

Capitolo 6

Calcolo tensoriale: introduzione elementare ed applicazione alla relatività speciale

Il calcolo tensoriale costituisce un capitolo della geometria differenziale, e potrebbe essere discusso in tale ambito mediante una trattazione sistematica di tipo geometrico. Qui ci limitiamo a una trattazione elementare, di tipo classico, proponendoci un obiettivo ben preciso: introdurre gli elementi del calcolo tensoriale che permettano di scrivere in forma tensoriale le equazioni di Maxwell e le equazioni di moto di una particella in campo elettromagnetico nell'ambito della relatività speciale (spaziotempo piatto, con metrica lorentziana). Prescindiamo quindi dagli elementi (derivata covariante e curvatura di una varietà) che sarebbero necessari per passare alla relatività generale.

Per non disperdersi converrà avere presente che, ai fini che ci proponiamo, bisognerà porre l'attenzione su un fatto centrale che riguarda gli spazi vettoriali, ovvero che quando si ha a che fare con uno spazio vettoriale V si viene immediatamente ad avere a che fare anche con un altro spazio vettoriale ad esso associato. Si tratta del duale V^* di V , i cui elementi sono chiamati **covettori**. Ma come mai avviene che di questo fatto non ci si accorge durante lo studio della fisica nei primi corsi universitari? La ragione è la seguente. Nella fisica elementare lo spazio vettoriale che si incontra è lo spazio euclideo tridimensionale, cioè lo spazio vettoriale \mathbb{R}^3 munito del ben noto prodotto scalare (euclideo). Ora, quello che succede in generale è che la presenza di un prodotto scalare introduce in maniera naturale un isomorfismo tra lo spazio vettoriale considerato V e il suo duale V^* , tra vettori e covettori. Inoltre, e questo è il punto rilevante, se il prodotto scalare è euclideo allora succede che l'isomorfismo è di natura tanto semplice che, con

una naturale scelta della base nello spazio vettoriale (una base ortonormale, come la familiare terna $\mathbf{i}, \mathbf{j}, \mathbf{k}$) un vettore e l'associato covettore sono individuati da componenti che addirittura coincidono. Questo è il motivo per cui, nello spazio ordinario, lavorando con basi ortonormali, non ci rendiamo mai conto del fatto che talvolta stiamo trattando con vettori e altre volte stiamo trattando con covettori.

Ma questo nascondimento non è più possibile in relatività speciale. Infatti in tal caso è pur vero che lo spaziotempo è ancora uno spazio vettoriale (si tratta ora di \mathbb{R}^4 anziché di \mathbb{R}^3), ma il prodotto scalare di cui esso è munito è ora pseudoeuclideo anziché euclideo: la lunghezza al quadrato di un vettore riferito a una base ortonormale ha la forma pseudoeuclidea

$$s^2 = c^2 t^2 - x^2 - y^2 - z^2$$

anziché l'analoga quadridimensionale della forma euclidea

$$l^2 = x^2 + y^2 + z^2 .$$

Questo fatto ha la conseguenza che il vettore e il covettore ad esso naturalmente associato non hanno più le medesime componenti: si trova che *le componenti spaziali del covettore hanno segno opposto a quelle del corrispondente vettore*. Quindi, anche se procediamo nella maniera più piatta possibile lavorando con le componenti in sistemi cartesiani ortogonali, siamo costretti a distinguere tra vettori e covettori. Questo fatto sembrerebbe costituire una complicazione.

Ma a questo punto interviene uno di quei fatti incredibili, apparentemente miracolosi, in cui quella che pareva una complicazione si tramuta in un pregio. È impossibile spiegare questo fatto senza che lo si veda. Per vederlo, occorre prima imparare a riconoscere la distinzione tra vettori e covettori, cioè (come si dice in gergo) occorre **imparare ad abbassare e ad alzare gli indici**. Allora si potrà vedere come le relazioni tra campi elettromagnetici e potenziali, che nello spazio ordinario hanno la forma altamente asimmetrica

$$\mathbf{H} = \text{rot } A , \quad \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \text{grad } \Phi ,$$

quando invece si passi allo spaziotempo e si introduca il covettore A_μ naturalmente associato al quadrivettore $A^\mu = (\Phi, \mathbf{A})$, assumono la semplicissima forma

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

con una ovvia identificazione degli elementi della matrice $F_{\mu\nu}$ (un tensore due volte covariante) con le componenti dei vettori \mathbf{E} e \mathbf{H} . Anzi, appare allora evidente che le quantità \mathbf{E} e \mathbf{H} , che nello spazio euclideo ordinario ci appaiono dei vettori, non sono veramente distinti, ma costituiscono una entità unica (questo, già lo sappiamo: la distinzione tra campo elettrico e

campo magnetico non è assoluta, ma dipende dal sistema di riferimento), e il loro confluire in un unico tensore (che costituisce una vera unità nello spaziotempo) è dunque la traduzione matematica di un fatto fisico. Di più: la natura tensoriale (due volte covariante) del tensore $F_{\mu\nu}$ determina automaticamente come si trasformano i campi sotto trasformazioni di Lorentz, proprio secondo la legge che Lorentz, Poincaré ed Einstein avevano dovuto stabilire con calcoli laboriosi di tipo forza bruta.

In conclusione, la natura pseudoeuclidea dello spaziotempo obbliga a distinguere tra vettori e covettori, e più in generale a identificare la natura tensoriale (nello spaziotempo) delle varie quantità considerate. Ma d'altra parte la geometria quadridimensionale dello spazio tempo permette di dare forma estremamente semplice ed elegante a formule che appaiono complicate quando ci si limiti alla geometria dello spazio ordinario e non si sia stati introdotti alla geometria dello spaziotempo. L'impatto con questa esperienza (semplificazione ed illuminamento delle formule nel passaggio dallo spazio allo spaziotempo) ha costituito un passaggio cruciale nella formazione culturale di gran parte degli scienziati. Quindi può forse valere la pena di compiere lo sforzo per familiarizzarsi con gli elementi del calcolo tensoriale che permettano di fare l'esperienza appena descritta.

6.1 Contravarianza: dalla misura delle grandezze fisiche alle componenti dei vettori

Nel calcolo tensoriale è essenziale distinguere tra quantità che variano in maniera covariante e altre che variano in maniera contravariante. Spesso, a un primo impatto con questa materia, questo fatto è causa di disagio. Pensiamo che il seguente esempio riguardante le misure delle grandezze fisiche, che prendiamo letteralmente dalle primissime pagine (anzi le primissime parole) del *Treatise* di Maxwell,¹ possa essere utile.

a) Misura delle grandezze fisiche. In sostanza si tratta di questo. Una grandezza fisica, quando venga fissata una unità (una ben definita grandezza della stessa natura), viene individuata mediante un numero positivo (il rapporto tra la grandezza considerata e la corrispondente unità²):

$$\text{grandezza} = \text{numero} \times \text{unita} . \quad (6.1.1)$$

Allora, quando si cambia unità, ad esempio per le lunghezze si passa dal metro al centimetro diminuendo di cento volte l'unità, si trova che la medesima grandezza (che non cambia, essendo, come si dice, un assoluto) viene rappresentata da un altro numero, che rispetto al primo cambia in maniera

¹J.C. Maxwell, *A treatise on electricity and magnetism*, Dover (New York, 1954), si veda pag. 1 e pag. 6.

²*How often the unit has to be taken*, nelle parole di Maxwell

inversa di come cambia l'unità. Infatti se si prende

$$\text{unita} = a \text{ unita}' , \quad (6.1.2)$$

sostituendo in (6.1.1) si trova

$$\text{grandezza} = \text{numero} \times \text{unita} = \text{numero} \times a \text{ unita}' ,$$

e per confronto con

$$\text{grandezza} = \text{numero}' \times \text{unita}'$$

si trova

$$\text{numero}' = a \text{ numero} \quad \text{o equivalentemente} \quad \text{numero} = a^{-1} \text{ numero}' .$$

Dunque abbiamo mostrato che

$$\text{se } \text{unita} = a \text{ unita}' \quad \text{allora} \quad \text{numero}' = a \text{ numero} ,$$

o equivalentemente

$$\text{se } \text{unita} = a \text{ unita}' \quad \text{allora} \quad \text{numero} = a^{-1} \text{ numero}' .$$

Diciamo che **il numero che rappresenta la grandezza contravaria rispetto a come varia l'unità**. Nelle parole di Maxwell (*Treatise*, pag. 6):

“In transforming the values of physical quantities determined in terms of one unit, so as to express them in terms of any other unit of the same kind, we have only to remember that *every expression for the quantity consists of two factors, the unit and the numerical part which expresses how often the unit is to be taken*. Hence the numerical part of the expression varies **inversely** as the magnitude of the unit, ...”

b) Vettori su una retta. Del tutto analoga è la situazione per quanto riguarda i vettori. Cominciamo a considerare il caso dei punti di una retta. Fissata una origine O , ogni punto P della retta è individuato da un vettore³ $\mathbf{x} = (P - O)$, e allora, fissato un vettore base \mathbf{e} (l'analogo della grandezza scelta come grandezza unitaria), ogni vettore può univocamente essere scritto come

$$\mathbf{x} = x \mathbf{e} .$$

Il numero x (che questa volta è un numero **reale**, cioè può essere anche negativo o nullo) che rappresenta il vettore \mathbf{x} rispetto al vettore base \mathbf{e}

³Maxwell, nelle prime pagine (pag 10) del *Treatise*, ricorda che il nome *vettore* proviene da *vehere*, che in latino significa trasportare (dove veicolo): “*The displacement of a point, represented by a straight line drawn from its original to its final position, may be taken as the typical vector quantity, from which indeed the name of Vector is derived.*”

viene detto **componente del vettore \mathbf{x} rispetto a quel vettore base**. Dunque, se si cambia vettore base introducendo un nuovo vettore base \mathbf{e}' , si ha anzitutto che anche il vecchio vettore base, come ogni altro vettore, potrà essere rappresentato nella nuova base, sicché esiste un unico numero reale a tale che

$$\mathbf{e} = a \mathbf{e}' , \quad (6.1.3)$$

e allora si ha

$$\mathbf{x} = x \mathbf{e} = xa \mathbf{e}' .$$

Pertanto, per confronto con $\mathbf{x} = x' \mathbf{e}'$, si trova che il cambiamento delle componenti segue la legge

$$x' = ax \quad \text{o equivalentemente} \quad x = a^{-1}x' . \quad (6.1.4)$$

Dunque, per confronto con la (6.1.3), abbiamo ancora che **la componente di una fissato vettore *contravaria* rispetto a come varia il vettore base**.

c) Componenti dei vettori in \mathbb{R}^n . L'ulteriore generalizzazione si ha poi quando si considerano i vettori $\mathbf{x} \in \mathbb{R}^n$. Sappiamo allora che si può scegliere in infiniti modi una base \mathbf{e}_i ($= 1, \dots, n$) sicché ogni vettore può essere univocamente rappresentato come combinazione lineare dei vettori base mediante opportuni coefficienti,⁴ detti le **componenti del vettore \mathbf{x}** :

$$\mathbf{x} = \sum_i x_i \mathbf{e}_i . \quad (6.1.5)$$

Si introduca ora una nuova base \mathbf{e}'_i . Allora ogni vettore \mathbf{e}_i della vecchia base, come qualsiasi altro vettore dello spazio \mathbb{R}^n , può essere rappresentato sulla nuova base mediante le sue n componenti. In altri termini, esiste una matrice A_{ik} tale che

$$\mathbf{e}_i = \sum_k A_{ik} \mathbf{e}'_k \quad (i = 1, \dots, n) . \quad (6.1.6)$$

Dunque, ancora per sostituzione nella relazione (6.1.5) e per confronto con la relazione

$$\mathbf{x} = \sum_i x'_i \mathbf{e}'_i \quad (6.1.7)$$

⁴Si noti bene che le componenti non hanno nulla a che fare con la nozione di ortogonalità. Ci muoviamo qui ancora al livello in cui uno spazio vettoriale è concepito in termini puramente algebrici, essendo definita la somma di due vettori e la moltiplicazione per uno scalare. sicché è definita la nozione di base vettoriale, ma non si introduce ancora la nozione di prodotto scalare. Se si vuole pensare in termini geometrici, le componenti di un vettore rispetto a una base in \mathbb{R}^3 possono essere pensate come ottenute mediante la regola del parallelogramma.

