

DIPARTIMENTO DI MATEMATICA – UNIVERSITÀ DI TORINO

CORSO DI STUDI IN FISICA

Elsa Abbena – Sergio Garbiero

**NOTE DI GEOMETRIA E
ALGEBRA LINEARE 2**

A.A. 2014-15

Indice

| | |
|--|----|
| Capitolo 1. Geometria differenziale delle curve | 1 |
| 1. Richiami sulle strutture di \mathbb{R}^3 | 1 |
| 2. Curve parametrizzate in \mathbb{R}^3 | 2 |
| 3. Cambiamenti di parametro e ascissa curvilinea | 4 |
| 4. Triedro e formule di Frenet: parte prima | 7 |
| 5. Caratterizzazione di alcune curve | 10 |
| 6. Triedro e formule di Frenet: parte seconda | 12 |
| 7. Eliche cilindriche | 14 |
| 8. Curve piane | 16 |
| 9. Esercizi svolti sulle curve | 19 |
| Capitolo 2. Geometria differenziale delle superfici | 23 |
| 1. Premesse di topologia | 23 |
| 2. Superfici parametrizzate in \mathbb{R}^3 | 25 |
| 3. Esempi di superfici regolari | 28 |
| 4. Piani tangenti ad una superficie | 31 |
| 5. Superfici orientabili | 32 |
| 6. Metrica su una superficie: prima forma quadratica | 35 |
| 7. Operatore forma: seconda forma quadratica | 39 |
| 8. Curvature principali, Gaussiana e media | 41 |
| 9. Simboli di Christoffel e Teorema Egregium di Gauss | 44 |
| 10. Curvature normali | 46 |
| 11. Esercizi svolti sulle superfici | 51 |
| Capitolo 3. Gruppi di matrici | 57 |
| 1. Strutture algebriche | 57 |
| 2. I gruppi lineari generali reali e complessi | 58 |
| 3. I quaternioni | 64 |
| 4. Gruppo lineare quaternionico | 67 |
| 5. Gli ottetti di Cayley | 69 |
| Capitolo 4. Forme bilineari e prodotti scalari | 71 |
| 1. Gruppi di matrici invarianti per forme bilineari | 71 |
| 2. Isometrie, similitudini | 72 |
| 3. Autovalori delle isometrie, similitudini | 80 |
| 4. Forme bilineari simmetriche non degeneri indefinite | 81 |
| 5. Forme simpletiche reali | 83 |
| 6. Matrici hermitiane e anti-hermitiane | 86 |
| 7. Spazi vettoriali hermitiani | 90 |
| 8. Trasformazioni unitarie | 97 |
| 9. Il teorema spettrale | 97 |

| | |
|--|------------|
| 10. Forme bilineari complesse | 103 |
| 11. Prodotto scalare quaternionico | 107 |
| Capitolo 5. Algebre e gruppi di Lie | 113 |
| 1. Algebre di Lie | 113 |
| 2. Sottoalgebre di Lie di $\mathfrak{gl}(n, \mathbb{R})$ | 115 |
| 3. Sottoalgebre di Lie di $\mathfrak{gl}(n, \mathbb{C})$ | 117 |
| 4. Sottoalgebre di Lie di $\mathfrak{gl}(n, \mathbb{H})$ | 118 |
| 5. Esponenziale di una matrice | 120 |
| Capitolo 6. Varietà differenziabili | 131 |
| 1. Definizione di varietà differenziabile | 131 |
| 2. Esempi di varietà differenziabili | 132 |
| 3. Funzioni differenziabili tra varietà differenziabili | 134 |
| 4. Vettori tangenti e spazio vettoriale tangente | 136 |
| 5. Differenziale di una funzione tra varietà differenziabili | 140 |
| 6. Fibrati tangenti e campi vettoriali | 143 |
| 7. Algebra di Lie di un gruppo di Lie | 146 |
| Bibliografia | 151 |

CAPITOLO 1

Geometria differenziale delle curve

In questo capitolo verranno date le definizioni e le proprietà fondamentali riguardanti la geometria differenziale delle curve. La maggior parte dell'esposizione avrà per oggetto curve nello spazio. La sezione finale sarà dedicata alle curve piane.

1. Richiami sulle strutture di \mathbb{R}^3

Consideriamo l'insieme dei punti dello spazio ordinario e fissiamo un riferimento cartesiano. Ogni punto P dello spazio è univocamente individuato dalle sue coordinate cartesiane (x, y, z) e, pertanto, lo spazio stesso si può identificare con \mathbb{R}^3 , insieme delle terne ordinate di numeri reali. Ricordiamo le più importanti strutture algebriche e geometriche di \mathbb{R}^3 .

- \mathbb{R}^3 è uno *spazio vettoriale*. In tal caso, i suoi elementi si dicono **vettori** e sono definite le operazioni di *somma* di vettori e di *prodotto* di un numero reale (detto *scalare*) per un vettore:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (a_1 + b_1, a_2 + b_2, a_3 + b_3), & c\mathbf{u} &= (ca_1, ca_2, ca_3), \\ \mathbf{u} &= (a_1, a_2, a_3) \in \mathbb{R}^3, & \mathbf{v} &= (b_1, b_2, b_3) \in \mathbb{R}^3, & c &\in \mathbb{R}.\end{aligned}$$

La **base standard** di \mathbb{R}^3 è costituita dai vettori:

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1).$$

Ogni vettore $\mathbf{u} = (a_1, a_2, a_3)$ si può esprimere come combinazione lineare dei vettori della base standard: $\mathbf{u} = a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}$. Gli scalari a_1, a_2, a_3 sono le **componenti** del vettore \mathbf{u} rispetto alla base $\mathbf{i}, \mathbf{j}, \mathbf{k}$.

- \mathbb{R}^3 è uno spazio *affine*. In questo contesto, i suoi elementi si dicono **punti**. Una coppia di punti $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ individua un vettore, che indicheremo con $\overrightarrow{P_1 P_2}$, oppure con $P_2 - P_1$, di componenti $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$. In particolare, dato il punto $P = (x, y, z)$, il **vettore posizione** di P è $\overrightarrow{OP} = x \mathbf{i} + y \mathbf{j} + z \mathbf{k}$, dove O indica l'origine del riferimento cartesiano.
- \mathbb{R}^3 è uno *spazio vettoriale euclideo* con il **prodotto scalare** dato da:

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= a_1 b_1 + a_2 b_2 + a_3 b_3, \\ \mathbf{u} &= (a_1, a_2, a_3) \in \mathbb{R}^3, & \mathbf{v} &= (b_1, b_2, b_3) \in \mathbb{R}^3.\end{aligned}$$

La **norma** (o modulo) del vettore \mathbf{u} e l'**angolo** fra i vettori \mathbf{u} e \mathbf{v} sono definiti, rispettivamente, da:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{a_1^2 + a_2^2 + a_3^2},$$

$$\cos(\widehat{\mathbf{u}\mathbf{v}}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \sqrt{b_1^2 + b_2^2 + b_3^2}}.$$

- \mathbb{R}^3 è uno *spazio metrico* (per la relativa definizione si veda [5]) in cui la **distanza euclidea** tra i punti $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ è:

$$d(P_1, P_2) = \|\overrightarrow{P_1 P_2}\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

La distanza euclidea induce la topologia standard su \mathbb{R}^3 (vedi la sezione 1 del Capitolo seguente). Nel seguito, supporremo sempre che \mathbb{R}^3 sia dotato di tale topologia.

2. Curve parametrizzate in \mathbb{R}^3

In prima approssimazione, possiamo pensare ad una curva come alla traiettoria descritta da un punto materiale che si muove nello spazio. Le definizioni seguenti hanno lo scopo di rendere precisa tale idea intuitiva.

DEFINIZIONE 1.1. *Sia I un intervallo aperto di \mathbb{R}^3 . Una **curva parametrizzata** in \mathbb{R}^3 , definita sull'intervallo I , è una funzione*

$$\alpha: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3, \quad t \in I \mapsto \alpha(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3,$$

dove t è il **parametro** della curva e $x(t)$, $y(t)$, $z(t)$ sono tre funzioni reali di variabile reale definite sull'intervallo I .

OSSERVAZIONE 1.1. Una curva può anche essere individuata mediante le sue **equazioni parametriche**

$$\begin{cases} x = x(t) \\ y = y(t), & t \in I, \\ z = z(t) \end{cases}$$

che danno le coordinate del generico punto $P(t)$ che percorre la curva al variare del "tempo" t .

ESEMPIO 1.1. È importante osservare che una curva non è semplicemente un sottoinsieme di punti di \mathbb{R}^3 (cioè la traiettoria descritta dal punto $P(t)$) ma è una *funzione* di $I \subseteq \mathbb{R}$ in \mathbb{R}^3 . Due equazioni parametriche diverse, per esempio

$$\begin{cases} x = t \\ y = t^2, & t \in (0, \infty), \\ z = \ln t \end{cases} \quad \begin{cases} x = e^{t'} \\ y = e^{2t'}, & t' \in \mathbb{R}, \\ z = t' \end{cases}$$

individuano due curve parametrizzate distinte, pur descrivendo la stessa traiettoria (come si vede facilmente sostituendo t con $e^{t'}$). Intuitivamente, possiamo pensare alla stessa curva percorsa in tempi diversi.

DEFINIZIONE 1.2. Una curva parametrizzata $\alpha: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$ si dice **differenziabile di classe C^k** ($k = 1, 2, \dots$) se le funzioni $x(t)$, $y(t)$, $z(t)$ hanno derivate continue (rispetto a t) fino all'ordine k .

La curva α si dice **differenziabile** se è di classe C^∞ , ossia $x(t)$, $y(t)$, $z(t)$ hanno derivate di ordine qualsiasi.

α è di **classe C^0** se le funzioni $x(t)$, $y(t)$, $z(t)$ sono continue.

ESEMPIO 1.2. La curva $\alpha(t) = (t, t^2, \sqrt[3]{t})$, $t \in \mathbb{R}$, è di classe C^0 ma non di classe C^1 (perchè?). Di che classe è la curva α se $t \in (0, \infty)$?

Invece la curva $\beta(t) = (t, t^2, t\sqrt[3]{t})$, $t \in \mathbb{R}$, è di classe C^1 ma non di classe C^2 (perchè?). Ovviamente, la curva $\gamma(t) = (t, t^2, t^3)$, $t \in \mathbb{R}$, è differenziabile.

Nel seguito, salvo avviso contrario, tutte le curve parametrizzate saranno differenziabili.

