (1750-1910).
  Il secondo periodo, che ebbe inizio verso il 1750 ad opera di diversi filosofi (E. Swedenborg 1688-1772; J. H: Lambert 1728-1777; I. Kant 1724-1804) fu dominato dall'idea che *la maggior parte delle stelle osservate ad occhio nudo o telescopicamente, appartenessero ad un insieme insulare, grande sì, ma finito, costituente la nostra Galassia (la Via Lattea); e che al di fuori di questo insieme altri ne dovessero esistere, più o meno simili ed ugualmente popolati di stelle, assai più lontani e diffusi ovunque nello spazio infinito.* A tale concezione seguì rapidamente la scoperta effettiva di oggetti celesti (**nebule o nebulose**) di apparenza diversa dalle stelle e perciò giudicati extragalattici. Nel 1783 C. Messier (1730-1817) ne aveva già contate 103. Successivamente, William Herschel (1738-1822), mediante il suo telescopio di 48 pollici costruito da lui stesso, arrivò ad identificarne più di 2500.
  La teoria dell'universo plurinsulare, dopo alterne fortune, fu definitivamente confermata il 1° gennaio 1925 quando Edwin Powell Hubble (1889-1953), il più famoso cosmologo dell'epoca, mediante il telescopio di 100 pollici di monte Wilson, riuscì a risolvere alcune nebule nelle stelle componenti e ad accertarne la distanza extragalatica.[4]

- **III° Periodo** (1910-1929).
  La terza tappa fondamentale ebbe inizio nel 1912 con W. M. Slipher che fu il primo ad osservare che gli spettri di "quasi tutte le nebulose extragalattiche" presentavano un sistematico spostamento verso il rosso di alcune righe spettrali.[5] Tale tappa culminò nel 1929, con le osservazioni di Edwin Hubble (1889-1953) e del suo assistente Milton Humason (1891-1972) che portarono alla scoperta che: *la stragrande maggioranza delle nebulose extragalattiche presentano un sistematico spostamento verso il rosso di alcune righe tipiche del loro spettro. Questo fatto, attribuito ad effetto Doppler, fu unanimemente interpretato come allontanamento sistematico (la cosiddetta "recessione") di tutte le galassie dell'universo visibile.* Osservazioni quantitative condussero Hubble ad enunciare, nel 1929, la famosa legge:
  Legge di Hubble: *Tutte le galassie dell'universo si allontanano le une dalle altre con velocità proporzionali alla loro distanza, cioè*:
  $$v = Hr. \tag{14.1}$$
  Il coefficiente di proporzionalità H viene oggi chiamato la **costante di Hubble**, benché non sia affatto esclusa una sua possibile dipendenza dal tempo, perciò esso va considerato come una *costante* soltanto rispetto alle variabili spaziali. Essendo $H = v/r$ la sua unità

[4] Ciò fu possibile per la presenza in tali nebule delle cosiddette **"cefeidi"**, stelle di altissima luminosità periodica (5 o 6 giorni), dal cui periodo si può risalire alla distanza.

[5] Abbiamo detto di "quasi tutte le nebulose extragalattiche" perché la prima nebulosa che Slipher osservò fu la famosa nebulosa di Andromeda - la più vicina alla nostra galassia (distando solo 10^6 anni luce) - e trovò, per ironia della sorte, che per essa le righe spettrali avevano uno spostamento verso il blu e che perciò si avvicinava a noi ad una velocità di 200 km/s anziché allontanarsi come le altre (*continua*). La spiegazione di questo fatto fu capita solo più tardi. Ora si sa che, mentre *gli ammassi di galassie* si allontanano gli uni dagli altri, le galassie contenute in ciascun ammasso hanno una configurazione relativamente stabile, eccetto che per quelli che sono detti i loro "moti peculiari." Questi movimenti peculiari sono regolati da campi gravitazionali locali, e non hanno niente a che fare con l'espansione dell'universo. Per i membri del nostro stesso ammasso, il cosiddetto Gruppo Locale (a cui appartiene sia la nostra galassia sia la nebulosa di Andromeda), le velocità peculiari possono essere sia di avvicinamento che di allontanamento e così non vi è nulla di sorprendente che alcuni spettri mostrino spostamenti verso il blu o il violetto. Per galassie più distanti in altri ammassi, le velocità peculiari non sono abbastanza grandi da annullare la velocità di espansione cosmica (che per la legge di Hubble, cresce con la distanza), così il moto che ne risulta è un moto recessivo, cioè di allontanamento.

di misura sarà una velocità divisa per una distanza, precisamente gli astronomi usano i chilometri al secondo per Megaparsec (1Mpc $= 10^6$pc), perciò l'unità di misura è km $\cdot$ s$^{-1} \cdot$ Mpc^{-1}.[6]
Sul valore della costante di Hubble vi sono delle incertezze, il valore oggi più attendibile è compreso nel seguente intervallo:

$$H = 50 \text{ km s}^{-1} \text{Mpc}^{-1} - 100 \text{ km s}^{-1} \text{Mpc}^{-1}, \tag{14.2}$$

cioè una galassia che è 100 milioni di parsec distante da noi, ha una velocità di allontanamento (detta anche *velocità di recessione*) compresa fra 5.000 *e* 10.000 km/s.

I tre periodi sopra descritti, come abbiamo visto, non sono divisi in compartimenti stagni ma si sovrappongono e si accavallano talvolta in modo tumultuoso con altri periodi caratterizzati da studi teorici e formulazione di modelli di universo piuttosto che dall'osservazione del cielo. Le principali tappe del successivo cammino possono essere sintetizzate come segue:

- **IV° Periodo** (1917 - 1930).
  Le premesse ai risultati di questo periodo furono poste da Albert Einstein che nel periodo 1905-1916 formulò le due teorie fondamentali: Teoria della Relatività Ristretta e Teoria della Relatività Generale. I fatti principali del periodo 1917-1927 furono:
  Formulazione dei primi modelli cosmologici relativistici da parte di A. Einstein (1917); Willem de Sitter (1917); Alexander Alexandrovic Friedmann (1922-1924); Georges Lemaitre (1927) e Arthur Stanley Eddington (1930).

Nei 30 anni che vanno dal 1930 al 1960, nonostante le moltissime osservazioni fatte soprattutto con i grandi telescopi americani di Monte Wilson e monte Palomar, non ci sono stati grandi avvenimenti di rilievo.

- **V° Periodo** (1960 - 2020).
  I fatti più importanti di quest'ultimo periodo sono:

  a) 1960: Scoperta delle **QUASAR (QUASi-stellAR radio source)** da parte di Mattews e Sandage;
  b) 1964-65: Scoperta della **Radiazione cosmica di fondo a microonde (Cosmic Microwave Background Radiation-CMBR)** da parte di Penzias ed Wilson;
  c) 1967: Scoperta delle **PULSAR (PULsating radio stAR)** da parte di ;
     Hewish e Bell;
  d) 1971: Osservazione in Cignus X-1 del primo candidato come **buco nero** (**black hole**).
     I buchi neri erano già stati ipotizzati teoricamente dopo la scoperta della metrica di Schwarzschild nel 1916;
  e) 1980: Scoperta delle **Lenti gravitazionali (Gravitational lenses)**;

[6] In Astronomia per **paralasse di una stella** s'intende l'angolo, espresso in secondi d'arco, sotto il quale si vedrebbe dalla stella il semidiametro maggiore a dell'orbita terrestre, corrispondente alla distanza media della Terra dal Sole ($a \approx 150 \times 10^6$ km). Se tale angolo è di $1''$ d'arco si dice che la stella si trova alla distanza di **1 parsec** (abbreviato in **pc**) e si dimostra che tale distanza e di circa 2,26 anni luce. Naturalmente si usano anche i multipli del **parsec**: 1 kiloparsec (1kpc) $= 10^3$ pc, 1 Megaparsec (1Mpc) $= 10^6$ pc, 1 Gigaparsec (1Gpc) $= 10^9$ pc.
Per chi voglia approfondire le questioni che riguardano l'Astronomia e l'Astrofisica si consigliano i seguenti volumi: Cesare Barbieri, *Lezioni di Astronomia*, Zanichelli, Bologna (1999); Hannu Kartunnen and Al. *Fundamental Astronomy*, 550 pages, Springer Nature (2016); Bradley W. Carroll and Dale A. Ostlie, *An Introduction to Modern Astrophysics*, 1359 pages, Cambridge University Press, second edition (2017).

f) Anni 1990: Discussioni attorno alla **materia oscura (dark matter)** già ipotizzata nel 1933 dall'astronomo svizzero Fritz Zwicky;

g) 1998: Scoperta **dell'accelerazione dell'Universo** associata all'esistenza **dell'energia oscura (dark energy)** da parte degli americani Saul Perlmutter e Adam Riess e all'australiano Brian Schmidt, premi Nobel 2011 per questa scoperta. Della materia oscura e dell'energia oscura che, si badi bene, sono due cose completamente diverse, se ne è parlato moltissimo, ma a tutt'oggi non si sa quasi nulla

h) 2015: Scoperta sperimentale delle **Onde gravitazionali (Gravitational waves)** ipotizzate teoricamente da Einstein nel 1916 sono state rivelate sperimentalmente il 14 Settembre 2015 con l'esperimento interferometrico statunitense LIGO e successivamente in Italia con un analogo esperimento chiamato VIRGO eseguito vicino a Pisa.

Per avere un'idea di tutti questi oggetti astronomici di notevole fascino si consiglia il lettore di consultare i vari siti Internet ad essi dedicati, accertandosi accuratamente della loro attendibilità.

14.3 Alcuni dati numerici su stelle e galassie

Stelle.
Secondo le conoscenze attuali, le stelle sono enormi **reattori termonucleari** che passano attraverso cicli di vita ben definiti, con una evoluzione che dipende in modo essenziale dalla loro massa. Il nostro Sole, che può essere considerato come un esempio tipico di stella media, ha un diametro dell'ordine di 14×10^{10} cm, una massa di 2×10^{33} g, ed una luminosità intrinseca totale di 4×10^{33} erg/s. Le sue dimensioni sono tali da contenere l'orbita lunare attorno alla Terra.
La Cosmologia non si interessa della costituzione interna delle stelle che è invece campo d'investigazione dell'Astrofisica.
Galassie.

Le Galassie sono degli aggregati che mediamente contengono circa da 10^{11} (100 miliardi) a 2×10^{11} (200 miliardi) di stelle, tenute assieme dalla loro mutua attrazione gravitazionale: di esse circa 6.000 sono visibili ad occhio nudo nella nostra Galassia (**la Via Lattea**) che viene considerata un esemplare tipico. Nella Fig.14.1 abbiamo riportato una sua rappresentazione schematica, molto semplificata, vista di profilo cioè secondo il cosiddetto *piano galattico*, ed in questa prospettiva essa appare come un disco con un *rigonfiamento centrale*. Nella vista di fronte invece, cioè nella direzione dell'asse di rotazione passante per il suo centro e perpendicolare al piano galattico essa appare come una tipica Galassia a spirale con alcuni bracci: per avere una visione più concreta si consiglia il lettore di consultare qualcuno dei tanti siti Internet che riportano splendide fotografie di diverse porzioni di Universo, specialmente quelle scattate dal famoso Hubble Telescope.
Le stelle che compongono la Galassia distano l'una dall'altra, in media, circa 10 anni luce. Una galassia ha generalmente forma lenticolare con diametro di circa 100.000 anni luce e spessore 10 volte più piccolo, cioè dell'ordine di 10.000 anni luce. Rappresentando le stelle come "teste di spillo", la distanza media di due stelle vicine diverrebbe di circa 50 km; il diametro della galassia sarebbe di circa 300.000 km e lo spessore di circa 30.000 km. Nella stessa scala il sistema solare, compresa l'orbita di Plutone, starebbe in un cerchio di 20 m di diametro.
La distanza media tra i centri di due galassie vicine è di $\sim 3 \times 10^6$ anni luce e cioè ~ 100 volte il

Figura 14.1: Rappresentazione schematica (non in scala) della nostra Galassia vista di profilo.

raggio di una galassia. Rappresentando le galassie come monete da 20 centesimi di euro, esse sarebbero intervallate di $\sim 1m$ l'una dall'altra.
Il numero delle galassie visibili al telescopio di 200 pollici di Monte Palomar (Stati Uniti) viene valutato a $\sim 10^{11}$ (100 miliardi). La più vicina (nebulosa di Andromeda) dista da noi $\sim 3 \times 10^6$ anni luce; attualmente la galassia più lontana è la GN-z11 che si trova da noi a circa $13,4 \times 10^9$ anni luce, che nella scala delle monete significa $\sim 4,5$ km. Per questi oggetti estremamente lontani, l'osservazione spettrale assegna una velocità di allontanamento pari a $\sim 4/5$ della velocità della luce. Partendo da questo dato ed accettando la legge di Hubble, non è possibile osservare galassie che si trovano al di là di un **orizzonte** posto a $\sim 2,25 \times 10^{10}$ anni luce, che nella scala delle monete significherebbe a $\sim 7,5$ km.
Ammassi e super ammassi (clusters and superclusters).
Le galassie non sono distribuite in modo rigorosamente uniforme nello spazio ma nemmeno in maniera del tutto casuale: esse piuttosto tendono ad aggregarsi in gruppi chiamati **ammassi (clusters)**, ciascun ammasso contiene da poche galassie ad alcune migliaia. Ci sono alcune evidenze sperimentali dell'esistenza di **ammassi di ammassi (o super-ammassi)** ma non gerarchie superiori.
Le osservazioni indicano che le galassie, in media, sono, ad ogni epoca (ovvero ad un dato istante), **sparpagliate uniformemente nell'Universo (omogeneità)**. Questo significa che se consideriamo una porzione di universo abbastanza grande in confronto delle distanze fra tipiche galassie vicine (queste distanze sono dell'ordine di 1 milione di anni luce), allora il numero di galassie in questa porzione è, in media, lo stesso di ogni altra porzione avente lo stesso volume ad ogni istante di tempo. La condizione "ad ogni istante di tempo" riferita alla distribuzione uniforme delle galassie è importante perché l'universo è in uno stato dinamico e perciò il numero di galassie in un dato volume cambia col tempo. La distribuzione delle galassie ci appare inoltre **isotropa**, cioè, è la stessa, in media, in tutte le direzioni rispetto a noi. Se ammettiamo che noi non occupiamo una speciale posizione fra le altre galassie, concludiamo che la loro distribuzione è **isotropa** rispetto ad ogni galassia, ciò implica necessariamente che le galassie sono **distribuite uniformemente in tutto l'universo.** Il significato ed il contenuto di queste ultime considerazioni sull'omogeneità ed isotropia dell'universo è stato compendiato in un principio generale che il famoso astrofisico inglese Edward Arthur Milne ha definito **Prin-**

cipio Cosmologico. Tale principio può dirsi alla base di tutta l'attuale Cosmologia e di esso ci occuperemo dettagliatamente nel prossimo paragrafo.

14.4 Principio Cosmologico

La presunzione che il nostro punto di osservazione, la nostra Galassia (in particolare la nostra Terra), non si trovi in posizione privilegiata ma al contrario rappresenti un **osservatorio tipico** del nostro universo, conduce. con una certa verosimiglianza (che, si badi bene, non è certezza), a ritenere che anche da ogni altro punto di osservazione (si pensi ad osservatori situati su altre galassie comunque lontane) l'universo appaia isotropo ed identico a quello che vediamo noi dalla Terra. Ciò conduce a postulare per l'universo un principio di simmetria che si può concisamente enunciare così:
Principio Cosmologico (PC)(1ª formulazione:)
Non esistono nell'universo punti di osservazione privilegiati,
ovvero con più precisione,
"in un dato momento", l'universo deve presentare lo stesso aspetto a tutti gli ipotetici osservatori che lo scrutino da tutte le "galassie tipiche", qualunque sia la direzione in cui si rivolge lo sguardo.[7]
In altri termini:
Principio Cosmologico (PC)(2^a formulazione:)
In scala sufficientemente ampia, l'universo è omogeneo ed isotropo.
Non è inutile qualche commento del precedente enunciato inteso a chiarirne l'esatto contenuto fisico.
Consideriamo allo scopo tanti osservatori situati sulle singole galassie, che diremo **osservatori fondamentali**. Abbastanza semplice è il concetto di **isotropia**: considerato un osservatore fondamentale, ad ogni istante il suo intorno deve apparirgli perfettamente uguale in tutte le possibili direzioni. E ciò deve accadere per ogni osservatore fondamentale.
Meno semplice è il concetto di **omogeneità**. Se l'universo, come l'osservazione lascia ben prevedere, non è statico, l'intorno di un osservatore fondamentale, pur restando isotropo, può differire profondamente da quello che era molti milioni di anni fa. Analogamente sono variabili gli intorni di tutti gli altri osservatori fondamentali. Che cos'è allora che questi devono confrontare, per giungere alla conclusione di omogeneità? Secondo A. G. Walker l'omogeneità va precisata nel modo seguente. Si supponga che ogni osservatore fondamentale tenga il *"giornale di bordo"* delle osservazioni da lui fatte sull'universo dal proprio punto di vista, assieme ai tempi in cui ha fatto le singole osservazioni, misurati con orologi standard (indicanti il suo tempo proprio). Dire che l'universo è omogeneo vuol dire che, previo opportuno aggiustamento dell'istante iniziale per ciascun osservatore, **"tutti i giornali di bordo risultano identici."** In altri termini: la storia dell'universo, vista da qualunque osservatore fondamentale può essere comunque complicata e varia, essa è però la stessa per tutti gli osservatori.
Come importante conseguenza del PC si ha *l'esistenza di un tempo cosmico* τ: è il *"tempo proprio"* di ciascuna galassia segnato dai vari orologi standard, aggiustati su situazioni identiche, in modo che i giornali di bordo risultano identici.
L'ipotesi contenuta nel PC è una ipotesi di simmetria dell'universo che in parte è suggerita dall'osservazione, ma che l'osservazione stessa non è ancora in grado di confermare completamente e senza incertezze. Essa va quindi considerata come *"un'ipotesi di lavoro"*, molto comoda perché molto semplificatrice, che è giustificato mantenere fino a che osservazioni più sicure non riusciranno a confermarla o a negarla definitivamente. In conclusione possiamo dire che **il PC non è un risultato sperimentale ma un'esigenza della ragione umana** che nasce dalla posizione tipica della scienza moderna, che, pur risalendo all'esperimento come criterio fon-

[7] L'espressione *"galassie tipiche"* significa galassie che non abbiano alcun apprezzabile moto peculiare ma siano semplicemente trasportate nel generale flusso cosmico.

damentale, riconosce che l'esperimento stesso non dice nulla senza l'intervento della ragione umana: è *"l'esperienza sensata"* di Galileo!

14.5 Difficoltà nel modello cosmologico newtoniano: paradosso di Olbers

È abbastanza strano che, finché la meccanica newtoniana ha dominato incontrastata, l'unico modello cosmologico comunemente ammesso sia stato quello, già suggerito da Newton, di un universo *infinito*, *omogeneo* e *statico*, oltre che *infinitamente antico*, senza mutue azioni di contatto e con sole interazioni gravitazionali. Inspiegabilmente altri modelli non furono mai proposti. Eppure il modello sopra nominato presenta non lievi difficoltà in teoria newtoniana.
Prima fra tutte il fatto che se la materia è distribuita in maniera omogenea nello spazio infinito, allora in ogni punto dello stesso si dovrebbe avere un potenziale gravitazionale anch'esso infinito, quindi inutilizzabile per i calcoli. D'altra parte, anche ammettendo, su basi intuitive, che le masse più lontane possano in qualche modo essere ignorate (come si fa abitualmente, per esempio, quando si studia il movimento del sistema solare) e che il potenziale gravitazionale seguiti a soddisfare localmente all'equazione di Poisson:

$$\Delta_2 \Phi = 4\pi G \rho,$$

appare chiaro che, in assenza di azioni di contatto l'equilibrio implicherebbe grad $\Phi = 0$, quindi essendo $\Delta_2 \Phi \equiv \text{divgrad}\, \Phi$, ciò implicherebbe che fosse nulla la densità della materia, $\rho = 0$, cioè assenza di materia, risultato manifestamente in contrasto con la realtà sperimentale[8].
Un'altra difficoltà, ancora più grave, consiste nel cosiddetto famosissimo **Paradosso di Olbers**[9], macroscopicamente contraddetto dall'osservazione, che consiste nella semplice, e apparentemente banale, domanda: **perché di notte è buio? In altre parole, perché il cielo non è illuminato di notte come di giorno?**
Più precisamente il paradosso può essere così enunciato:
Paradosso di Olbers:
Se l'universo fosse infinito, statico e uniformemente riempito di stelle e galassie, allora il cielo notturno dovrebbe apparire luminosissimo, più ancora di quello diurno, nel quale il Sole stesso diverrebbe indistinguibile dal fondo.
Ciò può essere dimostrato con un calcolo semplicissimo. Infatti, immaginiamo di tracciare, con centro nel punto da noi occupato, un *guscio sferico* di spessore infinitesimo dr, cioè lo spazio compreso fra due sfere concentriche di rispettivi raggi r ed $r + dr$, con r arbitrario; il volume di questo guscio, a meno di infinitesimi di ordine superiore, sarà allora $dV = 4\pi r^2 dr$. Se ora indichiamo con n il numero di stelle per unità di volume e con L la luminosità media di ciascuna stella, avremo che la luminosità totale delle stelle contenute nel guscio sferico sarà $nLdV = nL4\pi r^2 dr$, perciò il flusso luminoso $d\phi$ proveniente dal guscio sferico all'occhio dell'osservatore sarà

$$d\phi = \frac{nL4\pi r^2 dr}{4\pi r^2} = nLdr, \tag{14.3}$$

da cui integrando su di una sfera di raggio infinito, cioè sull'intero spazio, avremo

$$\phi = \int_0^{+\infty} nLdr = nL \int_0^{+\infty} dr = +\infty, \tag{14.4}$$

[8]Avvertiamo il lettore che finora abbiamo sempre usato la lettera μ per indicare la densità di massa e la lettera ρ per indicare la densità di carica elettrica. In Cosmologia però quasi tutti gli autori indicano con ρ la densità di massa, perciò in questo capitolo, anche noi ci uniformiamo a tale uso.

