CAPITOLO 11

La derivata covariante

Nello studio dei campi vettoriali su varietà non euclidee (ad es. se si vogliono trattare le equazioni di Maxwell in uno spazio-tempo curvo) è necessario introdurre una generalizzazione delle operazioni differenziali su campi vettoriali (gradiente, divergenza, ecc.) Nasce però subito un problema.

L'idea di derivare un campo vettoriale implica il confronto tra i valori (vettoriali) del campo in due punti vicini, confronto che si fa senza difficoltà in una varietà euclidea, dove ha un senso immediato dire se un campo $\mathbf{u}$ ha lo stesso valore in due punti distinti A e B: basta ad es. verificare se le componenti cartesiane del vettore sono o no le stesse in A e in B. Ma se siamo in una varietà non euclidea le coordinate cartesiane non esistono: dovremmo allora ricorrere a un altro sistema di coordinate? Si otterrebbe un confronto — e quindi una derivata — dipendente dal sistema di coordinate scelto, ossia *non covariante*. Occorre invece definire il confronto in modo *covariante*: l'idea di Levi-Civita è che ciò possa farsi attraverso un'opportuna definizione del *trasporto parallelo* di un vettore.

Come vedremo procedendo nel discorso, ci sono tre concetti strettamente connessi tra loro:
- la derivata covariante
- il trasporto parallelo dei vettori
- la definizione di geodetica.

Ciascuno dei tre può essere preso come primitivo, deducendone gli altri. Va anche aggiunto che mentre le idee introdotte nei capitoli precedenti sono intrinseche alla struttura di varietà differenziabile (o C^∞), questo non è più vero per i nuovi concetti: da un lato, il fatto di possedere una derivazione covariante è un *arricchimento della struttura*; dall'altro, *non c'è un unico modo* d'introdurre tale proprietà in una varietà generica.

Definizione astratta

Un primo approccio possibile consiste nel dare una definizione astratta di derivazione covariante, e nel dedurne le definizioni di trasporto parallelo e di geodetica. Si procede così:

Definizione: Diremo *derivazione covariante* (sottinteso *simmetrica*) un'applicazione bilineare
$$\mathcal{V} \times \mathcal{V} \to \mathcal{V}, \quad (\mathbf{u}, \mathbf{v}) \mapsto \nabla_{\mathbf{u}} \mathbf{v}$$
che soddisfi i seguenti assiomi:

$$\begin{array}{lll} 1. & \nabla_{\mathbf{u}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{u} = [\mathbf{u}, \mathbf{v}] & \text{(simmetria)} \\ 2. & \nabla_{\mathbf{u}}(f\mathbf{v}) = f\, \nabla_{\mathbf{u}} \mathbf{v} + (\partial_{\mathbf{u}} f)\, \mathbf{v} & \text{(derivazione).} \end{array}$$

Osservazione 1: È possibile anche definire una derivazione covariante non simmetrica, ma questa estensione non occorre per i nostri scopi.

Osservazione 2: Dagli assiomi si deduce facilmente

$$\nabla_{f\mathbf{u}}\mathbf{v} = f\,\nabla_{\mathbf{u}}\mathbf{v}$$

il che significa che $\nabla_{\mathbf{u}}\mathbf{v}$ non è un'operazione differenziale su $\mathbf{u}$. Si potrebbe perciò fare a meno dell'ipotesi che $\mathbf{u}$ sia un campo, e limitarsi a supporlo un vettore definito in un punto A; la derivata covariante risulterà anch'essa un vettore definito in A.

Ogni volta che si sia scelta una base $\{\mathbf{e}_\alpha\}$ (più esattamente un campo di basi) il vettore $\nabla_{\mathbf{e}_\alpha}\mathbf{e}_\beta$ si può scrivere come combinazione lineare degli $\{\mathbf{e}_\alpha\}$:

$$\nabla_{\mathbf{e}_\alpha}\mathbf{e}_\beta = \Gamma^\gamma{}_{\beta\alpha}\,\mathbf{e}_\gamma \tag{11–1}$$

(attenzione all'ordine degli indici!) I coefficienti $\Gamma^\gamma{}_{\beta\alpha}$ si chiamano *coefficienti di connessione*; chiaramente la conoscenza di tutti i coefficienti di connessione determina la derivata covariante sulla varietà, e si capisce perciò come una varietà possa essere dotata in infiniti modi di una derivazione covariante (o come anche si dice, di una *connessione affine*).