DEFINIZIONE 1.3. Un **campo vettoriale** X lungo una curva parametrizzata $\alpha: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$ è una funzione che assegna ad ogni punto $\alpha(t)$ della curva un vettore $X(t)$, avente estremo iniziale in $\alpha(t)$. Dare un campo vettoriale lungo α equivale a dare tre funzioni $a_1(t)$, $a_2(t)$ e $a_3(t)$ tali che:

$$X(t) = a_1(t)\mathbf{i} + a_2(t)\mathbf{j} + a_3(t)\mathbf{k}, \quad t \in I.$$

Il campo vettoriale X si dice **differenziabile** se le funzioni $a_1(t)$, $a_2(t)$, $a_3(t)$ sono di classe C^∞ .

Il campo vettoriale X si dice **parallelo** (o, anche, **costante**) se le funzioni $a_1(t)$, $a_2(t)$, $a_3(t)$ sono costanti.

DEFINIZIONE 1.4. Sia $\alpha(t) = (x(t), y(t), z(t))$, $t \in I$, una curva differenziabile. Il **campo tangente** alla curva α è il campo vettoriale:

$$\alpha'(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}, \quad t \in I,$$

che ha come componenti le derivate prime delle funzioni che individuano la curva.

$\alpha(t)$ si dice **punto regolare** se $\alpha'(t) \neq \mathbf{0}$. La curva è **regolare** se ogni suo punto è regolare.

Un punto $\alpha(t_0)$ in cui $\alpha'(t_0) = \mathbf{0}$ è un **punto singolare** della curva.

ESEMPIO 1.3. Fissiamo un punto $P_0 = (x_0, y_0, z_0)$ ed un vettore non nullo $\mathbf{a} = l\mathbf{i} + m\mathbf{j} + n\mathbf{k}$. La curva

$$\alpha(t) = P_0 + t\mathbf{a} = (x_0 + lt, y_0 + mt, z_0 + nt), \quad t \in \mathbb{R},$$

descrive la **retta** passante per P_0 e parallela al vettore \mathbf{a} . La retta è una curva regolare poichè $\alpha'(t) = \mathbf{a} \neq \mathbf{0}$, per ogni $t \in \mathbb{R}$.

ESEMPIO 1.4. Fissiamo due numeri reali $a > 0$ e $b \neq 0$. Le curve

$$\alpha(t) = (a \cos t, a \sin t, bt), \quad t \in \mathbb{R},$$

si dicono **eliche circolari**. Dato che $\alpha'(t) = (-a \sin t, a \cos t, b) \neq \mathbf{0}$, le eliche sono curve regolari.

Notare che $\alpha'(t) \cdot \mathbf{k} = b$, per ogni $t \in \mathbb{R}$; ciò significa che il campo tangente forma un angolo costante con l'asse z . Inoltre, se $P = (x, y, z)$ è il generico punto dell'elica, si ha che $x^2 + y^2 = a^2$, vale a dire che l'elica è contenuta sulla superficie di un cilindro di raggio a . La costante $2\pi|b|$ si

dice **passo** dell'elica: è la distanza tra due punti successivi che stanno sulla stessa generatrice del cilindro.

DEFINIZIONE 1.5. *Siano α una curva differenziabile e $P_0 = \alpha(t_0)$ un punto regolare. La **retta tangente** alla curva α nel punto P_0 è la retta passante per P_0 , parallela al vettore $\alpha'(t_0)$.*

ESEMPIO 1.5. Date le curve:

$$\alpha(t) = (t, t^2, t^3), \quad t \in \mathbb{R}; \quad \beta(t) = (t^2, t^3, t^4), \quad t \in \mathbb{R},$$

si ha: $\alpha'(t) = (1, 2t, 3t^2)$ e $\beta'(t) = (2t, 3t^2, 4t^3)$. Di conseguenza, α è una curva regolare in ogni punto, mentre β è regolare tranne che nell'origine.

Fissato il punto $P_0 = \alpha(1) = (1, 1, 1)$, la retta tangente alla curva α in P_0 è: $\gamma(t) = (1+t, 1+2t, 1+3t)$. La retta tangente alla curva β nel punto $P_1 = \beta(-1) = (-1, -1, 1)$ è $\delta(t) = (1-2t, -1+3t, 1-4t)$.

3. Cambiamenti di parametro e ascissa curvilinea

Si è visto nell'Esempio 1.1 che curve parametrizzate diverse possono descrivere la stessa traiettoria. Vogliamo ora precisare meglio cosa significa cambiare parametro ad una curva.

DEFINIZIONE 1.6. *Sia $\alpha: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$ una curva differenziabile. Data una funzione differenziabile $t = t(s)$, definita sull'intervallo J di \mathbb{R} ed a valori nell'intervallo I , diciamo che la curva*

$$\beta = \alpha \circ t: J \subseteq \mathbb{R} \rightarrow \mathbb{R}^3, \quad s \in J \mapsto \beta(s) = \alpha(t(s)) \in \mathbb{R}^3,$$

è una **riparametrizzazione** di α .

La riparametrizzazione è **regolare** se t è suriettiva, ossia $t(J) = I$, e $t'(s) = \frac{dt}{ds} \neq 0$, per ogni $s \in J$.

OSSERVAZIONE 1.2. Dato che ogni intervallo aperto della retta reale è connesso (vedi il Capitolo sulla Topologia), se la riparametrizzazione è regolare allora $t'(s)$ ha segno costante nell'intervallo J . Ciò implica che la funzione t è strettamente monotona e quindi è una biiezione tra gli intervalli J e I . In particolare, t è invertibile: indicheremo con $s = s(t)$ la funzione inversa, definita sull'intervallo I e a valori in J .

ESEMPIO 1.6. Consideriamo la curva

$$\alpha(t) = (\sqrt{t}, t\sqrt{t}, 1-t), \quad t \in I = (0, \infty),$$

e la funzione differenziabile

$$t: J = (0, \infty) \rightarrow I = (0, \infty), \quad s \mapsto t = s^2.$$

La curva riparametrizzata è:

$$\beta(s) = \alpha(t(s)) = (\sqrt{s^2}, s^2\sqrt{s^2}, 1-s^2) = (s, s^3, 1-s^2).$$

Dato che $t'(s) = 2s \neq 0$, se $s \neq 0$, la riparametrizzazione è regolare. Ovviamente, si passa dalla curva $\beta(s)$ alla curva $\alpha(t)$ mediante la funzione inversa $s = \sqrt{t}$.

PROPRIETÀ 1.1. Sia $\beta = \alpha \circ t$ una riparametrizzazione della curva α mediante la funzione $t = t(s)$. Allora:

$$\beta'(s) = \frac{dt}{ds} \alpha'(t(s)).$$

DIMOSTRAZIONE. Supponiamo che $\alpha(t) = (x(t), y(t), z(t))$, $t \in I$. Allora $\beta(s) = (x(t(s)), y(t(s)), z(t(s)))$. Per la regola di derivazione delle funzioni composte, si ha:

$$\frac{dx(t(s))}{ds} = \frac{dx}{dt} \Big|_{t(s)} \frac{dt}{ds}, \quad \text{ecc.}$$

e quindi

$$\beta'(s) = \frac{dt}{ds} \left(\frac{dx}{dt} \Big|_{t(s)}, \frac{dy}{dt} \Big|_{t(s)}, \frac{dz}{dt} \Big|_{t(s)} \right) = \frac{dt}{ds} \alpha'(t(s)).$$

□

Dalla proprietà precedente si vede che il vettore tangente in un punto ad una curva riparametrizzata è *proporzionale* al vettore tangente alla curva iniziale. In particolare, una riparametrizzazione regolare manda punti regolari in punti regolari.

Nel seguito, verranno considerate *solo riparametrizzazioni regolari*.

OSSERVAZIONE 1.3. Consideriamo l'insieme di tutte le curve differenziabili regolari e la relazione che associa due curve se la prima è una riparametrizzazione regolare della seconda. Si vede facilmente che questa relazione è di equivalenza. Da un punto di vista un po' più astratto, possiamo definire una curva come una classe di equivalenza rispetto a tale relazione. In termini intuitivi, ciò significa pensare una curva come una traiettoria nello spazio, percorsa in tutti i modi possibili.

Come si è visto nei corsi di Analisi Matematica, la lunghezza di un arco di curva è l'integrale del modulo del campo tangente. Più precisamente, vale la

DEFINIZIONE 1.7. Sia $\alpha(t) = (x(t), y(t), z(t))$ una curva differenziabile definita su un intervallo I . Supponiamo che $[a, b] \subseteq I$. La **lunghezza dell'arco di curva** compreso fra gli estremi $\alpha(a)$ e $\alpha(b)$ è:

$$\mathcal{L}_a^b(\alpha) = \int_a^b \|\alpha'(t)\| dt = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2} dt.$$

Verifichiamo innanzitutto che la definizione è sensata dal punto di vista geometrico, vale a dire che non dipende dal parametro usato.

PROPRIETÀ 1.2. Sia β una riparametrizzazione regolare della curva α mediante la funzione $t = t(s)$. Supponiamo che $a = t(a_1)$ e $b = t(b_1)$. Allora:

$$\int_a^b \|\alpha'(t)\| dt = \int_{a_1}^{b_1} \|\beta'(s)\| ds.$$

DIMOSTRAZIONE. Dato che $t'(s)$ ha segno costante (vedi l'Osservazione di pag. 4), supponiamo che $t'(s) > 0$. Ciò implica che la funzione $t(s)$ è crescente e, quindi, $a_1 < b_1$. Tenuto conto della Proprietà 1.1 e della formula del cambiamento di variabile negli integrali, si ha:

$$\begin{aligned} \int_{a_1}^{b_1} \|\beta'(s)\| ds &= \int_{a_1}^{b_1} \left\| \frac{dt}{ds} \alpha'(t(s)) \right\| ds = \int_{a_1}^{b_1} \|\alpha'(t(s))\| \left| \frac{dt}{ds} \right| ds \\ &= \int_{a_1}^{b_1} \|\alpha'(t(s))\| \frac{dt}{ds} ds = \int_a^b \|\alpha'(t)\| dt. \end{aligned}$$

La stessa conclusione vale anche nel caso in cui $t'(s) < 0$, tenuto conto del fatto che la funzione $t(s)$ è decrescente e, quindi, $a_1 > b_1$ e $|t'(s)| = -t'(s)$ \square

Una conseguenza importante della proprietà precedente è che si può sempre riparametrizzare una curva regolare in modo che il campo tangente abbia *lunghezza unitaria*.

DEFINIZIONE 1.8. Sia α una curva definita su un intervallo I . Fissiamo un valore $a \in I$. La funzione

$$s(t) = \int_a^t \|\alpha'(u)\| du$$

si dice **ascissa curvilinea** (oppure **parametro lunghezza d'arco**) della curva α , con punto iniziale $\alpha(a)$.