[9]Heinrich Wilhelm Olbers (1758-1840) è stato un medico tedesco, molto conosciuto per l'attività di astronomo amatoriale; il paradosso in questione venne enunciato da Olbers, allora settantenne, nel 1826 in un suo articolo apparso nell'Annuario astronomico di Bode. Esso rimase insoluto per quasi un secolo in quanto la soluzione si ebbe soltanto dopo la scoperta da parte di Hubble, nel 1929, dell'espansione dell'universo.

ovvero il flusso luminoso complessivo proveniente dallo spazio dovrebbe essere infinito. Dunque si dovrebbe avere sia di giorno che di notte una luminosità infinita, insomma non dovrebbe esistere la notte!
Nessuna spiegazione plausibile può essere data per colmare le precedenti difficoltà, se non abbandonando le ipotesi newtoniane. Queste ipotesi, ripetiamolo ancora una volta in modo esplicito, sono:

a) lo spazio, a grande scala, è euclideo ed infinito;

b) lo spazio è uniformemente riempito di stelle (omogeneità) (noi diremo piuttosto di galassie ma ai fini del ragionamento il risultato è lo stesso);

c) non vi sono movimenti sistematici delle stelle o meglio delle galassie (staticità);

d) l'universo è esistito "da sempre" (infinitamente antico).

Osserviamo, per inciso, che queste ipotesi sono, in sostanza, quello che ci si dovrebbe aspettare in base al Principio Cosmologico.
La prima spiegazione del paradosso che potrebbe venire in mente è che ci sia qualcosa nello spazio che ci nasconde la luce delle stelle molto lontane, un materiale assorbente, come c'è effettivamente nell'interno della nostra Galassia. Ma questa spiegazione non può reggersi, perché è necessario considerare il bilancio energetico del fenomeno.
La materia interstellare galattica che assorbe la luce delle stelle riemette in tutte le direzioni una quantità di energia uguale a quella assorbita, altrimenti si riscalderebbe. Noi non possiamo vedere l'energia riemessa, perché la materia interstellare è fredda e la sua radiazione si svolge nell'infrarosso; ma se sulla materia assorbente arrivasse tutta la luce degli infiniti gusci pieni di stelle, essa si riscalderebbe e con il tempo diventerebbe luminosa come le stelle. E siccome per l'ipotesi d) l'universo è esistito da sempre, c'è stato tutto il tempo necessario per questo riscaldamento.
D'altra parte si può pensare di lasciar cadere l'ipotesi a). Uno spazio non-euclideo e chiuso in sé stesso, come la superficie di una sfera, non può essere diviso in un numero infinitamente grande di gusci, e quindi si potrebbe pensare che la somma della luce di tutte le stelle potrebbe non essere infinita. Ma anche questo non regge, benché non si possa dimostrarlo in modo elementare. Sostanzialmente si può dire che se lo spazio fosse l'analogo 3-dimensionale della superficie di una sfera, la luce di ogni galassia potrebbe giungerci più volte: direttamente e dopo aver fatto uno o più giri dello spazio, e la somma sarebbe sempre infinitamente grande.
La forza del paradosso di Olbers è maggiore di quanto possa sembrare a prima vista! Si sono perciò tentate altre molteplici spiegazioni, sulle quali non possiamo qui soffermarci, ma che hanno condotto tutte ad un risultato negativo. Il paradosso di Olbers sembra dunque suggerire una rinuncia al modello newtoniano *infinito, omogeneo, statico ed infinitamente antico.* Il paradosso non si presenterebbe più ove si ammettesse una conveniente espansione dell'universo, rinunciando così alla sua staticità. Ciò perché la luce proveniente dalle stelle e galassie più lontane arriverebbe a noi tanto indebolita (de-energizzata) per effetto Doppler da rendere a noi avvertibile soltanto il contributo delle galassie più vicine. **Come abbiamo detto in precedenza le cose stanno proprio così in quanto Hubble nel 1929 ha dimostrato in maniera incontrovertibile che il nostro universo è in espansione e che le galassie più lontane si allontanano da noi a velocità prossime alla velocità della luce, quindi esiste un** *orizzonte* **oltre il quale non possiamo ricevere informazioni ed il Paradosso di Olbers svanisce**.
Concludendo, osserviamo un fatto piuttosto curioso e cioè: ignorando tutte le critiche mosse al modello statico newtoniano, anche il primo tentativo relativistico, compiuto con successo dallo stesso Einstein nel 1917 sia stato ancora quello di ricercare un modello statico. Il modello statico einsteiniano ha comunque un grande valore storico per essere stato il primo modello

relativistico ed è quindi da annoverare come uno dei primi successi della teoria gravitazionale einsteiniana. Di esso avremo occasione di far cenno in seguito.

14.6 Postulato di Weyl

Come abbiamo accennato alla fine del paragrafo precedente, il primo a descrivere un modello relativistico di universo è stato proprio Einstein nel 1917 col suo **modello statico**. Qualche anno dopo, nel 1922, il matematico e cosmologo russo Alexander Alexandrovic Friedmann fu praticamente il primo ad applicare la Relatività Generale all'universo nella sua globalità, trovando anche modelli non statici che sfuggirono ad Einstein. D'altra parte, nel 1923, il grande matematico tedesco Hermann Weyl, al fine di poter applicare la Relatività Generale all'intero universo, tenuto conto che al secondo membro delle equazioni gravitazionali di Einstein compare il tensore energetico-materiale T^{ik} (che rende conto della distribuzione e del moto della materia) *si domandò* **quale tipo di struttura materiale** *si debba considerare per avere un modello di universo il più realistico possibile.*
Weyl, considerò allora la totalità delle galassie dell'universo reale, come se esse fossero una porzione discreta delle particelle di un **fluido perfetto**, che chiamò **"substratum"**, che pervade l'intero universo; le altre particelle rappresentano dunque un semplice tessuto connettivo puramente ideale. Si possono allora distinguere due casi di effettiva importanza pratica:

1) **materia fluida**: il **"substratum"**, essendo un **fluido perfetto**, è caratterizzato, come abbiamo visto nel paragrafo 11.17, dal tensore:

$$T^{ik} = \left(\rho + \frac{p}{c^2}\right) U^i U^k - p\, g^{ik} = (\varepsilon + p)\, u^i u^k - p\, g^{ik}, \tag{14.5}$$

dove ρ rappresenta la densità propria di massa propria, p la pressione del fluido, le $U^i = \dfrac{dx^i}{d\tau}$ sono le componenti covarianti della quadrivelocità della generica particella, mentre le $u^i = \dfrac{dx^i}{ds}$ sono le corrispondenti 4-velocità adimensionali, ovvero le componenti del 4-vettore unitario, versore di U^i, essendo $u^i = U^i/c$ ed $\varepsilon = c^2\rho$ la densità propria di energia.
Nel caso particolare che la pressione sia trascurabile, possiamo porre nella precedente relazione $p = 0$, ottenendo allora lo schema cosiddetto di:

2) **materia disgregata o materia incoerente non interagente o materia pura o polvere**, *cioè un sistema materiale per il quale siano trascurabili, rispetto alle forze di volume, le forze di contatto, cioè in sostanza gli sforzi.*
Lo schema materia disgregata, anch'esso di uso corrente in Cosmologia, è caratterizzato da un tensore energia-impulso del tipo

$$T^{ik} = \rho U^i\, U^k = c^2 \rho u^i u^k = \varepsilon\, u^i u^k. \tag{14.6}$$

Tenuto conto delle considerazioni precedenti, possiamo enunciare il seguente fondamentale:
Postulato di Weyl:
Le linee di universo delle particelle del **"substratum"**, *e quindi delle galassie, formano* **una congruenza normale e geodetica del genere tempo** *divergente da un punto nel passato (a distanza finita o infinita) o convergente in tale punto nel futuro, o entrambe le cose.*

Osserviamo che il postulato di Weyl è ovviamente valido sia per **materia fluida** sia per **materia disgregata.**
Come sappiamo, dire che le linee di universo formano una **congruenza** significa che per ogni

Figura 14.2: Illustrazione schematica del Postulato di Weyl.

punto regolare (cioè un punto che non sia una singolarità) dello spaziotempo, passa una ed una sola di tali linee. Osserviamo che ciò è conforme al fatto che in RG il più generale riferimento fisico ammissibile è costituito proprio da una congruenza di linee del genere tempo, come abbiamo visto nel paragrafo 11.7.
Dire poi che si tratta di una **congruenza normale** significa che esiste una famiglia ∞^1 di ipersuperficie (del genere spazio) tali che le linee di universo (del genere tempo) della congruenza siano tutte ortogonali a tali ipersuperficie, o, in altri termini che le linee della congruenza sono le *traiettorie ortogonali* della famiglia di ipersuperficie (Fig.14.2).
Infine dire che **la congruenza è geodetica** significa, ovviamente, che essa è formata da linee che sono tutte geodetiche di $\mathcal{V}^4$.

14.7 Definizione generale di modello cosmologico relativistico

La **Cosmologia relativistica** si basa sull'assunzione che, a larga scala, nel nostro universo siano valide le tre ipotesi fondamentali:

1. *Il Principio Cosmologico*
2. *Il Postulato di Weyl*
3. *La Relatività Generale*

Con più precisione, possiamo porre la seguente:
Definizione:
Dicesi **modello cosmologico relativistico** una varietà pseudo-riemanniana $\mathcal{V}^4$, che sia:

1) munita di un opportuno tensore metrico g_{ik} basato sulla validità del PC, cioè sul fatto che, a larga scala, l'universo è omogeneo ed isotropo;

2) munita di un tensore energia-impulso del tipo (14.5) o (14.6) caratterizzante il tipo di materia che riempie l'universo: **"materia fluida"** oppure **"materia disgregata"**;

3) In $\mathcal{V}^4$ siano soddisfatte le equazioni gravitazionali di Einstein:

$$R^{ik} - \frac{1}{2}Rg^{ik} + \Lambda g^{ik} = \kappa T^{ik}, \qquad (i,k = 0,1,2,3). \tag{14.7}$$

Qualunque sia la struttura materiale assunta, la materia costituisce il riferimento universale per eccellenza, dato che nessun riferimento più realistico si può concepire. Ad ogni elemento di materia si può idealmente pensare associato un osservatore che chiameremo **osservatore fondamentale.** Ogni osservatore fondamentale si penserà munito di un orologio standard misurante lo scorrere del suo tempo proprio τ. In ciascuno di questi orologi è naturalmente disponibile l'istante assunto come iniziale ($\tau = 0$). Le linee orarie delle particelle verranno chiamate **linee fondamentali.**
Un modello cosmologico si dirà possedere un **tempo universale**, o **tempo cosmico** t, se il tempo τ degli osservatori fondamentali può essere assunto come una coordinata di $\mathcal{V}^4$ (*tempo pantopico* legato all'insieme delle ∞^3 particelle fondamentali, analogo al *tempo pantopico* che in RR viene associato ad ogni riferimento galileiano). Ciò equivale a dire, come sappiamo, che detto $\mathbf{u}(x)$ il campo dei vettori unitari, del genere tempo, tangenti alle linee fondamentali, la forma differenziale $\frac{1}{c}u_i dx^i$ - intervallo elementare di *tempo relativo standard* riferito al riferimento fondamentale individuato dal campo $\mathbf{u}(x)$ - sia un differenziale esatto. Affinché questo succeda si richiede l'annullarsi in $\mathcal{V}^4$ del cosiddetto *tensore vortice spazio-temporale* $\Omega_{ik} = \partial_i u_k - \partial_k u_i$, annullarsi che, a sua volta, implica la duplice circostanza:

1) le linee orarie fondamentali sono geodetiche;

2) esse costituiscono una congruenza normale.

In un modello cosmologico dotato di tempo universale si ha quindi la proprietà notevole che la congruenza delle linee orarie fondamentali ammette ∞^1 **sezioni normali** 3-dimensionali $\mathcal{W}^3$. Disponendo opportunamente dell'origine temporale dei singoli osservatori fonfamentali si può far sì che su ogni $\mathcal{W}^3$ il tempo universale t assuma dappertutto uno stesso valore: $t = cost$. Diremo allora che il **tempo cosmico** t **è normalizzato**, e si potrà dire che ogni sezione normale $t = cost$. rappresenta lo **spazio fisico** ad un istante cosmico assegnato.
L'adozione di ct (t tempo cosmico normalizzato) come coordinata x^0 e di coordinate spaziali x^α **comobili**, *cioè costanti lungo ogni linea oraria fondamentale*, riduce notoriamente (per la (11.81)) la metrica di $\mathcal{V}^4$ alla forma seguente

$$\mathbf{ds}^2 = c^2 dt^2 - g_{\alpha\beta} dx^\alpha dx^\beta = d(x^0)^2 - g_{\alpha\beta} dx^\alpha dx^\beta, \tag{14.8}$$

dove le $g_{\alpha\beta}$ sono generalmente funzioni delle x^α e di x^0.
La forma quadratica

$$d\sigma^2 = g_{\alpha\beta} dx^\alpha dx^\beta, \tag{14.9}$$

certamente definita positiva, rappresenta la metrica spaziale, cioè la metrica dello spazio fisico $\mathcal{W}^3$. L'eventuale indipendenza delle $g_{\alpha\beta}$ da x^0 ($\partial_0 g_{\alpha\beta} = 0$) caratterizza i cosiddetti **"modelli statici"**; la dipendenza delle $g_{\alpha\beta}$ da x^0 è propria invece dei **"modelli dinamici"**.
Si noti come l'esistenza di un tempo pantopico fondamentale non si concili con qualunque struttura materiale, ma imponga l'adozione di "materia fluida" o di "materia disgregata". Solo in tal modo infatti le traiettorie fondamentali possono risultare geodetiche di $\mathcal{V}^4$.
Si può dimostrare rigorosamente che nel PC, congiunto all'ipotesi di "materia fluida" o di "materia disgregata", è implicita l'ipotesi dell'esistenza di un tempo universale.[10]

[10]Più precisamente, secondo il linguaggio di Cattaneo: **l'isotropia** implica l'annullamento del **tensore vortice spaziale** $\widetilde{\Omega}_{ik} \stackrel{\text{def}}{=} \mathcal{P}_{\Sigma\Sigma}(\Omega_{ik}) = 0$, dove $\Omega_{ik} \equiv \partial_i u_j - \partial_j u_i$, e quindi il carattere **normale** della congruenza fondamentale. Dall'ipotesi di **"materia fluida" o di "materia disgregata"** discende poi la natura **geodetica** della congruenza medesima. Ne segue l'annullarsi del **tensore vortice spazio-temporale** $\Omega_{ik} \stackrel{\text{def}}{=} \widetilde{\Omega}_{ik} + C_i u_j - C_j u_i$ ($\Omega_{ik} = 0$), (dove $C_i \stackrel{\text{def}}{=} \frac{Du_i}{ds} = u^r \nabla_r u_i$, è il vettore di curvatura della linea $x^0 = var$. (che essendo una geodetica richiede che sia $C_i = 0$) e quindi l'esistenza di un **tempo universale**.

14.8 Metrica di Friedmann-Robertson-Walker (FRW)

Ricerchiamo ora il più generale modello di universo soddisfacente al **Principio Cosmologico**: *cioè un universo omogeneo ed isotropo*.
Questo principio, come già abbiamo visto, implica l'esistenza di un tempo cosmico, ciò che impone alla metrica di $\mathcal{V}^4$, in coordinate comobili e tempo cosmico normalizzato t, la forma (14.8), cioè:

$$\mathbf{ds}^2 = c^2 dt^2 - g_{\alpha\beta} dx^\alpha dx^\beta. \tag{14.10}$$

Isotropia ed omogeneità del modello implicano in particolare isotropia ed omogeneità delle sezioni normali $\mathcal{W}^3$ in quanto varietà 3-dimensionali. Ne segue, per un noto **teorema di Schur**, che queste sono necessariamente varietà a curvatura costante.[11]
Naturalmente la curvatura medesima, costante su ogni $\mathcal{W}^3$, potrà variare nel tempo.
Si ricordi che, almeno localmente, i punti di una varietà differenziabile $\mathcal{W}^3$ sono in corrispondenza biunivoca con i punti di uno spazio euclideo E^3; se i punti di quest'ultimo sono individuati da tre coordinate, per es. cartesiane x, y, z, queste stesse coordinate vengono attribuite ai corrispondenti punti di $\mathcal{W}^3$: insomma la corrispondenza trascina le coordinate di E^3 su $\mathcal{W}^3$.
In virtù di quanto affermato più sopra, nella teoria delle varietà a curvatura costante, si dimostra che la metrica di $\mathcal{W}^3$, in coordinate cartesiane, assume la forma:

$$d\sigma^2 = \frac{ds^2 \text{ euclideo}}{\left(1 + \frac{1}{4}Kr_1^2\right)^2} = \frac{dx^2 + dy^2 + dz^2}{\left(1 + \frac{1}{4}Kr_1^2\right)^2} \tag{14.11}$$

dove $r_1 = x^2 + y^2 + z^2$ è la distanza dell'origine di E^3 da un suo punto generico[12], e K è la curvatura costante della varietà $\mathcal{W}^3$ che può essere positiva, nulla o negativa. Nel caso che $K = 0$ il $d\sigma^2$ si riduce a quello euclideo, cioè la varietà $\mathcal{W}^3$ si identifica con uno spazio euclideo.
Pensiamo ora di riferire simultaneamente E^3 e $\mathcal{W}^3$ ad un sistema di coordinate r_1, θ, φ che siano polari in E^3, cioè tali che $x = r_1 \sin\theta \cos\varphi, y = r_1 \sin\theta \sin\varphi, z = r_1 \cos\theta$, e che, con leggero abuso di linguaggio, conservino lo stesso nome in $\mathcal{W}^3$. La metrica di $\mathcal{W}^3$ assume la forma:

$$d\sigma^2 = \frac{dr_1^2 + r_1^2(d\theta^2 + \sin^2\theta d\varphi^2)}{\left(1 + \frac{1}{4}Kr_1^2\right)^2}, \tag{14.12}$$

Nell'ipotesi $K \neq 0$, si può porre

$$R = \frac{k}{K}, \qquad k = \frac{K}{|K|}, \tag{14.13}$$

e scrivere la metrica (14.12) in forma appena diversa:

$$d\sigma^2 = \frac{dr_1^2 + r_1^2(d\theta^2 + \sin^2\theta d\varphi^2)}{\left(1 + \frac{k}{4}\frac{r^2}{R^2}\right)^2}, \qquad (k = -1, +1). \tag{14.14}$$

R è allora una costante positiva $(R > 0)$ e k una costante che può assumere soltanto i valori -1 oppure $+1$. Alla costante k si dà il nome di **indice di curvatura**. La curvatura è data in valore

[11] Ricordiamo allo studioso che nella geometria differenziale delle varietà si pone la seguente definizione: *una varietà $\mathcal{V}^n$ si dice localmente isotropa se in ciascun punto P della varietà la curvatura riemanniana K_π secondo la 2-giacitura π è indipendente da tale giacitura. La varietà è omogenea se tale curvatura è indipendente dal posto.*
A tal proposito, si dimostra il seguente:
Teorema di Schur: *L'isotropia locale di una varietà $\mathcal{V}^n$ $(n > 2)$ implica anche la sua omogeneità, si tratta perciò di una varietà a curvatura costante.*

[12] Abbiamo usato r_1 e non r, come sarebbe più naturale, per indicare la quantità $x^2 + y^2 + z^2$, per il semplice motivo che riserviamo la lettera r ad un'altra variabile che introdurremo un po' più avanti.

e segno da $K = \dfrac{k}{R^2}$; se K è positiva, la quantità R si può interpretare come **raggio di curvatura** della varietà $\mathcal{W}^3$. Il nome può essere mantenuto, per ragioni dimensionali, anche nel caso di curvatura negativa, senza peraltro attribuirgli alcun significato geometrico.
In qualche caso, può essere utile introdurre in luogo della coordinata r_1 una coordinata radiale $\bar{r}$ adimensionale così definita:

$$\bar{r} = \frac{r_1}{R}. \tag{14.15}$$

la metrica (14.14) si scriverà allora nella forma:

$$d\sigma^2 = R^2 \frac{d\bar{r}^2 + \bar{r}^2(d\theta^2 + \sin^2\theta d\varphi^2)}{\left(1 + \frac{k}{4}\bar{r}^2\right)^2}, \qquad (k = -1, 0, +1), \tag{14.16}$$

che diremo *metrica di una $\mathcal{W}^3$ a curvatura costante $K = \dfrac{k}{R^2}$ in coordinate polari adimensionali.* Le forme (14.14) e (14.16) sono state desunte nell'ipotesi $K \neq 0$, quindi per $k = -1$ ovvero $k = +1$. La (14.16) conserva peraltro significato anche per $k = 0$, con R lunghezza costante assegnata, rappresentando in tal caso una ordinaria metrica euclidea in coordinate polari. Il raggio vettore r_1 essendo dato da $r_1 = \bar{r}R$.
La (14.16) si presta pertanto a rappresentare la metrica di una varietà 3-dimesionale a curvatura costante del tutto generale, positiva, negativa, o nulla, a seconda che l'indice di curvatura k assuma il valore $+1, -1, 0$.
Se ora, come abbiamo detto più sopra, teniamo conto del fatto che la curvatura, costante su ogni $\mathcal{W}^3$, potrà variare nel tempo, dobbiamo nella (14.16), considerare la curvatura e quindi R come funzioni del tempo, pertanto la metrica spazio-temporale (14.10) assumerà la forma universalmente nota come **metrica di Robertson-Walker**:

$$\mathbf{ds}^2 = c^2dt^2 - R^2(t)\frac{d\bar{r}^2 + \bar{r}^2(d\theta^2 + \sin^2\theta d\varphi^2)}{\left(1 + \frac{k}{4}\bar{r}^2\right)^2}, \qquad (R(t) > 0, \;\; k = +1, -1, 0), \tag{14.17}$$

La (14.17), da noi ricavata con ragionamenti largamente euristici, è la metrica più generale compatibile col **Principio Cosmologico**, cioè con un universo isotropo ed omogeneo, con l'intesa che $\bar{r}, \theta, \varphi$ siano coordinate comobili. Come già abbiamo detto nel paragrafo 14.6, essa è stata trovata, anche se in maniera poco rigorosa, per la prima volta nel 1922 dal russo A. A. Friedmann che è stato anche il primo ad utilizzarla con successo nello studio delle equazioni gravitazionali di Einstein trovando così quei modelli di universo che ora portano il suo nome (**modelli di Friedmann**).
Tale metrica è stata ricavata in maniera del tutto rigorosa più tardi, tra il 1935 ed il 1936, dall'americano Howard Percy Robertson e dall'inglese Arthur Geoffrey Walker, indipendentemente l'uno dall'altro e costituisce una pietra miliare nello sviluppo teorico della moderna Cosmologia. Noi la chiameremo, com'è giusto e doveroso seguendo Rindler, **metrica di Friedmann-Robertson-Walker (FRW)**.[13]
Un'altra forma della metrica (14.17), molto comoda perché semplifica un po' i calcoli che faremo nei prossimi paragrafi, si ottiene assumendo come nuova variabile radiale la quantità

$$r = \frac{\bar{r}}{1 + \frac{k}{4}\bar{r}^2}. \tag{14.18}$$

[13]Per una dimostrazione rigorosa della **metrica di Friedmann-Robertson-Walker**, lo studioso può consultare i seguenti libri: Jamal Nazrul Islam, *An Introduction to Mathematical Cosmology*, second edition, Chapter 3, pages 37-59, Cambridge University Press, 2002; Steven Weinberg, *Gravitation and Cosmology*, Chap. 14, pag. 412-415, John Wiley & Sons, New York 1972.