(perché anche \mathbf{x} può essere rappresentato sulla nuova base), si trova

$$x'_k = \sum_i x_i A_{ik} , \quad (6.1.8)$$

Per confronto con la relazione (6.1.6) si vede che anche in questo caso si ha una forma di contravarianza, ma ora un poco più complicata. Questa volta anzitutto non interviene solo un numero, ma una matrice A , che esprime la base vecchia sulla nuova. Si trova poi che anche le vecchie componenti del vettore \mathbf{x} si esprimono in termini di quelle nuove mediante una matrice, che risulta anzitutto essere la matrice inversa A^{-1} di A . Se però si osserva attentamente la disposizione degli indici e si vuole continuare ad usare la convenzione consueta dell'algebra per la moltiplicazione di matrici, si troverà che occorre fare intervenire la trasposta della matrice inversa. Per fortuna esiste una notazione particolarmente conveniente, che consente di procedere in una maniera automatica senza dovere tener conto di tutti questi fatti coinvolgenti matrici inverse o trasposte. Questa notazione verrà introdotta subito sotto. Osserviamo comunque che anche nel caso dei vettori in \mathbb{R}^n abbiamo trovato che le componenti dei vettori *contravariano*, cioè variano in una ben definita opportuna maniera che non coincide con il modo in cui variano i vettori base (questi ultimi essendo gli analoghi delle unità di misura delle quantità fisiche considerate da Maxwell).

6.2 Un approccio più generale: i campi vettoriali su varietà. Una notazione più conveniente: *divertissement* sulla *chain rule*

Esistono situazioni fortunate in cui la considerazione di un caso più generale risulta semplificare le cose rispetto alla considerazione di un caso particolare. Nella situazione presente, il caso particolare che abbiamo considerato è quello di uno spazio vettoriale $V = \mathbb{R}^n$. Il caso più generale che consideriamo ora è invece quello di una varietà M , in cui lo spazio $T_P M$ tangente ad M in un punto P prende il posto dello spazio vettoriale V . Si tratta di una situazione che già avevamo considerato nel capitolo sulle equazioni di Lagrange, in cui la varietà M era lo spazio delle configurazioni di un sistema meccanico. Consideriamo dunque una varietà (differenziabile) M . In ogni suo punto $P \in M$ esiste allora lo spazio tangente, che denotiamo con $T_P M$. Scegliamo una carta per descrivere localmente (in un intorno di P) la varietà. Questo significa che sulla carta sono assegnate delle coordinate x^1, \dots, x^n ed è implicitamente assegnata una "funzione di immersione" che ad ogni punto x^1, \dots, x^n della carta associa un punto P della varietà M . Nel caso della relatività speciale, in cui la varietà M è lo spazio vettoriale \mathbb{R}^4 che può essere rappresentato anche da una sola carta, un punto generico

P della varietà può pensarsi rappresentato dalle sue coordinate cartesiane in una carta, diciamole X^μ .⁵

Notazioni. Rispetto al capitolo sulle equazioni di Lagrange, abbiamo denotato con x^1, \dots, x^n anziché con q^1, \dots, q^n le coordinate sulla carta. Oltre a questa differenza di nome, vi è anche la differenza che ora, seguendo un uso ormai universale, abbiamo **posto in alto l'indice per le componenti delle coordinate nella carta**,

$$x^1, \dots, x^n ,$$

mentre riserveremo l'indice in basso per individuare i vettori base (questo comunque è un fatto che, come si vedrà sotto, verrà prodotto “gratis” come conseguenza della convenzione di porre in alto l'indice per le coordinate). Avremo quindi **indici in alto (indici di contravarianza) e indici in basso (indici di covarianza)**.

Aggiungiamo subito l'altra convenzione (ormai universalmente accettata) cui ci atterremo, ovvero la cosiddetta **convenzione di Einstein: si sottintende la sommatoria su due indici ripetuti, uno in alto e uno in basso**. Ad esempio,

$$\alpha_i v^i \equiv \sum_i \alpha_i v^i .$$

Ora, siamo già familiari col fatto che la scelta di una carta (cioè una scelta delle coordinate) comporta spontaneamente la scelta di una base in ogni spazio tangente $T_P M$. Si tratta infatti di quella che abbiamo chiamato **base coordinata** $\mathbf{e}_i \in T_P M$, definita da

$$\mathbf{e}_i = \frac{\partial P}{\partial x^i} , \quad (i = 1, \dots, n) , \quad (6.2.1)$$

in cui ogni vettore della base è tangente a una delle linee coordinate. Si noti in particolare come l'indice in alto al denominatore di $\frac{\partial}{\partial x^i}$ corrisponda, coerentemente, all'indice in basso per \mathbf{e}_i .

In effetti, la notazione comunemente usata ai nostri giorni è quella in cui il vettore base \mathbf{e}_i è addirittura denotato semplicemente con

$$\mathbf{e}_i = \frac{\partial}{\partial x^i} \equiv \partial_i , \quad (i = 1, \dots, n) . \quad (6.2.2)$$

Si tratta in sostanza dell'isomorfismo che associa a un vettore \mathbf{v} la derivata direzionale nella direzione di quel vettore.⁶

⁵Nel caso generale, scriveremo talvolta la formula di immersione nella forma $P = P(x^1, \dots, x^n)$. Ad esempio, nel suo libro sulla relatività generale, Dirac denota con \mathbf{y} un vettore dello spazio ambiente, e la formula di immersione con $\mathbf{y}(x^1, \dots, x^n)$.

⁶Ad ogni vettore \mathbf{v} si può associare la derivata direzionale nella direzione \mathbf{v} , che denotiamo con $\partial_{\mathbf{v}}$, che agisce su una funzione f come

$$\partial_{\mathbf{v}} f = v^i \frac{\partial f}{\partial x^i} \equiv v^i \partial_i f .$$

Ora, per \mathbf{v} fissato, la derivata direzionale è un operatore differenziale lineare che agisce su tutte le funzioni (differenziabili, a valori reali) definite sulla varietà che si considera.

Mostriamo ora come questa ambientazione, che si riferisce allo spazio tangente a una varietà (con la varietà descritta in una carta), anziché a un generico spazio vettoriale V , sia non soltanto più significativa,⁷ ma addirittura più conveniente anche quando si abbia a che fare con uno spazio vettoriale (è questo il caso particolare in cui $M = V$, e si usa un sistema di coordinate arbitrario). Il motivo è il seguente. Un cambiamento di coordinate (ovvero un cambiamento di carta) è descritto da funzioni

$$x^i = x^i(x'^1, \dots, x'^n), \quad (i = 1, \dots, n), \quad (6.2.3)$$

e dunque la relazione che esprime la base vecchia in funzione di quella nuova si ottiene dalla (6.2.3) semplicemente con la regola di derivata di funzione composta (o *chain rule*):

$$\mathbf{e}_i = \frac{\partial P}{\partial x^i} = \frac{\partial P}{\partial x'^k} \frac{\partial x'^k}{\partial x^i},$$

ovvero

$$\mathbf{e}_i = \frac{\partial x'^k}{\partial x^i} \mathbf{e}'_k. \quad (6.2.4)$$

Da questa relazione si ottiene poi immediatamente la legge di trasformazione delle componenti dei vettori (o meglio, dei campi vettoriali), esattamente come nel caso considerato più sopra. Basta esprimere un generico vettore \mathbf{v} sulla base vecchia tramite le sue componenti, sostituire poi per ogni vettore della base vecchia la sua espressione in termini della base nuova, e infine confrontare con la espressione del vettore \mathbf{v} sulla base nuova. Dalla *chain rule* si ha infatti

$$\mathbf{v} = v^i \frac{\partial P}{\partial x^i} = v^i \frac{\partial x'^k}{\partial x^i} \frac{\partial P}{\partial x'^k},$$

e allora per confronto con

$$\mathbf{v} = v'^k \frac{\partial P}{\partial x'^k}$$

si ottiene la legge di trasformazione delle componenti di un campo vettoriale $v'^k = \frac{\partial x'^k}{\partial x^i} v^i$, o equivalentemente (poiché il nome degli indici è irrilevante)

$$v'^i = \frac{\partial x'^i}{\partial x^k} v^k. \quad (6.2.5)$$

Viceversa, è facile verificare (si veda ad esempio A. Mischenko, A. Fomenko, *A course on differential geometry and topology*, MIR (Mosca, 1988)) che ogni operatore lineare che soddisfi la regola di Leibniz è la derivata direzionale lungo un opportuno vettore \mathbf{v} . Dunque i vettori tangenti a una varietà possono essere identificati con gli operatori di derivata direzionale (agenti sulle funzioni reali definite su quella varietà). D'altra parte, come ricordato sopra, ogni tale operatore differenziale ha come base gli operatori di derivata parziale $\frac{\partial}{\partial x^i}$.

⁷Perché questa è la situazione che si incontra ad esempio in relatività generale, in cui lo spaziotempo non è lineare, e quindi non è uno spazio vettoriale.

La trasformazione inversa, dalle componenti vecchie alle nuove, si scrive immediatamente senza alcun calcolo. Basta semplicemente scambiare il posto degli apici, e si ottiene

$$v^k = \frac{\partial x^k}{\partial x'^i} v'^i. \quad (6.2.6)$$

Esercizio (abbastanza significativo). Si ottenga la relazione (6.2.6) dalla (6.2.5) eseguendo esplicitamente la inversione (cioè considerando la (6.2.5) come un sistema di n equazioni lineari nelle n incognite v^k).

Suggerimento. Si tenti. Si tratta di usare la relazione

$$\frac{\partial x^l}{\partial x'^k} \frac{\partial x'^k}{\partial x^i} = \delta_i^l,$$

che comunque verrà spiegata più avanti.