OSSERVAZIONE 1.4. Per definizione, la funzione $s(t)$ rappresenta la lunghezza (con segno) dell'arco di curva compresa tra l'estremo fisso $\alpha(a)$ e l'estremo variabile $\alpha(t)$. Cambiando il punto iniziale, si ottiene un'ascissa curvilinea che differisce dalla precedente per una costante additiva.

PROPRIETÀ 1.3. È sempre possibile riparametrizzare una curva regolare mediante l'ascissa curvilinea. In tal caso, il campo tangente della curva riparametrizzata ha modulo costante uguale ad uno.

DIMOSTRAZIONE. Basta osservare che l'ascissa curvilinea consente di effettuare una riparametrizzazione regolare. Dato che la curva è regolare, $s'(t) = \|\alpha'(t)\| > 0$ e, quindi, $s(t)$ è invertibile. Detta $t = t(s)$ la funzione inversa, sia $\beta(s) = \alpha(t(s))$ la riparametrizzazione di α mediante l'ascissa curvilinea. Allora:

$$\|\beta'(s)\| = \left| \frac{dt}{ds} \right| \|\alpha'(t)\| = \frac{1}{|s'(t)|} \|\alpha'(t)\| = \frac{1}{\|\alpha'(t)\|} \|\alpha'(t)\| = 1.$$

\square

ESEMPIO 1.7. Consideriamo la curva $\alpha(t) = (2t, t^2, \frac{1}{3}t^3)$, $t \in \mathbb{R}$. Dato che $\alpha'(t) = (2, 2t, t^2) \neq \mathbf{0}$, la curva è ovunque regolare.

La lunghezza dell'arco di curva compreso fra i punti $P_0 = (0, 0, 0)$ e $P_1 = (2, 1, \frac{1}{3})$ vale:

$$\mathcal{L}_0^1(\alpha) = \int_0^1 \|\alpha'(u)\| du = \int_0^1 \sqrt{4 + 4u^2 + u^4} du = \frac{7}{3}$$

(tenere conto che $P_0 = \alpha(0)$, ecc.).

Riparametrizziamo la curva mediante la funzione $t = -2s$ e calcoliamo la lunghezza dell'arco della curva riparametrizzata tra gli estremi P_0 e P_1 . La curva riparametrizzata è $\beta(s) = (-4s, 4s^2, -\frac{8}{3}s^3)$ e $\beta'(s) = (-4, 8s, -8s^2)$. Notiamo che la funzione $t = -2s$ è decrescente e gli estremi $P_0 = \beta(0)$ e $P_1 = \beta(-\frac{1}{2})$ vanno presi nell'ordine inverso. Pertanto la lunghezza dell'arco di curva è data da:

$$\mathcal{L}_{-\frac{1}{2}}^0(\beta) = \int_{-\frac{1}{2}}^0 \|\beta'(s)\| ds = \int_{-\frac{1}{2}}^0 \sqrt{16 + 64s^2 + 64s^4} ds = \frac{7}{3},$$

in accordo con la Proprietà 1.2.

Calcoliamo infine l'ascissa curvilinea della curva α a partire da P_0 :

$$s(t) = \int_0^t \|\alpha'(u)\| du = \int_0^t (2 + u^2) du = \frac{t^3}{3} + 2t.$$

Anche se la funzione s è invertibile, non è facile scrivere esplicitamente l'espressione della funzione inversa, necessaria per riparametrizzare α .

ESEMPIO 1.8. Consideriamo nuovamente l'elica circolare dell'Esempio 1.4. L'ascissa curvilinea, a partire dal punto $\alpha(0) = (a, 0, 0)$, è

$$s(t) = \int_0^t \|\alpha'(u)\| du = \int_0^t \sqrt{a^2 + b^2} du = \sqrt{a^2 + b^2} t. \quad (*)$$

Posto per semplicità $c = \sqrt{a^2 + b^2}$, le equazioni dell'elica riparametrizzata mediante l'ascissa curvilinea sono:

$$\beta(s) = \left(a \cos\left(\frac{s}{c}\right), a \sin\left(\frac{s}{c}\right), \frac{b}{c} s \right), \quad s \in \mathbb{R}.$$

Ad esempio, da (*) si ricava immediatamente che la lunghezza dell'arco di elica $\alpha(t) = (\cos t, \sin t, t)$, compreso tra i punti $P(1, 0, 0)$ e $Q(0, 1, \frac{\pi}{2})$, vale $\frac{\sqrt{2}}{2}\pi$.

4. Triedro e formule di Frenet: parte prima

Iniziamo questa sezione con alcune definizioni e proprietà dei campi vettoriali che saranno utilizzate più volte nel seguito.

Sia α una curva definita in un intervallo I . Dati due campi vettoriali lungo α , $X(t) = a_1(t)\mathbf{i} + a_2(t)\mathbf{j} + a_3(t)\mathbf{k}$ e $Y(t) = b_1(t)\mathbf{i} + b_2(t)\mathbf{j} + b_3(t)\mathbf{k}$, il loro prodotto scalare è la funzione:

$$(X \cdot Y)(t) = X(t) \cdot Y(t) = a_1(t)b_1(t) + a_2(t)b_2(t) + a_3(t)b_3(t).$$

Due campi vettoriali X e Y si dicono **ortogonali** (o perpendicolari) se:

$$X \cdot Y = 0.$$

Il modulo (o norma) di un campo X è la funzione:

$$\|X(t)\| = \sqrt{X(t) \cdot X(t)} = \sqrt{(a_1(t))^2 + (a_2(t))^2 + (a_3(t))^2}.$$

Il **derivato n -simo** ($n = 1, 2, \dots$) di un campo differenziabile X è il campo che ha come componenti le derivate n -sime delle componenti di X . Indicheremo con X', X'', X''' i primi tre campi derivati di X . In particolare, α'' (derivato del campo tangente ad una curva) viene detto **campo accelerazione**. Si noti che un campo differenziabile X è parallelo (vedi Definizione 1.3) se e solo se $X' = \mathbf{0}$.

Nel seguito si supporrà che *tutti i campi vettoriali siano differenziabili*.

LEMMA 1.4. *Sia X un campo che ha modulo costante. Allora il campo derivato X' è ortogonale ad X .*

DIMOSTRAZIONE. Se $X(t) = a_1(t)\mathbf{i} + a_2(t)\mathbf{j} + a_3(t)\mathbf{k}$ ha modulo costante c , si ha: $c^2 = X(t) \cdot X(t) = (a_1(t))^2 + (a_2(t))^2 + (a_3(t))^2$. Derivando tale uguaglianza, si ottiene:

$$\begin{aligned} 0 &= \frac{d}{dt} [X(t) \cdot X(t)] = 2a_1(t)a_1'(t) + 2a_2(t)a_2'(t) + 2a_3(t)a_3'(t) \\ &= 2X(t) \cdot X'(t). \end{aligned}$$

□

La proprietà seguente, di dimostrazione immediata, caratterizza le curve che si riducono ad un punto e le rette.

PROPRIETÀ 1.5. *Sia $\alpha: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$ una curva. Allora:*

- (1) α è una curva costante (la sua traiettoria si riduce ad un punto) se e solo se $\alpha' = \mathbf{0}$.
- (2) α è (parte di) una retta se e solo se $\alpha'' = \mathbf{0}$.

Vogliamo ora costruire, in ogni punto di una curva, un riferimento cartesiano che consenta di studiare le proprietà geometriche della curva. Dato che tale riferimento varia da punto a punto, esso sarà individuato da tre campi vettoriali unitari, a due a due ortogonali.

Data una curva qualsiasi, si può sempre supporre che sia parametrizzata mediante l'ascissa curvilinea (Proprietà 1.3). Sia $\beta(s)$, $s \in J$ una tale curva. Indichiamo con $T(s) = \beta'(s)$ il suo campo tangente. Poichè $T(s)$ ha modulo uguale ad uno, il campo derivato $T'(s) = \beta''(s)$ sarà ortogonale a $T(s)$ (Lemma 1.4). In generale, non è detto che T' abbia a sua volta modulo costante. Consideriamo perciò la funzione

$$k: J \rightarrow \mathbb{R}, \quad s \in J \mapsto k(s) = \|T'(s)\|.$$

Si noti che $k(s) \geq 0$. La funzione $k(s)$ viene detta **curvatura** della curva β . Definiamo ora il **campo normale** alla curva:

$$N(s) = \frac{1}{\|T'(s)\|} T'(s) = \frac{1}{k(s)} T'(s).$$

Per costruzione, N ha modulo unitario ed è ortogonale a T . Infine, definiamo il **campo binormale** come il prodotto vettoriale:

$$B(s) = T(s) \times N(s).$$

Notare che, per costruzione, in ogni punto $\beta(s)$ della curva, i tre vettori $\{T(s), N(s), B(s)\}$ formano una *base ortonormale* di \mathbb{R}^3 .

DEFINIZIONE 1.9. *Sia β una curva parametrizzata mediante l'ascissa curvilinea. La terna di campi vettoriali $\{T, N, B\}$ si dice **triedro di Frenet** della curva β .*

*Fissato un punto $\beta(s)$, le rette passanti per $\beta(s)$ e parallele, rispettivamente, ai vettori $T(s)$, $N(s)$ e $B(s)$ si dicono: **retta tangente**, **retta normale** e **retta binormale**.*

ESEMPIO 1.9. Troviamo la curvatura ed il riferimento di Frenet dell'elica circolare (Esempio 1.8)

$$\beta(s) = \left(a \cos\left(\frac{s}{c}\right), a \sin\left(\frac{s}{c}\right), \frac{b}{c}s \right), \quad \text{dove } c = \sqrt{a^2 + b^2}.$$

Derivando due volte si trova:

$$\begin{aligned} T(s) = \beta'(s) &= \left(-\frac{a}{c} \sin\left(\frac{s}{c}\right), \frac{a}{c} \cos\left(\frac{s}{c}\right), \frac{b}{c} \right), \\ T'(s) = \beta''(s) &= \left(-\frac{a}{c^2} \cos\left(\frac{s}{c}\right), -\frac{a}{c^2} \sin\left(\frac{s}{c}\right), 0 \right). \end{aligned}$$

La curvatura dell'elica vale:

$$k(s) = \|T'(s)\| = \frac{a}{c^2}$$

(notare che è costante). Infine, per completare il triedro di Frenet, calcoliamo:

$$\begin{aligned} N(s) &= \frac{1}{k(s)} T'(s) = \left(-\cos\left(\frac{s}{c}\right), -\sin\left(\frac{s}{c}\right), 0 \right), \\ B(s) &= T(s) \times N(s) = \left(\frac{b}{c} \sin\left(\frac{s}{c}\right), -\frac{b}{c} \cos\left(\frac{s}{c}\right), \frac{a}{c} \right). \end{aligned}$$

Il lettore verifichi direttamente che $\{T, N, B\}$ è, in ogni punto, una base ortonormale. Si noti anche che $T(s) \cdot \mathbf{k} = b/c$ è una costante non nulla. Questo fatto implica che, in ogni punto dell'elica, la retta tangente forma un angolo costante con l'asse z . Inoltre $N(s) \cdot \mathbf{k} = 0$: il campo normale è sempre ortogonale all'asse z .