Apparirà ora chiaro perché l'assioma 1 sia stato denominato "simmetria": infatti grazie ad esso *è sempre* $\Gamma^\gamma{}_{\beta\alpha} = \Gamma^\gamma{}_{\alpha\beta}$ *tutte le volte che* $\{\mathbf{e}_\alpha\}$ *è una base coordinata* (la dimostrazione è ovvia).

Poniamo $\mathbf{v} = v^\alpha \mathbf{e}_\alpha$: allora avremo

$$\begin{aligned}\nabla_{\mathbf{e}_\beta}\mathbf{v} &= \nabla_{\mathbf{e}_\beta}(v^\alpha\mathbf{e}_\alpha) = v^\alpha\,\nabla_{\mathbf{e}_\beta}\mathbf{e}_\alpha + (\partial_\beta v^\alpha)\,\mathbf{e}_\alpha\\ &= v^\alpha\,\Gamma^\gamma{}_{\alpha\beta}\,\mathbf{e}_\gamma + (\partial_\beta v^\gamma)\,\mathbf{e}_\gamma = (v^\gamma{}_{,\beta} + \Gamma^\gamma{}_{\alpha\beta}\,v^\alpha)\,\mathbf{e}_\gamma.\end{aligned}$$

Se definiamo

$$v^\gamma{}_{;\beta} \overset{\text{def}}{=} v^\gamma{}_{,\beta} + \Gamma^\gamma{}_{\alpha\beta}\,v^\alpha \tag{11–2}$$

potremo scrivere:

$$\nabla_{\mathbf{e}_\beta}\mathbf{v} = v^\gamma{}_{;\beta}\,\mathbf{e}_\gamma.$$

Dunque $v^\gamma{}_{;\beta}$ è la componente γ della derivata covariante di $\mathbf{v}$ rispetto a $\mathbf{e}_\beta$.

Trasporto parallelo

Diremo che il campo vettoriale $\mathbf{v}$ è *parallelo* rispetto al vettore $\mathbf{u}$ nel punto P se $\nabla_{\mathbf{u}}\mathbf{v} = 0$ in P. Ciò posto, diremo che il campo vettoriale $\mathbf{v}$ è *trasportato parallelamente lungo una curva* γ se $\mathbf{v}$ è parallelo in ogni punto di γ rispetto al suo vettore tangente. Si vede subito che in termini di componenti ciò richiede

$$u^\beta\,v^\alpha{}_{;\beta} = 0 \quad \text{su } \gamma.$$

Se la curva è descritta dalle equazioni parametriche $x^\alpha = x^\alpha(\lambda)$, abbiamo $u^\beta = dx^\beta/d\lambda$, e quindi

$$\frac{dv^\alpha}{d\lambda} + \Gamma^\alpha{}_{\gamma\beta}\,\frac{dx^\beta}{d\lambda}\,v^\gamma = 0.$$

Quest'ultimo è un sistema di equazioni differenziali del primo ordine per le componenti di $\mathbf{v}$: dunque il trasporto parallelo determina $\mathbf{v}$ su tutta γ a partire da un suo punto.

Osservazione: Non si deve ritenere che se due curve γ_1 e γ_2 hanno in comune gli estremi A e B, il trasporto da A a B debba dare lo stesso risultato su entrambi i cammini: su questo punto ritorneremo.

Se si assume come primitiva la nozione di trasporto parallelo di un vettore lungo una curva, si può definire la derivata covariante come segue. Dato un vettore $\mathbf{u}$ in A e un campo vettoriale $\mathbf{v}$, prendiamo una curva γ (con parametro λ) che abbia $\mathbf{u}$ come vettore tangente in A. Indichiamo con $\mathbf{v}'$ il risultato del trasporto parallelo di $\mathbf{v}$ da un generico punto P di γ in A, e poniamo

$$\nabla_{\mathbf{u}}\mathbf{v} \stackrel{\mathrm{def}}{=} \lim_{\lambda\to 0} \frac{1}{\lambda}\bigl(\mathbf{v}' - \mathbf{v}(A)\bigr). \tag{11–3}$$

Occorre naturalmente dimostrare che la definizione è corretta, ossia: *a*) che il risultato dipende solo da $\mathbf{u}$, e non da γ; *b*) che ha le proprietà richieste dalla definizione di derivata covariante. Ciò richiede anzitutto di definire un opportuno insieme di assiomi per il trasporto parallelo: tralasciamo qui di dare tutti i necessari dettagli.