Si ha allora:

$$\mathbf{ds}^2 = c^2dt^2 - R^2(t)\Big[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\Big] \qquad (k = +1, -1, 0). \tag{14.19}$$

Si noti che anche r, θ, φ sono coordinate comobili, con r coordinata radiale adimensionale.

14.9 Alcune proprietà cinematiche della metrica FRW

Esaminiamo ora la metrica FRW, (14.19), da un punto di vista cinematico. A questo scopo consideriamo due particelle fondamentali infinitamente vicine la cui distanza $d\sigma$ è fornita dalla formula

$$d\sigma^2 = R^2(t)\Big[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2)\Big] \qquad (k = +1, -1, 0). \tag{14.20}$$

e valutiamo la variazione che subisce nel tempo questa distanza.
Derivata membro a membro la (14.20), rispetto a t, tenendo costanti sia le coordinate comobili r, θ, φ che i loro differenziali, avremo:

$$2d\sigma d\dot{\sigma} = 2R\dot{R}\Big[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2)\Big]$$

se dividiamo membro a membro quest'ultima per la (14.20) medesima, si ricava

$$\frac{d\dot{\sigma}}{d\sigma} = \frac{\dot{R}}{R} = H(t). \tag{14.21}$$

Il rapporto tra $d\dot{\sigma}$ e $d\sigma$ è quindi una funzione del solo tempo, che non dipende dalla particolare coppia di particelle fondamentali presa in considerazione. Ciò significa che lo spazio fisico, o meglio la materia che lo riempie, subisce una dilatazione o una contrazione **spazialmente isotropa ed uniforme**, la cui rapidità $H(t) = \dot{R}/R$ può variare da un istante all'altro. In termini equivalenti si può dire che la velocità di allontanamento di due particelle fondamentali infinitamente vicine è, istante per istante, proporzionale alla distanza fra le particelle medesime

$$d\dot{\sigma} = H(t)d\sigma. \tag{14.22}$$

Naturalmente si tratterà di una vera dilatazione soltanto se $H > 0$. Supposta verificata, per t corrispondente all'era attuale, quest'ultima condizione, il nostro modello realizza un universo in espansione rispettante la legge di Hubble. La **costante di Hubble** (costante nello spazio ma non generalmente nel tempo) è data da $H(t) = \dot{R}/R$.
Come si vede le ipotesi ammesse sono ipotesi molto forti che già bastano ad individuare quasi completamente il modello. In esso due soli elementi rimangono indeterminati: L'indice di curvatura k suscettibile di soli tre valori $(+1, 0, -1)$ e la funzione positiva $R(t)$. A quest'ultima, che soltanto in caso di curvatura positiva $(k = +1)$ è interpretabile come raggio di curvatura dell'universo, viene dato il nome più generico di **funzione di espansione o fattore di scala**. Il nome è ben giustificato in base alla (14.20) ed è appropriato a qualsiasi valore di k.
Altro elemento ancora indeterminato nel modello, che nelle considerazioni cinematiche svolte finora non ha avuto motivo di intervenire, è la densità di massa propria ρ della materia che riempie l'universo. Di essa si può soltanto dire, per il momento, che stante la supposta omogeneità del modello, deve essere la stessa per tutte le particelle e quindi ridursi ad una funzione della sola t.
Nel prossimo paragrafo esamineremo quali ulteriori condizioni siano imposte al modello se richiediamo che esso soddisfi alle equazioni gravitazionali di Einstein.

14.10 Condizioni dinamiche. Equazioni di Friedmann

Nella definizione di modello cosmologico sono inclusi, come sappiamo, la scelta della costituzione della materia riempiente l'universo e la condizione di soddisfare alle equazioni gravitazionali. In virtù del **Principio di Weyl** assumeremo per la materia lo schema **fluido perfetto** ovvero **materia disgregata**, essi sono rappresentati, come sappiamo, rispettivamente dai seguenti tensori energia-impulso:

$$T^{ik} = \left(c^2\rho + p\right) u^i u^k - p\, g^{ik} = (\varepsilon + p)\, u^i u^k - p\, g^{ik}, \tag{14.23}$$

$$T^{ik} = c^2 \rho u^i u^k = \varepsilon\, u^i u^k, \tag{14.24}$$

col consueto significato dei simboli. Il secondo tensore è deducibile dal primo nel caso che la pressione del fluido cosmico sia trascurabile $(p = 0)$, perciò per avere la massima generalità, possiamo fare i calcoli con il tensore relativo a materia fluida.
I due schemi materiali sopra adottati sono giustificati, almeno in prima approssimazione, dal fatto che la grande distanza tra le galassie rende assai probabile la prevalenza dell'interazione gravitazionale su ogni altro tipo di interazione. Essi presuppongono inoltre che anche l'eventuale materia intergalattica sia sensibilmente priva di sforzi interni. Imponiamo ora la validità della RG, ossia richiediamo che siano verificate le equazioni gravitazionali (14.7), che, per un motivo che vedremo fra poco, scriviamo nella forma mista:

$$G^i_k + \Lambda g^i_k = \kappa T^i_k, \qquad (i, k = 0, 1, 2, 3), \tag{14.25}$$

dove

$$G^i_k \equiv R^i_k - \frac{1}{2} R g^i_k, \qquad (i, k = 0, 1, 2, 3), \tag{14.26}$$

è il tensore di Einstein-Levi Civita, g^i_k sono i coefficienti della metrica dello spaziotempo, R^i_k ed $R = R^i_i$ rispettivamente il tensore di Ricci e lo scalare di curvatura (o scalare di Ricci). Nella (14.25) abbiamo incluso anche il termine cosmologico Λg^i_k, per avere la più generale equazione di campo.
Il tensore energia-impulso, nella forma (14.23) o (14.24) figurerà al secondo membro delle (14.25) (espresso ovviamente in forma mista) inoltre Λ e κ sono due costanti universali, **costante cosmologica** la prima, **costante di gravitazione einsteiniana** la seconda. Quest'ultima è legata alla **costante di gravitazione newtoniana** G dalla relazione

$$\kappa = \frac{8\pi G}{c^4}, \tag{14.27}$$

dove, ricordiamolo, si ha $G = (6,6726 \pm 0,0008) \times 10^{-8}\,\mathrm{dyn} \times \mathrm{cm}^2 \times \mathrm{g}^{-2}$ e $\kappa = 2,08 \times 10^{-48}\mathrm{cm}^{-1} \times \mathrm{g}^{-1} \times \mathrm{s}^2$, inoltre la costante cosmologica ha un valore estremamente piccolo $|\Lambda| \leq 10^{-55}\mathrm{cm}^{-2}$.
Torneremo in seguito sul significato del termine cosmologico che ha grande importanza concettuale in Cosmologia. L'inclusione di tale termine nelle equazioni di campo non deve considerarsi artificiosa: il fatto che nell'ambito del sistema solare i dati di osservazione si spieghino perfettamente bene senza di esso, deve soltanto intendersi come trascurabilità del termine stesso di fronte agli altri nei problemi di scala planetaria. Questa sua trascurabilità viene meno su scala cosmologica, soprattutto quando si considera l'universo nella sua totalità.
Nel sistema di coordinate comobili $(x^0, x^1, x^2, x^3) \equiv (ct, r, \theta, \varphi)$ impiegato nel paragrafo precedente, è facile calcolare le componenti non nulle della metrica (14.19) di Friedmann- Robertson-

Walker, avremo:

$$\begin{cases} g_{00} = 1, \quad g_{11} = -\dfrac{R^2}{1-kr^2}, \quad g_{22} = -r^2R^2, \quad g_{33} = -r^2R^2\sin^2\theta, \\ \\ g^{00} = 1, \quad g^{11} = -\dfrac{1-kr^2}{R^2}, \quad g^{22} = -\dfrac{1}{r^2R^2}, \quad g^{33} = -\dfrac{1}{r^2R^2\sin^2\theta}, \end{cases} \tag{14.28}$$

Inoltre, essendo la metrica diagonale, il determinante $g = \det(g_{ik})$ dei suoi coefficienti è uguale al prodotto degli elementi della diagonale principale, cioè:

$$g = \det(g_{ik}) = -\frac{r^4R^6\sin^2\theta}{1-kr^2}. \tag{14.29}$$

Per calcolare le componenti del tensore di Einstein G_{ik} è necessario calcolare prima il tensore R_{ik} di Ricci e lo scalare $R = g^{ik}R_{ik} = R^i_i$ di curvatura e per far ciò bisogna calcolare preliminarmente tutti i simboli di Christoffel di 1ª specie e poi quelli di 2ª specie

$$\Gamma^i_{jk} = g^{ih}\Gamma_{jk,h} = g^{ih}\frac{1}{2}(\partial_j g_{kh} + \partial_k g_{jh} - \partial_h g_{jk}), \qquad (i,j,h,k = 0,1,2,3).$$

Questi calcoli, si capisce, non sono difficili ma piuttosto laboriosi. Si può constatare che i soli simboli di Christoffel di 2ª specie che non si annullano sono i seguenti che abbiamo disposto in quattro gruppi a seconda dei valori $0,1,2,3$ dell'indice in alto:

$$\begin{cases} \Gamma^0_{11} = c^{-1}R\dot{R}/(1-kr^2), \quad \Gamma^0_{22} = c^{-1}r^2R\dot{R}, \quad \Gamma^0_{33} = c^{-1}r^2R\dot{R}\sin^2\theta, \\ \\ \Gamma^1_{01} = c^{-1}\dot{R}/R, \ \ \Gamma^1_{11} = kr/(1-kr^2), \ \ \Gamma^1_{22} = -r(1-kr^2), \ \ \Gamma^1_{33} = -r(1-kr^2)\sin^2\theta, \\ \\ \Gamma^2_{02} = c^{-1}\dot{R}/R, \quad \Gamma^2_{12} = 1/r, \quad \Gamma^2_{33} = -\sin\theta\cos\theta, \\ \\ \Gamma^3_{03} = c^{-1}\dot{R}/R, \quad \Gamma^3_{13} = 1/r, \quad \Gamma^3_{23} = \cot\theta. \end{cases} \tag{14.30}$$

Ora è necessario calcolare il tensore di Ricci, e per far ciò è conveniente usare la formula (10.177) del Cap.10:

$$R_{ik} = \partial_l\Gamma^l_{ik} - \partial_{ik}\log\sqrt{|g|} + \partial_m\log\sqrt{|g|}\Gamma^m_{ik} - \Gamma^l_{nk}\Gamma^n_{il}, \tag{14.31}$$

nella quale si deve porre

$$\sqrt{|g|} = \sqrt{-g} = \frac{r^2R^3\sin\theta}{\sqrt{1-kr^2}}. \tag{14.32}$$

Le sole componenti non nulle sono quelle di indici uguali, precisamente si trova:

$$\begin{cases} R_{00} = -3\ddot{R}/R, \\ \\ R_{11} = (R\ddot{R} + 2\dot{R}^2 + 2c^2k)/(1-kr^2), \\ \\ R_{22} = r^2(R\ddot{R} + 2\dot{R}^2 + 2c^2k), \\ \\ R_{33} = r^2\sin^2\theta(R\ddot{R} + 2\dot{R}^2 + 2c^2k). \end{cases} \tag{14.33}$$

Le corrispondenti componenti miste del tensore di Ricci sono invece:

$$\begin{cases} R^0_0 = \dfrac{-3}{c^2}\dfrac{\ddot{R}}{R}, \\ \\ R^1_1 = R^2_2 = R^3_3 = -\dfrac{1}{c^2}\Big(\dfrac{\ddot{R}}{R} + \dfrac{2\dot{R}^2 + 2kc^2}{R^2}\Big). \end{cases} \tag{14.34}$$

Usando le (14.33) o le (14.34) non è difficile calcolare lo scalare di curvatura, si ottiene:

$$R = \frac{-6}{c^2}\Big(\frac{\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{kc^2}{R^2}\Big). \tag{14.35}$$

Sfortunatamente, per lo scalare di curvatura, normalmente si usa la stessa lettera che si usa per la funzione di espansione $R(t)$, ma dal contesto non c'è pericolo di confusione.
Infine, siamo ora in grado di calcolare il tensore di Einstein-Levi Civita e lo facciamo scrivendolo nella forma tensoriale mista che comporta calcoli più semplici, basta sostituire nella (14.26) le (14.34), (14.35), otteniamo:
1°) tutte le componenti G^i_k di indici diversi sono nulle:

$$G^i_k = 0, \quad per \;\; i \neq k, \quad (i, k = 0, 1, 2, 3); \tag{14.36}$$

2°) la componente G^0_0, cioè per $i = k = 0$, è data da:

$$G^0_0 = -\frac{3\dot{R}^2}{c^2R^2} - \frac{3k}{R^2}; \tag{14.37}$$

3°) le componenti G^α_β $(con\ \alpha = \beta = 1, 2, 3)$, sono tutte e tre uguali ed hanno la seguente espressione:

$$G^1_1 = G^2_2 = G^3_3 = -\frac{2\ddot{R}}{c^2R} - \frac{\dot{R}^2}{c^2R^2} - \frac{k}{R^2}. \tag{14.38}$$

Per scrivere le equazioni gravitazionali ci mancano ancora le componenti del tensore energia-impulso che ora scriviamo, per materia fluida, nella forma mista

$$T^i_k = (c^2\rho + p)u^iu_k - pg^i_k, \tag{14.39}$$

dove u^i e u_k sono rispettivamente le componenti controvarianti e covarianti del vettore unitario 4-velocità di una generica particella del fluido, g^i_k il tensore metrico in forma mista che, com'è noto, coincide con il delta di Kronecker: $g^i_k \equiv \delta^i_k$. Nel sistema di coordinate comobili ct, r, θ, φ, utilizzato più volte, le componenti del vettore unitario, sia nella forma controvariante che in quella covariante assumono gli stessi valori valori: $u^i = (1, 0, 0, 0)$, $u_k = (1, 0, 0, 0)$. Perciò avremo:

$$\begin{cases} T^0_0 = c^2\rho = \varepsilon, \\ \\ T^1_1 = T^2_2 = T^3_3 = -p, \end{cases} \tag{14.40}$$

tutte le altre componenti del tensore energia-impulso sono nulle.
Se ora sostituiamo le (14.38) e (14.40) nelle equazioni gravitazionali di Einstein (14.30) che, ricordiamolo, sono in numero di 10, si riconosce che una parte delle equazioni stesse (quelle per cui $i \neq k$) sono soddisfatte identicamente. Sussistono come effettive condizioni soltanto le equazioni di indici uguali $(i = k = 0, 1, 2, 3)$, le ultime tre, cioè quelle aventi gli indici spaziali

$1, 2, 3$ identificandosi in una medesima equazione, la prima, quella di indice 0, essendo invece distinta:

$$\begin{cases} \dfrac{\dot{R}^2}{c^2R^2} + \dfrac{k}{R^2} - \dfrac{\Lambda}{3} = \dfrac{8\pi G}{3c^2}\rho, \\ \\ \dfrac{2\ddot{R}}{c^2R} + \dfrac{\dot{R}^2}{c^2R^2} + \dfrac{k}{R^2} - \Lambda = -\dfrac{8\pi G}{c^4}p. \end{cases} \tag{14.41}$$

Sono queste le celebri equazioni differenziali ricavate, per la prima volta da **Friedmann** tra il 1922 ed il 1924.
Da adesso in poi supporremo che la pressione sia trascurabile, $p = 0$, perciò considereremo soltanto **materia disgregata**, il precedente sistema si presenterà allora nella forma semplificata:

$$\begin{cases} \dfrac{\dot{R}^2}{c^2R^2} + \dfrac{k}{R^2} - \dfrac{\Lambda}{3} = \dfrac{8\pi G}{3c^2}\rho, \\ \\ \dfrac{2\ddot{R}}{c^2R} + \dfrac{\dot{R}^2}{c^2R^2} + \dfrac{k}{R^2} - \Lambda = 0. \end{cases} \tag{14.42}$$

Le (14.42) sono tutte e sole le condizioni che le equazioni gravitazionali impongono agli elementi che tuttora sono rimasti indeterminati nel modello: le due funzioni $R(t)$ e $\rho(t)$ entrambe positive, e l'indice di curvatura costante k, per il quale tre valori soltanto sono a priori possibili: $-1, 0, +1$. Il numero delle equazioni è uguale al numero delle funzioni incognite; è pertanto presumibile che le condizioni medesime non soltanto risultino compatibili ma altresì conducano ad una determinazione univoca del modello, non appena esse siano completate da opportune condizioni iniziali.
ll sistema (14.42) non si presenta facilmente integrabile in termini espliciti con funzioni d'impiego ordinario, si può comunque integrare con l'uso di funzioni ellittiche. Sarà preferibile però, come spesso si fa, un'appropriata discussione qualitativa.
A tal fine è facile determinare in termini espliciti un integrale primo del sistema in esame. Infatti, se moltiplichiamo la 1ª delle (14.42) per R^3 e dell'equazione ottenuta ne deriviamo 1° e 2° membro rispetto a t ed aggiustiamo opportunamente i termini del primo membro in modo da poter raccogliere a fattore comune la quantità $R^2\dot{R}$ otteniamo

$$R^2\dot{R}\left[\frac{2\ddot{R}}{c^2R} + \frac{\dot{R}^2}{c^2R^2} + \frac{k}{R^2} - \Lambda\right] = \frac{1}{c^2}\frac{d}{dt}\left(\frac{8\pi}{3}G\rho R^3\right).$$

Tenendo conto della 2^a delle (14.42) si ricava l'integrale primo annunciato:

$$\frac{8\pi}{3}G\rho R^3 = C, \tag{14.43}$$

C essendo una costante. La sua interpretazione fisica è immediata; esso esprime la conservazione della massa propria $\frac{4\pi}{3}\rho(\epsilon R)^3$ contenuta in una sfera di raggio infinitesimo ϵR, comobile con la materia del modello. Questo risultato era facilmente prevedibile, infatti tenuto conto dell'espressione (14.24) del tensore energia-impulso, dall'equazione gravitazionale (14.25) discende, come già vedemmo nel paragrafo 11.17, che $\nabla_i T^i_k = 0$ che esprime le leggi di conservazione della materia-energia e della quantità di moto. La legge di conservazione della massa si ottiene precisamente per $k = 0$, cioè si traduce nell'equazione $\nabla_i(\rho u^i) = 0$ che esprime appunto la conservazione della massa propria contenuta in un volume comobile infinitesimo.
Dell'integrale primo (14.43) ci si puo' valere per ricavare una equazione nella sola incognita $R(t)$. Basta allo scopo eliminare $\rho(t)$ tra la (14.43) e la 1ª delle (14.42), si ottiene così:

$$\dot{R}^2 = \frac{C}{R} + \frac{\Lambda c^2R^2}{3} - kc^2, \qquad (k = -1, 0, +1) \tag{14.44}$$

che è la celebre **equazione differenziale di Friedmann**.
Come si è visto, le equazioni originarie (14.42) implicano la (14.43) e la (14.44) come conseguenze necessarie. Viceversa dalle equazioni (14.43), (14.44) segue immediatamente la 1ª delle (14.42); da questa, derivando rispetto a t si ottiene anche la 2ª delle (14.42), purché sia $\dot{R} \neq 0$. Pertanto la discussione del sistema (14.42) può generalmente essere sostituita dalla discussione dell'equazione di Friedmann (14.44) associata all'integrale primo (14.43), cioè del sistema:

$$\begin{cases} \dot{R}^2 = \dfrac{C}{R} + \dfrac{\Lambda c^2 R^2}{3} - kc^2, \qquad (k = -1, 0, +1) \\ \\ \dfrac{8\pi}{3} G\rho R^3 = C. \end{cases} \tag{14.45}$$

Soltanto nei casi statici $(\dot{R} \equiv 0)$ o negli istanti di stazionarietà $(\dot{R} = 0)$, venendo meno la completa equivalenza dei due sistemi, si dovrà ricorrere anche al sistema originario (14.42).
Ai modelli di universo ammettenti, in coordinate comobili ct, r, θ, φ la metrica (14.17) di **Friedmann- Robertson-Walker (FRW)** e soddisfacenti alle equazioni dinamiche suddette si suol dare, a buon diritto, il nome di **modelli di Friedmann**: li chiameremo anche, più brevemente, **F-modelli**.