Il triedro di Frenet varia generalmente da punto a punto, in base alla geometria della curva. È quindi importante poter valutare tale variazione: per questo motivo i derivati dei tre campi del triedro verranno espressi come combinazione lineare della base $\{T, N, B\}$.

TEOREMA 1.6 (Formule di Frenet per l'ascissa curvilinea). *Sia β una curva regolare, parametrizzata mediante l'ascissa curvilinea e definita su un intervallo J . Supponiamo che $k(s) > 0$, per ogni $s \in J$. Allora:*

$$\begin{cases} T'(s) &= k(s)N(s) \\ N'(s) &= -k(s)T(s) + \tau(s)B(s), \\ B'(s) &= -\tau(s)N(s) \end{cases}$$

dove $\tau(s) = -B'(s) \cdot N(s)$ è la funzione **torsione** della curva.

DIMOSTRAZIONE. La prima formula di Frenet segue immediatamente dalla definizione di campo normale (vedi pag. 8).

Dimostriamo la terza formula di Frenet. Il campo binormale B è ortogonale al campo tangente T , quindi $B(s) \cdot T(s) = 0$. Derivando tale uguaglianza, si trova: $B'(s) \cdot T(s) + B(s) \cdot T'(s) = 0$ e quindi, per la prima formula di Frenet:

$$B'(s) \cdot T(s) = -B(s) \cdot T'(s) = -k(s) B(s) \cdot N(s) = 0,$$

ossia B' è ortogonale a T . Poiché B' è anche ortogonale a B (Lemma 1.4), B' deve essere necessariamente parallelo a N . In altri termini, esiste una funzione $\tau(s)$ tale che:

$$B'(s) = -\tau(s) N(s)$$

(il segno meno non è essenziale; serve solo a semplificare certe formule).

Calcoliamo infine N' . Poichè N' è ortogonale a N , N' deve essere combinazione lineare di T e B . Dato che la terna di campi T , N , e B è ortonormale, possiamo scrivere:

$$N'(s) = [N'(s) \cdot T(s)]T(s) + [N'(s) \cdot B(s)]B(s).$$

Troviamo $[N'(s) \cdot T(s)]$. Derivando l'uguaglianza $N(s) \cdot T(s) = 0$, si ha: $N'(s) \cdot T(s) + N(s) \cdot T'(s) = 0$, da cui si ricava: $N'(s) \cdot T(s) = -N(s) \cdot T'(s) = -k(s)N(s) \cdot N(s) = -k(s)$.

In modo del tutto analogo, partendo dal fatto che N e B sono ortogonali e tenuto conto della terza formula di Frenet, si trova che $N'(s) \cdot B(s) = \tau(s)$. \square

5. Caratterizzazione di alcune curve

Le proprietà seguenti hanno lo scopo di caratterizzare alcuni tipi di curve e di mostrare il significato geometrico delle funzioni curvatura e torsione.

PROPRIETÀ 1.7. *Sia β una curva regolare parametrizzata mediante l'ascissa curvilinea. La curvatura k è identicamente nulla se e solo se β è parte di una retta.*

DIMOSTRAZIONE. Supponiamo che $\beta(s) = (x(s), y(s), z(s))$, con $s \in J$, e che $k(s) = 0$, per ogni $s \in J$. Dato che $k(s) = \|T'(s)\|$, abbiamo che $T'(s) = \beta''(s) = \mathbf{0}$ e quindi: $x''(s) = y''(s) = z''(s) = 0$. Di conseguenza, x , y e z devono essere funzioni lineari di s , ossia β è parte di una retta.

Il viceversa è ovvio. \square

OSSERVAZIONE 1.5. Se β è una retta, il suo triedro di Frenet è indeterminato. Questa è la ragione per cui è stata fatta l'ipotesi $k(s) > 0$ nell'enunciato del Teorema 1.6.

PROPRIETÀ 1.8. *Sia β una curva regolare parametrizzata mediante l'ascissa curvilinea, definita su un intervallo J e tale che $k(s) > 0$, per ogni $s \in J$. La torsione τ è identicamente nulla se e solo se β è una curva piana.*

DIMOSTRAZIONE. Supponiamo che β sia una curva piana, contenuta nel piano passante per un punto P_0 e perpendicolare ad un vettore \mathbf{a} . Ciò significa che, per ogni $s \in J$, i vettori $\overrightarrow{P_0\beta(s)} = \beta(s) - P_0$ ed \mathbf{a} sono ortogonali. Derivando l'uguaglianza $(\beta(s) - P_0) \cdot \mathbf{a} = 0$, si trova che: $\beta'(s) \cdot \mathbf{a} = 0$, ossia \mathbf{a} è ortogonale a $\beta'(s) = T(s)$. Derivando un'altra volta, si vede che \mathbf{a} è ortogonale a $\beta''(s) = T'(s) = k(s)N(s)$ e, quindi, anche a $N(s)$. Di conseguenza, \mathbf{a} deve essere parallelo a $B(s)$, vale a dire: $B(s) = c\mathbf{a}$, per un certo scalare c . Ciò implica che B è un campo vettoriale costante. Dalla terza formula di Frenet, si ha: $\mathbf{0} = B'(s) = -\tau(s)N(s)$, da cui: $\tau(s) = 0$.

Viceversa, supponiamo che la torsione τ sia identicamente nulla e proviamo che la curva β è contenuta in un piano. Dalla terza formula di Frenet ricaviamo che $B' = \mathbf{0}$, ossia $B(s) = \mathbf{b}$ è un campo costante. Fissato un punto qualsiasi $\beta(s_0)$ della curva, consideriamo la funzione $g(s) = (\beta(s) - \beta(s_0)) \cdot \mathbf{b}$. Derivando, si trova:

$$g'(s) = \beta'(s) \cdot \mathbf{b} = T(s) \cdot B(s) = 0.$$

Essendo $g(s_0) = 0$, la funzione g è identicamente nulla: questo significa che i vettori $\beta(s) - \beta(s_0)$ e \mathbf{b} sono ortogonali, per ogni $s \in J$. In altri termini, ogni punto della curva sta sul piano passante per il punto $\beta(s_0)$ e perpendicolare al vettore \mathbf{b} . \square

ESEMPIO 1.10. La torsione delle eliche circolari dell'Esempio 1.8 vale:

$$\tau(s) = -B'(s) \cdot N(s) = \frac{b}{c^2} = \frac{b}{a^2 + b^2} \neq 0.$$

Ciò conferma che le eliche non sono curve piane.

ESEMPIO 1.11. Troviamo il triedro di Frenet della curva

$$\beta(s) = \left(\frac{4}{5} \cos s, 1 - \sin s, -\frac{3}{5} \cos s \right), \quad s \in [0, 2\pi].$$

Calcoliamo i primi due campi derivati:

$$\beta'(s) = \left(-\frac{4}{5} \sin s, -\cos s, \frac{3}{5} \sin s \right), \quad \beta''(s) = \left(-\frac{4}{5} \cos s, \sin s, \frac{3}{5} \cos s \right).$$

Dato che $\|T(s)\| = \|\beta'(s)\| = 1$, si vede che s è l'ascissa curvilinea. Inoltre $k(s) = \|T'(s)\| = \|\beta''(s)\| = 1$ e, quindi, la curva non è una retta, pur essendo la curvatura costante. Inoltre, $N(s) = T'(s) = \beta''(s)$ e

$$B(s) = T(s) \times N(s) = -\frac{3}{5} \mathbf{i} - \frac{4}{5} \mathbf{k}.$$

Quindi $B' = \mathbf{0}$, $\tau(s) = 0$ e la curva è piana. La curva è contenuta nel piano passante per il punto $\beta(\frac{\pi}{2}) = (0, 0, 0)$ e perpendicolare al vettore $\mathbf{b} = -\frac{3}{5} \mathbf{i} - \frac{4}{5} \mathbf{k}$, di equazione: $3x + 4z = 0$. Dalle equazioni parametriche di β si vede che se un punto $P(x, y, z)$ sta sulla curva, le sue coordinate verificano l'equazione $x^2 + (y - 1)^2 + z^2 = 1$. Quindi, ogni punto della curva sta sulla sfera di centro $C(0, 1, 0)$ e raggio 1. Possiamo allora concludere che β è una circonferenza.

PROPRIETÀ 1.9. Sia β una curva regolare parametrizzata mediante l'ascissa curvilinea. β è (parte di) una circonferenza se e solo se ha curvatura costante $k > 0$ e torsione $\tau = 0$. Inoltre, il raggio della circonferenza vale $1/k$.

DIMOSTRAZIONE. Supponiamo $k > 0$ costante e $\tau = 0$. Per la Proprietà 1.8, β è una curva piana. Consideriamo la curva

$$\gamma(s) = \beta(s) + \frac{1}{k} N(s).$$

Derivando ed applicando la seconda formula di Frenet, si ha:

$$\gamma'(s) = \beta'(s) + \frac{1}{k} N'(s) = T(s) + \frac{1}{k} (-kT(s)) = \mathbf{0}.$$

Quindi $\gamma(s) = C$ è una curva costante e

$$d(\beta(s), C) = \|\beta(s) - C\| = \left\| \frac{1}{k} N(s) \right\| = \frac{1}{k},$$

vale a dire che β è la circonferenza di centro C e raggio $1/k$.

Viceversa, se β è la circonferenza di centro C e raggio r , si ha subito che $\tau = 0$ e $\beta(s) = C - rN(s)$ (tener conto del fatto che $N(s)$ sta sul piano che

contiene la circonferenza ed è perpendicolare alla retta tangente; inoltre la terna $\{T, N, B\}$ è orientata positivamente). Essendo:

$$T(s) = \beta'(s) = -rN'(s) = r(k(s)T(s)),$$

si ha che $rk(s) = 1$ e quindi $k(s) = \frac{1}{r}$ è costante. \square

OSSERVAZIONE 1.6. Se β è una curva qualsiasi, fissato un punto $\beta(s_0)$, si consideri la circonferenza di

$$\text{centro: } \beta(s_0) + \frac{1}{k(s_0)}N(s_0) \quad \text{e raggio: } \frac{1}{k(s_0)}.$$

Tale circonferenza viene detta **cerchio osculatore** e costituisce una buona approssimazione locale della curva. Il centro del cerchio osculatore è il **centro di curvatura**; il **piano osculatore** è il piano passante per il punto $\beta(s_0)$ e contenente il cerchio osculatore. Dal punto di vista geometrico, si prova che il cerchio osculatore è la posizione limite a cui tende la circonferenza passante per tre punti “vicini” $\beta(s_0)$, $\beta(s_1)$ e $\beta(s_2)$, quando $\beta(s_1)$ e $\beta(s_2)$ tendono a $\beta(s_0)$ lungo la curva.