Nota: modo diretto per ottenere la legge di trasformazione delle componenti dei vettori. Si noti che la relazione (6.2.5) può essere anche ottenuta direttamente pensando i vettori come vettori velocità,⁸ cioè con le componenti (rispetto alla base coordinata) ottenute derivando un movimento $x^i = x^i(t)$, ossia come $v^i = \frac{dx^i}{dt}$ (oppure come $u^\mu = \frac{dx^\mu}{ds}$ in relatività, dove s è il tempo proprio). Dunque dalla *chain rule* si ha direttamente

$$v^i = \frac{dx^i}{dt} = \frac{\partial x^i}{\partial x'^k} \frac{dx'^k}{dt},$$

ovvero $v^i = \frac{\partial x^i}{\partial x'^k} v'^k$, cioè proprio la (6.2.6).

Equivalentemente, si potrebbe intuitivamente pensare a un vettore come a uno spostamento infinitesimo, ottenuto differenziando le componenti che esprimono la posizione, e allora dalla *chain rule* per il differenziale si avrebbe

$$dx^i = \frac{\partial x^i}{\partial x'^k} dx'^k,$$

ovvero la legge di trasformazione (6.2.6). Vedremo in seguito che in effetti i differenziali dx^i devono piuttosto essere pensati come costituenti una base per i campi covettoriali.

Confrontiamo ora la legge di trasformazione (6.2.4) dei vettori base con la corrispondente legge di trasformazione (6.2.5) delle componenti di un vettore. Si vede allora che non abbiamo avuto bisogno di fare riferimento a inverse o trasposte di matrici. In ogni caso la matrice di trasformazione è la matrice jacobiana del cambiamento di variabili considerato, e quello che resta da stabilire è se le coordinate vecchie x^i siano al numeratore o al denominatore, cioè se intervenga la matrice $\frac{\partial x^i}{\partial x'^k}$ oppure la matrice $\frac{\partial x'^k}{\partial x^i}$, e con quali indici. La scelta naturalmente viene fissata dal procedimento seguito. Ma una **regola mnemonica** molto comoda, che si può usare anche in tutti i casi che incontreremo in seguito, è la seguente.

⁸È questa in effetti la definizione più profonda di vettore.

Regola mnemonica per le leggi di trasformazione. Guardiamo ad esempio la formula (6.2.4) per il cambiamento della base. A sinistra si ha una quantità non primata, con indice i libero (cioè su cui non si somma), che si trova in basso? Allora a destra la quantità che va in basso è x^i (cioè ancora la quantità non primata), e proprio con lo stesso indice, e al numeratore si va poi in conseguenza. Guardiamo ora la legge di trasformazione (6.2.5) delle componenti dei vettori (o dei vettori, come più semplicemente si dice). A sinistra si ha una quantità primata con un indice k libero in alto? Allora a destra la quantità che va in alto è ancora quella primata, cioè x'^k , e proprio con lo stesso indice, e al denominatore si va poi in conseguenza.

Le leggi prototipo di trasformazione sono dunque:

- 1) covarianza:

$$\partial_i = \frac{\partial x'^k}{\partial x^i} \partial'_k$$

- 2) contravarianza

$$\frac{dx^i}{dt} = \frac{\partial x^i}{\partial x'^k} \frac{dx'^k}{dt} \quad \text{o equivalentemente} \quad dx^i = \frac{\partial x^i}{\partial x'^k} dx'^k.$$

In conclusione, un campo vettoriale su una varietà M viene in linea di principio assegnato in maniera intrinseca mediante una funzione che ad ogni punto $P \in M$ associa un vettore $\mathbf{v} \in T_P M$. Concretamente però si sceglie una carta, ovvero un sistema locale di coordinate $\mathbf{x} \equiv (x^1, \dots, x^n)$, che si trasporta dietro implicitamente la corrispondente base coordinata $\mathbf{e}_i = \frac{\partial P}{\partial x^i}$. Dunque il campo vettoriale viene definito concretamente assegnando la n -upla v^i delle componenti (rispetto alla base coordinata) in funzione del punto della carta:

$$v^i = v^i(x^1, \dots, x^n). \quad (6.2.7)$$

Ovviamente si ha allora che, se si sceglie un'altra carta con coordinate (x'^1, \dots, x'^n) , il medesimo campo vettoriale sarà rappresentato da un'altra n -upla

$$v'^i = v'^i(x'^1, \dots, x'^n)$$

che sappiamo essere connessa alla vecchia dalla relazione di contravarianza (6.2.5). Nei testi classici un campo vettoriale veniva definito mediante l'assegnazione di una n -upla in una carta, con la condizione che al variare della carta (cioè delle coordinate) la n -upla variasse secondo la relazione di contravarianza. In effetti, come abbiamo appena mostrato, questa condizione sul modo di cambiare della n -upla al variare della carta è proprio la condizione che garantisce (quando si tenga conto del modo di variare della base coordinata), che il vettore in tal modo definito sia una quantità assoluta, cioè indipendente dalla carta scelta, perché si ha allora

$$v^i \mathbf{e}_i = v'^i \mathbf{e}'_i.$$

6.3 Covarianza e campi covettoriali. I funzionali lineari e le 1-forme differenziali

a) I campi covettoriali definiti nella maniera classica. Nei testi classici i campi covettoriali venivano definiti in completa analogia con i campi vettoriali. In una assegnata carta (o sistema di coordinate), un campo covettoriale viene definito assegnando una n -upla di componenti (rispetto alla base coordinata)

$$\alpha_1, \dots, \alpha_n$$

(**indice in basso**) in funzione del punto della carta:

$$\alpha_i = \alpha_i(x^1, \dots, x^n) \quad (i = 1, \dots, n) . \quad (6.3.1)$$

Si richiede però ora la condizione di covarianza, ovvero si richiede che, se si sceglie un'altra carta con coordinate (x'^1, \dots, x'^n) , allora il medesimo campo covettoriale debba essere rappresentato da un'altra n -upla

$$\alpha'_{i} = \alpha'_{i}(x'^1, \dots, x'^n)$$

connessa alla vecchia dalla **relazione di covarianza**, ovvero

$$\alpha'_{i} = \frac{\partial x^k}{\partial x'^i} \alpha_k . \quad (6.3.2)$$

b) Loro significato assoluto. È importante ora comprendere quale è il significato assoluto (cioè indipendente dalla carta) dei campi covettoriali. Questo si capisce quando si consideri, oltre al campo covettoriale α_i , anche un arbitrario campo vettoriale v^i , perché si trova allora immediatamente che vale la seguente

Proposizione. Si ha l'uguaglianza

$$\alpha'_{i} v'^i = \alpha_i v^i . \quad (6.3.3)$$

Lemma. Si ha

$$\frac{\partial x^k}{\partial x'^i} \frac{\partial x'^i}{\partial x^m} = \delta_m^k \quad (6.3.4)$$

Dimostrazione. Basta pensare x^k come funzione delle x'^i e queste a loro volta come funzioni delle x^l , ovvero

$$x^k = x^k(x'^i(x^m)) ;$$

si calcola poi $\frac{\partial x^k}{\partial x^l}$ mediante la *chain rule* e si osserva che

$$\frac{\partial x^k}{\partial x^m} = \delta_m^k .$$

Dimostrazione della proposizione. Per le assegnate proprietà di covarianza di α_i e di contravarianza di v^i si ha

$$\alpha'{}_i v'^i = \frac{\partial x^k}{\partial x'^i} \frac{\partial x'^i}{\partial x^m} \alpha_k v^m ,$$

e dunque, facendo uso del lemma,

$$\alpha'{}_i v'^i = \delta_m^k \alpha_k v^m = \alpha_k v^k$$

Dunque **la quantità $\alpha_i v^i$ è un assoluto**, è un numero che non dipende dalle coordinate scelte, ma soltanto dal covettore e dal vettore considerati. Possiamo quindi prendere una carta qualsiasi. Pensiamo ora fissato il covettore $\alpha \equiv \{\alpha_i\}$, e guardiamo al valore che il numero $\alpha_i v^i$ assume al variare del vettore $v \equiv \{v^i\}$. Quello che osserviamo è che questo numero dipende linearmente dal vettore $v \equiv \{v^i\}$, perché evidentemente si ha

$$\alpha_i (v^i + w^i) = \alpha_i v^i + \alpha_i w^i .$$

Dunque il numero che otteniamo è una funzione (a valori reali) avente per dominio lo spazio vettoriale di tutti i vettori. In altri termini, in ogni fissato punto della varietà M il covettore α è un funzionale lineare sullo spazio vettoriale $T_P M$. A questo punto è pertanto opportuno fare una digressione sui funzionali lineari.

c) Digressione sui funzionali lineari.⁹ Dato uno spazio vettoriale V di dimensione n , si dice **funzionale lineare su V** , o **covettore su V** , ogni funzione

$$\alpha : V \rightarrow \mathbb{R}$$

che abbia la proprietà di linearità, cioè la proprietà

$$\alpha (a\mathbf{v} + b\mathbf{w}) = a \alpha(\mathbf{v}) + b \alpha(\mathbf{w})$$

per ogni coppia di vettori $\mathbf{v}, \mathbf{w} \in V$ e ogni coppia di numeri reali a, b . Data una base \mathbf{e}_i , per ogni covettore α sono pertanto definiti i numeri

$$\alpha_i := \alpha(\mathbf{e}_i) ,$$

e questi individuano completamente il covettore α stesso, perché permettono di conoscere il risultato della sua azione su qualunque vettore $\mathbf{v} \in V$. Infatti, per la assunta proprietà di linearità, si ha

$$\alpha(\mathbf{v}) = \alpha(v^i \mathbf{e}_i) = v^i \alpha(\mathbf{e}_i) = v^i \alpha_i .$$

⁹Questa parte ha un carattere un poco più astratto delle parti precedenti, e pertanto, ai fini di una introduzione elementarissima, potrebbe anche essere tralasciata.

I covettori, per come sono stati definiti, hanno un aspetto in qualche modo astratto. Ma in effetti essi sono oggetti concretissimi, come ora mostriamo. Si tratta di riguardare alla definizione delle componenti di un vettore con altri occhi. Finora, data una base \mathbf{e}_i , abbiamo pensato ad un fissato vettore \mathbf{v} e abbiamo preso in considerazione le corrispondenti componenti v^i , ad esempio la prima componente v^1 . Ma riguardiamo ora alla cosa in un altro modo. Fissata la base, riguardiamo alla determinazione della prima componente come una operazione, una funzione, che agisce su **tutti** i vettori, estraendo, per ogni vettore, la sua prima componente. Quindi l'operazione di estrarre la prima componente dei vettori è una funzione avente per dominio lo spazio vettoriale V , ed è evidentemente lineare. In altri termini, **l'operazione di estrarre la prima componente è un funzionale lineare sullo spazio dei vettori, ovvero un covettore**. Anzi, avendo fissato una base in V , abbiamo in tal modo definito n covettori, che denoteremo con \mathbf{e}^i , che corrispondono ad estrarre la i -esima componente di ogni vettore.:

$$\mathbf{e}^i(v^k \mathbf{e}_k) = v^i \quad \text{o equivalentemente} \quad \mathbf{e}^i(\mathbf{e}_k) = \delta_k^i . \quad (6.3.5)$$

Si mostra facilmente che i covettori (con una naturale definizione di somma e di moltiplicazione per gli scalari)¹⁰ costituiscono uno spazio vettoriale di dimensione n , che viene detto **spazio duale** a V e viene denotato con V^* . Si mostra anche¹¹ che il duale del duale coincide con lo spazio vettoriale di partenza: $(V^*)^* = V$. Infine, covettori \mathbf{e}^i definiti dalla (6.3.5) sono una base di V^* e vengono detti costituire la **base duale** della base \mathbf{e}_i che era stata fissata in V .