14.11 F-modelli statici: Modello di Einstein

Nonostante oggi sappiamo con certezza, dopo le osservazioni di Edwin Hubble, che il nostro universo è in espansione, in questo paragrafo ci soffermiamo brevemente su una classe di modelli di Friedmann statici non fosse altro per il fatto che Einstein fu il primo, nel 1917, a considerare un tale modello.
Un modello statico deve necessariamente soddisfare alla condizione $(\dot{R} \equiv 0)$. In questa ipotesi, occorre, come abbiamo detto nel precedente paragrafo, ricorrere al sistema originario di Friedmann, (14.42), il quale fornisce per la densità ρ e le costanti in gioco R, Λ, k, a priori non conosciute, le semplici relazioni:

$$\Lambda = \frac{4\pi G\rho}{c^2}, \qquad \Lambda = \frac{k}{R^2}. \tag{14.46}$$

Per la costante k sono a priori possibili, come sappiamo, i valori $-1, 0, +1$.
La prima delle (14.46) implica la costanza di ρ, ciò che d'altronde per un modello statico è necessario a priori. Fissato per ρ il valore oggi ritenuto abbastanza attendibile in base a stime varie, $\rho \approx 10^{-30}\text{g} \times \text{cm}^3$ se ne desume per Λ il valore positivo $\Lambda \approx 10^{-57}\text{cm}^{-2}$. Dalla seconda delle (14.46) si trae allora, senza incertezza $k = +1$, $R = \dfrac{1}{\Lambda} \approx 10^{28}\text{cm}$. Il modello è dunque caratterizzato dai seguenti valori dei parametri:

$$\rho \approx 10^{-30}\text{g} \times \text{cm}^3, \qquad \Lambda \approx 10^{-57}\text{cm}^{-2}, \qquad R \approx 10^{28}\text{cm}. \tag{14.47}$$

Questo modello, corrispondente ad una varietà a curvatura uniforme e costante nel tempo, non è altro che **il modello statico scoperto da Einstein** e qui ottenuto in un contesto più generale. *Si noti come sia stata essenziale, per l'ottenimento di un universo statico, la presenza di un pur debole termine cosmologico con $\Lambda > 0$. Questo ha, per così dire, un* **effetto repulsivo** *capace di neutralizzare l'azione gravitazionale attrattiva.*
Il solo altro modo per soddisfare coerentemente alle equazioni (14.46) è di porre

$$k = 0, \quad \Lambda = 0, \quad \rho = 0. \tag{14.48}$$

Il modello diventa allora una varietà piatta minkowskiana riempita di materia con massa infinitesima. Questa infinitesimalità rende nulle le azioni gravitazionali da essa generate (in accordo con il carattere minkowskiano di $\mathcal{V}^4$); la massa riempiente l'universo si riduce a semplice *"materia di prova"*.

14.12 Osservazioni sull'equazione di Friedmann. Soluzioni normalizzate

Lasciati da parte i modelli statici perché non conformi all'evidenza osservativa, possiamo ormai ritenere, almeno generalmente, $\dot{R} \neq 0$ e quindi riferirci al sistema (14.46) che ora riscriviamo nella forma più concisa

$$\begin{cases} \dot{R}^2 = F(R), \\ \\ \dfrac{8\pi}{3} G\rho R^3 = C, \end{cases} \tag{14.49}$$

dove si è posto

$$F(R) \equiv \frac{C}{R} + \frac{\Lambda c^2 R^2}{3} - kc^2, \quad (k = -1, 0, +1). \tag{14.50}$$

Su queste equazioni si possono fare le seguenti osservazioni preliminari.

a) La $(14.49)_2$ mostra che ferma lasciando ρ e cambiando $R(t)$ in $-R(t)$, C si muta in $-C$. Ne consegue che di fronte allo stesso cambiamento la funzione $F(R)$, e con essa l'equazione di Friedmann $(14.49)_1$, rimangono invariate. Pertanto ad ogni soluzione $R(t)$ del sistema (14.49), corrispondente ad un valore assegnato della costante d'integrazione C, fa riscontro la soluzione $-R(t)$ dello stesso sistema corrispondente al valore opposto di C. Se d'altra parte si osserva che nella metrica FRW, (14.19), $R(t)$ compare al quadrato, nulla vieta di identificare le due predette soluzioni $(R(t), C)$, $(-R(t), -C)$ o meglio di limitarsi a considerare soluzioni tali che

$$\rho \geq 0, \quad R \geq 0, \quad C \geq 0. \tag{14.51}$$

Sono queste d'altronde, le sole soluzioni di effettivo interesse fisico.

b) L'equazione di Friedmann rimane invariata sia cambiando t in $t + cost.$, sia cambiando t in $-t$.

Figura 14.3: Rappresentazione grafica dell'invarianza dell'equazione di Friedmann per traslazione o inversione temporale.

Perciò, se pensiamo di rappresentare su di un piano cartesiano la funzione $R(t)$, riportando sull'asse delle ascisse il tempo t e sull'asse delle ordinate i valori di $R(t)$ ogni soluzione dell'equazione $(14.49)_1$ si può comunque traslare nella direzione dell'asse t ovvero ribaltare attorno

ad una qualunque parallela all'asse R, dando ancora luogo ad una soluzione (Fig.14.3). Potremo allora unificare tutte queste soluzioni assumendo come funzione rappresentativa di una stessa classe quella soddisfacente alle seguenti condizioni iniziali

$$R(0) = 0, \quad \dot{R}(0) > 0. \tag{14.52}$$

Una soluzione soddisfacente sia alle condizioni (14.51) sia alle (14.52) si dirà una **soluzione normalizzata**.
Passiamo ora alla discussione effettiva del sistema di equazioni (14.49), discussione che divideremo in tre parti:

1) modelli vuoti ($\rho = 0, \quad C = 0$);

2) modelli con materia senza azione cosmologica ($\rho > 0, \quad \Lambda = 0$);

3) modelli generali con azione cosmologica ($\rho > 0, \quad C > 0, \quad \Lambda \neq 0$).

14.13 F-modelli vuoti

Tenuto conto della loro semplicità, esaminiamo per primi i modelli vuoti. Sebbene essi costituiscano una classe di modelli evidentemente poco realistici, essi non sono completamente privi di interesse fisico non fosse altro per il fatto che in questa classe cadono due famosi fra i primi modelli studiati: **l'universo di De Sitter** (1917) ed il **modello di Milne** (1932).
Supposto $\rho = 0$, la $(14.49)_2$ dà $C = 0$ e l'equazione di Friedmann $(14.49)_1$ diviene

$$\dot{R}^2 = \frac{\Lambda c^2 R^2}{3} - kc^2, \qquad (k = -1, 0, +1). \tag{14.53}$$

Di qui, separando le variabili t ed R ed integrando si ottiene

$$t = \frac{1}{c} \int \frac{dR}{\sqrt{\frac{\Lambda}{3} R^2 - k}}. \tag{14.54}$$

Le soluzioni normalizzate possono essere classificate in sei tipi distinti:

1) $\Lambda = 0, \quad k = 0$;
La (14.53) dà immediatamente $R = cost.arb. > 0$. *Si tratta di un modello statico piatto* che coincide col modello statico considerato alla fine del paragrafo 14.11.

2) $\Lambda = 0, \quad k = -1$;
La (14.54) dà subito $R = ct$. La metrica FRW (14.19) si riduce in tal caso alla cosiddetta metrica di Milne:

$$\mathbf{ds}^2 = c^2 dt^2 - R^2(t) \Big[\frac{dr^2}{1 + r^2} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \Big], \qquad (k = -1). \tag{14.55}$$

Lo spaziotempo è lo spazio piatto di Minkowski e siamo perciò nell'ambito della RR. Esso (lo spaziotempo) ospita materia di prova mobile secondo l'eplosione di Milne (analoga ad un big bang). Questo modello formulato, intorno al 1930, dal celebre astronomo britannico Edward Arthur Milne, è molto interessante ma poco realistico. Oggi il **modello di Milne** ha soltanto un'importanza storica.

3) $\Lambda > 0, \quad k = 0;$
Dalla (14.54) si deduce

$$R = R_0 e^{c\sqrt{\Lambda/3}\, t}.$$

Lo spazio 3-dimensionale $\mathcal{W}^3$ è **piatto ed esponenzialmente espansivo**. Lo spaziotempo tuttavia non è piatto e si identifica con una soluzione delle equazioni gravitazionali, senza secondo membro ma con termine cosmologico, già data dall'astronomo olandese Willem de Sitter nel 1917. Il modello in questione va perciò sotto il nome di **Universo di de Sitter**. A questo punto, se poniamo $x = aR$, con $a = \sqrt{\Lambda/3}$, si avrà $dx = adR$, e la (14.54) può essere riscritta nella forma più conveniente:

$$t = \frac{1}{ac} \int \frac{dx}{\sqrt{x^2 - k}}, \qquad \left(a = \sqrt{\frac{\Lambda}{3}} > 0\right). \tag{14.56}$$

Possiamo ora discutere i tre rimanenti casi di modelli vuoti.

4) $\Lambda > 0, \quad k = 1;$
L'integrale che figura nella (14.56), per $k = 1$ diventa

$$\int \frac{dx}{\sqrt{x^2 - 1}},$$

si tratta dunque di un integrale noto.[14] Tenuto conto dell'espressione del 1° integrale dato nella nota a piè di pagina, la (14.56), risolta rispetto ad R, ci fornisce la soluzione:

$$R = \frac{\cosh\left(c\sqrt{\Lambda/3}\; t\right)}{\sqrt{\Lambda/3}}.$$

Si tratta ancora di un modello indefinitamente espansivo nell'intervallo temporale $(0, +\infty)$ che tuttavia ha avuto una fase di contrazione nell'intervallo temporale $(-\infty, 0)$. Lo spazio 3-dimensionale è una varietà a curvatura costante positiva. Lo spaziotempo è ancora uno spaziotempo di de Sitter.

5) $\Lambda > 0, \quad k = -1;$
L'integrale che figura nella (14.56), per $k = -1$ diventa ora

$$\int \frac{dx}{\sqrt{x^2 + 1}}.$$

Quest'ultimo integrale è anch'esso riportato nella nota a piè di pagina, ed in virtù di esso la (14.56), risolta rispetto ad R, ci fornisce la soluzione:

$$R = \frac{\sinh\left(c\sqrt{\Lambda/3}\, t\right)}{\sqrt{\Lambda/3}}.$$

[14] Nei corsi di Analisi Matematica si dimostra che

$$\int \frac{dx}{\sqrt{x^2 - 1}} = \cosh^{-1} x = \log\left(x + \sqrt{x^2 - 1}\right),$$

$$\int \frac{dx}{\sqrt{x^2 + 1}} = \sinh^{-1} x = \log\left(x + \sqrt{x^2 + 1}\right).$$

Si tratta di un modello indefinitamente espansivo in $(0, +\infty)$ analogo al modello di Milne (spazio 3-dimensionale a curvatura costante negativa) immerso in uno spaziotempo di de Sitter.

6) $\Lambda < 0, \quad k = -1$;
Dalla (14.54), con ragionamenti analoghi fatti nei due casi precedenti, otteniamo la soluzione
$$R = \frac{\sinh\left(c\sqrt{-\Lambda/3}\, t\right)}{\sqrt{-\Lambda/3}}.$$
Lo spazio tridimensionale è una varietà a curvatura costante negativa in espansione come nella soluzione precedente. Lo spaziotempo ambiente è però ora del tipo di de Sitter con $\Lambda < 0$.

Possiamo osservare che tutti questi modelli vuoti pur essendo poco realistici hanno in comune la caratteristica desiderabile di essere, quasi tutti, espansivi.

14.14 F-modelli senza azione cosmologica: i tre modelli standard

Passiamo ora ai modelli più realistici con materia disgregata soddisfacenti al Principio Cosmologico, cioè omogenei ed isotropi, trovati da Friedmann e compendiati nella sua celebre equazione:
$$\dot{R}^2 = \frac{C}{R} + \frac{\Lambda c^2 R^2}{3} - kc^2, \qquad (k = -1, 0, +1). \tag{14.57}$$
Supponiamo dunque $\rho > 0$, $\Lambda = 0$, inoltre, per semplificare un pò i calcoli, supponiamo di lavorare con le cosiddette unità relativistiche in cui *la velocità della luce si suppone unitaria*: $c = 1$. In tali ipotesi il sistema di Friedmann (14.49) si specializza nella forma seguente:
$$\begin{cases} \dot{R}^2 = F(R), \\ \\ F(R) \equiv \dfrac{C}{R} - k, \\ \\ \dfrac{8\pi}{3} G\rho R^3 = C, \qquad (C > 0,\ k = -1, 0, +1). \end{cases} \tag{14.58}$$
L'equazione differenziale $(14.58)_1$ nella quale il 2° membro sia espresso dalla $(14.58)_2$ è integrabile per separazione delle variabili e da essa si deduce subito
$$t = \int \frac{dR}{\sqrt{F(R)}} = \int \frac{dR}{\sqrt{\dfrac{C}{R} - k}}. \tag{14.59}$$
Sono possibili tre situazioni tipiche che danno luogo ad altrettante **soluzioni esatte dell'equazione di Friedmann**.

i) $\Lambda = 0,\ k = 0$.
L'equazione di Friedmann dà luogo alla soluzione normalizzata
$$R = \left(\frac{9}{4}C\right)^{1/3} t^{2/3}. \tag{14.60}$$

Figura 14.4: I tre modelli standard senza azione cosmologica.

Si tratta di uno spazio 3-dimensionale piatto, **indefinitamente espansivo** immerso in una varietà spaziotemporale curva. Il modello va sotto il nome di **universo (aperto) di Einstein-de Sitter** (Fig.14.4). [15]

ii) $\Lambda = 0,\ k = 1.$
L'equazione di Friedmann diventa allora:

$$\dot{R}^2 = \frac{C}{R} - 1. \tag{14.61}$$

Per risolverla è opportuno fare il seguente cambio di variabile:

$$u^2 = \frac{R}{C}. \tag{14.62}$$

Derivando rispetto al tempo, 1° e 2° membro della (14.62), avremo $2u\dot{u} = \dot{R}/C$, e sostituendo nella (14.61) avremo

$$\dot{u}^2 = \frac{\dot{R}^2}{4C^2u^2} = \frac{1}{4C^2u^2}\left(\frac{C}{R} - 1\right) = \frac{1}{4C^2u^2}\left(\frac{1}{u^2} - 1\right). \tag{14.63}$$

Dalla precedente equazione, estraendo la radice quadrata del 1° e 2° membro, si ha

$$\dot{u} \equiv \frac{du}{dt} = \pm\frac{1}{2Cu}\sqrt{\frac{1}{u^2} - 1}.$$

Siccome stiamo studiando *soluzioni normalizzate* nel senso spiegato alla fine del paragrafo 14.12, possiamo considerare soltanto la radice positiva ottenendo la seguente equazione differenziale a variabili separate

$$\frac{1}{C}dt = 2\frac{u^2}{\sqrt{1-u^2}}du,$$

ed integrando con la condizione iniziale del big bang ($R = 0$ ovvero $u = 0$), otteniamo

$$\frac{1}{C}\int_0^t dt = 2\int_0^u \frac{u^2}{\sqrt{1-u^2}}du. \tag{14.64}$$

Per calcolare l'integrale che figura al secondo membro effettuiamo un nuovo cambio di variabile ponendo $u = \sin\alpha$, avremo

$$2\int_0^u \frac{u^2}{\sqrt{1-u^2}}du = 2\int_0^\alpha \frac{\sin^2\alpha\cos\alpha}{\sqrt{1-\sin^2\alpha}}d\alpha = 2\int_0^\alpha \sin^2\alpha\, d\alpha.$$

[15] Questo modello fu studiato in collaborazione da Einstein e De Sitter che lo pubblicarono nel 1932.

Applicando all'ultimo integrale della precedente relazione la formula si duplicazione del coseno si ottiene

$$2\int_0^\alpha \sin^2\alpha\, d\alpha = \int_0^\alpha (1-\cos 2\alpha)\, d\alpha = \alpha - \frac{1}{2}\sin\alpha\cos\alpha.$$

Tornando alla vecchia variabile u, la (14.64) ci dà:

$$\frac{1}{C}t = \sin^{-1} u - u\sqrt{1-u^2},$$

e ricordando che $u = \left(\frac{R}{C}\right)^{1/2}$, la soluzione in funzione di R è:

$$t = C\Big[\sin^{-1}\left(\frac{R}{C}\right)^{1/2} - \left(\frac{R}{C} - \frac{R^2}{C^2}\right)^{1/2}\Big]. \tag{14.65}$$

Se ora poniamo

$$\frac{R}{C} = \sin^2\frac{\varphi}{2}, \tag{14.66}$$

con φ parametro arbitrario è possibile esprimere la soluzione (14.65) in forma parametrica. Si ha infatti $\left(\frac{R}{C}\right)^{1/2} = \sin\frac{\varphi}{2}$, che sostituita nella (14.65) ci dà:

$$t = C\Big[\sin^{-1}(\sin\frac{\varphi}{2}) - \left(\sin^2\frac{\varphi}{2} - \sin^4\frac{\varphi}{2}\right)^{1/2}\Big] = C\Big[\frac{\varphi}{2} - \sin\frac{\varphi}{2}\cos\frac{\varphi}{2}\Big] = C\Big[\frac{\varphi}{2} - \frac{1}{2}\sin\varphi\Big]$$

ovvero

$$t = \frac{C}{2}(\varphi - \sin\varphi).$$

Tenuto conto della (14.66) e dell'identità trigonometrica $\sin^2\frac{\varphi}{2} = \frac{1}{2}(1-\cos\varphi)$, otteniamo

$$R = \frac{C}{2}(1-\cos\varphi).$$

Riunendo le due formule precedenti abbiamo la soluzione in forma parametrica:

$$R = \frac{C}{2}(1-\cos\varphi), \qquad t = \frac{C}{2}(\varphi - \sin\varphi). \tag{14.67}$$

Queste equazioni mostrano chiaramente che il grafico della funzione $R(t)$ è una cicloide e il modello dicesi infatti **cicloidale** (Fig.14.4). Lo spazio 3-dimensionale, a curvatura positiva costante, subisce una fase espansiva cui segue una fase di collasso. Se si ammette che a collasso avvenuto ($R = 0$), $\dot{R}$ cambi bruscamente di segno (rimbalzo delle particelle collassate, il cosiddetto **Big Bounce**) ha inizio un nuovo ciclo identico al precedente. Si tratta dunque di un modello periodico pulsante che alterna fasi di espansione a fasi di contrazione. Esso va sotto il nome di **modello (chiuso) di Friedmann-Einstein**.
Se a collasso avvenuto ($R = 0$), $\dot{R}$ non dovesse cambiare bruscamente di segno si avrebbe l'opposto del **Big Bang**, il cosiddetto **Big Crunch**, e dunque la morte dell'universo.

iii) $\Lambda = 0,\ k = -1.$
In quest'ultimo caso procediamo in maniera del tutto analoga a quella del caso precedente, l'equazione di Friedmann diventa:

$$\dot{R}^2 = \frac{C}{R} + 1. \tag{14.68}$$

Per risolverla è opportuno fare lo stesso cambio di variabile fatto nel caso precedente:

$$u^2 = \frac{R}{C}. \tag{14.69}$$

Derivando rispetto al tempo, 1° e 2° membro della (14.69), avremo $2u\dot{u} = \dot{R}/C$, e sostituendo nella (14.68) otteniamo

$$\dot{u}^2 = \frac{\dot{R}^2}{4C^2u^2} = \frac{1}{4C^2u^2}\Big(\frac{C}{R}+1\Big) = \frac{1}{4C^2u^2}\Big(\frac{1}{u^2}+1\Big). \tag{14.70}$$

Dalla precedente equazione, estraendo la radice quadrata del 1° e 2° membro, si ha

$$\dot{u} \equiv \frac{du}{dt} = \pm\frac{1}{2Cu}\sqrt{\frac{1}{u^2}+1}.$$

Considerando la radice positiva otteniamo la seguente equazione differenziale a variabili separate

$$\frac{1}{C}dt = 2\frac{u^2}{\sqrt{1+u^2}}du,$$

ed integrando con la condizione iniziale del big bang ($R = 0$ ovvero $u = 0$), otteniamo

$$\frac{1}{C}\int_0^t dt = 2\int_0^u \frac{u^2}{\sqrt{1+u^2}}du. \tag{14.71}$$

Per calcolare l'integrale che figura al secondo membro effettuiamo un nuovo cambio di variabile ponendo $u = \sinh\,\alpha$, avremo[16]

$$2\int_0^u \frac{u^2}{\sqrt{1+u^2}}du = 2\int_0^\alpha \frac{\sinh^2\alpha\,\cosh\alpha}{\sqrt{1+\sinh^2\alpha}}d\alpha = 2\int_0^\alpha \sinh^2\alpha\,d\alpha.$$

Applicando all'ultimo integrale della precedente relazione la formula si duplicazione del coseno iperbolico si ottiene

$$2\int_0^\alpha \sinh^2\alpha\,d\alpha = \int_0^\alpha (\cosh\,2\alpha - 1)\,d\alpha = \frac{1}{2}\sinh\,2\alpha - \alpha = \sinh\,\alpha\,\cosh\,\alpha - \alpha.$$

Tornando alla vecchia variabile u, essendo $\sinh\,\alpha = u$, $\alpha = \sinh^{-1}u$, e $\cosh\,\alpha = \sqrt{1+\sinh^2\,\alpha} = \sqrt{1+u^2}$, il precedente integrale, che altro non è che (14.71), ci dà:

$$\frac{1}{C}t = u\sqrt{1+u^2} - \sinh^{-1}u,$$

e ricordando che $u = \Big(\frac{R}{C}\Big)^{1/2}$, la soluzione in funzione di R è:

$$t = C\Big[\Big(\frac{R}{C}+\frac{R^2}{C^2}\Big)^{1/2} - \sinh^{-1}\Big(\frac{R}{C}\Big)^{1/2}\Big]. \tag{14.72}$$

Se ora poniamo

$$\frac{R}{C} = \sinh^2\frac{\varphi}{2}, \tag{14.73}$$

[16]Per comodità del lettore ricordiamo qui di seguito alcune proprietà delle funzioni iperboliche utilizzate nei calcoli sopra riportati:
$\cosh^2 x - \sinh^2 x = 1$, $\sinh(2x) = 2\sinh x\cosh x$, $\cosh(2x) = \cosh^2 x + \sinh^2 x$, $\sinh^2 x = \frac{1}{2}[\cosh(2x) - 1]$, $D\sinh x = \cosh x$, $D\cosh x = \sinh x$.

con φ parametro arbitrario è possibile esprimere la soluzione (14.72) in forma parametrica. Si ha infatti $\left(\frac{R}{C}\right)^{1/2} = \sinh\frac{\varphi}{2}$, che sostituita nella (14.72) ci dà:

$$t = C\Big[\Big(\sinh^2\frac{\varphi}{2} + \sinh^4\frac{\varphi}{2}\Big)^{1/2} - \sinh^{-1}\Big(\sinh\frac{\varphi}{2}\Big)\Big],$$

ovvero

$$t = C\Big[\sinh\frac{\varphi}{2}\cosh\frac{\varphi}{2} - \frac{\varphi}{2}\Big],$$

cioè

$$t = \frac{C}{2}(\sinh\,\varphi - \varphi).$$

Tenuto conto della (14.73) e dell'identità $\sinh^2\frac{\varphi}{2} = \frac{1}{2}(\cosh\varphi - 1)$, otteniamo

$$R = \frac{C}{2}(\cosh\,\varphi - 1).$$

Riunendo le due formule precedenti abbiamo la soluzione in forma parametrica:

$$R = \frac{C}{2}(\cosh\,\varphi - 1), \qquad t = \frac{C}{2}(\sinh\,\varphi - \varphi). \tag{14.74}$$

Lo spazio 3-dimensionale è una varietà a curvatura costante negativa, indefinitamente espansiva, con velocità $\dot{R}$ di espansione decrescente verso il valore asintotico c come si vede facilmente dall'equazione di Friedmann (14.57) per $R \to +\infty$, ove si tenga conto che $\Lambda = 0$ e $k = -1$.