Geodetiche

Si definisce *geodetica* una curva lungo la quale si trasporta parallelamente il vettore tangente. In formule, se $\mathbf{u}$ è il vettore tangente alla curva γ, questa è una geodetica sse in ogni suo punto

$$\nabla_{\mathbf{u}}\mathbf{u} = 0$$

ovvero in componenti:

$$u^\beta\,u^\alpha{}_{;\beta} = 0.$$

In termini delle equazioni parametriche della curva, questa si scrive:

$$\frac{d^2x^\gamma}{d\lambda^2} + \Gamma^\gamma{}_{\alpha\beta}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\beta}{d\lambda} = 0. \tag{11–4}$$

Da un punto di vista intuitivo, questa definizione di geodetica formalizza l'idea di una curva che procede "sempre diritta," per quanto è possibile in uno spazio curvo. Da un punto di vista formale, è importante osservare che l'equazione (11–4) delle geodetiche è del secondo ordine: ne segue che *una geodetica è completamente caratterizzata assegnando un punto e il vettore tangente.* Da un punto partono perciò infinite geodetiche (tante quanti vettori tangenti).

Osservazione: Occorrerà dimostrare che la definizione di geodetica data in precedenza tramite un principio variazionale coincide con quella attuale; ma in realtà ciò è vero solo sotto le seguenti condizioni:

– che la varietà possieda una metrica

– che la connessione affine sia *compatibile* con la metrica.

Vedremo più avanti l'esatto significato di tutto ciò.

Come abbiamo già detto, è possibile definire la struttura affine a partire dalle geodetiche: per dimostrarlo, basta far vedere come si definisca il trasporto parallelo, supposte date le geodetiche. Questo si fa con la seguente *costruzione di Schild* (fig. 11–1). Sia γ una curva qualsiasi, $A = \gamma(0)$ e $B = \gamma(1)$ due suoi punti, $\mathbf{v}(A)$ il vettore che si vuole trasportare in B.

1. S'individuano su γ i punti $A_k = \gamma(k/n)$ $(k = 0 \ldots n)$.
2. Si fa partire da $A_0 = A$ la geodetica tangente a $\mathbf{v}(A)$, e si prende su di essa il punto C_0 corrispondente a un certo $\bar{\lambda}$.
3. Si costruisce la geodetica che unisce C_0 con A_1, e tale che su di essa $\lambda = 0$ in C_0 e $\lambda = \bar{\lambda}$ in A_1 (si può dimostrare che tale geodetica esiste ed è unica se i punti A_1 e C_0 sono "abbastanza vicini").
4. Si prende su questa geodetica il punto D_0 corrispondente a $\lambda = \bar{\lambda}/2$.
5. Si costruisce la geodetica che unisce A_0 con D_0, e tale che su di essa $\lambda = 0$ in A_0 e $\lambda = \bar{\lambda}/2$ in D_0 (v. in 3 circa esistenza e unicità).
6. Si prende su questa geodetica il punto C_1 corrispondente a $\lambda = \bar{\lambda}$.
7. Si costruisce la geodetica che unisce A_1 con C_1, e tale che su di essa $\lambda = 0$ in A_1 e $\lambda = \bar{\lambda}$ in C_1.
8. Si ripetono i passi da 3 a 7 per $k = 1 \ldots n - 1$.
9. Il vettore tangente in $B = A_n$ all'ultima geodetica costruita è il risultato *approssimato* del trasporto parallelo di $\mathbf{v}(A)$ in B.
10. Si fa il limite della costruzione precedente per $n \to \infty$, $\bar{\lambda} \to 0$.

Com'è ovvio, il valore della costruzione di Schild non sta nella sua realizzabilità pratica, ma nella dimostrazione che le geodetiche sono un punto di partenza possibile per definire una connessione affine.

Tensori

Oltre ai vettori (campi) e alle forme differenziali, faremo uso in seguito di altri enti, che occorre definire: i tensori di diversi tipi (ranghi).