6. Triedro e formule di Frenet: parte seconda

Nelle sezioni precedenti abbiamo sempre considerato curve parametrizzate con l'ascissa curvilinea. Anche se questa ipotesi non è restrittiva dal punto di vista logico, tuttavia presenta molti inconvenienti pratici quando si vogliono studiare delle curve con un parametro qualsiasi. Vediamo dunque come si modificano le formule di Frenet per curve con un parametro arbitrario.

Siano $\alpha(t)$, $t \in I$, una curva regolare, $s = s(t)$ l'ascissa curvilinea di α (a partire da un punto fissato) e $\beta(s) = \alpha(t(s))$ la riparametrizzazione di α mediante l'ascissa curvilinea. Indichiamo con: $\bar{k}(s)$, $\bar{\tau}(s)$ e $\{\bar{T}(s), \bar{N}(s), \bar{B}(s)\}$ la curvatura, la torsione ed il triedro di Frenet di β , rispettivamente.

Il triedro di Frenet della curva α è dato da:

$$T(t) = \bar{T}(s(t)), \quad N(t) = \bar{N}(s(t)), \quad B(t) = \bar{B}(s(t)),$$

mentre la curvatura e la torsione di α sono definite da:

$$k(t) = \bar{k}(s(t)), \quad \tau(t) = \bar{\tau}(s(t)).$$

OSSERVAZIONE 1.7. Notare che, in ogni punto della curva, il triedro, la curvatura e la torsione sono gli stessi sia per α che per β ; ciò che cambia è solo il parametro con cui viene individuato il punto sulla curva. Questa osservazione implica che *tutte le proprietà della sezione 5 valgono per curve con un parametro qualsiasi*.

TEOREMA 1.10 (Formule di Frenet per un parametro qualsiasi). *Sia $\alpha(t)$ una curva regolare definita su un intervallo I . Supponiamo che $k(t) > 0$, per ogni $t \in I$. Allora:*

$$\begin{cases} T'(t) &= & k(t)\dot{\alpha}(t)N(t) \\ N'(t) &= & -k(t)\dot{\alpha}(t)T(t) & + \tau(t)\dot{\alpha}(t)B(t), \\ B'(t) &= & -\tau(t)\dot{\alpha}(t)N(t) \end{cases}$$

dove $\dot{\alpha}(t) = \|\alpha'(t)\|$ è la funzione **velocità** della curva.

DIMOSTRAZIONE. Deriviamo $T(t) = \bar{T}(s(t))$ rispetto a t ed applichiamo il Teorema 1.6:

$$T'(t) = \left. \frac{d\bar{T}}{ds} \right|_{s(t)} \left. \frac{ds}{dt} \right|_t = [\bar{k}(s(t))\bar{N}(s(t))] \|\alpha'(t)\| = \dot{\alpha}(t)k(t)N(t).$$

Le altre due formule si dimostrano in modo analogo. \square

Vogliamo ora determinare le relazioni che permettono di trovare il triedro di Frenet, la curvatura e la torsione senza passare attraverso l'ascissa curvilinea. Prima però è necessario un lemma. La seconda formula del lemma si può interpretare come la decomposizione dell'accelerazione di un punto mobile nelle componenti tangenziale e centripeta.

LEMMA 1.11. *Sia $\alpha(t)$, $t \in I$ una curva regolare. Allora:*

$$\alpha'(t) = \dot{\alpha}(t)T(t) \quad (1.1)$$

$$\alpha''(t) = \ddot{\alpha}(t)T(t) + k(t)[\dot{\alpha}(t)]^2N(t), \quad (1.2)$$

dove $\ddot{\alpha}(t) = \frac{d\dot{\alpha}(t)}{dt} = \frac{d}{dt}\|\alpha'(t)\|$.

DIMOSTRAZIONE. Deriviamo $\alpha(t) = \beta(s(t))$, rispetto a t :

$$\alpha'(t) = \frac{d\beta}{ds} \frac{ds}{dt} = \bar{T}(s(t))\dot{\alpha}(t) = \dot{\alpha}(t)T(t).$$

Deriviamo nuovamente ed applichiamo la prima formula di Frenet:

$$\alpha''(t) = \ddot{\alpha}(t)T(t) + \dot{\alpha}(t)T'(t) = \ddot{\alpha}(t)T(t) + \dot{\alpha}(t)[k(t)\dot{\alpha}(t)N(t)].$$

\square

TEOREMA 1.12. *Sia $\alpha: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$ una curva regolare. Allora:*

$$T(t) = \frac{\alpha'(t)}{\|\alpha'(t)\|}, \quad N(t) = B(t) \times T(t), \quad B(t) = \frac{\alpha'(t) \times \alpha''(t)}{\|\alpha'(t) \times \alpha''(t)\|}.$$

La curvatura e la torsione di α sono date da:

$$k(t) = \frac{\|\alpha'(t) \times \alpha''(t)\|}{\|\alpha'(t)\|^3}, \quad \tau(t) = \frac{\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t)}{\|\alpha'(t) \times \alpha''(t)\|^2}.$$

DIMOSTRAZIONE. Allo scopo di semplificare le notazioni, la dipendenza da t delle varie funzioni e campi viene omessa.

Se α è regolare si ha che $\dot{\alpha} = \|\alpha'\| > 0$. Dalla (1.1) si ricava:

$$T = \frac{1}{\dot{\alpha}} \alpha' = \frac{1}{\|\alpha'\|} \alpha'.$$

Inoltre, applicando la (1.2), si ha:

$$\alpha' \times \alpha'' = (\dot{\alpha}T) \times (\ddot{\alpha}T + k\dot{\alpha}^2N) = \dot{\alpha}\ddot{\alpha}T \times T + k\dot{\alpha}^3T \times N = k\dot{\alpha}^3B$$

da cui si ricava $\|\alpha' \times \alpha''\| = k\dot{\alpha}^3$ e quindi:

$$B = \frac{1}{k\dot{\alpha}^3} \alpha' \times \alpha'' = \frac{1}{\|\alpha' \times \alpha''\|} \alpha' \times \alpha'', \quad k = \frac{\|\alpha' \times \alpha''\|}{\dot{\alpha}^3} = \frac{\|\alpha' \times \alpha''\|}{\|\alpha'\|^3}.$$

Per trovare la torsione, occorre derivare ancora una volta la (1.2) e usare le formule di Frenet:

$$\begin{aligned}\alpha''' &= \ddot{\alpha} T + \ddot{\alpha} T' + (k\dot{\alpha}^2)' N + k\dot{\alpha}^2 N' \\ &= \ddot{\alpha} T + k\dot{\alpha}\ddot{\alpha} N + (k\dot{\alpha}^2)' N + k\dot{\alpha}^2(-k\dot{\alpha} T + \tau\dot{\alpha} B) \\ &= (\dots) T + (\dots) N + k\tau\dot{\alpha}^3 B.\end{aligned}$$

La formula della torsione segue da:

$$\alpha' \times \alpha'' \cdot \alpha''' = (k\dot{\alpha}^3 B) \cdot \alpha''' = \tau (k\dot{\alpha}^3)^2 = \tau \|\alpha'(t) \times \alpha''(t)\|^2.$$

□

ESEMPIO 1.12. Troviamo il triedro di Frenet, la curvatura e la torsione della curva

$$\alpha(t) = (2t, t^2, \frac{1}{3}t^3), \quad t \in \mathbb{R},$$

già considerata nell'Esempio 6.9. Avevamo notato allora che non era possibile riparametrizzare esplicitamente la curva con l'ascissa curvilinea e, pertanto, non si possono usare le formule del Teorema 1.6 ma occorre applicare il teorema precedente. I primi tre campi derivati sono:

$$\alpha'(t) = (2, 2t, t^2), \quad \alpha''(t) = (0, 2, 2t), \quad \alpha'''(t) = (0, 0, 2).$$

da cui si ricava:

$$\begin{aligned}\|\alpha'(t)\| &= 2 + t^2, & \alpha'(t) \times \alpha''(t) &= (2t^2, -4t, 4), \\ \|\alpha'(t) \times \alpha''(t)\| &= 2(2 + t^2), & \alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) &= 8,\end{aligned}$$

e, quindi:

$$\begin{aligned}T(t) &= \frac{1}{2+t^2} (2, 2t, t^2), & N(t) &= \frac{1}{2+t^2} (-2t, 2-t^2, 2t), \\ B(t) &= \frac{1}{2+t^2} (t^2, -2t, 2), & k(t) = \tau(t) &= \frac{2}{(2+t^2)^2}.\end{aligned}$$

Notare che la curvatura e la torsione non sono costanti ma è costante il loro rapporto.

7. Eliche cilindriche

Abbiamo visto nell'Esempio 1.8 che i tre campi del triedro di Frenet di un'elica circolare formano un angolo costante con l'asse z . Questa proprietà caratterizza un'importante famiglia di curve di cui le eliche circolari sono un caso particolare.

DEFINIZIONE 1.10. *Un'elica cilindrica è una curva non piana le cui tangenti formano un angolo costante con una direzione fissa.*

Se $\alpha(t)$, $t \in I$, è un'elica cilindrica, esiste un versore (vettore di modulo uno) \mathbf{u} tale che:

$$T(t) \cdot \mathbf{u} = \cos \theta, \quad t \in I,$$

dove θ è un angolo costante. Il nome di queste curve è dovuto al fatto che la loro traiettoria è contenuta nel cilindro generalizzato (vedi la Definizione 2.16 a pagina 37) formato dalle rette passanti per i punti di α e parallele ad \mathbf{u} . Il versore \mathbf{u} si dice **asse** dell'elica.

Nel seguito, supporremo che la curva sia riparametrizzata mediante l'ascissa curvilinea e , con abuso di notazioni, indicheremo con $\{T, N, B\}$ sia il triedro di Frenet della curva data che quello della riparametrizzazione (vedi l'Osservazione di pag. 12).

La proprietà seguente mostra come tutti i campi del triedro di Frenet formino un angolo costante con \mathbf{u} .