Intermezzo: i covettori, gli iperpiani e i fronti d'onda. Un covettore α agente sullo spazio vettoriale V ha una interessantissima interpretazione geometrica nello spazio V stesso. Infatti pensiamo allo spazio vettoriale come a uno spazio di punti nella maniera familiare (della geometria affine), in cui ogni punto è l'estremo di un vettore spiccato da un altro punto fissato (l'origine). Consideriamo allora il

¹⁰La somma $\alpha + \beta$ dei covettori α e β è definita da

$$(\alpha + \beta)(\mathbf{v}) := \alpha(\mathbf{v}) + \beta(\mathbf{v}) \quad \forall \mathbf{v}$$

(il secondo membro è conosciuto, perché sono conosciuti i covettori α e β , e il primo membro definisce il covettore $\alpha + \beta$). Analogamente, la moltiplicazione per uno scalare a è definita da

$$(a\alpha)(\mathbf{v}) = a\alpha(\mathbf{v}) .$$

¹¹Infatti, si tratta di riguardare al numero $\alpha_i v^i$ in un nuovo modo. Sopra avevamo fissato α_i , e quindi $\alpha_i v^i$ era un'operazione che forniva un numero per ogni vettore v^i . Se invece si fissa v^i , allora $\alpha_i v^i$ associa un numero ad ogni covettore α_i , e questa operazione è evidentemente lineare. Dunque ogni fissato vettore $v \equiv \{v^i\} \in V$ agisce come un funzionale lineare sullo spazio V^* dei covettori. Dunque V è il duale di V^* , ovvero si ha $(V^*)^* = V$.

sottoinsieme $\Pi \in V$ dei punti definiti da

$$\alpha_i x^i = 0 \quad \text{ovvero} \quad \alpha(\mathbf{x}) = 0$$

(in termini operatoriali, si tratta del *nucleo* (*ingl. kernel*) del funzionale lineare α). Evidentemente si tratta di un iperpiano passante per l'origine. Dunque a ogni funzionale lineare corrisponde un iperpiano passante per l'origine. Viceversa, si vede immediatamente che un covettore è univocamente individuato (cioè è conosciuta la sua azione su qualsiasi vettore) quando sia assegnato un iperpiano passante per l'origine (pensato come suo nucleo) e un altro iperpiano ad esso parallelo, sui punti del quale il covettore assume il valore 1.

Si capisce così come i covettori siano connessi alla descrizione dei fronti d'onda. Ad esempio, nella funzione

$$u(\mathbf{x}) = \cos(k_i x^i) \quad \mathbf{x} \in \mathbb{R}^3$$

le quantità k_i sono le componenti di un covettore. Di solito questa relazione viene scritta nella forma

$$u(\mathbf{x}) = \cos(\mathbf{k} \cdot \mathbf{x})$$

in cui figura il prodotto scalare tra il vettore \mathbf{x} e un opportuno altro vettore \mathbf{k} . In effetti, avviene che, se lo spazio vettoriale V è munito di un prodotto scalare, allora esiste una corrispondenza biunivoca tra covettori e vettori, proprio definita dalla relazione

$$k_i x^i = \mathbf{k} \cdot \mathbf{x} .$$

Di questa corrispondenza biunivoca tra vettori e covettori che viene definita quando si ha a disposizione un prodotto scalare, ci occuperemo nel prossimo paragrafo. Osserviamo tuttavia che gli iperpiani di un generico spazio vettoriale sono definiti come abbiamo fatto sopra, mediante covettori, e questo è particolarmente significativo nel caso generale in cui non sia assegnato un prodotto scalare "naturale".

d) Le 1-forme differenziali, e i differenziali di funzioni, come campi covettoriali. Nelle trattazioni più elementari (riguardanti lo spazio fisico \mathbb{R}^3 munito del suo naturale prodotto scalare) abbiamo imparato che il gradiente di una funzione è un campo vettoriale. Data una funzione $f = f(\mathbf{x})$ (dove $\mathbf{x} = (x, y, z)$, con riferimento a coordinate cartesiane ortogonali), abbiamo definito il campo vettoriale (un vettore in ogni punto)

$$\text{grad } f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

e abbiamo definito il differenziale della funzione f come

$$df = \text{grad } f \cdot d\mathbf{x} .$$

Più in generale, abbiamo considerato le 1-forme differenziali ω definite da $\omega = \sum_i a_i(\mathbf{x}) dx_i$, che ora con le nostre convenzioni riscriviamo come

$$\omega(\mathbf{x}) = a_i(\mathbf{x}) dx^i .$$

Vogliamo ora convincere il lettore che, in effetti, in entrambi i casi si avrebbe a che fare con campi covettoriali anziché con campi vettoriali. Risulta tuttavia che nel caso dello spazio ordinario munito del suo naturale prodotto scalare esiste una corrispondenza biunivoca naturale tra covettori e vettori per cui risulta giustificata la familiare trattazione elementare. Ma già in relatività, con la sua geometria pseudoeuclidea dello spaziotempo, tale identificazione non è più possibile. Convieni quindi attenersi al caso più generale.

Per i coefficienti a_i della forma differenziale $\alpha_i dx^i$ abbiamo posto l'indice in basso perché essi sono proprio le componenti di un covettore. Questo è evidente nel caso particolare del differenziale di una funzione f , in cui si ha

$$a_i = \partial_i f ,$$

perché abbiamo ripetutamente osservato che l' n -upla $\partial_i f$ è covariante. Ma questo si capisce anche nel caso generale. Ciò è dovuto al fatto che nella definizione stessa di forma differenziale è sottinteso che si richiede che l'espressione $a_i dx^i$ abbia carattere assoluto, ovvero indipendente dalle coordinate.¹² Allora, poiché sappiamo che le quantità dx^i sono contravarianti, dalla condizione di absolutezza dell'espressione $a_i dx^i$ segue che le quantità a_i devono essere covarianti, cioè sono le componenti di un covettore (o di un campo covettoriale, cioè una legge che assegna un covettore in ogni punto).

Un punto sottile: i differenziali dx^i come covettori base. Per i campi vettoriali, abbiamo distinto tra componenti da una parte e vettori base dall'altra, e abbiamo introdotto la decomposizione su una base

$$\mathbf{v}(\mathbf{x}) = v^i(\mathbf{x}) \frac{\partial P}{\partial x^i} .$$

Vogliamo ora fare osservare che del tutto analoga è la decomposizione

$$\alpha(\mathbf{x}) = \alpha_i(\mathbf{x}) dx^i$$

di una forma differenziale, in quanto campo covettoriale. Ora, il posto del campo vettoriale $\mathbf{v}(\mathbf{x})$ è preso dal campo covettoriale $\alpha(\mathbf{x})$, e le quantità dx^i ne costituiscono gli elementi della base (si tratta della base duale (che abbiamo denotato con \mathbf{e}^i) rispetto alla base coordinata $\mathbf{e}_i = \frac{\partial P}{\partial x^i}$), mentre l' n -upla α_i definisce le componenti del covettore α su quella base. Occorre un poco di concentrazione per convincersi che le singole quantità dx^i sono covettori (e non componenti di vettori), analogamente a come le singole quantità $\frac{\partial P}{\partial x^i}$ sono vettori (e non componenti di covettori). Ci si convince allora che ogni covettore base dx^i risulta definito dalla proprietà

$$dx^i \left(v^k \frac{\partial P}{\partial x^k} \right) = v^i \quad \text{o equivalentemente} \quad dx^i \left(\frac{\partial P}{\partial x^k} \right) = \delta_k^i .$$

¹²Nel caso del differenziale df , si ha anzitutto che la funzione f stessa è un assoluto (cioè il suo valore in un punto non dipende dalle coordinate); poi si considera la differenza del valore della funzione f in due punti vicini (anche questo è un assoluto) e se ne prende la parte lineare nell'incremento. Del tutto analoga è la situazione per la forma differenziale $\alpha_i dx^i$. Essa viene infatti definita come una quantità assoluta (ad esempio il lavoro che si compie passando da un punto a un punto vicino), e si deve pensare che questa quantità sia definita intrinsecamente, cioè indipendentemente dalle coordinate.

Dunque dovremmo avere convinto il lettore che la n -upla $\partial_i f$ definisce le componenti di un covettore. Vedremo tuttavia nel prossimo paragrafo come mai avvenga che nello spazio ordinario \mathbb{R}^3 munito del noto prodotto scalare sia consistente (**quando si considerino coordinate cartesiane ortogonali**) pensare al gradiente di una funzione come definente un campo vettoriale.

6.4 Il prodotto scalare come funzionale bilineare. L'isomorfismo da esso indotto tra vettori e covettori; abbassamento e innalzamento gli indici.

Veniamo dunque ad illustrare il fatto che esiste un isomorfismo naturale (corrispondenza biunivoca, compatibile con la struttura lineare) tra V e V^* , cioè tra vettori e covettori, quando sia assegnato un prodotto scalare nello spazio vettoriale V .

a) Il prodotto scalare come funzionale bilineare; le sue componenti g_{ik} . Cominciamo dunque a ricordare che cosa è un prodotto scalare nello spazio vettoriale V . Per definizione, un prodotto scalare è anzitutto un'applicazione bilineare $g : V \times V \rightarrow \mathbb{R}$, ovvero una "macchinetta" con due entrate, diciamo $g(\cdot, \cdot)$, la quale produce un numero, $g(\mathbf{v}, \mathbf{w})$, per ogni coppia di vettori \mathbf{v}, \mathbf{w} ; la "macchinetta" deve essere lineare in \mathbf{v} se si fissa \mathbf{w} , e lineare in \mathbf{w} se si fissa \mathbf{v} . Inoltre, deve essere simmetrica ($g(\mathbf{v}, \mathbf{w}) = g(\mathbf{w}, \mathbf{v})$) e nondegenere.¹³ Vediamo anzitutto come questa proprietà di bilinearità porta spontaneamente ad associare al prodotto scalare una matrice con due indici di covarianza quando sia assegnata una base \mathbf{e}_i in V . Si tratta di quelle che si chiamano "le componenti della metrica", che sono definite semplicemente dai prodotti scalari tra i vettori base, ovvero da

$$g_{ik} := g(\mathbf{e}_i, \mathbf{e}_k) . \quad (6.4.1)$$

Queste vengono introdotte in maniera del tutto naturale. Infatti, considerando il prodotto scalare $g(\mathbf{v}, \mathbf{w})$ tra i due vettori $\mathbf{v} = v^i \mathbf{e}_i$ e $\mathbf{w} = w^k \mathbf{e}_k$, per la bilinearità del prodotto scalare g si ha

$$g(\mathbf{v}, \mathbf{w}) = g(v^i \mathbf{e}_i, w^k \mathbf{e}_k) = v^i w^k g(\mathbf{e}_i, \mathbf{e}_k) = g_{ik} v^i w^k .$$

Dunque per le componenti della metrica si trovano spontaneamente degli indici in basso, e questo corrisponde al fatto che tali componenti hanno proprio carattere "due volte covariante" (cioè si trasformano come il prodotto

¹³Ovvero, l'unico vettore ortogonale a tutti i vettori deve essere il vettore nullo; in formule

$$"g(\mathbf{v}, \mathbf{w}) = 0 \quad \forall \mathbf{w}" \quad \text{comporta} \quad \mathbf{v} = 0 .$$

$\alpha_i \alpha_k$). In altri termini, se si cambiano coordinate (nel modo consueto delle carte di una varietà) e si passa per i vettori base alla corrispondente base coordinata si trova

$$g_{ik} = \frac{\partial x'^l}{\partial x^i} \frac{\partial x'^m}{\partial x^k} g'_{lm} \quad (6.4.2)$$

Esercizio: carattere due volte covariante (delle componenti) del prodotto scalare. La dimostrazione della legge di trasformazione (6.4.2) è un utile esercizio, che può essere compiuto in due modi:

- 1) Si considera la quantità

$$g_{ik} v^i w^k$$

di cui si sa che ha carattere assoluto (è il prodotto scalare di due vettori, che non dipende dalla base). Ma sappiamo come si trasformano le componenti dei vettori v^i e w^k (contravariano), e allora l'assolutezza del risultato impone la doppia covarianza per le componenti g_{ik} .