Un tensore di rango $\binom{p}{q}$ può essere definito in diversi modi: ad es. come un'applicazione lineare

$$\mathbf{T} : \mathcal{T}_A^{*\times p} \times \mathcal{T}_A^{\times q} \to \mathrm{I\!R}, \quad (\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_p, \mathbf{v}_1 \ldots \mathbf{v}_q) \mapsto \mathbf{T}(\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_p, \mathbf{v}_1 \ldots \mathbf{v}_q). \quad (11\text{–}5)$$

In particolare, un tensore di rango $\binom{1}{1}$ è un'applicazione (sott. lineare)

$$\mathbf{T} : \mathcal{T}_A^* \times \mathcal{T}_A \to \mathbb{R}, \quad (\boldsymbol{\sigma}, \mathbf{v}) \mapsto \mathbf{T}(\boldsymbol{\sigma}, \mathbf{v}). \tag{11–6}$$

Se però leggiamo la (11–6) tenendo fisso $\boldsymbol{\sigma}$, essa definisce un'applicazione lineare in $\mathcal{T}_A$, ossia una forma differenziale $\boldsymbol{\sigma}'$, che dipende linearmente da $\boldsymbol{\sigma}$: in altre parole, facendo ora variare $\boldsymbol{\sigma}$ in $\mathcal{T}_A^*$, abbiamo un'applicazione lineare $\mathcal{T}_A^* \to \mathcal{T}_A^*$. Questa è un'altra possibile definizione di un tensore di rango $\binom{1}{1}$.

Possiamo anche operare al rovescio, fissando $\mathbf{v}$: arriviamo così a un'applicazione lineare $\mathcal{T}_A \to \mathcal{T}_A$, che dà una nuova definizione dello stesso oggetto. Si capisce quindi che per un tensore di rango $\binom{p}{q}$ con p e q generici le possibilità si moltiplicano, e potranno occasionalmente riuscirci utili.

Osserviamo che *anche vettori e forme possono essere visti come tensori*: i vettori sono tensori di rango $\binom{1}{0}$, le forme sono di rango $\binom{0}{1}$. Lo si vede facilmente applicando la definizione generale di tensore.

Dai tensori definiti in un punto A si passa senza difficoltà ai *campi tensoriali*: non occorre entrare in dettagli.

Prodotto tensoriale, componenti

A partire da due tensori $\mathbf{T}$, $\mathbf{U}$, risp. di rango $\binom{p}{q}$, $\binom{r}{s}$, se ne può costruire un terzo, detto *prodotto tensoriale* $\mathbf{W} = \mathbf{T} \otimes \mathbf{U}$, al modo seguente: $\mathbf{W}$ ha rango $\binom{p+r}{q+s}$ e

$$\begin{aligned}\mathbf{W}(\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_p, \boldsymbol{\sigma}_{p+1} \ldots \boldsymbol{\sigma}_{p+r}, \mathbf{v}_1 \ldots \mathbf{v}_q, \mathbf{v}_{q+1} \ldots \mathbf{v}_{q+s}) = \\ \mathbf{T}(\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_p, \mathbf{v}_1 \ldots \mathbf{v}_q)\, \mathbf{U}(\boldsymbol{\sigma}_{p+1} \ldots \boldsymbol{\sigma}_{p+r}, \mathbf{v}_{q+1} \ldots \mathbf{v}_{q+s}).\end{aligned}$$

La definizione di prodotto tensoriale permette di costruire in modo canonico una base nello spazio dei tensori di dato rango a partire dalle basi $\{\mathbf{e}_\alpha\}$, $\{\boldsymbol{\omega}^\beta\}$ di $\mathcal{T}_A$, $\mathcal{T}_A^*$: la base consiste dei tensori

$$\mathbf{e}_{\alpha_1} \otimes \ldots \otimes \mathbf{e}_{\alpha_p} \otimes \boldsymbol{\omega}^{\beta_1} \otimes \ldots \otimes \boldsymbol{\omega}^{\beta_q}.$$

Un generico tensore di rango $\binom{p}{q}$ è combinazione lineare di questi:

$$\mathbf{T} = T^{\alpha_1 \ldots \alpha_p}{}_{\beta_1 \ldots \beta_q}\, \mathbf{e}_{\alpha_1} \otimes \ldots \otimes \mathbf{e}_{\alpha_p} \otimes \boldsymbol{\omega}^{\beta_1} \otimes \ldots \otimes \boldsymbol{\omega}^{\beta_q}.$$