PROPRIETÀ 1.13. *Se α è un'elica cilindrica, allora:*

- (1) *il campo normale N è ortogonale a \mathbf{u} ,*
- (2) *il campo binormale B forma angolo costante con \mathbf{u} .*

DIMOSTRAZIONE. Sia $\beta(s)$ la riparametrizzazione di α mediante l'ascissa curvilinea. Derivando $T(s) \cdot \mathbf{u} = \cos \theta$, si trova

$$T'(s) \cdot \mathbf{u} = k(s)N(s) \cdot \mathbf{u} = 0.$$

Dato che $k(s) > 0$, si può concludere che N è ortogonale a \mathbf{u} .

Per provare la (2), si osservi che $N(s) \cdot \mathbf{u} = 0$ implica

$$[-\tau(s)N(s)] \cdot \mathbf{u} = B'(s) \cdot \mathbf{u} = 0,$$

da cui si deduce $[B(s) \cdot \mathbf{u}]' = 0$ e, quindi, la tesi. \square

Le eliche cilindriche sono caratterizzate dal rapporto tra curvatura e torsione.

PROPRIETÀ 1.14. *Una curva non piana è un'elica cilindrica se e solo se il rapporto tra torsione e curvatura è costante.*

DIMOSTRAZIONE. Supponiamo che $\beta(s)$ sia la riparametrizzazione della curva data mediante l'ascissa curvilinea.

Se β è un'elica cilindrica, allora $T(s) \cdot \mathbf{u} = \cos \theta$, per un certo versore \mathbf{u} che risulta essere ortogonale ad $N(s)$. Quindi, \mathbf{u} è combinazione lineare di $T(s)$ e $B(s)$: $\mathbf{u} = \cos \theta T(s) + \sin \theta B(s)$. Derivando si trova:

$$\begin{aligned} \mathbf{0} &= \cos \theta T'(s) + \sin \theta B'(s) = \cos \theta (k(s)N(s)) + \sin \theta (-\tau(s)N(s)) \\ &= (k(s) \cos \theta - \tau(s) \sin \theta) N(s). \end{aligned}$$

Quindi: $k(s) \cos \theta - \tau(s) \sin \theta = 0$ ed il rapporto $\tau/k = \cotg \theta$ è costante.

Viceversa, supponiamo che il rapporto torsione curvatura sia costante. Esiste allora un angolo θ tale che:

$$\frac{\tau(s)}{k(s)} = \frac{\cos \theta}{\sin \theta}. \quad (*)$$

Consideriamo il campo vettoriale unitario (a priori, dipendente da s):

$$\mathbf{u}(s) = \cos \theta T(s) + \sin \theta B(s).$$

Tenuto conto delle formule di Frenet e della condizione (*), si ha:

$$\mathbf{u}'(s) = \cos \theta T'(s) + \sin \theta B'(s) = (\cos \theta k(s) - \sin \theta \tau(s)) N(s) = \mathbf{0}.$$

Di conseguenza, $\mathbf{u}(s) = \mathbf{u}$ è un campo costante e

$$T(s) \cdot \mathbf{u} = T(s) \cdot [\cos \theta T(s) + \sin \theta B(s)] = \cos \theta.$$

Pertanto, l'angolo tra $T(s)$ e \mathbf{u} è costante e la curva è un'elica cilindrica. \square

ESEMPIO 1.13. Consideriamo la curva dell'Esempio 1.12. Adesso siamo in grado di dire che si tratta di un'elica cilindrica. Poichè $\tau/k = 1$, l'angolo θ vale $\pi/4$. Dalla dimostrazione del teorema precedente, vediamo che l'asse dell'elica è dato da:

$$\mathbf{u} = \cos \theta T(0) + \sin \theta B(0) = \frac{\sqrt{2}}{2}(1, 0, 0) + \frac{\sqrt{2}}{2}(0, 0, 1) = \frac{\sqrt{2}}{2}(1, 0, 1).$$

Un'elica si dice **circolare** se la sua traiettoria è contenuta in un cilindro circolare retto. Le eliche circolari sono caratterizzate dalla seguente

PROPRIETÀ 1.15. *Una curva non piana è un'elica circolare se e solo se la curvatura e la torsione sono costanti.*

Si prova anche che, in un opportuno sistema di coordinate, le equazioni di un'elica circolare sono del tipo: $\alpha(t) = (a \cos t, a \sin t, bt)$, $a > 0, b \neq 0$.

ESEMPIO 1.14. Data l'elica circolare (è l'Esempio 1.9, con $a = b = 1$)

$$\beta(s) = \left(\cos \left(\frac{s}{\sqrt{2}} \right), \sin \left(\frac{s}{\sqrt{2}} \right), \frac{s}{\sqrt{2}} \right), \quad s \in \mathbb{R},$$

si consideri la curva

$$\bar{\beta}(s) = \beta(s) + \frac{1}{k(s)} N(s), \quad s \in \mathbb{R}.$$

$\bar{\beta}$ è il luogo dei centri di curvatura di β e si dice **evolva** di β .

Verifichiamo che $\bar{\beta}$ è ancora un'elica circolare. Dall'Esempio 1.9, si ha subito

$$k(s) = \frac{1}{2}, \quad N(s) = \left(-\cos \left(\frac{s}{\sqrt{2}} \right), -\sin \left(\frac{s}{\sqrt{2}} \right), 0 \right)$$

e quindi

$$\bar{\beta}(s) = \left(-\cos \left(\frac{s}{\sqrt{2}} \right), -\sin \left(\frac{s}{\sqrt{2}} \right), \frac{s}{\sqrt{2}} \right).$$

La curvatura e la torsione di $\bar{\beta}$ valgono entrambe $1/2$ e quindi $\bar{\beta}$ è un'elica circolare.

8. Curve piane

Anche se le curve piane sono un caso particolare delle curve nello spazio e possono quindi essere studiate con i metodi precedenti, la loro importanza storico - pratica è tale che esse meritano una trattazione particolare. Ci limiteremo ad alcuni cenni ed a pochi esempi che mettono in evidenza le differenze rispetto alle curve nello spazio. In particolare, la definizione di curvatura è in parte diversa da quella vista prima: *la curvatura di una curva piana ha un segno.*

Sia $\beta(s) = (x(s), y(s))$ una curva piana parametrizzata mediante l'ascissa curvilinea. Come in precedenza, il campo tangente è dato da

$$T(s) = \beta'(s) = (x'(s), y'(s)).$$

Invece il campo normale è definito dalla condizione:

$$N(s) = iT(s) = (-y'(s), x'(s)),$$

dove i significa che il vettore $T(s)$ è ruotato di $\pi/2$ in senso antiorario. La scelta del simbolo i non è casuale: se si identificano i vettori del piano con

i numeri complessi, la rotazione di $\pi/2$ in senso antiorario corrisponde alla moltiplicazione di un numero complesso per l'unità immaginaria.

Dato che il campo T' è ortogonale a T (Lemma 1.4) e quindi parallelo ad N , esiste una funzione $k(s)$ tale che:

$$T'(s) = k(s) N(s). \quad (*)$$

Questa è la prima (e, come vedremo, anche l'unica) formula di Frenet per le curve piane. La funzione $k(s)$, detta **curvatura**, può assumere sia valori positivi che negativi. Vediamo il significato geometrico del suo segno. Indicato con $\theta(s)$ l'angolo fra il vettore $T(s)$ e l'asse x , si ha: $\cos \theta(s) = T(s) \cdot \mathbf{i}$ e $T(s) = (\cos \theta(s), \sin \theta(s))$ (ricordare che $T(s)$ è un versore). Allora

$$T'(s) = (-\theta'(s) \sin \theta(s), \theta'(s) \cos \theta(s)) = \theta'(s) N(s).$$

Dal confronto con la (*), si vede che

$$k(s) = \frac{d\theta(s)}{ds},$$

in altri termini: *la curvatura misura la variazione dell'angolo che la retta tangente forma con l'asse delle x .*

Negli intervalli in cui $k(s) > 0$ (oppure, $k(s) < 0$), l'angolo $\theta(s)$ è crescente (decrescente) e la curva ha la concavità nella stessa direzione (nella direzione opposta a quella) del campo N .

Derivando $N(s) = iT(s)$ si ottiene:

$$N'(s) = \frac{d(iT(s))}{ds} = iT'(s) = i(k(s) N(s)) = -k(s) T(s),$$

dato che $iN = i(iT) = i^2T = -T$. Da questa espressione si vede che la seconda formula di Frenet non contiene altre informazioni rispetto alla prima.

Data una curva piana con parametro qualsiasi, adattando la dimostrazione del Teorema 1.12, si prova il seguente

TEOREMA 1.16. *Sia $\alpha(t) = (x(t), y(t))$ una curva piana regolare. Allora:*

$$\begin{aligned} T(t) &= \frac{\alpha'(t)}{\|\alpha'(t)\|}, & N(t) &= \frac{i\alpha'(t)}{\|\alpha'(t)\|}, \\ k(t) &= \frac{\alpha''(t) \cdot (i\alpha'(t))}{\|\alpha'(t)\|^3}, & T'(t) &= k(t)\|\alpha'(t)\| N(t). \end{aligned}$$

ESEMPIO 1.15. Troviamo il riferimento di Frenet e la curvatura della curva $\alpha(t) = (t^2 - 1, 2t)$. Dato che:

$$\alpha'(t) = (2t, 2), \quad \alpha''(t) = (2, 0), \quad i\alpha'(t) = (-2, 2t),$$

applicando le formule del teorema precedente, si ha:

$$T(t) = \frac{1}{\sqrt{1+t^2}}(t, 1), \quad N(t) = \frac{1}{\sqrt{1+t^2}}(-1, t), \quad k(t) = \frac{-1}{2\sqrt{(1+t^2)^3}}.$$

Non è difficile riconoscere che α è la parabola di vertice $V(-1, 0)$, con asse parallelo all'asse x . Il fatto che la curvatura sia negativa significa che il campo normale N punta verso la regione "esterna" della parabola.

ESEMPIO 1.16. *Evolventi.* L'**evolvente** di una data curva α è una curva che interseca ortogonalmente le rette tangenti di α . Si dimostra che, se β è la riparametrizzazione di α mediante l'ascissa curvilinea, le evolventi di β sono date da:

$$\bar{\beta}(s) = \beta(s) - (s + c)T(s),$$

dove c è una costante arbitraria. Le evolventi di una curva sono quindi infinite.