- 2) Si usa la definizione (6.4.1) delle componenti g_{ik} , e a secondo membro si esprimono i vettori base \mathbf{e}_i in termini di quelli nuovi.

b) L'isomorfismo tra vettori e covettori indotto dal prodotto scalare. Mostriamo ora come il prodotto scalare introduce un naturale isomorfismo tra vettori e covettori, procedendo dapprima nella maniera tradizionale. Passeremo poi a un procedimento "più sofisticato e intrinseco". L'osservazione base è che le quantità

$$g_{ik} v^k \quad (i = 1, \dots, n),$$

costituiscono evidentemente una n -upla covariante (dovremmo averlo ormai imparato; si tratta di un altro esempio di quella che si chiama **regola della traccia** che illustreremo in generale nel prossimo paragrafo). Quindi abbiamo un covettore, che è univocamente individuato dal vettore $v \equiv \{v^i\}$. Anzi, per mettere ancor più in evidenza il fatto che questo covettore è l'immagine in V^* del vettore $v \equiv \{v^i\}$ in V , introduciamo la convenzione di denotare tale covettore con la stessa lettera del vettore di cui è immagine, solo con l'attenzione di mettere l'indice in basso, per rammentarci che si tratta ora di un covettore. È questa l'operazione di **abbassamento dell'indice**: si fa corrispondere al vettore con componenti v^i (in breve, al vettore v^i) il covettore con componenti (in breve il covettore) v_i definito da

$$v_i = g_{ik} v^k. \quad (6.4.3)$$

Ovviamente, esiste anche l'operazione inversa di **innalzamento dell'indice**. A tal fine si introduce la matrice inversa della matrice della metrica. L'uso universale è di denotarla semplicemente come g^{ik} (con due indici in

alto, perché si tratta di due indici di contravarianza, come abbastanza facilmente si verifica). Dunque la matrice g^{ik} , come inversa della matrice della metrica g_{ik} , è definita dalla proprietà

$$g^{ik} g_{km} = \delta_m^i . \quad (6.4.4)$$

Questa matrice ovviamente stabilisce la corrispondenza inversa a quella di abbassamento dell'indice, ovvero fa corrispondere a un covettore α_i il vettore (che denoteremo con la stessa lettera del covettore, solo con l'indice in alto, per rammentarci che è un vettore)

$$\alpha^i = g^{ik} \alpha_k . \quad (6.4.5)$$

Il fatto che questa sia la funzione inversa di quella dell'abbassamento dell'indice si riflette naturalmente nel fatto che dalla (6.4.3), usando la (6.4.4), si trova

$$g^{ik} v_k = v^i .$$

Esercizio: Interpretazione geometrica della componente v_i . Dovrebbe essere evidente che la componente covariante di un vettore, $v_i = g_{ik} v^k$, può essere equivalentemente definita da¹⁴

$$v_i = g(\mathbf{e}_i, \mathbf{v}) .$$

Dunque la componente covariante v_i di un vettore $v \equiv \{v^k\}$ è il prodotto scalare del vettore con il vettore base i -esimo. In particolare, se i vettori base sono normalizzati, allora v_i è la proiezione ortogonale del vettore \mathbf{v} sulla direzione di quel vettore base.

c) Il caso dello spazio ordinario (metrica euclidea) e quello dello spaziotempo (metrica pseudoeuclidea). Si noti in particolare come, in presenza di una metrica euclidea, esistono le basi ortonormali, cioè basi tali che $g_{ik} \equiv g(\mathbf{e}_i, \mathbf{e}_k) = \delta_{ik}$, e dunque in tali basi l'operazione di abbassamento od innalzamento di un indice non produce alcuna variazione: **Negli spazi euclidei le componenti di vettori e covettori isomorfi, se riferite a una base ortonormale, sono le medesime.** È questo il motivo per cui nella fisica e nella geometria elementari si può tralasciare di parlare di covettori (almeno se ci si riferisce a basi ortonormali).

Ma ciò non è più possibile in relatività speciale, in cui si considera lo spaziotempo riferito a una metrica pseudoeuclidea. Infatti, ciò significa che nelle basi ortonormali (corrispondenti dunque a sistemi inerziali in cui per la parte spaziale ci si riferisce a coordinate cartesiane ortogonali) la matrice

¹⁴Infatti si ha

$$g(\mathbf{e}_i, \mathbf{v}) = g(\mathbf{e}_i, v^k \mathbf{e}_k) = v^k g(\mathbf{e}_i, \mathbf{e}_k) = g_{ik} v^k .$$

della metrica e la sua inversa hanno la forma (con la consueta notazione per gli indici, $\mu, \nu = 0, 1, 2, 3$.)

$$\{g_{\mu\nu}\} = \{g^{\mu\nu}\} = \text{diag}(1, -1, -1, -1),$$

ovvero

$$g_{00} = g^{00} = 1, \quad g_{11} = g^{11} = g_{22} = g^{22} = g_{33} = g^{33} = -1,$$

mentre tutte le altre componenti sono nulle. **Dunque nello spaziotempo l'innalzamento o l'abbassamento di un indice spaziale comporta un cambiamento di segno, mentre non cambia il segno se si abbassa o alza l'indice temporale.**

Esempio fondamentale. Se si considera il “vettore-evento” di componenti

$$\{x^\mu\} = (ct, \mathbf{x}) \equiv (ct, x, y, z)$$

si ha che il covettore ad esso corrispondente secondo la metrica lorentziana ha componenti

$$\{x_\mu\} = (ct, -\mathbf{x}) \equiv (ct, -x, -y, -z).$$

Altro esempio. Ammettiamo di avere stabilito che in elettromagnetismo il potenziale scalare Φ e il potenziale vettore \mathbf{A} siano rispettivamente la componente temporale e quella spaziale di un quadivettore A^μ , ovvero si abbia

$$\{A^\mu\} = (\Phi, \mathbf{A}) \equiv (\Phi, A_x, A_y, A_z),$$

o più concretamente

$$A^0 = \Phi, \quad A^1 = A_x, \quad A^2 = A_y, \quad A^3 = A_z. \quad (6.4.6)$$

Allora si ha $\{A_\mu\} = (\Phi, -\mathbf{A})$, ovvero

$$A_0 = \Phi, \quad A_1 = -A_x, \quad A_2 = -A_y, \quad A_3 = -A_z. \quad (6.4.7)$$

d) Modo intrinseco di riguardare all'isomorfismo tra vettori e covettori. Abbiamo già ricordato che il prodotto scalare $g(\cdot, \cdot)$ è un funzionale bilineare, una “macchinetta” che per ogni coppia di vettori \mathbf{v}, \mathbf{w} produce un numero, ed è lineare nel secondo se si fissa il primo, lineare nel primo se si fissa il secondo. Dunque, se si fissa \mathbf{v} , allora $g(\mathbf{v}, \cdot)$ definisce un funzionale lineare su V , cioè un covettore. Risulta in tal modo definita un'applicazione naturale da V in V^* , che ad ogni vettore $v \in V$ associa il covettore $g(\mathbf{v}, \cdot)$, e si mostra facilmente che tale corrispondenza è biunivoca¹⁵. È questo l'isomorfismo naturale tra V e V^* indotto dal prodotto scalare. Detto in altri

¹⁵Questa è la traduzione della proprietà di nondegenerazione.

termini: dato un prodotto scalare, a ogni vettore \mathbf{v} corrisponde un covettore (funzionale lineare $\alpha_{\mathbf{v}}$) che è semplicemente l'operazione (la funzione) "prendere il prodotto scalare con il fissato vettore \mathbf{v} ": in formule

$$\alpha_{\mathbf{v}}(\mathbf{w}) = g(\mathbf{v}, \mathbf{w}) \quad \text{per ogni } \mathbf{w} \in V .$$

Questa formula si scrive anche

$$\alpha_{\mathbf{v}}(\cdot) = g(\mathbf{v}, \cdot) .$$

Viceversa, un vettore \mathbf{v} è conosciuto se è dato il suo prodotto scalare $g(\mathbf{v}, \mathbf{w})$ con ogni altro vettore \mathbf{w} .

Esercizio: Mostrare che l'isomorfismo così definito coincide con quello definito dall'innalzamento e dall'abbassamento dell'indice.

Svolgimento. Sappiamo già che si ha $g(\mathbf{v}, \mathbf{w}) = g_{ik}v^i w^k$. D'altra parte, se si pensa \mathbf{v} fissato, è naturale scrivere

$$g_{ik}v^i w^k = \alpha_k w^k, \quad \text{con } \alpha_k = g_{ik}v^i$$

ed è chiaro che $\{\alpha_k\}$ è proprio il covettore $g(\mathbf{v}, \cdot)$ che si ottiene fissando \mathbf{v} nel prodotto scalare. Per questo motivo, come abbiamo già osservato, le componenti del covettore $\alpha \equiv \alpha_{\mathbf{v}}$ vengono denotate addirittura con v_k , perché sono le componenti del covettore univocamente associato al vettore \mathbf{v} , e dunque si scrive $v_k := g_{ik}v^i$. Anzi, per la simmetria del prodotto scalare, $g(\mathbf{v}, \mathbf{w}) = g(\mathbf{w}, \mathbf{v})$, da cui $g_{ik} = g_{ki}$, si scrive anche $v_k = g_{ki}v^i$ o equivalentemente

$$v_i = g_{ik}v^k .$$

È questa l'operazione di *abbassamento dell'indice*.