Se poi abbiamo

$$\boldsymbol{\sigma}_i = \sigma_{i\alpha}\, \boldsymbol{\omega}^\alpha, \quad \mathbf{v}_j = v_j^\beta\, \mathbf{e}_\beta \qquad (i = 1, \ldots, p \quad j = 1, \ldots, q)$$

ne segue

$$\mathbf{T}(\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_p, \mathbf{v}_1 \ldots \mathbf{v}_q) = T^{\alpha_1 \ldots \alpha_p}{}_{\beta_1 \ldots \beta_q}\, \sigma_{1\alpha_1} \ldots \sigma_{p\alpha_p}\, v_1^{\beta_1} \ldots v_q^{\beta_q}.$$

Una volta introdotte le componenti, è facile mostrare che se $\mathbf{W} = \mathbf{T} \otimes \mathbf{U}$, le componenti di $\mathbf{W}$ sono

$$W^{\alpha_1\ldots\alpha_p,\alpha_{p+1}\ldots\alpha_{p+r}}{}_{\beta_1\ldots\beta_q,\beta_{q+1}\ldots\beta_{q+s}} = T^{\alpha_1\ldots\alpha_p}{}_{\beta_1\ldots\beta_q} U^{\alpha_{p+1}\ldots\alpha_{p+r}}{}_{\beta_{q+1}\ldots\beta_{q+s}}.$$

Derivazione di tensori

La derivazione covariante si estende facilmente alle forme differenziali, e poi ai tensori, al modo seguente. Data un forma $\boldsymbol{\sigma}$ (più esattamente un campo) e un vettore $\mathbf{u}$, definiamo $\nabla_{\mathbf{u}}\boldsymbol{\sigma}$ come una forma tale che sia per qualsiasi campo vettoriale $\mathbf{v}$:

$$\partial_{\mathbf{u}}\langle\boldsymbol{\sigma},\mathbf{v}\rangle = \langle\nabla_{\mathbf{u}}\boldsymbol{\sigma},\mathbf{v}\rangle + \langle\boldsymbol{\sigma},\nabla_{\mathbf{u}}\mathbf{v}\rangle, \tag{11–7}$$

ossia che valga anche per $\langle\ ,\ \rangle$ la regola di derivazione del prodotto.
Nota: Non è ovvio che la (11–7) sia un definizione ben posta: bisogna dimostrare che $\nabla_{\mathbf{u}}\boldsymbol{\sigma}$ è una forma, ossia un funzionale lineare su $\mathcal{V}$. Lasciamo la dimostrazione per esercizio.

Ciò posto, l'estensione a tensori (campi) qualsiasi è immediata, usando la (11–5). Definiamo $\nabla_{\mathbf{u}}\mathbf{T}$ come il tensore tale che $\forall\ \boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_p, \mathbf{v}_1\ldots\mathbf{v}_q$:

$$\begin{aligned}
&\partial_{\mathbf{u}}\mathbf{T}(\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_p,\mathbf{v}_1\ldots\mathbf{v}_q)\\
&\qquad= (\nabla_{\mathbf{u}}\mathbf{T})(\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_p,\mathbf{v}_1\ldots\mathbf{v}_q)\\
&\qquad\quad+ \mathbf{T}(\nabla_{\mathbf{u}}\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_p,\mathbf{v}_1\ldots\mathbf{v}_q) + \cdots + \mathbf{T}(\boldsymbol{\sigma}_1\ldots\nabla_{\mathbf{u}}\boldsymbol{\sigma}_p,\mathbf{v}_1\ldots\mathbf{v}_q)\\
&\qquad\quad+ \mathbf{T}(\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_p,\nabla_{\mathbf{u}}\mathbf{v}_1\ldots\mathbf{v}_q) + \cdots + \mathbf{T}(\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_p,\mathbf{v}_1\ldots\nabla_{\mathbf{u}}\mathbf{v}_q).
\end{aligned} \tag{11–8}$$

Lasciamo al lettore
– di verificare che la (11–8) definisce effettivamente un tensore, ovviamente ancora di rango $\binom{p}{q}$
– di estendere a un tensore qualsiasi l'espressione (11–2) per le componenti della derivata covariante.