Troviamo, ad esempio, le evolventi della circonferenza di centro l'origine e raggio a : $\alpha(t) = (a \cos t, a \sin t)$. Con un semplice calcolo si trova che l'ascissa curvilinea vale $s = at$ e quindi la riparametrizzazione della circonferenza è

$$\beta(s) = \left(a \cos \left(\frac{s}{a} \right), a \sin \left(\frac{s}{a} \right) \right).$$

da cui si deducono le equazioni delle evolventi:

$$\bar{\beta}(s) = \left(a \left[\cos \left(\frac{s}{a} \right) + \frac{s+c}{a} \sin \left(\frac{s}{a} \right) \right], a \left[\sin \left(\frac{s}{a} \right) - \frac{s+c}{a} \cos \left(\frac{s}{a} \right) \right] \right).$$

Calcolare la curvatura di $\bar{\beta}$.

ESEMPIO 1.17. *Evolute.* L'**evoluta** di una curva α data è l'unica curva $\bar{\alpha}$ che ha α come evolvente. Si dimostra che l'evoluta di una curva è il *luogo dei centri di curvatura* e quindi:

$$\bar{\alpha}(t) = \alpha(t) + \frac{1}{k(t)} N(t).$$

L'evoluta della parabola $\alpha(t) = (t^2 - 1, 2t)$ vista nell'Esempio 1.15 è $\bar{\alpha}(t) = (3t^2 + 1, 2t(1 - t - t^3))$.

Trovare l'evoluta della curva $\alpha(t) = (\cos t + t \sin t, \sin t - t \cos t)$ e verificare che si tratta di una circonferenza.

ESEMPIO 1.18. La **cicloide** è la curva descritta da un punto solidale con una circonferenza che rotola, senza strisciare, lungo un asse orizzontale.

Consideriamo un riferimento cartesiano in cui l'asse x coincide con l'asse del moto ed indichiamo con P il generico punto della cicloide. Supponiamo che la circonferenza abbia raggio 1 e che il suo centro, all'inizio del moto, stia sull'asse y . Detto C il centro della circonferenza che ruota, prendiamo come parametro t l'angolo formato dai vettori \overrightarrow{CP} e \mathbf{j} . Quindi: $\overrightarrow{OC} = t\mathbf{i} + \mathbf{j}$ e

$$\overrightarrow{CP} = \cos \left(-t - \frac{\pi}{2} \right) \mathbf{i} + \sin \left(-t - \frac{\pi}{2} \right) \mathbf{j} = -\sin t \mathbf{i} - \cos t \mathbf{j}.$$

Tenuto conto del fatto che $\overrightarrow{OP} = \overrightarrow{OC} + \overrightarrow{CP}$, le equazioni della cicloide sono:

$$\alpha(t) = (t - \sin t, 1 - \cos t).$$

Verificare che ci sono infiniti punti singolari e che la curvatura della cicloide vale:

$$k(t) = \frac{-1}{2\sqrt{2(1 - \cos t)}}.$$

9. Esercizi svolti sulle curve

ESERCIZIO 1.19. Data la curva nello spazio $\alpha(t) = (at - t^3, 3t^2, 3t + t^3)$,

- a) stabilire per quale valore di $a \in \mathbb{R}$ la curva è piana. Assegnato ad a tale valore, trovare l'equazione del piano che contiene la curva.
 b) Posto $a = 3$, determinare la curvatura, la torsione e il triedro di Frenet in un generico punto di α . Di quale curva si tratta?

SOLUZIONE. a) I primi tre campi derivati della curva sono:

$$\alpha'(t) = (a - 3t^2, 6t, 3 + 3t^2), \quad \alpha''(t) = (-6t, 6, 6t), \quad \alpha'''(t) = (-6, 0, 6).$$

La curva è piana se e solo se la sua torsione è nulla (Proprietà 1.8). Applicando le formule di pag. 13, si vede che basta cercare i valori di a che annullano $\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t)$. Dato che

$$\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) = \begin{vmatrix} a - 3t^2 & 6t & 3 + 3t^2 \\ -6t & 6 & 6t \\ -6 & 0 & 6 \end{vmatrix} = 36(a + 3),$$

α è piana se $a = -3$.

Dalla dimostrazione della Proprietà 1.8, si deduce che il piano contenente la curva passa per un punto (arbitrariamente scelto) della curva ed è perpendicolare al campo binormale. Tenuto conto delle espressioni del Teorema 1.12, il piano cercato passa per il punto $\alpha(0) = (0, 0, 0)$ ed è perpendicolare ad un vettore proporzionale a

$$\alpha'(t) \times \alpha''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 - 3t^2 & 6t & 3 + 3t^2 \\ -6t & 6 & 6t \end{vmatrix} = 18(t^2 - 1)(\mathbf{i} + \mathbf{k}).$$

L'equazione del piano è $x + z = 0$.

b) Per trovare il triedro di Frenet, la curvatura e la torsione si usano le formule del Teorema 1.12. Dalle espressioni precedenti, riscritte ponendo $a = 3$, si ricava:

$$\|\alpha'(t)\| = 3\sqrt{2}(1 + t^2), \quad \alpha'(t) \times \alpha''(t) = 18[(t^2 - 1)\mathbf{i} - 2t\mathbf{j} + (1 + t^2)\mathbf{k}],$$

$$\|\alpha'(t) \times \alpha''(t)\| = 18\sqrt{2}(1 + t^2), \quad \alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) = 216,$$

da cui:

$$T(t) = \frac{\sqrt{2}}{2(1 + t^2)} (1 - t^2, 2t, 1 + t^2), \quad N(t) = \frac{1}{1 + t^2} (-2t, 1 - t^2, 0),$$

$$B(t) = \frac{\sqrt{2}}{2(1 + t^2)} (t^2 - 1, -2t, t^2 + 1); \quad k(t) = \tau(t) = \frac{1}{3(1 + t^2)^2}.$$

Dato che il rapporto $\tau(t)/k(t)$ è costante, la curva è un'elica cilindrica (Proprietà 1.14).

ESERCIZIO 1.20. Data la curva nello spazio $\alpha(t) = (3t^2, 1 + 3t, at^3)$,

- a) stabilire per quale valore di $a \in \mathbb{R}$ la curva α è, rispettivamente: i) piana, ii) un'elica cilindrica.
 b) Posto $a = 2$, determinare la curvatura, la torsione e il triedro di Frenet in un generico punto di α .

- c) Calcolare la lunghezza dell'arco di curva compreso fra i punti $A(0, 1, 0)$ e $B(3, -2, -2)$.

SOLUZIONE. a) I primi tre campi derivati della curva sono:

$$\alpha'(t) = (6t, 3, 3at^2), \quad \alpha''(t) = (6, 0, 6at), \quad \alpha'''(t) = (0, 0, 6a).$$

Procedendo come nell'esercizio precedente, si trova che la curva è piana se $a = 0$.

Per vedere se α è un'elica, dopo aver calcolato la curvatura e la torsione

$$k(t) = \frac{2}{3} \sqrt{\frac{a^2 t^4 + a^2 t^2 + 1}{(a^2 t^4 + 4t^2 + 1)^3}}, \quad \tau(t) = -\frac{a}{3(a^2 t^4 + a^2 t^2 + 1)},$$

cerchiamo per quali valori di a il rapporto

$$\frac{\tau(t)}{k(t)} = -\frac{a}{2} \left(\frac{a^2 t^4 + 4t^2 + 1}{a^2 t^4 + a^2 t^2 + 1} \right)^{\frac{3}{2}}$$

è costante. Il valore $a = 0$ va scartato perchè, in tal caso, la curva è piana. Il numeratore ed il denominatore sono polinomi di quarto grado: l'unico modo per rendere costante il loro rapporto è che i due polinomi coincidano. Confrontando i coefficienti, si vede che ciò accade se $a^2 = 4$, ossia per $a = \pm 2$.

b) La curvatura e la torsione si deducono dalle espressioni precedenti ponendo $a = 2$. Il triedro di Frenet, in un punto generico della curva, è:

$$T(t) = \frac{1}{1 + 2t^2} (2t, 1, 2t^2), \quad N(t) = \frac{1}{2t^2 + 1} (1 - 2t^2, -2t, 2t),$$

$$B(t) = \frac{1}{1 + 2t^2} (2t, -2t^2, -1).$$

- c) Per calcolare la lunghezza dell'arco di curva si usa la formula di pag. 5. Dato che $A = \alpha(0)$ e $B = \alpha(-1)$, si ha:

$$\mathcal{L}(\alpha) = \int_{-1}^0 \|\alpha'(t)\| dt = \int_{-1}^0 3(1 + 2t^2) dt = 5.$$

ESERCIZIO 1.21. Data la curva nello spazio $\alpha(t) = (\cos t, \sin t, \cos at)$,

- a) stabilire per quali valori di $a \in \mathbb{R}$ la curva è piana. Quali curve si trovano in questo caso?
 b) Posto $a = 2$, determinare la curvatura e la torsione di α in un punto generico. Trovare il triedro di Frenet nel punto corrispondente a $t = \pi/2$.

SOLUZIONE. a) I primi tre campi derivati della curva sono:

$$\alpha'(t) = (-\sin t, \cos t, -a \sin at), \quad \alpha''(t) = (-\cos t, -\sin t, -a^2 \cos at),$$

$$\alpha'''(t) = (\sin t, -\cos t, a^3 \sin at).$$

Dato che $\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) = a(a^2 - 1) \sin at$, la curva è piana se $a = 0$ oppure $a = \pm 1$. Si noti che la funzione $\sin at$ è identicamente nulla solo per $a = 0$, valore già considerato.

Se $a = 0$, le equazioni parametriche di α sono:

$$\begin{cases} x = \cos t \\ y = \sin t \\ z = 1. \end{cases}$$

Dalle prime due si ricava $x^2 + y^2 = 1$: ciò significa che i punti della curva stanno su un cilindro circolare avente come asse l'asse z (vedi il successivo Esempio 2.13). La terza equazione dice che α è contenuta nel piano $z-1=0$, perpendicolare all'asse del cilindro. Si può allora concludere che la curva è una circonferenza di equazioni cartesiane:

$$\begin{cases} x^2 + y^2 = 1 \\ z - 1 = 0. \end{cases}$$

Se $a = \pm 1$, eliminando t dalle equazioni parametriche, si vede che la curva è intersezione del cilindro precedente col piano $x-z=0$. Questa volta si tratta di un'ellisse perchè il piano che contiene la curva non è nè perpendicolare nè parallelo all'asse del cilindro.

b) La curvatura e la torsione di α valgono:

$$k(t) = \sqrt{\frac{16 \cos^2 2t + 4 \sin^2 2t + 1}{(4 \sin^2 2t + 1)^3}}, \quad \tau(t) = \frac{6 \sin 2t}{16 \cos^2 2t + 4 \sin^2 2t + 1}.$$

Per trovare il triedro di Frenet nel punto $\alpha(\pi/2) = (0, 1, -1)$, conviene calcolare il valore dei vari campi vettoriali nel punto considerato e applicare il Teorema 1.12. Si trova:

$$T\left(\frac{\pi}{2}\right) = (-1, 0, 0), \quad N\left(\frac{\pi}{2}\right) = \frac{\sqrt{17}}{17} (0, -1, 4), \quad B\left(\frac{\pi}{2}\right) = \frac{\sqrt{17}}{17} (0, 4, 1).$$

ESERCIZIO 1.22. Data la curva nello spazio $\alpha(t) = (e^{at}, e^{-t}, \sqrt{2}t)$,

- a) stabilire per quali valori di $a \in \mathbb{R}$ la curva è piana.
 b) Posto $a = 1$, determinare la curvatura e la torsione di α in un punto generico. Di che curva si tratta?