6.5 Definizione generale dei campi tensoriali

. Da quanto precede dovrebbe apparire giustificata la definizione classica dei campi tensoriali, che veniva data nel modo seguente (si noti che un tensore viene definito mediante le sue componenti, ad esempio un vettore mediante le sue componenti v^i , sottintendendo che ci si riferisce alla base coordinata individuata da una scelta delle coordinate)

Definizione: Data una varietà M e una sua carta locale con coordinate x^1, \dots, x^n , un tensore (o meglio un campo tensoriale) di tipo r, s (ovvero r volte contravariante, s volte covariante) è individuato da componenti $T_{i_1, \dots, i_s}^{j_1, \dots, j_r}$ con la proprietà che, al cambiare della carta (ovvero sotto trasformazione delle coordinate), le nuove componenti sono date da $T'_{i_1, \dots, i_s}^{j_1, \dots, j_r}$, dove

$$T'_{i_1 \dots i_s}^{j_1 \dots j_r} = \frac{\partial x'^{j_1}}{\partial x^{l_1}} \dots \frac{\partial x'^{j_r}}{\partial x^{l_r}} \frac{\partial x^{m_1}}{\partial x'^{i_1}} \dots \frac{\partial x^{m_s}}{\partial x'^{i_s}} T_{m_1 \dots m_s}^{l_1 \dots l_r} \quad (6.5.1)$$

Questa proprietà, abbastanza complicata a scriversi, rispecchia la più significativa definizione geometrica seguente: un tensore di tipo r, s su uno

spazio vettoriale V è nient'altro che un funzionale multilineare

$$T : \underbrace{V^* \times \dots \times V^*}_{r \text{ volte}} \times \underbrace{V \times \dots \times V}_{s \text{ volte}} \rightarrow \mathbb{R} .$$

Si noti che, conformemente alla osservazione fatta sopra in una nota, questa definizione geometrica comprende come caso particolare anche il vettore, pensato come funzionale lineare su V^* , ovvero $\mathbf{v} : V^* \rightarrow \mathbb{R}$.

Nel caso delle varietà M , il posto dello spazio vettoriale V viene preso dalla famiglia di spazi tangenti $T_x M$ con $x \in M$ e si ha una corrispondente definizione intrinseca per i tensori. Si verifica allora immediatamente che sotto cambiamento di coordinate le componenti del tensore T si trasformano nel modo suddetto (si ripensi alla definizione della metrica g e alla legge di trasformazione delle sue componenti g_{ik}).

Si noti anche come questi tensori di tipo r, s sono una generalizzazione del “tensore prototipo” di ordine zero, cioè lo scalare.¹⁶

È interessante notare che vale anche l'inverso di quanto detto sopra: se un certo ente è definito mediante componenti che si trasformano nel modo sopra indicato, allora tale ente ha significato geometrico, cioè è un funzionale multilineare del tipo detto sopra. Un esempio fondamentale che illustra questo fatto è quello del covettore, che abbiamo considerato poco sopra (si veda la (6.3.3)), e che ammonta a quanto segue. Dato uno spazio vettoriale V , sappiamo che i vettori hanno componenti che si trasformano con la regola $v'^i = \frac{\partial x'^i}{\partial x^k} v^k$. Se ora ammettiamo di avere un ente definito da componenti α_i che si trasformano con la legge $\alpha'_i = \frac{\partial x^k}{\partial x'^i} \alpha_k$, allora possiamo dimostrare che le componenti $\{\alpha_i\}$ definiscono un covettore. Ciò vuol dire che, per ogni vettore \mathbf{v} , si deve avere un risultato $\alpha(\mathbf{v})$ che non dipende dalla base scelta, ovvero si deve avere

$$\alpha_i v^i = \alpha'_k v'^k ,$$

come in effetti abbiamo verificato che si ha, in virtù della identità

$$\frac{\partial x^l}{\partial x'^i} \frac{\partial x'^i}{\partial x^k} = \delta_k^l . \tag{6.5.2}$$

L'esempio appena illustrato è un caso particolare della fondamentale

Regola della traccia (o della contrazione o della saturazione): Quando in una espressione contenente delle componenti tensoriali si somma su un indice

¹⁶Si tratta di una funzione a valori reali definita sulla varietà, diciamo $F : M \rightarrow \mathbb{R}$. Essa definisce un numero reale per ogni punto della varietà; e allora la forma funzionale della funzione f che rappresenta F deve necessariamente variare al variare delle coordinate proprio in maniera tale che non cambi il valore di F in corrispondenza di un definito punto della varietà. Ad esempio, se M è la retta reale, e x una coordinata, allora una funzione scalare $F : M \rightarrow \mathbb{R}$ sarà rappresentata da una funzione reale di variabile reale, diciamo $f = f(x)$. Se poi si passa a un'altra coordinata $x' = x'(x)$, allora la medesima funzione $F : M \rightarrow \mathbb{R}$ sarà rappresentata da una diversa funzione f' definita da $f'(x') = f(x(x'))$.

ripetuto che si trova una volta in alto (indice di contravarianza) e una volta in basso (indice di covarianza) (o, come si dice, *si satura un indice in alto con uno in basso*) si ottiene un tensore di due ordini in meno, in cui “sono scomparsi quei due indici”. Così le quantità $\alpha_i v^k$ individuano un tensore doppio $1 - 1$ (una volta contravariante, una volta covariante), ma se si esegue la saturazione dei due indici si ha la quantità $\alpha_i v^i$ che è uno scalare. Analogamente, se T_i^{kl} sono le componenti di un tensore di tipo $2 - 1$, allora T_i^{il} (avendo sommato su i) sono le componenti di un tensore di tipo $1 - 0$, ovvero di un vettore. Analogamente $g_{ik} v^k$ “è un covettore”, mentre $g^{ik} \alpha_k$ “è un vettore”.

Questa proprietà è una immediata conseguenza della definizione (6.5.1) di tensore e della identità (6.5.2).

6.6 Gli operatori differenziali e il problema della derivata covariante

Veniamo infine all'ultimo argomento di questa introduzione elementare al calcolo tensoriale: esso riguarda gli *operatori differenziali*. Cominciamo richiamando un esempio già illustrato: Data una funzione (scalare) f sulla varietà M , le quantità

$$\partial_i f := \frac{\partial f}{\partial x^i}$$

sono le componenti di un covettore (o meglio, di un campo covettoriale) ovvero si trasformano secondo la legge

$$\partial'_i f' = \frac{\partial x^k}{\partial x'^i} \partial_k f . \quad (6.6.1)$$

Abbiamo qui denotato con f' la funzione f in cui si è eseguito il cambiamento di variabili, e anche $\partial'_i f' \equiv \frac{\partial}{\partial x'^i} f'$.

Si potrebbe allora pensare ingenuamente che eseguendo un'operazione di derivazione su un tensore di tipo (r, s) si ottenga ancora un tensore di tipo $(r, s + 1)$, cioè che ogni operazione di derivazione aggiunga un indice di covarianza. Ma ciò non è vero, come mostra il seguente

Esempio. Per le derivate di un covettore α_m vale la legge di trasformazione¹⁷

$$\partial'_i \alpha'_k = \frac{\partial x^l}{\partial x'^i} \frac{\partial x^m}{\partial x'^k} \partial_l \alpha_m + \frac{\partial^2 x^m}{\partial x'^i \partial x'^k} \alpha_m . \quad (6.6.2)$$

Dimostrazione. Per il teorema di derivata di una funzione composta (6.6.1), e per la legge di trasformazione dei covettori, si ha

$$\partial'_i \alpha'_k = \frac{\partial x^l}{\partial x'^i} \partial_l \left(\frac{\partial x^m}{\partial x'^k} \alpha_m \right) = \frac{\partial x^l}{\partial x'^i} \frac{\partial x^m}{\partial x'^k} \partial_l \alpha_m + \frac{\partial x^l}{\partial x'^i} \partial_l \left(\frac{\partial x^m}{\partial x'^k} \right) \alpha_m .$$

¹⁷È significativo anche il caso particolare del gradiente, in cui $\alpha_m = \partial_m f$.

In tal modo è stato ottenuto il primo termine a secondo membro della (6.6.2). Nel secondo termine si usa poi

$$\frac{\partial x^l}{\partial x'^i} \partial_l = \partial'_i \equiv \frac{\partial}{\partial x'^i} .$$

Q.E.D.

Tuttavia la situazione è molto più semplice se ci si limita a considerare trasformazioni di coordinate che siano lineari. Si ha infatti il

Corollario. Se ci si limita a trasformazioni di coordinate $x'^i = x'^i(x^1, \dots, x^n)$ lineari, le quantità $\partial_i \alpha_k$ (in particolare le quantità $\partial_i \partial_k f$) si comportano come le componenti di un tensore due volte covariante.

Dimostrazione. Poiché anche la trasformazione inversa è lineare, si ha

$$\frac{\partial^2 x^m}{\partial x'^i \partial x'^k} = 0 .$$

Dunque nella legge di trasformazione (6.6.2) si annulla il secondo termine, e ci si riduce alla legge di trasformazione dei tensori due volte covarianti. **Q.E.D.**

Ci si rende conto immediatamente che una situazione analoga si presenta quando si considerano le derivate di un campo tensoriale.

Esiste però un modo generale per modificare l'ordinaria operazione di derivazione in modo che essa produca quantità geometriche (cioè aventi carattere tensoriale) agendo su dei tensori. Tale modificazione utilizza in maniera essenziale la presenza di una metrica. Questo procedimento fu inventato da Levi Civita nel 1916 e fu poi generalizzato da H. Weyl in caso di assenza di metrica (si tratta della **derivata covariante**)¹⁸

Dimenticandoci ora del problema generale della derivata covariante, ci basta qui avere constatato (nel caso delle derivate seconde, ma si vede subito che il risultato è generale) che, **se ci si limita a trasformazioni lineari (come le rotazioni nello spazio euclideo, e le trasformazioni di Lorentz nello spaziotempo), allora è vero che le operazioni di derivazione aggiungono altrettanti indici di covarianza.** Ad esempio, se ci si limita a trasformazioni lineari, allora $\partial_i \partial_k f$ (dove f è uno scalare) si comporta come un tensore due volte covariante (di tipo 0 – 2); analogamente, se v^i sono le componenti di un vettore (o meglio, di un campo vettoriale), allora $\partial_k v^i$ si comporta come un tensore di tipo 1 – 1, e $\partial_i v^i$ come uno scalare (la divergenza del campo vettoriale \mathbf{v}) e così via.

Si ha tuttavia una proprietà, riguardante l'operazione di derivazione, che è completamente indipendente dall'introduzione della derivata covariante. Si tratta della seguente

¹⁸Noi abbiamo implicitamente usato tale metodo nella deduzione dell'equazione di Lagrange.

Proposizione. Siano α_k le componenti di un covettore. Allora le quantità

$$F_{ik} = \partial_i \alpha_k - \partial_k \alpha_i$$

sono le componenti di un tensore doppio due volte covariante.

Dimostrazione. Se si scrive l'analogia della formula (6.6.2) per le quantità F'_{ik} , si vede subito che i due termini "impropri" contenenti le derivate seconde si cancellano, e si resta con la corretta formula di trasformazione per i tensori due volte covarianti.