L'andamento della funzione di espansione in vicinanza di $R = 0$ (cioè in vicinanza del **Big Bang**) (Fig.14.4) è lo stesso nei tre modelli i), ii), iii), perché quando $R \to 0$ nell'equazione di Friedmann (14.57) il termine kc^2 diviene, qualunque sia k, trascurabile rispetto a $C/R \to +\infty$.
I tre modelli, senza azione cosmologica ($\Lambda = 0$), studiati in questo paragrafo rientrano nella cosiddetta **Teoria del Big Bang** (lo stato singolare $R = 0$) e sono noti come **i tre Modelli Standard della Cosmologia**: è ad essi che gli studiosi rivolgono di più la loro attenzione.

14.15 Alcune considerazioni sulla costante cosmologica

Prima di trattare dei modelli di Friedmann con azione cosmologica, riteniamo opportuno fare alcune necessarie considerazioni sulla costante cosmologica.
Come sappiamo la teoria della Relatività Generale formulata da Einstein fu pubblicata nei primi mesi del 1916 e Karl Schwarzschild, alcuni mesi dopo, fornì la prima soluzione esatta delle equazioni di campo

$$R^{ik} - \frac{1}{2}Rg^{ik} = \frac{8\pi G}{c^4}T^{ik}, \qquad (i,k = 0,1,2,3). \tag{14.75}$$

applicandole ad un corpo a simmetria sferica come il nostro Sole. Einstein stesso si impegnò presto nel compito di applicare la RG all'universo nella sua totalità. A quel tempo si pensava che, salvo moti locali minori, l'Universo fosse statico, ed Einstein cercava naturalmente una soluzione statica delle sue equazioni di campo (14.75). Non ne trovò nessuna. Sembrava che la Relatività Generale non si applicasse all'Universo nella sua totalità. Einstein perciò modificò le originali equazioni di campo (14.75) aggiungendovi un certo termine, conosciuto come termine lambda, dalla lettera greca Λ che usò per denotarlo e le equazioni di campo diventarono:

$$R^{ik} - \frac{1}{2}Rg^{ik} + \Lambda g^{ik} = \frac{8\pi G}{c^4}T^{ik}. \qquad (i,k = 0,1,2,3). \tag{14.76}$$

La modifica doveva essere tale da non alterare l'accordo della teoria originale con i fatti astronomici del sistema solare. Questa condizione era soddisfatta dal termine Λ che è molto piccolo (il suo modulo è attualmente valutato in $|\Lambda| \leq 10^{-55}\,\mathrm{cm}^{-2}$). Sicché Λ ha un effetto trascurabile sul campo gravitazionale per piccole distanze (quali sono quelle del sistema solare, o distanze nell'ambito della Via Lattea). Tuttavia, Λ ha un'influenza notevole per distanze grandi dal punto di vista cosmologico. **L'equazione di campo modificata** (14.76) fornì ad Einstein quello che cercava, un modello di universo statico, come abbiamo visto in dettaglio nel paragrafo 14.11. Nello stesso anno 1917 de Sitter prese in considerazione le equazioni del campo modificate e ne trovò un'importante soluzione che abbiamo già visto nell'esempio 3) del paragrafo 14.13 nell'ambito dei modelli vuoti. L'importanza della soluzione consisteva nel fatto che essa si riferiva ad un universo in espansione (anche se vuoto), nel senso che se vi viene introdotta una piccola particella, essa si allontana dall'osservatore proprio come fanno le galassie. Questa fu la prima indicazione precisa che l'universo poteva essere in espansione. Va sottolineato che il modello apparve molti anni prima che le osservazioni mostrassero chiaramente che vi è una recessione generale delle galassie (espansione).
Possiamo ben affermare che la RG ha fatto in questo caso una previsione teorica della più grande importanza. È appunto ciò che dovrebbe fare ogni buona teoria: dire agli osservatori che cosa cercare, invece di avere il senno di poi, e presentare una spiegazione ad osservazioni avvenute. Questo fu il secondo successo della RG (il primo fu la soluzione di Schwarzschild).
È chiaro che per Einstein, la scoperta di de Sitter dev'essere stata un vero shock, e fu una delle ragioni che più tardi lo indussero ad abbandonare il termine Λ, e a ritornare alle equazioni di campo originali. Sembra che Einstein abbia dapprima cercato la soluzione che poi de Sitter trovò, ma senza successo. La ripugnanza di Einstein per la soluzione di de Sitter sorse perché egli era profondamente contrario all'idea di un universo che, pur essendo vuoto, contenga tuttavia campi di forza (che ci siano tali campi di forza nel modello di de Sitter si deduce dal moto delle piccole particelle di prova). I campi di forza, pensava Einstein, non possono esserci senza la materia (Principio di Mach), e respinse pertanto fermamente come non realistico il modello di de Sitter, e con esso il termine Λ che lo aveva originato. Nel 1922 però si ebbe un importante progresso quando Friedmann trovò **i tre modelli in espansione** che abbiamo studiato nel paragrafo 14.14 i quali sono soluzioni esatte delle equazioni di campo non modificate, e pertanto **dimostrano che modelli in espansione si possono ottenere anche senza il termine Λ.** Poco dopo la scoperta di Friedmann, gli astronomi cominciarono a capire che l'universo è in espansione, e ad Einstein ed a molti altri parve non esserci più ragione per mantenere il termine Λ, cioè il termine cosmologico nelle equazioni di campo che era stato introdotto per ottenere un modello statico di universo, ora eliminato dalle osservazioni. Tutto ciò indusse Einstein ad esclamare: *"l'introduzione del termine cosmologico fu l'errore più grande della mia vita!"*
Ciò nondimeno, molti cosmologi hanno studiato modelli che si ottengono soltanto mantenendo il termine Λ, come vedremo nel prossimo paragrafo. Alcuni di questi modelli sono di grande interesse tanto che due cosmologi importanti come Eddington e Lemaitre, hanno fermamente sostenuto che il termine cosmologico deve essere conservato nelle equazioni di campo di Einstein.
Il modello statico di Einstein ci fornisce una chiave per capire fisicamente il termine Λ. *Perché un universo statico dovrebbe essere possibile con la presenza, nelle equazioni di campo, del termine Λ, ma non senza?* La ragione è che le equazioni non modificate descrivono solo la gravitazione, e non possiamo quindi aspettarci l'esistenza di un modello statico, perché tutta la materia precipiterebbe sotto l'effetto del suo stesso campo gravitazionale. Il termine Λ impedisce la caduta perché rappresenta una forza di repulsione opposta alla gravitazione che tiene il modello in equilibrio.
La forza di repulsione è di un tipo piuttosto insolito. Le forze di repulsione sono ben note in fisica: per esempio, due particelle entrambe cariche positivamente o negativamente si respingono

a vicenda con una forza direttamente proporzionale al prodotto delle cariche e inversamente proporzionale al quadrato della loro distanza. Tale forza dipende da alcune proprietà specifiche della materia, in questo esempio la carica elettrica. Ma la repulsione cosmica è tale che la materia è respinta da ogni punto O dell'universo con una forza proporzionale alla sua distanza da O. Pertanto, *maggiore è la distanza, più grande è la forza di repulsione.* La forza repulsiva di una galassia è proporzionale alla sua massa, ma del tutto indipendente dal fatto che ci sia o meno una galassia in O.
Quest'idea, *che le galassie abbiano un effetto maggiore l'una sull'altra quanto più sono distanti suona molto strana per noi che siamo cresciuti con la credenza che le cose interagiscono più fortemente se sono più vicine.* È ora chiaro perché il termine Λ è importante solo in Cosmologia, perché per distanze piccole come quelle all'interno della nostra galassia, la repulsione è molto piccola, ma per distanze cosmologiche diventa apprezzabile.
Possiamo ora formulare una spiegazione fisica del perché *il modello di universo di Einstein, come ha dimostrato Eddington, è instabile.* Nella condizione non perturbata l'attrazione gravitazionale e la repulsione cosmica si bilanciano esattamente: se ammettiamo però che per qualche ragione sconosciuta una parte dell'universo si rarefaccia, cioè che le distanze fra le galassie aumentino, allora la gravitazione diminuisce e la repulsione aumenta. L'espansione delle galassie interessate ne risulta favorita, e il campo gravitazionale indebolito non può più fermarla. Così una parte dell'universo comincia ad espandersi, e questo presumibilmente è sufficiente ad innescare un'espansione generale. Con un ragionamento simile si può vedere che un addensamento iniziale dà origine ad una contrazione, dal momento che, in questo caso, è l'accresciuta attrazione gravitazionale che diventa la forza non bilanciata.
Quando, alla fine degli anni venti del novecento, l'espansione dell'Universo venne definitivamente accertata da Hubble con l'osservazione, gli astronomi credettero naturalmente che fosse connessa con la repulsione cosmica introdotta nelle equazioni di campo dal termine Λ. Fu solo più tardi, quando si comprese il lavoro teorico di Friedmann, che si capì che i modelli in espansione erano possibili anche senza il termine cosmologico come succede nei modelli **con la grande esplosione: il Big Bang**. Se si mantiene il termine Λ, si aumenta la velocità di espansione come effettivamente le osservazioni astronomiche ultimamente hanno rilevato.

14.16 F-modelli generali con azione cosmologica

Riprendiamo l'equazione (14.57) di Friedmann che ora riscriviamo così:

$$\begin{cases} \dot{R}^2 = F(R), \\ \\ F(R) \equiv \dfrac{C}{R} + \dfrac{\Lambda c^2 R^2}{3} - kc^2, \quad (k = -1, 0, +1) \end{cases} \tag{14.77}$$

dove si è posto

$$\frac{8\pi}{3} G\rho R^3 = C, \tag{14.78}$$

con $C > 0$ costante d'integrazione.
Veniamo ora alla trattazione del caso generale $\rho > 0,\ \Lambda \neq 0$. Si osservi anzitutto che la presenza nell'equazione di Friedmann $(14.77)_1$ di parametri indeterminati - siano essi costanti strutturali a priori non conosciute, come Λ, oppure costanti d'integrazione come C - rende molto più complicata la trattazione matematica di tale equazione e molto ampia la classe dei modelli possibili.
È interessante osservare che l'equazione di Friedmann $(14.77)_1$, che descrive l'evoluzione di tutti i possibili modelli di universo, **ha la stessa identica struttura matematica dell'equazione**

differenziale che regola il movimento di un pendolo, *ovvero, più generalmente, dell'equazione che regge il moto di un sistema meccanico ad un solo grado di libertà q.* Tale equazione si presenta generalmente nella forma

$$\dot{q}^2 = F(q), \tag{14.79}$$

ed è una delle equazioni differenziali più importanti dell'intera Meccanica.
È ben noto, dai corsi di Meccanica Razionale, che la trattazione matematica dell'equazione differenziale del primo ordine (14.79) alla quale si associano le condizioni iniziali:

$$\text{per } \ t = 0, \quad q = q_0, \ \ \dot{q} = \dot{q}_0, \qquad \text{con } \ |\dot{q}_0| = \sqrt{F(q_0)}, \tag{14.80}$$

si può dominare qualitativamente mediante la cosiddetta **discussione di Weierstrass** della quale intendiamo avvalerci per discutere l'equazione di Friedmann. [17]
Il contenuto di tale discussione può essere così sintetizzato:
all'infuori dei moti oscillatori periodici (fra due radici semplici di $F(R)$) sono possibili soltanto moti che presentano al più una inversione di senso (corrispondente ad una radice semplice di $F(R)$) e ammettono una meta asintotica a distanza finita (in una radice multipla di $F(R)$) oppure una meta asintotica a distanza infinita (quando il sistema nel suo moto non incontra alcuna radice di $F(R)$). Infine, il sistema non può mai passare nel corso della sua evoluzione per una posizione corrispondente ad una radice multipla di $F(R)$: può soltanto tendere ad essa con moto asintotico, oppure restarvi indefinitamente, quando, per caso, vi si trovi all'inizio.

Fissato il valore di $k = -1, 0, +1$ ed il valore di $C > 0$, conviene lasciare completamente indeterminato il valore di Λ e fare una rappresentazione quotata della quantità $F_k(R, \Lambda)$ sul piano cartesiano (R, Λ)[18].

Figura 14.5: Rappresentazione grafica delle curve di livello per F- modelli generali con azione cosmologica.

Si disegneranno allo scopo sul piano suddetto (Fig.14.5) le curve di livello $F_k(R, \Lambda) = cost$. che costituiscono palesemente una famiglia di ∞^1 cubiche tutte aventi l'asse R e l'asse Λ come

[17] Per l'enunciato del teorema di Weierstrass, a cui si allude più sopra, si rimanda il lettore all'APPENDICE alla fine di questo capitolo.
Una buona trattazione alternativa dei modelli di Friedmann, che non utilizza il Teorema di Weierstrass come abbiamo fatto noi, si può trovare sia in Wolfgang Rindler, *Relativity, Special, General and Cosmological*, Second Edition, Oxford University Press, 2008, Chap. 18, pagg. 403-406; sia in Ray d'Inverno, *Introduzione alla Relatività di Einstein*, CLUEB, Bologna, 2001, Cap. 23, pagg. 409-441.

[18] Indichiamo con F_k il secondo membro della $(14.77)_2$ corrispondente al fissato valore di k. Pertanto con F_{-1}, F_0, F_{+1}, indicheremo ciò che diviene la F ponendo rispettivamente $-1, 0, +1$ in luogo di k.

asintoti. Per semplificare un po' i calcoli conviene, da questo punto in poi, usare le cosiddette unità relativistiche considerando unitaria la velocità della luce $(c = 1)$, sicché l'equazione della famiglia delle curve di livello, $F_k(R, \Lambda) = cost.$, ora si scriverà in maniera leggermente più semplice:

$$F_k(R, \Lambda) \equiv \frac{C}{R} + \frac{\Lambda R^2}{3} - k = cost. \qquad (k = -1, 0, +1). \tag{14.81}$$

Poiché k figura in F, nella (14.81), come costante additiva, le tre famiglie di linee $F_{-1} = cost.$, $F_0 = cost.$, $F_{+1} = cost.$, che abbiamo rappresentato nello stesso disegno (Fig.14.5), coincidono in quanto famiglie, anche se individualmente una linea $F_{-1} = cost.$ sarà diversa dalla linea $F_0 = cost.$ e dalla linea $F_{+1} = cost.$ corrispondenti ad un medesimo valore della costante a secondo membro.
In particolare saranno tra loro distinte le curve di equazione cartesiana:

$$F_{+1}(R, \Lambda) \equiv \frac{C}{R} + \frac{\Lambda R^2}{3} - 1 = 0, \tag{14.82}$$

$$F_0(R, \Lambda) \equiv \frac{C}{R} + \frac{\Lambda R^2}{3} = 0, \tag{14.83}$$

$$F_{-1}(R, \Lambda) \equiv \frac{C}{R} + \frac{\Lambda R^2}{3} + 1 = 0, \tag{14.84}$$

che scritte in forma esplicita sono rispettivamente:

$$\Lambda(R) = \frac{3}{R^2} - \frac{3C}{R^3}, \tag{14.85}$$

$$\Lambda(R) = -\frac{3C}{R^3}, \tag{14.86}$$

$$\Lambda(R) = -\frac{3}{R^2} - \frac{3C}{R^3}, \tag{14.87}$$

corrispondenti al valore zero della costante al secondo membro $(cost = 0)$ nella (14.81).
Le curve $F_{-1} = 0$ ed $F_0 = 0$ si mantengono completamente al disotto dell'asse R con andamento monotòno crescente. Diverso è invece l'andamento della curva $F_{+1} = 0$. Essa ha dapprima andamento crescente; dopo aver tagliato l'asse R nel punto di ascissa $R = C$, raggiunge un punto di massimo nel punto M_* di coordinate

$$\begin{cases} R_* = \dfrac{3C}{2}, \\[2ex] \Lambda_* = \dfrac{4}{9C^2} = \dfrac{1}{R_*^2}, \end{cases} \tag{14.88}$$

decresce poi indefinitamente, sempre rimanendo al disopra dell'asse R cui si avvicina asintoticamente[19] (Fig.14.5).
Pe ogni valore di k la regione $F_k < 0$, situata sotto la curva $F_k = 0$, è fisicamente inaccessibile (infatti si avrebbero valori immaginari per $\dot{R}$, dato che è $\dot{R}^2 = F_k(R, \Lambda)$). È invece accessibile la regione $F_k > 0$, situata al disopra della curva $F_k = 0$.

[19]Infatti, i massimi ed i minimi relativi della funzione $\Lambda(R) = \frac{3}{R^2} - \frac{3C}{R^3}$ si ottengono ponendo $\Lambda'(R) = 0$, cioè $\frac{d}{dR}\left(3R^{-2} - 3CR^{-3}\right) = 0$, che è soddisfatta per $R_* = \frac{3C}{2}$ a cui corrisponde $\Lambda_* = \frac{4}{9C^2} = \frac{1}{R_*^2}$, ed è facile constatare che si tratta di un massimo relativo proprio; l'altra soluzione è $R = +\infty$ che equivale a dire che l'asse R è un asintoto orizzontale della curva.

Ogni modello di Friedmann è caratterizzato da un valore di k e da un valore di Λ (a parità della costante $C > 0$). Quindi per ogni modello di Friedmann gli "stati" attraversati durante la sua evoluzione si ottengono muovendosi lungo una ben determinata retta parallela all'asse R nella (Fig.14.5) ed applicando il citato teorema di Weierstrass. Tale retta può, oppure no, intersecare la linea $F_k = 0$: restano in tal modo determinate quali parti della retta sono fisicamente accessibili: ovviamente quelle che sono contenute nella regione $F_k > 0$.
Ciò premesso, discutiamo i diversi modelli tipici, in corrispondenza ai possibili valori dell'indice di curvatura k e della costante cosmologica Λ (che, per fissare le idee, supporremo sempre positiva $\Lambda > 0$) determinando qualitativamente il grafico di R in funzione di t dalle proprietà della funzione $\Lambda = \Lambda(R)$ (in particolare dalla natura sei suoi zeri).

I - MODELLI ILLIMITATAMENTE ESPANSIVI

$$(k = +1, \quad \Lambda > \Lambda_*, \ \ oppure \quad k = 0, -1, \ \ \Lambda > 0)$$

In tutti questi casi, la retta $\Lambda = cost.$, corrispondente al valore prescelto di Λ non taglia nessuna delle curve $F_k = 0 \quad (k = +1, 0, -1)$ in quanto queste rimangono tutte completamente al disotto di tale retta come mostrato in(Fig.14.6) e in (Fig.14.7). Supponendo allora che per $t = 0$ sia $R_0 = 0$ ed $\dot{R}_0 > 0$ (potendo essere anche $\dot{R}_0 = +\infty$), si ha $F(R_0) > 0$ e perciò il punto rappresentativo nel piano (R, Λ), dello stato del modello, si muoverà indefinitamente verso destra in quanto, in tale direzione, non si incontra alcuna radice di $F(R)$ (caso A_2 del teorema di Weierstrass enunciato nell'APPENDICE alla fine del capitolo). Di conseguenza anche il punto rappresentativo del sistema nel piano (t, R) si muoverà sulla curva $R = R(t)$ indefinitamente verso destra con $R(t)$ funzione crescente di t.
Per maggiore chiarezza, distinguiamo i due casi seguenti:

I_a - MODELLI ILLIMITATAMENTE ESPANSIVI (Universo di Lemaitre)

$$\text{con } (k = +1, \quad \Lambda > \Lambda_*)$$

Figura 14.6: Modelli illimitatamente espansivi con $k = +1$, $\Lambda > \Lambda_*$: Universo di Lemaitre.