Questa proprietà è particolarmente importante per l'elettromagnetismo. Infatti, considerando il covettore A_μ corrispondente al quadrivettore $A^\mu \equiv (\Phi, \mathbf{A})$ definito dal quadripotenziale, si viene a definire il tensore doppio due volte covariante

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (6.6.3)$$

Esso è evidentemente emisimmetrico ($F_{\mu\nu} = -F_{\nu\mu}$), sicché è individuato da 6 componenti. Verificheremo poco più sotto che le componenti indipendenti del tensore $F_{\mu\nu}$ sono proprio le componenti del campo elettrico \mathbf{E} e del campo magnetico (cambiato di segno) $-\mathbf{H}$. Il tensore $F_{\mu\nu}$ viene talvolta detto *Tensore di Faraday*.

Questo fatto ha un significato geometrico profondo, perché corrisponde al fatto che $F_{\mu\nu}$ rappresenta la 2-forma che è la derivata esterna (ovvero la generalizzazione del rotore) della 1-forma individuata da A_μ .

Possiamo infine concludere questo paragrafo con il seguente fondamentale esempio.

Esempio: Invarianza in forma dell'operatore dalembertiano sotto trasformazioni di Lorentz. Si osserva anzitutto che in un sistema inerziale (con coordinate spaziali cartesiane ortogonali), avendo la metrica $g_{\mu\nu}$ la forma diagonale $diag(1, -1, -1, -1)$, il dalembertiano (denotiamo $\partial_{tt}^2 = \partial_t \partial_t$ e così via)

$$\square := \partial_{tt}^2 - (\partial_{xx}^2 + \partial_{yy}^2 + \partial_{zz}^2)$$

si esprime nella forma

$$\square = g^{\mu\nu} \partial_\mu \partial_\nu .$$

Questa forma è quella buona, perchè satura due indici, uno in alto e uno in basso, e fornisce uno scalare. Sappiamo pertanto che, se ora passiamo ad un altro *arbitrario* sistema di coordinate¹⁹, si avrà

$$\square' = g'^{\mu\nu} \partial'_\mu \partial'_\nu$$

dove $g'^{\mu\nu}$ avrà una certa espressione fornita dalla regola

$$g'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\lambda} \frac{\partial x'^\nu}{\partial x^\sigma} g^{\lambda\sigma} .$$

¹⁹Sottintendiamo, ottenuto con trasformazioni lineari (come quelle di Lorentz), perchè altrimenti dovremmo introdurre la derivata covariante in luogo dell'ordinaria derivazione.

Tale espressione risulta in generale alquanto complicata. Ma se ci limitiamo a considerare trasformazioni di Lorentz, sappiamo che queste sono *isometrie*, ovvero sono tali che²⁰ $g'^{\mu\nu} = \text{diag}(1, -1, -1, -1)$, e dunque si ha ancora

$$\square' = \partial_{t't'}^2 - (\partial_{x'x'}^2 + \partial_{y'y'}^2 + \partial_{z'z'}^2) .$$

Si confronti ora questa dimostrazione con quella che abbiamo ottenuto nel capitolo precedente col metodo di forza bruta, in cui si compie materialmente il cambiamento di variabili dato dalla trasformazione di Lorentz, e si esprimono le derivate rispetto alle vecchie variabili in termini delle derivate rispetto alle nuove variabili. La semplicità della presente dimostrazione è impressionante.

Esercizio: Soluzioni dell'equazione di d'Alembert in forma di onde piane, ed effetto Doppler. Consideriamo l'equazione di d'Alembert (nell'incognita $u = u(t, \mathbf{x})$)

$$\partial^\mu \partial_\mu u = 0 \quad \text{ovvero} \quad g^{\mu\nu} \partial_\mu \partial_\nu u = 0 , \quad (6.6.4)$$

e cerchiamone una soluzione nella forma

$$u = A \exp[ik_\mu x^\mu] \quad (6.6.5)$$

con dei parametri k_μ liberi. Per ogni fissato k_μ si tratta di un' *onda piana* perché il luogo geometrico $u = \text{cost}$ è definito nello spaziotempo dalla condizione $k_\mu x^\mu = \text{cost}$, ovvero da un iperpiano. A sua volta (al modo solito), questo iperpiano nello spaziotempo corrisponde nello spazio ordinario a una famiglia di piani paralleli che traslano con una certa velocità. Questa viene determinata nel modo seguente.

Si osserva che la condizione che u soddisfi l'equazione di d'Alembert si traduce nella condizione $g^{\mu\nu} k_\mu k_\nu = 0$, ovvero

$$k^\mu k_\mu = 0 . \quad (6.6.6)$$

Dunque l'onda piana (6.6.5) soddisfa l'equazione di d'Alembert (6.6.4) soltanto se il quadrivettore k^μ è un *vettore nullo* (cioè ha pseudolunghezza nulla). La relazione (6.6.6) viene detta **relazione di dispersione** e la ragione è la seguente. Scriviamo $k_\mu x^\mu$ nella forma tradizionale

$$k_\mu x^\mu = \omega t - \mathbf{k} \cdot \mathbf{x} ,$$

il che vuol dire (ricordando $x^\mu = (ct, \mathbf{x})$) che il quadrivettore k^μ associato al covettore k_μ viene decomposto in parte temporale (frequenza angolare o pulsazione) e parte spaziale (vettore d'onda) come

$$\{k^\mu\} = \left(\frac{\omega}{c}, \mathbf{k}\right) .$$

Allora la (6.6.6) fornisce una relazione tra frequenza ω e vettore d'onda \mathbf{k} come avviene nella familiare *relazione di dispersione*. Nel caso qui considerato di soluzioni dell'equazione di d'Alembert nel vuoto, tale relazione ha dunque la forma $(\omega/c)^2 - \|\mathbf{k}\|^2 = 0$, ovvero

$$|\omega| = ck, \quad k = \|\mathbf{k}\| .$$

²⁰Più direttamente, avremmo $g'_{\mu\nu} = g_{\mu\nu}$, ma poi segue allora $g'^{\mu\nu} = g^{\mu\nu}$.

In altri termini, l'equazione di d'Alembert nel vuoto ha soluzioni della forma di onde piane normali al vettore d'onda \mathbf{k} (dunque con lunghezza d'onda $\lambda = 2\pi/k$) e con pulsazione ω se questi piani si spostano con velocità c e inoltre si ha $\omega = ck$.

Esercizio: Si deduca la formula per l'effetto Doppler relativistico per un "boost" di Lorentz (K' trasla con velocità v lungo l'asse x di K). Se (l, m, n) denotano i coseni direttori del vettore d'onda \mathbf{k} (ovvero si ha $k_x = l \omega/c$, $k_y = m \omega/c$, $k_z = n \omega/c$) allora si trovi

$$\omega' = \omega \gamma (1 - vl/c) . \quad (6.6.7)$$

Si ottengano anche le analoghe relazioni per i coseni direttori del vettore d'onda \mathbf{k} nel sistema K' . Si confronti l'articolo originale di Einstein, paragrafo 7. ²¹

6.7 Applicazione: L'elettromagnetismo in forma covariante (o tensoriale)

6.7.1 Forma covariante della relazione tra potenziali e campi: il tensore di Faraday.

La relazione tra potenziali e campi è l'ambito in cui il passaggio dal formalismo tridimensionale e quello quadridimensionale nello spaziotempo manifesta tutta la sua potenzialità; infatti la relazione tra potenziali e campi prende una forma di una semplicità e una simmetria stupefacenti.

Ricordiamo brevemente quanto avevamo già visto con il formalismo elementare tridimensionale. Avevamo introdotto i potenziali scalare Φ e vettore \mathbf{A} , che forniscono i campi \mathbf{E} e \mathbf{H} mediante le formule

$$\mathbf{E} = -\text{grad } \Phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} , \quad \mathbf{H} = \text{rot } \mathbf{A} , \quad (6.7.1)$$

come traduzione delle equazioni di Maxwell omogenee. Le equazioni inomogenee assumevano invece, nel gauge di Lorentz, caratterizzato da

$$\frac{1}{c} \frac{\partial \Phi}{\partial t} + \text{div } \mathbf{A} = 0 , \quad (6.7.2)$$

la forma di d'Alembert

$$\square \Phi = \rho , \quad \square \mathbf{A} = \mathbf{j}/c . \quad (6.7.3)$$

Osserviamo ora la potenza del formalismo quadridimensionale. Il primo passo consiste nell'osservare che la densità di carica ρ e la densità di corrente \mathbf{j} si mettono assieme a formare il quadrivettore densità di quadricorrente j^μ definito da²²

$$\{j^\mu\} \equiv (\rho c, \mathbf{j}) . \quad (6.7.4)$$

²¹Si faccia attenzione al fatto che nell'articolo originale Einstein denota con β il fattore di Lorentz che oggi tutti denotano con γ .

²²Si deve pensare che la densità di corrente \mathbf{j} associata a una particella coincide con $\rho \mathbf{v}$ dove \mathbf{v} è la velocità della particella e ρ la corrispondente densità di carica.

Tra l'altro, si ha allora che l'equazione di continuità $\frac{\partial \rho}{\partial t} + \text{div } j = 0$ assume la semplice forma

$$\partial_\mu j^\mu = 0 .$$

Più rilevante però è il fatto che segue allora che anche i potenziali costituiscono un quadrivettore, cioè si può porre

$$\{A^\mu\} \equiv (\Phi, \mathbf{A}) . \quad (6.7.5)$$

Questa è infatti coerente con le equazioni di Maxwell inomogenee (6.7.3), perché il dalembertiano è invariante.

La prima semplificazione di scrittura che si ottiene allora è che la condizione di Lorentz (6.7.2) viene scritta in forma quadridimensionale nella semplicissima e simmetrica forma

$$\partial_\mu A^\mu = 0 . \quad (6.7.6)$$

Inoltre, anche le relazioni tra potenziali e campi, dalla loro forma estremamente asimmetrica (6.7.1) vengono ad assumere una forma estremamente elegante e semplice nel formalismo quadridimensionale. Infatti si ha ad esempio

$$\begin{aligned} H_z &= \partial_1 A^2 - \partial_2 A^1 \\ E_z &= -\partial_0 A^3 - \partial_3 A^0 , \end{aligned}$$

e si osserva anzitutto che questa scrittura assume forma più simmetrica se si abbassano gli indici (e dunque cambiamo di segno alle componenti spaziali: $A^i = -A_i$, $i = 1, 2, 3$), perché allora si ha

$$\begin{aligned} -H_z &= \partial_1 A_2 - \partial_2 A_1 \\ E_z &= \partial_0 A_3 - \partial_3 A_0 . \end{aligned}$$

Considerando anche le altre relazioni in maniera analoga, è allora spontaneo introdurre le quantità $F_{\mu\nu}$ definite da

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (F_{\mu\nu} = -F_{\nu\mu}) , \quad (6.7.7)$$

che sappiamp costituire un tensore doppio due volte covariante. Esso è evidentemente antisimmetrico, e dunque è individuato da 6 componenti indipendenti (quante sono le componenti dei campi \mathbf{E} ed \mathbf{H}). Infatti si riconosce immediatamente che le componenti del tensore $F_{\mu\nu}$ sono date, in termini dei campi, da

$$\{F_{\mu\nu}\} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -H_z & H_y \\ -E_y & H_z & 0 & -H_x \\ -E_z & -H_y & H_x & 0 \end{pmatrix} . \quad (6.7.8)$$

Si vede dunque che il campo elettrico e il campo magnetico costituiscono una unità, il tensore doppio antisimmetrico $F_{\mu\nu}$ (detto talvolta **tensore di Faraday**). Un punto cruciale è che la struttura tensoriale del quadrivettore A^μ comporta automaticamente una struttura tensoriale per $F_{\mu\nu}$ (tensore due volte covariante), sicché viene automaticamente stabilito quale è la legge di trasformazione delle sue componenti (e quindi anche dei campi \mathbf{E} , \mathbf{H}) quando si compie una trasformazione di Lorentz sulle coordinate. Svolgeremo questo esercizio più sotto, verificando che si ottengono proprio le leggi di trasformazione già trovate nel precedente capitolo con il metodo elementare alla Lorentz, Poincaré ed Einstein (metodo forza bruta). Questo esempio dovrebbe illustrare in maniera sufficiente le parole di Einstein:

“Egli (Minkowski) invece riuscì ad introdurre un formalismo tale che *la forma matematica della legge garantisce di per sé l’invarianza della legge* stessa rispetto alle trasformazioni di Lorentz. Creando un calcolo tensoriale quadridimensionale, egli ottenne per lo spaziotempo ciò che il calcolo tensoriale aveva ottenuto per le tre dimensioni spaziali.”