La curva che descrive questo caso è quella di equazione (14.82), cioè $F_{+1}(R, \Lambda) = 0$, la cui forma esplicita è la (14.85). L'andamento di tale curva nel piano (R, Λ) è rappresentato nella Fig.14.6 e si può constatare che corrisponde al caso A_2 del teorema di Weierstrass.
Siccome, in questo caso, la $F(R)$ non si annulla mai ed anzi è sempre $F(R) > 0$, essendo inoltre $\dot{R}^2 = F(R)$ ed avendo supposto $\dot{R}_0 > 0$, per ragioni di continuità, consegue che $\dot{R}(t)$ ha sempre lo stesso segno, precisamente è sempre $\dot{R}(t) > 0$, dunque $R(t)$ è funzione crescente di t ed il

suo grafico è rappresentato dalla Fig.14.7.
Se studiamo anche l'andamento della funzione $\dot{R}(t)$ possiamo osservare che essa va decrescendo da valori infinitamente grandi, assunti in vicinanza di $R = 0$ (**Big Bang**), fino a raggiungere il suo valor minimo per $R = \left(\frac{3C}{2\Lambda}\right)^{1/3}$, in cui la curva di equazione $R = R(t)$ presenta un flesso a tangente obliqua, per poi riprendere a crescere indefinitamente.
Questo modello fu scoperto dal padre gesuita belga Georges Lemaitre ed è perciò noto come **universo di Lemaitre**. Lemaitre fu il primo a capire che lo spostamento verso il rosso della luce delle stelle era la prova dell'espansione dell'universo ed a proporre la relazione di proporzionalità fra la distanza delle galassie e la loro velocità di recessione, poi confermata dalle osservazioni sperimentali di Hubble. Nel 1927 pubblicò *L'ipotesi dell'atomo primigenio (o primevo)* oggi nota come **teoria del Big Bang**. Lemaitre fu sempre un sostenitore dell'espansione illimitata dell'universo ed a questo scopo conservò sempre nel suo modello la costante cosmologica, proposta da Einstein, ma abbandonata da lui e da quasi tutti i fisici e cosmologi dopo la scoperta dell'espansione dell'universo. L'idea della costante cosmologica ritornò solo dopo che fu scoperta, da parte di Saul Perlmutter, Brian Schmidt e Adam Riess, l'accelerazione dell'espansione dell'universo nel 1998, oltre trent'anni dopo la morte di Lemaitre. Tale scoperta fruttò ai tre scienziati nominati il premio Nobel per la fisica nel 2011.

I_b - MODELLI ILLIMITATAMENTE ESPANSIVI

con $(k = 0, -1, \ \Lambda > 0)$ Fig.14.7

Figura 14.7: Modelli illimitatamente espansivi con $k = 0, -1, \Lambda > 0$

Supponiamo ancora che le condizioni iniziali siano che per $t = 0$ sia $R_0 = 0$ ed $\dot{R}_0 > 0$ (potendo essere anche $\dot{R}_0 = +\infty$), si ha $F(R_0) > 0$ e perciò, come abbiamo già detto in precedenza, il punto rappresentativo, dello stato del modello, si muoverà, per il teorema di Weierstrass (caso A_2), indefinitamente verso destra sul piano (R, Λ); inoltre anche il punto rappresentativo del sistema nel piano (t, R) si muoverà sulla curva $R = R(t)$ indefinitamente verso destra con $R(t)$ funzione crescente di t. Tutto ciò risulta chiaro dalle Fig.14.6 e Fig.14.7.
Infine, val la pena osservare che: sui modelli di universo in questione, la R ha diverso significato nel caso $k = +1$, in cui è interpretabile come raggio di curvatura dell'universo e nei casi $k = 0, k - 1$ in cui ha soltanto il significato di "*funzione di espansione*".
Se Λ è molto prossima a Λ_* (per $k = +1$) o a zero (per $k = 0, \ k = -1$), per un lungo tratto, $\dot{R}(t)$ diviene pressapoco costante o, rispettivamente, tende asintoticamente a divenire costante.

II - Supponiamo ora

$$(k = +1, \quad 0 < \Lambda < \Lambda_*).$$

In queste ipotesi le parti accessibili della retta $\Lambda = cost.$ sono due, tra loro disgiunte: il segmento AB e la semiretta CD (Fig.14.8). Sono quindi possibili due tipi distinti di modelli, a seconda del valore iniziale di R; distinguiamo pertanto i due seguenti sottocasi.

Figura 14.8: Modelli pulsanti a periodo finito, con $k = +1, 0 < \Lambda < \Lambda_*$.

II_a - MODELLI PULSANTI A PERIODO FINITO (Fig. 14.8)

$$(k = +1, \quad 0 < \Lambda < \Lambda_*)$$

Se il valore iniziale di R appartiene al segmento AB, per es. $R_0 = 0$ con $\dot{R}_0 > 0$, l'evoluzione ha luogo verso destra ed $R(t)$ è crescente (Fig.14.8) fino a che non viene raggiunto il punto B in cui R assume il massimo valore R_1 e la velocità di espansione si annulla $\dot{R}_1 = 0$ (perché $F_1(R, \Lambda) = 0$). Dato che R_1 è una radice semplice di $F(R)$ (com'è facile constatare) il punto B viene raggiunto in un tempo finito (per il caso A_3 del teorema di Weierstrass) e precisamente all'istante $t_1 = \int_0^{R_1} \frac{dq}{\sqrt{F(q)}}$. In tale istante, in base all'equazione del 2° ordine $(14.42)_2$ (ove si pone come convenuto $c = 1$), avremo in forma esplicita:

$$\ddot{R} = \frac{R_1}{2}\left(\Lambda_1 - \frac{1}{R_1^2}\right),$$

e poiché $R_1 < R_*$ implica $\frac{1}{R_1^2} > \frac{1}{R_*^2}$, cioè $\frac{1}{R_1^2} > \Lambda_* > \Lambda_1$ si ha, nel punto B, $\ddot{R} < 0$. Ha quindi inizio, da tale istante, una fase di collasso durante la quale il modello torna a riassumere, in ordine inverso, tutti gli stati per i quali era passato in fase di espansione. Raggiunto il collasso completo ($R = 0$), ed ammesso un brusco cambiamento di segno di $\dot{R}$ (il cosiddetto **Big Bounce**), ha inizio un nuovo ciclo completo, identico al precedente. Si può quindi parlare di un **modello pulsante periodico a periodo finito** (Fig.14.8b) dato da:

$$T = 2\int_0^{R_1} \frac{dR}{F_{+1}(R)}.$$

II_b - MODELLI COLLASSO-ESPANSIVI (Fig.14.9)

Se il valore assunto da R per $t = 0$ appartiene alla semiretta CD con $\dot{R}_0 < 0$, il modello ha una prima fase di contrazione (che può aver avuto origine in tempi comunque lontani) fino a che R non abbia raggiunto in C il suo minimo valore $R_2 \neq 0$. Si può verificare che R_2 è ancora una radice semplice di $F(R)$ e che a partire da tale valore il modello inverte la sua evoluzione e prende ad espandersi indefinitamente ed oltre ogni limite per $t \to +\infty$ (caso A_2 del teorema di Weierstrass) (Fig.14.9b).

III - Supponiamo ora

$$(k = +1, \quad \Lambda = \Lambda_*).$$

Figura 14.9: Modelli collasso-espansivi.

L'ipotesi dà luogo a quatto sottocasi:

- III_a- MODELLO STATICO DI EINSTEIN (Fig.14.10)

Figura 14.10: Modello statico di Einstein.

Il valore iniziale di R è $R_0 = R_*$ (Fig.14.10a). Ciò implica per $t = 0$, $\dot{R}_0 = 0$ perché essendo R_* radice doppia di $F(R)$, l'equazione $\dot{R}^2 = F(R)$ diventa $\dot{R}_0^2 = 0$. Inoltre dato che $\ddot{R} = \frac{R}{2}\left(\Lambda - \frac{1}{R^2}\right)$, per $R = R_*$ essendo $\Lambda = \Lambda_* = \frac{1}{R_*^2}$ si ha anche $\ddot{R}_0 = 0$. Il modello (per il caso B_2 del teorema di Weierstrass) permane indefinitamente in tale stato, che è uno stato di equilibrio instabile, come è stato dimostrato da Eddington. Si tratta dunque di un modello statico che altro non è che **il modello statico di Einstein** (1917) già discusso nel paragrafo 14.11.

- III_b - MODELLO ESPANSIVO A META ASINTOTICA STATICA

  Il valore iniziale di R è compreso tra 0 ed R_* cioè $0 < R_0 < R_*$ ed inoltre supposto $\dot{R}_0 > 0$ il modello evolve verso il valore R_* che è radice multipla di $F(R)$ (Fig.14.11). Perciò (per il caso A_5 del teorema di Weierstrass) si ha allora un modello a curvatura costante positiva $(k = +1)$ indefinitamente espansivo tendente asintoticamente, cioè in un tempo infinito, al modello statico di Einstein.

- III_c - MODELLO IN CONTRAZIONE A META ASINTOTICA STATICA

Figura 14.11: Modello espansivo a meta asintotica statica.

Figura 14.12: Modello in contrazione a meta asintotica statica.

Le condizioni iniziali sono $R_0 > R_*$ e $\dot{R}_0 < 0$ (Fig.14.12). Per le stesse considerazioni fatte nel caso precedente, si ha un modello a curvatura positiva $(k = +1)$ in contrazione con tendenza asintotica al modello statico di Einstein.

- III_d - MODELLO ESPANSIVO ASINTOTICAMENTE EUCLIDEO

  Le condizioni iniziali sono $R_0 > R_*$ e $\dot{R}_0 > 0$ (Fig.14.13). Si ha in tal caso (per il punto A_2 del teorema di Weierstrass) un modello indefinitamente espansivo con $R \to +\infty$. La varietà 3-dimensionale è a curvatura costante positiva ed ha come limite asintotico uno spazio euclideo. Si tratta del cosiddetto **modello di Eddington-Lemaitre**.

Concludendo, possiamo dire che tutto ciò di cui si è discusso nei paragrafi precedenti ed in questo paragrafo costituisce il quadro generale dei modelli di Friedmann con o senza costante cosmologica. Come si è visto, oltre a molte diversità qualitative (modelli statici, modelli di età infinita o di età finita, modelli a curvatura positiva, nulla o negativa, modelli indefinitamente espansivi, modelli pulsanti, ecc.) nell'insieme dei modelli di Friedmann c'è una immensa varietà quantitativa dovuta alla presenza dei parametri k, C, R_0, Λ (con R_0 indichiamo il valore attuale di $R(t)$). La misura diretta di queste costanti appare, per il momento, impossibile. Con considerazioni matematiche ingegnose si è riconosciuto che la misura diretta delle costanti k, C, R_0, Λ può essere sostituita dalla misura di altre costanti, il cui valore sembra o più facilmente osservabile o più facilmente congetturabile (valori attuali della densità media di massa $\rho(t)$, dell'età dell'universo (misura della costante H di Hubble), e di recente l'accelera-

Figura 14.13: Modello espansivo asintoticamente euclideo: di Eddington-Lemaitre.

zione dell'espansione, ecc.). Ma, purtroppo, a detta degli Astrofisici e dei Cosmologi, anche la determinazione sperimentale di queste nuove costanti è per ora troppo incerta per arrivare a conclusioni attendibili. Anche ammesso che un modello di Friedmann possa avere una certa attendibilità, le osservazioni di cui disponiamo non permettono ancora di prevedere se l'attuale fase di espansione del nostro universo continuerà indefinitamente ovvero se essa è destinata ad attenuarsi, con tendenza all'inversione (collasso), oppure con tendenza asintotica verso una situazione statica, ovvero se continuerà l'attuale enigmatica fase di accelerazione.

14.17 Densità attuale e critica, costante di Hubble e parametro di decelerazione

Nello studio di un qualsiasi modello di universo si introducono diverse grandezze rilevanti, fra queste ce ne sono almeno quattro che sono, come vedremo, assolutamente importanti:

1) **l'attuale età dell'universo** che indicheremo con t_0;

2) **il valore attuale della costante di Hubble** indicata con $H_0 = H(t_0)$;

3) **il valore attuale della densità della materia che riempie l'universo** indicata con $\rho_0 = \rho(t_0)$;

4) **la densità critica** ρ_c detta anche **densità di chiusura** (perché caratterizza un universo chiuso).

Se combiniamo le due equazioni (14.77), (14.78), ponendovi $\Lambda = 0$ e $c = 1$, e ricaviamo l'indice di curvatura k avremo:

$$k = \frac{C}{R} - \dot{R}^2 = \frac{8}{3}\pi G\rho R^2 - \dot{R}^2, \qquad (k = -1, 0, +1)$$

ricordando poi la (14.21), cioè $H(t) = \dfrac{\dot{R}}{R}$ otteniamo:

$$k = \frac{8}{3}\pi G R^2 \Big[\rho(t) - \frac{3H^2}{8\pi G}\Big], \qquad (k = -1, 0, +1). \tag{14.89}$$

Valutiamo quest'ultima equazione all'epoca attuale, cioè per $t = t_0$, avremo:

$$k = \frac{8}{3}\pi G R_0^2 \Big[\rho_0 - \frac{3H_0^2}{8\pi G}\Big], \qquad (k = -1, 0, +1) \tag{14.90}$$

se inoltre definiamo come **densità critica** ρ_c l'espressione

$$\rho_c = \frac{3H_0^2}{8\pi G}, \tag{14.91}$$

e introduciamo **il parametro di densità attuale**

$$\Omega_0 = \frac{\rho_0}{\rho_c}, \tag{14.92}$$

la (14.90) può essere riscritta così:

$$k = \frac{8}{3}\pi G R_0^2 \Big[\rho_0 - \rho_c\Big] = \frac{8}{3}\pi G R_0^2 \rho_c \Big[\Omega_0 - 1\Big]. \qquad (k = -1, 0, +1) \tag{14.93}$$

Dalla relazione precedente si vede immediatamente che la densità attuale ρ_0 della materia che riempie l'universo (o se si vuole il parametro di densità attuale Ω_0) permette di discriminare tra i tre modelli standard nel modo seguente (Fig.14.14):

Figura 14.14: I tre modelli standard discriminati dal valore attuale Ω_0 del parametro di densità.

$$\begin{cases} \rho_0 < \rho_c \ (\Omega_0 < 1) \Longrightarrow k = -1 \quad \text{(modello con espansione illimitata)}, \\ \rho_0 = \rho_c \ (\Omega_0 = 1) \Longrightarrow k = 0 \quad \text{(modello di Einstein-de Sitter con espansione illimitata)}, \\ \rho_0 > \rho_c \ (\Omega_0 > 1) \Longrightarrow k = +1 \quad \text{(modello cicloidale oscillante)}. \end{cases} \tag{14.94}$$

Perciò possiamo dire (se $\Lambda = 0$) che *l'evoluzione futura dell'universo dipende sostanzialmente dalla attuale densità di materia ρ_0 e dall'attuale valore della costante di Hubble H_0*. Supponendo, come sembra attendibile dalle osservazioni, che attualmente sia $H_0 = 75\,\mathrm{km/s} \cdot Mpc^{-1} = (13,7 \times 10^9\,\text{anni})^{-1}$ la densità critica risulterebbe $\rho_c = 10^{-29}\,\mathrm{g/cm^3}$ per la quale se $\rho_0 \leq \rho_c$ l'espansione continuerà senza limiti, se $\rho_0 > \rho_c$ dopo l'espansione si avrà la contrazione. Dalle (14.94) si vede chiaramente che dal valore della densità di materia dipende anche un'altra importante proprietà dell'Universo: la sua struttura geometrica globale; **la chiusura (se è finito) o l'apertura (se è infinito)**. Se $\rho_0 > \rho_c$, l'Universo è finito sebbene illimitato come la superficie di una sfera che non ha limiti ma è finita. Se $\rho_0 \leq \rho_c$ l'Universo è infinito ed in questo senso non differisce dal classico spazio tridimensionale euclideo. **Quindi densità, futuro e struttura geometrica dell'Universo sono concetti tra loro connessi.**

Nell'ambito dei modelli FRW, con $\Lambda = 0$, c'è un'altra grandezza importante da considerare, è il cosiddetto **parametro di decelerazione** indicato con $q(t)$ e così definito:

$$q(t) = \frac{-R(t)}{\dot{R}^2(t)}\ddot{R}(t). \tag{14.95}$$

Il segno di $q(t)$ discrimina perciò il caso in cui la velocità di espansione dell'universo è in diminuzione o in aumento. Infatti se $q(t) < 0$ significa che $\ddot{R}(t) > 0$ cioè l'espansione sta accelerando, se $q(t) > 0$, cioè se $\ddot{R}(t) < 0$, l'espansione sta decelerando, infine se $\ddot{R}(t) = 0$ si ha un'espansione a velocità costante.
Nel caso in cui la costante cosmologica non sia nulla ($\Lambda \neq 0$), la situazione è più difficile: per esempio, per alcune relazioni tra $\Lambda \neq 0$ e ρ_0 è possibile che l'Universo sia finito, ma in tempi futuri esso presenti un'espansione illimitata. Purtroppo dobbiamo dire che la determinazione della densità media della materia dell'Universo, dall'osservazione diretta delle sue componenti, è un problema assai difficile dato che è molto complicato osservare tutte le possibili forme di materia.

14.18 Considerazioni conclusive

A conclusione di questo capitolo sulla Cosmologia vogliamo avvertire il lettore che attualmente ci sono molte teorie sul destino ultimo del nostro Universo a partire dall'origine con un **Big Bang caldo**: **Big Crunch** (Grande Schianto), **Big Rip** (Grande Strappo), **Big Bounce** (Grande Rimbalzo), **Big Freeze** (Grande Congelamento), ecc., con diverse varianti. Storicamente si sono formate ed affermate sostanzialmente soltanto due grandi teorie cosmologiche rivali: **la Teoria del Big Bang caldo** e **la Teoria dello Stato Stazionario**.
Oggigiorno però più del 99,9 per cento dei cosmologi, credono nella teoria del Big Bang caldo cioè di un evento, simile ad una grande esplosione, avvenuta circa 14 miliardi di anni fa e che ha dato origine all'Universo. Storicamente le cose sono andate pressapoco come cerchiamo di spiegare brevemente in ciò che segue.

1) **Teoria del Big Bang caldo**.
   Come, già abbiamo detto, George Lemaitre fu il primo a capire che lo spostamento verso il rosso delle righe spettrali della luce proveniente dalle galassie lontane fosse la prova dell'espansione dell'Universo e a proporre la relazione di proporzionalità fra la distanza delle galassie e la loro velocità di allontanamento, poi confermata dalle osservazioni di Edwin Hubble nel 1929. Nel 1927 pubblicò *l'ipotesi dell'atomo primevo o primigenio* dal quale scaturì l'Universo, oggi nota come *la Teoria del Big Bang.*
   La teoria più ambiziosa e dettagliata della *"grande esplosione"* è quella sviluppata, tra la fine degli anni 1940 e gli inizi degli anni 1950, dal fisico e cosmologo russo George Gamow e dai suoi collaboratori americani: Ralph Alpher e Robert Herman. Il suo fondamentale articolo sulla *cosmogenesi* scritto con l'americano Ralph Alpher fu pubblicato come teoria $\alpha\,\beta\,\gamma$ perché porta i nomi di *Alpher-Bethe-Gamow*, per analogia con le prime tre lettere dell'alfabeto greco. Gamow aggiunse volutamente il nome del grande fisico tedesco Hans Bethe, scopritore delle reazioni termonucleari che avvengono all'interno delle stelle (*il cosiddetto ciclo di Bethe*) e forniscono loro la grandissima energia che li tiene accese per miliardi di anni.
   Gamow e collaboratori fecero l'ipotesi che l'Universo abbia avuto inizio come una massa di neutroni ad altissima temperatura e densità. Questi decaddero presto a formare protoni ed elettroni, ma alcuni dei primi protoni si unirono coi neutroni che rimanevano in modo da produrre nuclei atomici di vari gradi di complessità. La temperatura, mentre ciò accadeva, era di circa 10 miliardi di gradi, e la maggior parte degli elementi pesanti si

formò nei primi 30 minuti dell'espansione! Questa è la teoria $\alpha\,\beta\,\gamma$, almeno nelle sue linee generali; ma sfortunatamente quando Gamow ed i suoi collaboratori passarono ai calcoli dettagliati, incontrarono delle difficoltà, che si dimostrarono insuperabili, perché i calcoli non riescono a render conto dell'origine e delle quantità degli elementi pesanti che pure esistono nell'Universo.[20]
La fabbrica chimica che stiamo cercando risulta essere la parte interna delle stelle. Lo si era sospettato negli anni cinquanta quando fu notato in alcune stelle lo spettro di un elemento detto **tecneto**. L'importanza della scoperta sta nel fatto che il tecneto va soggetto ad un decadimento radioattivo piuttosto rapido, e che, se si fosse formato al tempo del Big Bang, si sarebbe già disintegrato da molto tempo. Se lo si osserva ora vuol dire che la sua formazione è continua. In seguito, sforzi pazienti di astrofisici, quali il grande astronomo inglese Fred Hoyle, hanno fornito un fondamento teorico alla generazione di elementi pesanti nelle stelle. Oggi è generalmente riconosciuto che le stelle producono gli elementi più pesanti in quantità sufficiente a giustificarne l'abbondanza che si osserva. Di conseguenza la grande esplosione non deve più rendere conto della produzione degli elementi pesanti. È stata questa funzione che dapprima ha attratto i cosmologi verso i modelli con uno stato singolare iniziale, poiché a quel tempo il Big Bang sembrava il solo modo per spiegare l'abbondanza relativa degli atomi dei vari elementi. Una volta resisi conto che un'alternativa è possibile, si può lasciar cadere l'ipotesi della grande esplosione, a patto che i nostri modelli cosmologici possano essere alterati in modo tale da non presentare uno stato singolare.

2) **Teoria dello Stato Stazionario (Steady State Theory)**.
Fu nel clima intellettuale sopra descritto che nel 1948, tre giovanissimi astronomi di Cambridge, Hermann Bondi, Thomas Gold e Fred Hoyle proposero **la Teoria dello Stato Stazionario** in contrapposizione alla **Teoria del Big Bang**. Questa teoria aveva il grande vantaggio di superare la difficoltà della scala del tempo senza rinunciare al Principio Cosmologico, anzi introducendo una sua versione più forte il cosiddetto
Principio Cosmologico Perfetto (PCP): *In scala sufficientemente ampia, l'Universo non solo è omogeneo ed isotropo ma è* **essenzialmente lo stesso in ogni tempo.**
Perciò *l'Universo deve presentare lo stesso aspetto a tutti gli ipotetici osservatori galattici, qualunque sia la loro posizione, qualunque sia la direzione in cui si rivolge lo sguardo e qualunque sia l'epoca in cui si osserva.*
La Teoria dello Stato Stazionario sostanzialmente postula che l'Universo mantenga le stesse proprietà nello spazio e nel tempo e non abbia né inizio né fine. Essa non deve essere confusa con il modello di un **Universo Statico**: infatti la teoria descrive un Universo in espansione come la teoria del Big Bang. Ma poiché l'espansione fa diminuire la densità della materia nel corso del tempo, la Teoria dello Stato Stazionario postula **la creazione continua di materia** per mantenere la densità costante.

Pro e contro la teoria dello stato stazionario.
La teoria ha indubbiamente delle caratteristiche attraenti.
- In primo luogo, l'idea di un Universo che su larga scala non muta è semplice, e per molti studiosi anche esteticamente piacevole.
- In secondo luogo, i problemi dello stato singolare, che provocano tanto disturbo nei modelli relativistici, mancano del tutto. Poiché la densità e la temperatura media - come tutte le altre proprietà su larga scala del modello stazionario - sono immutabili, non sorge il problema di un'epoca in cui esse erano molto alte.

[20]Sull'origine dell'Universo, una lettura interessantissima che raccomandiamo fortemente al lettore, dovuta al premio Nobel per la fisica Steven Weinberg è: *I primi tre minuti, l'affascinante storia dell'origine dell'Universo,* Oscar Saggi Mondadori, Milano, 1980.

- La mancanza di uno stato singolare nella teoria dello stato stazionario fu in principio ritenuta un difetto, perché una grande esplosione era ritenuta necessaria per fornire le condizioni richieste per la formazione degli elementi pesanti. Ma dal momento in cui si scoprì che questi elementi possono formarsi all'interno delle stelle ha eliminato la necessità della grande esplosione, per cui questa obiezione alla teoria dello stato stazionario non sussiste più.
- Un altro difetto importante della teoria è il fatto che essa, ammettendo la creazione continua di materia, viola il principio di conservazione dell'energia ritenuto dai fisici sacrosanto.
- Un'altra obiezione importante alla teoria dello stato stazionario è che non è affatto una teoria fondamentale nel senso ordinario. Manca per questo la parte più importante: *Non è data alcuna spiegazione del perché l'Universo si espande.* Nella teoria relativistica questa spiegazione è contenuta nelle equazioni di campo di Einstein, con o senza la costante cosmologica Λ. *Nella teoria dello stato stazionario non ci sono equazioni di campo.* Il modello dello stato stazionario è stato proposto senza un'adeguata base matematica: è un modello fenomenologico, basato sul principio cosmologico perfetto, o poco più.
Nell'estate del 1964 Hoyle con il suo allievo Narlikar pubblicarono un lavoro nel quale proponevano una nuova ambiziosa teoria della gravitazione che doveva sostituire la relatività generale e costituire una base matematica per la cosmologia dello stato stazionario.[21]
In quello stesso anno, però, Penzias ed Wilson scoprirono la **radiazione cosmica di fondo a microonde (Cosmic Microwaves Background (CMB))** di circa $2,7\,K$ che, negli anni successivi, sancì in modo pressoché definitivo la validità ed il trionfo della Teoria del Big Bang. Da quell'anno si parlò sempre meno della teoria dello stato stazionario e, a quanto pare, la nuova teoria della gravitazione di Hoyle e Narlikar non ebbe grande risonanza nella comunità dei cosmologi.

14.19 APPENDICE AL CAP. 14

Nel paragrafo (14.16) abbiamo discusso i più generali modelli cosmologici di Friedmann che - fissato a priori entro la funzione $F_k(\Lambda, R)$ il valore della costante cosmologica Λ, il valore della costante d'integrazione $C > 0$ e l'indice di curvatura k - sono governati dall'omonima equazione differenziale del 1° ordine:

$$\dot{R}^2 = F(R),$$

dove R è la variabile indipendente, $\dot{R}$ la sua derivata rispetto al tempo cosmico, ed $F(R)$ una funzione assegnata di R.
Questa equazione è di un tipo notevolissimo che si presenta in diversi problemi di Meccanica e di Fisica Matematica, cioè è un'equazione del tipo:

$$\dot{q}^2 = F(q), \tag{14.96}$$

dove q è l'unica coordinata lagrangiana che caratterizza, istante per istante, la configurazione del sistema e va considerata come funzione del tempo $q = q(t)$.
Essa *regge il moto di un sistema con un solo grado di libertà che ammetta l'integrale primo dell'energia, espresso appunto dalla stessa equazione.* Esempi notevoli sono: l'equazione del moto di un pendolo semplice o composto, quella che descrive il moto di un corpo rigido con un asse fisso e quella del problema ristretto dei due corpi.
Per la realtà del moto, essendo $\dot{q}^2 \geq 0$, dobbiamo supporre che sia anche $F(q) \geq 0$, inoltre indicheremo con q_0 e $\dot{q}_0$ rispettivamente il valore di $q(t)$ e della sua derivata prima all'istante iniziale $t = 0$, cioè $q(0) = q_0$, $\dot{q}(0) = \dot{q}_0$ e supporremo sempre $\dot{q}_0 > 0$.
L'equazione (14.96) è facilmente risolubile con una semplice integrazione:

$$t = \int_{q_0}^{q} \frac{dq}{\sqrt{F(q)}}, \tag{14.97}$$

[21] F. Hoyle and J. V Narlikar, *A New Theory of Gravitation*, Proc R. Soc. London, A, **282**, 191-207.

la quale però, com'è noto, conduce generalmente a funzioni ellittiche. Per l'esattezza, nella (14.97), il denominatore dovrebbe comparire col doppio segno, $\pm\sqrt{F(q)}$, ma, per fissare le idee, avendo supposto $\dot{q}_0 > 0$, abbiamo considerato soltanto il segno $+$, in quanto una analoga discussione si può ripetere per il segno $-$ se si suppone che inizialmente sia $\dot{q}_0 < 0$.
Detto $\phi(q)$ l'integrale a secondo membro della (14.97) avremo:

$$t = \phi(q), \tag{14.98}$$

da cui, invertendo la $\phi(q)$, si ricava la q come funzione del tempo t:

$$q = \phi^{-1}(t). \tag{14.99}$$

Si ha pertanto il risultato che: *l'integrazione della* (14.96) *è, come si dice, sempre riconducibile alle quadrature*. Con ciò il problema è ben lungi dall'essere risolto, perché bisogna, in primo luogo, fissare quale segno si deve prendere davanti al radicale, in secondo luogo discutere la convergenza o meno dell'integrale a secondo membro della (14.97) quando la q raggiunge un valore per cui la $F(q)$ si annulli.
Tuttavia, è molto più semplice ed espressiva una discussione qualitativa della (14.96) basata sul seguente notevole[22]:
Teorema di Weierstrass:
L'andamento qualitativo del moto di un sistema dinamico governato dall'equazione differenziale del 1° ordine:

$$\dot{q}^2 = F(q),$$

consiste sostanzialmente nella discussione di due casi, dipendenti dalle condizioni iniziali, a seconda che sia:

A) $F(q_0) > 0$,
oppure

B) $F(q_0) = 0$.

La discussione è fondata sulla considerazione di cinque sottocasi per A) e due sottocasi per B), come segue:

ANALISI DEL PRIMO CASO: $F(q_0) > 0$.

A1 $F(q_0) > 0$
Se $F(q_0) > 0$ e a partire da q_0 fino a $+\infty$, nel verso della velocità iniziale $\dot{q}_0 > 0$ non si incontra mai *uno zero* di $F(q)$ (cioè se $F(q) > 0$ per $q_0 \leq q < +\infty$) (vedi Fig.14.15a) ed inoltre $\int_{q_0}^{q} \frac{dq}{\sqrt{F(q)}}$ è **convergente** per $q \to +\infty$, cioè

$$\int_{q_0}^{+\infty} \frac{dq}{\sqrt{F(q)}} = t_1 > 0, \tag{14.100}$$

allora la funzione $q(t)$ **è crescente indefinitamente**, cioè

$$\lim_{t \to t_1^-} q(t) = +\infty,$$

ed **il moto avviene senza inversioni in un tempo finito dato dalla** (14.100) (vedi Fig.14.15b).

[22]Il notevole teorema da noi citato è desunto da una Memoria, ormai classica, di Karl Weierstrass dal titolo *Ueber eine Gattung reel periodischen Functionem*, Monatsbericht der Akademie der Wissenschaften zu Berlin, 1886, tale Memoria è stata poi inserita nel Vol. II delle sue Opere.
Il lettore può trovare la dimostrazione del teorema di Weierstrass in T. Levi-Civita e U. Amaldi, *Lezioni di Meccanica Razionale*, Vol. II, parte prima, Cap. I, § 6, pagg. 21-30, Zanichelli Editore Bologna, 1974; oppure in G. Ferrarese *Lezioni di Meccanica Razionale*, Vol. I, 1ª Edizione, Pitagora Editrice Bologna, 1980, pagg. 267-271. Un'altra ottima trattazione si trova nel libro di Silvio Nocilla, *Meccanica Razionale per allievi ingegneri*, Libreria Editrice Universitaria Levrotto e Bella, Torino, 1982, pagg. 436-441.

Figura 14.15: Confronto fra i due grafici dell'andamento del moto nei due piani $(q, F(q))$ e $(t, q(t))$, nell'ipotesi $F(q_0) > 0$, $\dot{q}_0 > 0$ e l'integrale convergente.

A2 - Se $F(q_0) > 0$ e a partire da q_0 fino a $+\infty$, nel verso della velocità iniziale $\dot{q}_0 > 0$ non si incontra mai uno zero di $F(q)$ (cioè se $F(q) > 0$ per $q_0 \leq q < +\infty$) (vedi Fig.14.16a) e se inoltre $\int_{q_0}^{q} \frac{dq}{\sqrt{F(q)}}$ è **divergente** per $q \to +\infty$, cioè

Figura 14.16: Confronto fra i due grafici dell'andamento del moto nei due piani $(q, F(q))$ e $(t, q(t))$, nell'ipotesi $F(q_0) > 0$, $\dot{q}_0 > 0$ e l'integrale divergente.

$$\int_{q_0}^{+\infty} \frac{dq}{\sqrt{F(q)}} = +\infty, \tag{14.101}$$

allora la funzione $q(t)$ **è crescente indefinitamente**, cioè

$$\lim_{t \to +\infty} q(t) = +\infty,$$

ed **il moto avviene ancora senza inversioni ma in un tempo infinito** (vedi Fig.14.16b).

A3 - Se $F(q_0) > 0$ ed esiste nel verso indicato da $\dot{q}_0 > 0$ **uno zero semplice** q_1 di $F(q)$ (cioè se è $F(q_1) = 0$, $F'(q_1) \neq 0$, con $q_0 < q_1$ ed inoltre $F(q) > 0$ per $q_0 \leq q < q_1$) (vedi Fig.14.17a). **il sistema mobile va in un tempo finito** (dato da $t_1 = \int_{q_0}^{q_1} \frac{dq}{\sqrt{F(q)}}$), **sempre muovendosi in un medesimo verso, dalla posizione iniziale** q_0 **alla posizione finale** q_1**, dopodiché inverte il suo moto** (vedi Fig.14.17b).

A4 - Se $F(q_0) > 0$ e la posizione iniziale q_0 è compresa fra due zeri semplici consecutivi q_1, q_2 $[q_1 < q_0 < q_2]$ di $F(q)$ **il sistema mobile oscilla indefinitamente fra le posizioni estreme** q_1, q_2**, cioè il moto è periodico con periodo** (vedi Fig.14.18a e Fig.14.18b):

$$T = 2 \int_{q_1}^{q_2} \frac{dq}{\sqrt{F(q)}}. \tag{14.102}$$

Figura 14.17: Confronto fra i due grafici dell'andamento del moto nei due piani $(q, F(q))$ e $(t, q(t))$, nell'ipotesi $F(q_0) > 0$, $\dot{q}_0 > 0$ e q_1 zero semplice di $F(q)$.

Figura 14.18: Confronto fra i due grafici dell'andamento del moto nei due piani $(q, F(q))$ e $(t, q(t))$, nell'ipotesi $F(q_0) > 0$, $\dot{q}_0 > 0$, $q_1 < q_0 < q_2$ con q_1, q_2 due zeri semplici consecutivi di $F(q)$.

A5 - Se $F(q_0) > 0$ ed esiste nel verso della velocità iniziale $\dot{q}_0 > 0$ **uno zero multiplo** q_* (di ordine $n \geq 2$) di $F(q)$ (cioè se è $F(q_*) = 0, F'(q_*) = 0, F''(q_*) = 0, \ldots\ldots, F^{(n-1)}(q_*) = 0, F^{(n)}(q_*) \neq 0$), con $q_0 < q_*$) **il sistema mobile, pur sempre con moto progressivo, si avvicina indefinitamente alla posizione** q_* **senza mai raggiungerla (moto a meta asintotica)** (vedi Fig.14.19a e Fig.14.19b):

$$\lim_{t \to +\infty} q(t) = q_*.$$

Figura 14.19: Confronto fra i due grafici dell'andamento del moto nei due piani $(q, F(q))$ e $(t, q(t))$, nell'ipotesi $F(q_0) > 0$, $\dot{q}_0 > 0$ e q_* zero multiplo di $F(q)$: moto a meta asintotica.

ANALISI DEL SECONDO CASO: $F(q_0) = 0$.

B1 - Se $F(q_0) = 0$ con q_0 **zero semplice** di $F(q)$, l'equazione del 1° ordine $\dot{q}^2 = F(q)$ è insufficiente a discutere il moto e bisogna ritornare all'originaria equazione lagrangiana

del 2° ordine.[23] **Per la "legge del moto incipiente"[24] però, un istante dopo quello iniziale, si verifica uno dei cinque casi precedenti**.

B2 - Se $F(q_0) = 0$ con q_0 **zero multiplo** di $F(q)$, **il sistema mobile rimane indefinitamente in equilibrio nella posizione iniziale** q_0 **e questa è posizione di equilibrio instabile**.

[23]Nel nostro caso, dell'equazione di Friedmann, tale equazione del 2° ordine è la $(14.42)_2$, cioè

$$\frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{k}{R^2} - \Lambda = 0,$$

avendo posto $c = 1$.

[24]Ricordiamo al lettore il significato di "moto incipiente" che si introduce nella dinamica di un punto materiale ma che possiamo adattare a situazioni più generali come nel caso di un generico sistema dinamico. Supponiamo che il sistema dinamico in considerazione sia quello dell'enunciato del teorema di Weierstrass ad un solo grado di libertà definito dalla coordinata lagrangiana q. Supponiamo inoltre che ad un dato istante t_0, il sistema si trovi in quiete, sia cioè $\dot{q}(t_0) = 0$; se per $t > t_0$ il sistema inizierà a muoversi, questo moto, chiamato **moto incipiente** (*moto che sta per nascere*), avrà la direzione ed il senso determinati da $\dot{q}(t) \neq 0$. Nell'istante t_0, in cui la velocità lagrangiana è nulla, $\dot{q}(t_0) = 0$, viene a mancare questa norma; ma ammessa la continuità, la direzione ed il senso del moto si possono desumere come limite di $\dot{q}(t)$ negli istanti immediatamente consecutivi a t_0. Ora si ha che:

$$\lim_{t \to t_0^+} \frac{\dot{q}(t)}{t - t_0} = \ddot{q}(t_0),$$

onde si conclude che *la direzione ed il senso del moto nell'istante t_0 sono determinati dall'accelerazione lagrangiana $\ddot{q}(t_0)$: in ciò consiste* **la legge del moto incipiente**.

14.20 Epilogo

A parte l'introduzione di Fisica Classica, in questo libro, che abbiamo voluto intitolare "Relatività per Principianti ", ci siamo sforzati di presentare nel modo più semplice possibile i fondamenti scientifici e concettuali delle due grandi teorie dovute al genio di Albert Einstein: la Teoria della Relatività Ristretta e la Teoria della Relatività Generale.
Dopo oltre un secolo dalla nascita di tali teorie, esse sono state studiate e sviscerate in tutti i modi possibili da eminenti fisici e matematici che ne hanno sviluppato molteplici conseguenze e messo in luce questioni che lo stesso Einstein probabilmente non si sarebbe immaginato. Noi ci siamo sforzati di attenerci fedelmente al suo pensiero originale illustrando, fra le altre cose, dettagliatamente le tre classiche prove sperimentali e studiato nei dettagli due delle più importanti soluzioni esatte delle equazioni gravitazionali: la soluzione esterna di Schwarzschild con le notevoli implicazioni allo studio affascinante di oggetti strani del nostro universo quali sono i **buchi neri** e le equazioni di Friedmann con le notevoli implicazioni allo studio di molteplici possibili **modelli di universo**. Ci siamo occupati, anche se brevemente, nell'APPENDICE al Cap.12 della cosiddetta linearizzazione delle equazioni gravitazionali di Einstein utilizzando l'ipotesi di campo debole la quale conduce ancora ad un sistema di 10 equazioni alle derivate parziali per i potenziali incogniti che però è lineare e quindi molto più maneggevole dell'originario non lineare. Questo procedimento dovuto sostanzialmente ad Einstein (1916) fa intravvedere la possibilità, almeno teoricamente, dell'esistenza di onde gravitazionali anche se, bisogna dire onestamente, che la questione è molto più complessa di quanto non si possa pensare ed il cui esame dettagliato esce fuori dagli scopi del libro: a noi basta avervi accennato.[25] Questi studi teorici hanno in ogni caso stimolato, per più di 50 anni, la ricerca della loro rivelazione sperimentale che, come abbiamo detto, ha avuto pieno successo soltanto nel 2015.
Molti sono gli argomenti che abbiamo dovuto tralasciare: la soluzione di Kerr per lo studio dei buchi neri rotanti, lo studio dei buchi neri carichi con la soluzione di Reissner- Nordstrøm, lo studio dettagliato delle lenti gravitazionali a cui abbiamo peraltro accennato, gli enigmi inerenti la materia oscura, l'energia oscura e l'accelerazione dell'espansione dell'universo, lo studio dei suoi stadi primordiali, ecc., ecc., ciascuno dei quali ormai riempie interi libri. Ma tali limitazioni sono inerenti al carattere ed allo scopo di *questo libro che è stato pensato e scritto per studenti universitari che si accingono per la prima volta allo studio delle meravigliose teorie della Relatività.* Speriamo almeno di aver aperto alcune finestre sull'immenso orizzonte che le teorie relativistiche hanno spalancato sul nostro Universo.
Lasciamo il giudizio al lettore, terminando con una frase di Einstein: *"Possa questo libro procurare a qualcuno alcune ore felici di stimolante meditazione!"*[26]

[25]Come abbiamo già detto all'inizio dell'APPENDICE al Cap. 12 il trattato, oggi più importante, sulle onde gravitazionali è l'opera in 2 volumi di Michele Maggiore (Professore di Fisica Teorica all'Università di Ginevra), *Gravitational waves*, volume 1: Theory and Experiments, Oxford University Press, 2008 - 554 pages; *Gravitational waves*, volume 2: Astrophysics and Cosmology, Oxford University Press, 2018, 848 pages.
[26]Albert Einstein, *Relatività, esposizione divulgativa*, ultima frase della Prefazione, Boringhieri Torino, 1967.

Capitolo 15

Bibliografia Commentata

La bibliografia, sia sulla Relatività Ristretta sia sulla Relatività Generale è immensa, ci limiteremo perciò ad indicare alcune fra le opere che abbiamo avuto modo di consultare più spesso. I riferimenti pertinenti ad ogni capitolo, sono stati inseriti nelle note a piè di pagina nel testo man mano se ne presentava l'occasione.

15.1 Testi divulgativi

Citeremo soltanto quattro testi a carattere divulgativo, che riteniamo vadano letti dallo studente, prima di tutto per il loro inestimabile valore didattico, ma anche per la grande levatura scientifica dei loro autori:

D1 **Albert Einstein**, *Relatività, Esposizione divulgativa*, Universale scientifica Bollati Boringhieri, Torino, Ristampato nel 2006.
Un'esposizione della teoria della relatività che porta il nome di Einstein garantisce che si tratta di un libro eccezionale per densità di idee, rigore concettuale e chiarezza di linguaggio: è un classico.

D2 **L. D. Landau - G.B. Rumer**, *Che cos'è la Relatività*, Editori Riuniti, Roma, IV edizione, 1981.
Questo delizioso volumetto opera del grande fisico russo Lev Davidovic Landau, premio Nobel per la fisica, e di uno dei suoi collaboratori, G.B. Rumer, si può considerare un classico della divulgazione scientifica. Probabilmente si tratta del più chiaro ed elementare di tutti i libri pubblicati finora per spiegare al più vasto pubblico la teoria della Relatività Ristretta.

D3 **Tullio Regge**, *Spazio, Tempo, Relatività*, Loescher Editore-Torino, 1981.
Testo di livello un po' più avanzato, ma di notevole pregio didattico.