Mostriamo anche quale forma assumono le equazioni di Maxwell in termini del tensore di Faraday. Preliminarmente, osserviamo che il tensore antisimmetrico $F^{\mu\nu}$ associato ad $F_{\mu\nu}$ secondo le regole per l’innalzamento degli indici (cambiano di segno le componenti tempo–spazio, e non le componenti spazio–spazio) è dato da

$$\{F^{\mu\nu}\} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -H_z & H_y \\ E_y & H_z & 0 & -H_x \\ E_z & -H_y & H_x & 0 \end{pmatrix}. \quad (6.7.9)$$

6.7.2 Le equazioni di Maxwell in termini del tensore di Faraday F .

Si ha la

Proposizione. In termini del tensore di Faraday $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ (e del suo corrispondente tensore contravariante $F^{\mu\nu}$) le equazioni di Maxwell inomogenee si scrivono (in coordinate cartesiane ortogonali rispetto alla metrica di Lorentz) nella forma

$$\partial_\mu F^{\mu\nu} = j^\nu/c \quad (\nu = 0, 1, 2, 3), \quad (6.7.10)$$

mentre le equazioni omogenee assumono la forma

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0, \quad (\lambda, \mu, \nu = 0, 1, 2, 3). \quad (6.7.11)$$

Dimostrazione. Per le equazioni inomogenee la verifica è immediata. Per quanto riguarda quelle omogenee, osserviamo anzitutto che le equazioni che si ottengono in tal modo sono proprio in numero di 4, e corrispondono alle scelte possibili degli indici (λ, μ, ν) tutti diversi tra di loro, ovvero $(0, 1, 2)$, $(0, 1, 3)$, $(0, 2, 3)$,

(1, 2, 3) . Infatti si verifica facilmente che in tutti gli altri casi, per l'antisimmetria di F , $F_{\mu\nu} = -F_{\nu\mu}$, si ottiene l'identità $0 = 0$. Nei casi non banali si ha ad esempio:

$$(0, 1, 2) \rightarrow \partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} = 0$$

cioè la terza componente di $\frac{1}{c}\partial_t \mathbf{H} + \text{rot } \mathbf{E} = 0$. Analogamente per i casi (0, 1, 3) e (0, 2, 3). Infine si ha

$$(1, 2, 3) \rightarrow \partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12} = 0$$

ovvero $\text{div } \mathbf{H} = 0$.

Q.E.D.

Esercizio. Verificare che la legge generale di trasformazione delle componenti dei tensori, applicata al tensore di Faraday, fornisce per la trasformazione dei campi esattamente quella precedentemente trovata con il metodo di Lorentz, Poincaré ed Einstein.

Svolgimento. Si veda Landau Lifshitz, Teoria dei campi, paragrafo 24.

6.7.3 Particella in campo elettromagnetico: Equazioni di moto in forma covariante

Seguendo il procedimento induttivo sviluppato nel capitolo precedente, siamo già pervenuti all'assioma che l'azione hamiltoniana di una particella in campo elettromagnetico è data da

$$S = - \int (mc + \frac{e}{c}g(A, u))ds , \quad (6.7.12)$$

dove $g(A, u)$ è il prodotto scalare tra quadrivelocità u^μ e quadripotenziale A^μ ,

$$g(A, u) = g_{\mu\nu}A^\mu u^\nu = A_\mu u^\mu ,$$

ovvero

$$S = - \int (mc + \frac{e}{c}A_\mu u^\mu)ds . \quad (6.7.13)$$

Facendo uso della scrittura (6.7.12) per l'azione, abbiamo già ottenuto per una particella in campo elettromagnetico l'equazione di moto in forma tridimensionale, ovvero

$$\frac{d}{dt}(m\gamma\mathbf{v}) = e(\mathbf{E} + \frac{1}{c}\mathbf{v} \times \mathbf{H}) , \quad (6.7.14)$$

con il corrispondente teorema dell'energia

$$\frac{d}{dt}m\gamma c^2 = e\mathbf{E} \cdot \mathbf{v} , \quad (6.7.15)$$

e vogliamo ora scrivere le corrispondenti equazioni in forma covariante. Otterremo in tal modo quattro equazioni, di cui la componente spaziale coinciderà con la (6.7.14), mentre la componente temporale fornirà il teorema dell'energia (6.7.15).

Proposizione. L'equazione di moto per una particella in campo elettromagnetico, scritta in forma covariante, è data da

$$mc^2 a_\mu = eF_{\mu\nu}u^\nu \quad (\mu = 0, 1, 2, 3) \quad (6.7.16)$$

o equivalentemente da

$$mc^2 a^\mu = eF^{\mu\nu}u_\nu \quad (\mu = 0, 1, 2, 3) . \quad (6.7.17)$$

Dimostrazione. Nello svolgere i calcoli per determinare gli estremali dell'azione, può essere utile rappresentare i moti come curve (di tipo tempo) nello spazio-tempo, parametrizzate da un parametro generico λ anziché dal tempo proprio (o pseudolunghezza) s . I moti saranno dunque scritti nella forma $x^\mu = x^\mu(\lambda)$, e le corrispondenti velocità $\frac{dx^\mu}{d\lambda}$ saranno semplicemente denotate con un punto:

$$\dot{x}^\mu \equiv \frac{dx^\mu}{d\lambda} .$$

La ragione di questa scelta è il fatto che, quando si usa come parametro il tempo proprio, si sta imponendo implicitamente il vincolo che la curva sia percorsa con velocità di modulo uno. Se invece si usa un parametro generico, si è in una situazione in cui si cercano gli estremali “liberi” (cioè senza alcun vincolo). Ottenuta l'equazione per gli estremali, vedremo poi che essa assume una forma più semplice se si sceglie come parametro il tempo proprio, $\lambda = s$, sicché la velocità avrà la consueta espressione $u^\mu = \frac{dx^\mu}{ds}$.

Per quanto riguarda “l'elemento di linea” ds , ricordando la definizione $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$, se si usa un generico parametro λ si ha l'espressione

$$ds = \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} d\lambda .$$

Osservando inoltre che si ha

$$u^\mu ds = dx^\mu = \dot{x}^\mu d\lambda ,$$

l'azione hamiltoniana (6.7.13) prende la forma

$$S = \int L d\lambda$$

con lagrangiana L data da²³

$$L = -mc\sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} - \frac{e}{c}A_\mu\dot{x}^\mu .$$

Gli estremali dell'azione sono allora le soluzioni $x^\mu(\lambda)$ delle equazioni di Eulero-Lagrange

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0 . \quad (6.7.18)$$

Si calcola

$$\frac{\partial L}{\partial \dot{x}^\mu} = -mc \frac{\dot{x}_\mu}{\sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}} - \frac{e}{c}A_\mu ,$$

²³La lagrangiana $L^{(mecc)} = -mc\sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}$ è quella della particella libera.

$$\frac{\partial L}{\partial x^\mu} = -\frac{e}{c}(\partial_\mu A_\nu)\dot{x}^\nu .$$

Si vede dunque che è conveniente prendere $\lambda = s$, perchè in tal modo scompare il denominatore nell'espressione del momento, essendo $g_{\mu\nu}u^\mu u^\nu = 1$. Pertanto, sostituendo λ con s e quindi anche \dot{x}^μ con u^μ , il primo membro dell'equazione di Lagrange (6.7.18) prende la forma

$$\begin{aligned} \frac{d}{ds} \frac{\partial L}{\partial u^\mu} - \frac{\partial L}{\partial x^\mu} &= \\ -mca_\mu - \frac{e}{c}[(\partial_\nu A_\mu)u^\nu - (\partial_\mu A_\nu)u^\nu] &= \\ -mca_\mu + \frac{e}{c}(\partial_\mu A_\nu - \partial_\nu A_\mu)u^\nu &= \\ -mca_\mu + \frac{e}{c}F_{\mu\nu}u^\nu . \end{aligned}$$

Dunque le equazioni di Eulero-Lagrange hanno la forma

$$mc^2 a_\mu = eF_{\mu\nu}u^\nu .$$

Q.E.D.

Esercizio. Si verifichi che la parte spaziale delle equazioni di moto fornisce la nota equazione (6.7.14) di moto tridimensionale, che ha a secondo membro la forza di Lorentz. Invece, la parte temporale fornisce il teorema dell'energia (6.7.15). Si osservi poi come il teorema dell'energia sia conseguenza delle equazioni di moto tridimensionali, e come questo fatto corrisponda alla proprietà cinematica che la quadriaccelerazione è ortogonale alla quadrivelocità,²⁴ ovvero

$$g_{\mu\nu}a^\mu u^\nu \equiv a_\mu u^\mu = 0 .$$

Infine, si confronti la presente deduzione con il calcolo che era stato compiuto nel capitolo precedente per mostrare che la forza di Lorentz deriva da un potenziale dipendente dalla velocità. Si constaterà come qui la dimostrazione sia molto più semplice, non richiedendo di fare uso di una identità sugli operatori differenziali che lì interveniva.

Osservazione. Le medesime equazioni si ottengono prendendo la lagrangiana

$$L' = \frac{1}{2}m g_{\mu\nu}u^\mu u^\nu - \frac{e}{c}A_\mu u^\mu .$$

Ciò segue subito dalla osservazione già fatta nel capitolo precedente, a proposito delle due diverse azioni (e quindi delle due diverse lagrangiane) che si possono utilizzare per la particella libera.

²⁴Basta derivare rispetto ad s la relazione che esprime che la quadrivelocità ha lunghezza unitaria:

$$g_{\mu\nu}u^\mu u^\nu = 1 ,$$

usando $g_{\mu\nu} = g_{\nu\mu}$. Si ricordi anche che la proprietà della quadrivelocità di avere lunghezza unitaria è equivalente alla definizione stessa dell'elemento di linea,

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu$$

(formalmente, si divida tale relazione per ds^2).