D4 **E. F. Taylor and J.A. Wheeler**, *Spacetimes Physics*, Freeman, New York, 1992.
È una delle migliori introduzioni elementari alla Relatività Speciale.

15.2 Relatività Ristretta o Speciale

Fra i testi universitari che forniscono un'ottima trattazione ad un livello abbastanza elementare sono da citare:

RR1 **Wolfgang Rindler**, *Introduction to Special Relativity*, Oxford Science Publications, Clarendon Press-Oxford, second edition, 1991, reprinted, 1996.
Questo bellissimo libro è scritto da uno degli studiosi di Relatività fra i più conosciuti ed

apprezzati in campo internazionale. È un libro veramente adatto per un principiante, la trattazione è mantenuta ad un livello elementare ed è preoccupazione costante dell'autore presentare ogni argomento nella maniera più semplice possibile sia dal punto di vista concettuale sia da quello matematico.
Esiste una traduzione italiana di un precedente libro dello stesso autore, ancora valido, e cioè:
W. Rindler, *La Relatività Ristretta*, Casa editrice Cremonese, Roma, 1971; purtroppo non si trova più in commercio.

RR2 **Robert Resnick**, *Introduzione alla Relatività Ristretta*, Casa Editrice Ambrosiana, Milano, 1979.
Testo didatticamente ottimo ed estremamente chiaro. La trattazione è mantenuta ad un livello elementare, il libro ha però un limite in quanto si fa soltanto un cenno fugace allo spaziotempo di Minkowski, e perciò non si dà nessuna formulazione covariante, in termini tensoriali, sia della meccanica che dell'elettrodinamica. In compenso, il testo propone molti esercizi e quesiti da risolvere che costituiscono un ottimo stimolo per il lavoro individuale dello studente.

RR3 **Jürgen Freund**, *Special Relativity for Beginners*, World Scientific Publishing Co. Pte. Ltd. Singapore, 2008.
Testo didatticamente ottimo, quanto il precedente, la trattazione anche qui è mantenuta ad un livello elementare ed è corredata da svariati esercizi risolti; molti di quelli da risolvere sono corredati da soluzione riportata alla fine del volume. Certi risultati cruciali sono dimostrati più di una volta con approcci differenti: la trasformazione speciale di Lorentz in una dimensione spaziale è ricavata tre volte, la formula dell'effetto Doppler quattro volte, la trasformazione di Lorentz in 2 dimensioni spaziali è ricavata due volte; analogamente 2 volte l'unificazione delle forze elettriche e magnetiche. Seguire diverse strade per arrivare a scoprire certi fatti fondamentali può essere gratificante per i principianti.

I seguenti volumi sono di livello un po' più avanzato:

RR4 **Vincenzo Barone**, *Relatività*, Bollati Boringhieri-Torino, 2004.
Questo libro moderno offre un'ampia panoramica della Relatività Speciale con numerose applicazioni ed esercizi, la trattazione è didatticamente ottima, i riferimenti bibliografici sono dettagliati e molto aggiornati: si tratta di un libro fra i più completi sulla Relatività Speciale e contiene anche un capitolo introduttivo sulla Relatività Generale. Raccomandabile agli studenti che desiderano raggiungere una buona ed ampia preparazione.

RR5 **Nicholas M. J. Whoodhouse**, *Special Relativity*, Lecture Notes in Physics, Springer Verlag, Berlin, Heidelberg, Germany, 1992.
Questo volume di appena 86 pagine è un'ottima trattazione della Relatività Speciale svolta fin dall'inizio nell'ambito dello spaziotempo 4-dimensionale di Minkowski: è raccomandabile soprattutto per lettori impazienti di arrivare rapidamente al nocciolo dei concetti relativistici e che non amano una trattazione tradizionale fatta alla luce dell'evoluzione storica delle idee.

Terminiamo, questa breve carrellata bibliografica, citando alcuni testi ad un livello decisamente superiore a quello dei precedenti, e generalmente non adatti ad un principiante, ma che uno studioso serio di Relatività dovrebbe conoscere.
Questi libri sono considerati dagli esperti dei classici.

RR6 **L. D. Landau e E. M. Lifschitz**, *Teoria dei Campi*, Editori Riuniti, Roma (l'ultima ristampa è del 2010).

Questo libro è il 2° volume del famosissimo corso di *Fisica Teorica*, in 10 volumi, opera del grande fisico russo Lev Davidovic Landau (premio Nobel per la Fisica nel 1962) e del suo allievo Evgenij M. Lifschitz. I primi due capitoli del libro sono dedicati ad un'esposizione molto stringata, ma chiara, limpida ed efficace, della Relatività Ristretta di Einstein che viene utilizzata poi in tutto il libro.

RR7 **Wolfgang Pauli**, *Teoria della Relatività*, Boringhieri-Torino, ristampa del 2008.
È notevole il fatto che questo libro è stato scritto da Pauli, per la "Encyklopädie der Mathematischen Wissenschften", all'età di 21 anni, quando era ancora studente, ed è stato per lunghissimo tempo uno dei testi di riferimento sia sulla Relatività Speciale che Generale.

RR8 **Christian Møller**, *The Theory of Relativity*, Oxford University Press, Printed in India, New Delhi, second edition, 1972.
Ottimo manuale, l'esposizione è chiara ed esauriente; vi si tratta sia della Relatività Ristretta, sia della Relatività Generale.

RR9 **J. L. Synge**, *Relativity: The Special Theory*, North-Holland, Amsterdam, 1965.
Testo ad alto livello, che procede accuratamente lungo linee personali: bibliografia assai completa per l'epoca in cui è stato scritto.

Sono ottimi anche i seguenti volumi:

RR10 **J.L. Anderson**, *Principles of Relativity Physics*, Accademic Press, New York, 1967.

RR11 **H. Arzeliès**, *Relativistic Kinematics*, Pergamon Press, Oxford, 1966.

RR12 **H. Arzeliès**, *Relativistic Point Dynamics*, Pergamon Press, Oxford, 1972.

15.3 Calcolo Tensoriale e Geometria Riemanniana

Tutti i libri sulla **Relatività Generale** contengono le nozioni fondamentali di Calcolo Tensoriale e Geometria Differenziale, in particolare Riemanniana, necessari per la trattazione della teoria einsteiniana della gravitazione. Qui di seguito ci limiteremo ad indicare alcuni testi che abbiamo avuto modo di consultare con più frequenza.
Primi fra tutti, le dispense - ora non più reperibili - sulle quali uno di noi (F.B.) ha appreso, per la prima volta, il calcolo tensoriale ed i fondamenti della Relatività Generale:

T1 **Carlo Cattaneo**, *Introduzione alla teoria einsteiniana della gravitazione,* vol. 1° e vol.2°, Appunti dal Corso di Fisica Matematica, 1960-61, Libreria Eredi Virgilio Veschi, Università degli Studi di Roma.

T2 **Carlo Cattaneo**, *Appunti dal Corso di Teorie Relativistiche*, Anno Accademico 1969-70, La Goliardica società Coop. a.r.l. Città Universitaria - Roma.

Fra i classici, non possiamo non citare la bella esposizione di Tullio Levi-Civita, uno dei fondatori - assieme a Gregorio Ricci Curbastro suo maestro all'Università di Padova - del Calcolo Tensoriale. Le notazioni usate nel libro, ed in parte anche il linguaggio, sono ormai superati, ma il contenuto è ancora validissimo e l'esposizione di primo piano:

T3 **Tullio Levi-Civita**, *Lezioni di Calcolo Differenziale Assoluto,* Alberto Stock Editore, Roma 1925.

Un altro ottimo testo diventato anch'esso un classico è:

T4 **J. L. Synge and A. Schild**, *Tensor Calculus,* first published by University Toronto Press, General Publishing Company, Ltd., Ontario, Canada 1949; reprinted by Dover Publications, Inc., New York, 1969 and 1978 .

Fra i testi più moderni consigliamo:

T5 **Alexander Mishchenko and Anatoly Fomenko**, *A Course of Differential Geometry and Topology,* Mir Publishers Moscow, 1988.

T6 **Boris A. Dubrovin, Sergej P. Novikov and Anatolij T. Fomenko**, *Geometria Contemporanea,* Vol. 1°, Editori Riuniti, Roma - Edizioni Mir, Mosca, 1987.

Per la Geometria differenziale Classica sono ottimi i seguenti volumi:

T7 **Dirk J. Struik**, *Lectures on Classical Differential Geometry,* Second Edition, Dover Publications, Inc., New York, 1988.

T8 **James J. Stoker**, *Differential Geometry,* Wiley-Interscience, Wiley Classic Edition Published in 1989, New York.

T9 **Erwin Kreyszig**, *Differential Geometry,* Dover Publications, Inc., New York, 1991.

Infine, un'ampio trattato, in due volumi, di circa 1000 pagine, moderno, ottimo sotto tutti i punti di vista, con un'amplissima raccolta di bellissime figure a colori, ottenute con Mathematica, ed una bibliografia pressoché completa è quello di:

T10 **Renzo Caddeo e Alfred Gray**, *Lezioni di Geometria Differenziale su curve e superfici,* C.U.E.C. (*Cooperativa Universitaria Editrice Cagliaritana*), vol. I, 2001, vol. II, 2002.

Infine un'ottima e molto concisa introduzione al Calcolo Tensoriale e alla Geometria riemanniana, sufficiente per lo studio della Relatività Generale, si può trovare in

T11 **Moshe Carmeli**, *Classical Fields, General Relativity and Gauge Theory,* Capitolo 2, pagg. 20-83 John Wiley & Sons, New York, 1982.

15.4 Relatività Generale

Citiamo subito due testi adatti a principianti, la cui lettura è raccomandata agli studenti:

RG1 **Ray d'Inverno**, *Introduzione alla Relatività di Einstein*, Cooperativa Libraria Universitaria Editrice Bologna (CLUEB), 2001.
Si tratta della traduzione italiana di un corso che l'autore ha tenuto per parecchi anni all'Università di Southampton (Gran Bretagna) e tratta soprattutto della Relatività Generale, ma contiene anche una succinta ma buona esposizione della Relatività Speciale in 4 capitoli. È uno dei pochissimi libri che introduce la Relatività Speciale usando il cosiddetto calcolo k ideato dal famoso cosmologo inglese Hermann Bondi. La trattazione è didatticamente molto appropriata, lucida ed efficace soprattutto per studenti principianti, si procede per piccoli passi ma si arriva fino in fondo al nocciolo degli argomenti trattati. Contiene una amplissima bibliografia. Fortemente raccomandato.

RG2 **Bernard F. Schutz**, *A first Course in General Relativity*, second edition, Cambridge University Press, 2009.
L'autore è uno degli specialisti di Relatività fra i più apprezzati; attualmente direttore del Max Planck Institute for Gravitational Physics (Albert Einstein Institute). Il libro, molto

moderno, è caratterizzato da chiarezza e ottima leggibilità, combinate con appropriato rigore. Molto completo anche per quel che riguarda le applicazioni recenti della Relatività Generale, quali: buchi neri, lenti e onde gravitazionali, pulsars, ecc.

Una referenza fondamentale è il già citato:

RG3 **L.D. Landau e E. Lifschitz**, *Teoria dei Campi*, Editori Riuniti, Roma, 1976, ristampato nel 2010.
Oltre le cose già dette, riguardanti il capitolo sulla Relatività Ristretta, aggiungiamo ora che dal capitolo 3° al 9° vi si tratta in modo completo la Teoria del Campo Elettromagnetico; i capitoli dal 10° al 13° trattano esaurientemente, ma sempre in maniera piuttosto stringata ed essenziale, la teoria einsteiniana del Campo Gravitazionale, cioè la Relatività Generale. L'ultimo capitolo, il 14°, fornisce una buona introduzione alla Cosmologia Relativistica. Nell'esposizione si bada più alla fisica che alla matematica con uno stile incantevole caratteristico di tutti i libri di Landau: è unanimemente considerato un classico della letteratura mondiale sulla Teoria dei Campi.

RG4 **Ronald Adler, Maurice Bazin and Menahem Schiffer**, *Introduction to General Relativity*, Second Edition, McGraw-Hill, Inc. 1975.
Questo volume, veramente notevole dal punto di vista didattico, crediamo che si possa considerare, ancora oggi, come una delle migliori introduzioni alla Relatività Generale. Questa seconda edizione è stata notevolmente migliorata ed aggiornata rispetto alla prima del 1965. Purtroppo questo libro non è più reperibile nelle librerie ma si può trovare ancora nelle biblioteche universitarie più fornite.

RG5 **Hans C. Ohanian e Remo Ruffini**, *Gravitazione e Spazio-tempo*, prima edizione italiana condotta sulla seconda edizione americana, Zanichelli Editore, Bologna, 1997.
L'esposizione è chiara ed esauriente, vi si discutono anche di molteplici applicazioni della Relatività Generale all'Astrofisica ed alla Cosmologia. Nel 1° capitolo gli autori si occupano della Teoria gravitazionale di Newton in maniera piuttosto dettagliata, cosa che tutti gli autori a noi noti non fanno, ma che noi crediamo essenziale se si vuole fare un serio confronto con la teoria einsteiniana. Il testo è dotato di una notevole ed aggiornata bibliografia commentata e di svariati interessanti esercizi.

Altri due ottimi manuali moderni sulla Relatività generale sono:

RG6 **Sean M. Carroll**, *An Introduction to General Relativity, Spacetime and Geometry*, Addison Wesley, 2004.

RG7 **James B. Hartle**, *Gravity: An Introduction to Einstein's General Relativity*, Pearson New International Edition, 2014.

Terminiamo, citando alcuni testi ad un livello decisamente superiore a quello dei precedenti, e generalmente non adatti ad un principiante, ma che uno studioso di relatività dovrebbe conoscere.

RG8 **Misner C. W., Thorne K. S., John Archibald Wheeler**, *Gravitation*, Freeman, San Francisco, 1973.
Per molti anni, ed è ancor oggi, questo testo è (stato) considerato la Bibbia della Relatività Generale.

Altri ottimi testi canonici (anche se gli ultimi quattro ormai invecchiati) sono:

RG9 **Wolfgang Rindler**, *Relativity, Special, General, and Cosmological*, second edition, Oxford University Press, Oxford 2006.

RG10 **Hans Stephani**, *Relativity*, third edition, Cambridge University Press, 2004.

RG11 **Steven Weinberg**, *Gravitation and Cosmology*, John Wiley & Sons, New York 1972.

RG12 **Christian Møller**, *The Theory of Relativity*, Oxford University Press, Printed in India, New Delhi, second edition, 1972.

RG13 **J. L. Synge**, *Relativity: The General Theory*, North-Holland, Amsterdam, 1965.

RG14 **Richard C.Tolman**, *Relativity, Thermodynamics and Cosmology*, Oxford University Press, Oxford, England, 1934.

15.5 Autobiografia, Opere scelte e Biografie di Einstein

A1 **Albert Einstein**, *Autobiografia scientifica*, Universale Bollati Boringhieri, Torino, 2007.

A2 **Albert Einstein**, *Opere scelte*, A cura di Enrico Bellone, Universale Bollati Boringhieri, Torino, 1988.
Questo volume contiene gli scritti più importanti di Einstein, si raccomanda vivamente al lettore di leggere almeno i due articoli più significativi, quello del 1905 sulla Relatività Ristretta: *Sull'elettrodinamica dei corpi in movimento*, e quello del 1916 sulla Relatività Generale: *I fondamenti della teoria della Relatività Generale*.
Questi due articoli meravigliosi costituiscono ciascuno una pietra miliare nella storia della fisica.

Come si può immaginare, ci sono moltissime biografie di Einstein, ci limitiamo a citare quelle che a noi ci appaiono le migliori:

A3 **Abraham Pais**, *Sottile è il Signore....; La scienza e la vita di Albert Einstein*, Bollati Boringhieri-Torino, 1991. Si trova anche nella collana Universale scientifica del 2008.
Quest'opera di Pais è unanimemente considerata la migliore biografia non solo della vita di Einstein ma soprattutto della sua opera scientifica di cui discute minuziosamente l'evoluzione attraverso l'esame di moltissimi scritti originali. Molto bella.

A4 **Banesh Hoffmann**, *Albert Einstein Creatore e Ribelle*, Editore Bompiani-Milano, 1977.
Questa è un'ottima biografia fatta da un collega ed amico di Einstein ma con carattere molto diverso dall'opera di Pais. Scritto in modo piano e accessibile anche al lettore non specializzato, il libro illustra l'opera scientifica di Einstein senza trascurare il lato umano, il mondo degli affetti e dei suoi rapporti sociali. La lettura del libro è molto piacevole.

15.6 Cosmologia

Capitoli di Cosmologia, più o meno estesi, si possono trovare in moltissimi libri sulla Relatività Generale, in particolare buone esposizioni si trovano nei libri, già citati in precedenza, dei seguenti autori: **Landau-Lifschitz, d'Inverno, Adler-Bazin-Schiffer, Ohanian-Ruffini, Tolman, Weinberg, Rindler, Stephani, Schutz.** Ci limitiamo perciò ad aggiungere alcune opere notevoli esclusivamente dedicate alla Cosmologia come scienza a sé stante.

C1 **Dennis W. Sciama**, *Cosmologia moderna*, Oscar Mondadori, Milano, 1999.
Questo è un libro a carattere divulgativo, ma non banale, dovuto ad uno dei più eminenti cosmologi inglesi, che nonostante sia abbastanza datato, contiene un'esposizione di primo piano della Cosmologia, e perciò la sua lettura può essere ancora consigliata ad un principiante.

C2 **Jakov B. Zeldovic e Igor D. Novikov**, *Struttura ed evoluzione dell'Universo*, in 2 volumi, Editori Riuniti-Edizioni Mir, Mosca 1982.
Esposizione molto estesa ed approfondita, di livello universitario, dovuta a due eminenti cosmologi russi: purtroppo i libri sulla Cosmologia invecchiano molto rapidamente e questi due volumi non sfuggono a questa legge. Tuttavia la loro lettura pensiamo che sia ancora molto utile ad uno studioso serio che voglia occuparsi di problemi di Cosmologia.

Due volumi più moderni e, nel loro genere, molto approfonditi sono i seguenti:

C3 **Jamal Nazrul Islam**, *An Introduction to Mathematical Cosmology*, Second Edition, Cambridge University Press, 2002.

C4 **Jayant Vishnu Narlikar**, *An Introduction to Cosmology*, Third Edition, Cambridge University Press, 2002.
Gli autori di questi volumi (il primo del Bangladesh, il secondo Indiano di Bombay) sono stati tutti e due allievi della rinomata scuola britannica di cosmologia di Cambridge (U.K.). Narlikar è stato allievo del grande astronomo, astrofisico e cosmologo Fred Hoyle ed è uno dei più strenui sostenitori della *Teoria dello stato stazionario*, in contrapposizione con la, quasi universalmente accettata, *Teoria del Big Bang*.

15.7 Libri di Esercizi

La maggior parte dei testi sia di Relatività Ristretta sia di Relatività Generale, specialmente quelli americani, tradizionalmente propongono alla fine di ogni capitolo degli esercizi da risolvere, ma sono veramente pochi i testi contenenti raccolte di esercizi completamente risolti. Uno di questi è il libro di Lightman et al. che è stato il testo di riferimento per parecchi decenni. Contiene gli enunciati e le soluzioni di molti esercizi sulla Cinematica e sulla Dinamica relativistica, sulle trasformazioni di coordinate, sugli invarianti e sui tensori, sull'elettromagnetismo, su materia e radiazione, ecc. ecc. fino alla Relatività Generale ed alla Cosmologia.

E1 **Alan P. Lightman, William H. Press, Richard H. Price, Saul A. Teukolsky**, *Problem book in relativity and gravitation*, Princeton University Press, Princeton, New Jersey, 1975.

Altri due buoni testi di esercizi, di autori italiani, che possono essere consultati con profitto, sono:

E2 **Paolo Rossi**, *Meccanica Relativistica e Analitica, Compendio con 350 esercizi svolti,* Pisa University Press, 2006.

E3 **Fernando De Felice e Giovanni Preti**, *Esercizi di Relatività Generale*, Bollati Boringhieri, 2007.

Finito di stampare nel mese di febbraio 2021

Francesco Borghero, professore associato di Fisica Matematica (attualmente fuori ruolo) presso il Dipartimento di Matematica e Informatica dell'Università di Cagliari ha tenuto, per più di trent'anni, numerosi corsi di Meccanica Razionale per matematici, fisici ed ingegneri, Istituzioni di Fisica Matematica, Fisica Matematica, Relatività Speciale e Relatività Generale per matematici. Nella sua attività di ricerca si è occupato di problemi inversi della Dinamica Lagrangiana, della Meccanica dei fluidi, dell'Ottica Geometrica con applicazione a microonde superficiali. Si è anche occupato della struttura matematica delle equazioni gravitazionali di Einstein. Infine, nell'ambito della Termodinamica Estesa, si è interessato allo studio della propagazione di onde nei gas relativistici ed ultrarelativistici. I risultati scientifici ottenuti in questi campi sono tutti pubblicati su riviste nazionali ed internazionali.

Francesco Demontis è professore associato di Fisica Matematica presso il Dipartimento di Matematica e Informatica dell'Università di Cagliari dove attualmente insegna Meccanica 2 e Relatività presso il corso di studio in Matematica. La sua attività di ricerca è rivolta principalmente allo studio di equazioni integrabili nonlineari di evoluzione tramite la Trasformata Inversa Spettrale. In particolare, in questo contesto, si è occupato della determinazione delle soluzioni solitoniche dei modelli studiati. In misura minore, si è interessato di propagazione ondosa nei gas ultrarelativistici nell'ambito della Termodinamica Estesa e di problemi inversi in Ottica Geometrica. I risultati scientifici ottenuti in questi campi sono tutti pubblicati su riviste nazionali ed internazionali.

euro 30,00

ISBN 978-88-3312-030-0 (versione cartacea)
978-88-3312-031-7 (versione online)