SOLUZIONE. a) I primi tre campi derivati della curva sono:

$$\alpha'(t) = (ae^{at}, -e^{-t}, \sqrt{2}), \quad \alpha''(t) = (a^2e^{at}, e^{-t}, 0), \quad \alpha'''(t) = (a^3e^{at}, -e^{-t}, 0).$$

Poichè $\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) = -\sqrt{2}a^2(a+1)e^{(a-1)t}$, la curva è piana se $a = 0$ oppure $a = -1$.

Se $a = 0$, eliminando t dalle equazioni parametriche, si vede che α è la curva intersezione del cilindro generalizzato $y = e^{-\frac{\sqrt{2}}{2}z}$ col piano $x-1=0$. Dato che il piano è perpendicolare all'asse del cilindro, la curva è una copia della direttrice. Analogamente, per $a = -1$ la curva è l'intersezione del cilindro precedente con il piano $x-y=0$.

b) La curvatura e la torsione di α valgono:

$$k(t) = \frac{\sqrt{2}e^{2t}}{(1+e^{2t})^2}, \quad \tau(t) = -\frac{\sqrt{2}e^{2t}}{(1+e^{2t})^2}.$$

La curva è un'elica cilindrica.

Geometria differenziale delle superfici

In questo capitolo verranno date le definizioni e le proprietà fondamentali riguardanti la geometria differenziale delle superfici.

1. Premesse di topologia

In questa sezione sono elencate le nozioni e le proprietà topologiche che saranno usate successivamente. Per maggiori dettagli, si veda [5].

Si consideri \mathbb{R}^n dotato della struttura di spazio metrico. Dati un punto $P = (p_1, \dots, p_n) \in \mathbb{R}^n$ e un numero reale positivo r , l'insieme

$$B(P; r) = \{(x_1, \dots, x_n) \in \mathbb{R}^n / (x_1 - p_1)^2 + \dots + (x_n - p_n)^2 < r^2\}$$

si dice **intorno sferico** di P di raggio r . Un **intorno** di P è un qualsiasi sottoinsieme $N \subseteq \mathbb{R}^n$ tale che:

$$i) P \in N, \quad ii) B(P; r) \subseteq N, \text{ per un certo } r > 0.$$

In particolare, ogni intorno sferico è un intorno di P .

DEFINIZIONE 2.1. *Un sottoinsieme A di \mathbb{R}^n è **aperto** (rispetto alla topologia standard di \mathbb{R}^n) se, per ogni punto $P \in A$, esiste un intorno di P contenuto in A .*

*Un sottoinsieme C di \mathbb{R}^n è **chiuso** se il suo insieme complementare $\{Q \in \mathbb{R}^n / Q \notin C\}$ è aperto.*

In altri termini, un insieme aperto è unione di intorni sferici. Invece, un insieme che contiene un solo punto è chiuso.

Si verifica facilmente che valgono le seguenti

- PROPRIETÀ 2.1.** (1) \mathbb{R}^n e \emptyset sono aperti;
 (2) l'unione di un numero qualsiasi di aperti è aperta;
 (3) l'intersezione di un numero finito di aperti è aperta.

Si noti che l'intersezione di un numero infinito di aperti può non essere aperta. Ad esempio, l'intersezione degli insiemi aperti

$$B(O; \frac{1}{n}) = \left\{ (x_1, x_2) \in \mathbb{R}^2 / x_1^2 + x_2^2 < \frac{1}{n} \right\}, \quad n = 1, 2, 3, \dots,$$

è costituita dall'origine, che è un chiuso.

Vale una proprietà analoga alla precedente per gli insiemi chiusi.

- PROPRIETÀ 2.2.** (1) \mathbb{R}^n e \emptyset sono chiusi;
 (2) l'unione di un numero finito di chiusi è chiusa;
 (3) l'intersezione di un numero qualsiasi di chiusi è chiusa.

In questo caso, l'unione di un numero infinito di chiusi può non essere chiusa. Ad esempio, l'unione degli insiemi chiusi

$$\overline{B(O; \frac{1}{n})} = \left\{ (x_1, x_2) \in \mathbb{R}^2 / x_1^2 + x_2^2 \leq \frac{n+1}{n+2} \right\}, \quad n = 0, 1, 2, \dots,$$

coincide con l'insieme $B(0; 1)$, che è un aperto.

La Proprietà 2.1 motiva la seguente

DEFINIZIONE 2.2. *Sia X un insieme qualsiasi. Una **topologia** su X è una collezione $\mathcal{T} = \{A_i, i \in I\}$ di sottoinsiemi di X , detti **aperti**, che verifica le seguenti condizioni:*

- (1) X e \emptyset sono aperti, cioè $X \in \mathcal{T}$ e $\emptyset \in \mathcal{T}$;
- (2) l'unione di un numero qualsiasi di aperti è aperta;
- (3) l'intersezione di un numero finito di aperti è aperta.

La coppia (X, \mathcal{T}) si dice **spazio topologico**. Un sottoinsieme C di X è **chiuso** se il suo complementare $X - C$ è aperto.

ESEMPIO 2.1. Dato un qualsiasi insieme X , $\mathcal{T} = \{X, \emptyset\}$ è una topologia su X detta *topologia banale*.

ESEMPIO 2.2. Dato un qualsiasi insieme X , $\mathcal{T} = \{A / A \subseteq X\}$ è una topologia su X detta *topologia discreta*. In questo caso, ogni sottoinsieme di X è sia aperto che chiuso.

ESEMPIO 2.3. Siano (X, \mathcal{T}) uno spazio topologico e S un sottoinsieme di X . La *topologia indotta* su S ha come aperti i sottoinsiemi del tipo $S \cap A$, dove $A \in \mathcal{T}$ è un qualsiasi aperto di X .

DEFINIZIONE 2.3. *Siano (X, \mathcal{T}) e (X', \mathcal{T}') due spazi topologici.*

- (1) Una funzione $f: X \rightarrow X'$ è **continua** se, per ogni aperto A' di X' , la sua controimmagine $f^{-1}(A') = \{x \in X / f(x) \in A'\}$ è aperta in X .
- (2) Una funzione $f: X \rightarrow X'$ è un **omeomorfismo** se f è biettiva e se f e f^{-1} sono continue.

DEFINIZIONE 2.4. *Uno spazio topologico (X, \mathcal{T}) è **connesso** se non esistono due aperti A_1 e A_2 tali che:*

$$A_1 \neq \emptyset, \quad A_2 \neq \emptyset, \quad A_1 \cap A_2 = \emptyset, \quad X = A_1 \cup A_2.$$

OSSERVAZIONE 2.1. Intuitivamente, uno spazio topologico è connesso se è "fatto da un solo pezzo". \mathbb{R}^n , gli intervalli della retta reale, gli interni sferici di \mathbb{R}^n sono tutti connessi. Uno spazio topologico discreto che contiene almeno due punti distinti non è connesso.

Questa sezione termina con alcune nozioni necessarie per definire le varietà differenziabili nel Capitolo 6.

DEFINIZIONE 2.5. *Sia (X, \mathcal{T}) uno spazio topologico.*

- (1) Un sottoinsieme $\mathcal{B} \subseteq \mathcal{T}$ è una **base** per la topologia \mathcal{T} se ogni aperto di X è unione di aperti che appartengono a \mathcal{B} .
- (2) X è a **base numerabile** se esiste una base per la sua topologia formata da un insieme numerabile (cioè in biiezione con \mathbb{N}) di elementi.

(3) X è uno spazio di **Hausdorff** se, per ogni coppia di punti $P, Q \in X$, esistono due aperti U e V tali che:

$$P \in U, \quad Q \in V, \quad U \cap V = \emptyset.$$

ESEMPIO 2.4. Si consideri \mathbb{R}^n con la topologia standard. L'insieme \mathcal{B} formato da tutti gli intorni sferici $B(P; r)$, aventi centro in un punto P con coordinate *razionali* e di raggio $r \in \mathbb{Q}$, è una base numerabile di \mathbb{R}^n . \mathbb{R}^n è anche uno spazio di Hausdorff.

2. Superfici parametrizzate in \mathbb{R}^3

Per ragioni di carattere topologico, la definizione di superficie è più complessa di quella data per le curve: mentre in quest'ultimo caso una sola funzione era sufficiente per descrivere la curva, vedremo che, in generale, per individuare una superficie è necessaria più di una parametrizzazione.

Nel seguito si supporrà sempre che gli spazi \mathbb{R}^3 e \mathbb{R}^2 siano dotati della topologia standard. Si ricordi che un sottoinsieme S di \mathbb{R}^3 (oppure di \mathbb{R}^2) ha la topologia indotta (vedi l'Esempio 2.3) se i suoi aperti sono del tipo: $S \cap A$, dove A è un aperto di \mathbb{R}^3 (oppure di \mathbb{R}^2).

DEFINIZIONE 2.6. Una **superficie parametrizzata** è un sottoinsieme M di \mathbb{R}^3 tale che, per ogni punto $P \in M$, esistono un aperto connesso D di \mathbb{R}^2 , un aperto U di M contenente P ed una funzione

$$\varphi: D \subseteq \mathbb{R}^2 \rightarrow U \subseteq \mathbb{R}^3, \quad (u, v) \in D \mapsto \varphi(u, v) = (x(u, v), y(u, v), z(u, v)) \in U,$$

che verifica le seguenti condizioni:

- (1) φ è un omeomorfismo tra D e U , ossia φ è biettiva e, inoltre, φ e φ^{-1} sono continue;
- (2) φ è differenziabile, cioè le funzioni $x(u, v)$, $y(u, v)$ e $z(u, v)$ possiedono derivate parziali di ordine qualsiasi.

La coppia (D, φ) si dice **parametrizzazione locale** della superficie M . Le **coordinate locali** (o anche **coordinate curvilinee**) di un punto $Q \in U$ sono (u, v) , dove $\varphi(u, v) = Q$.

OSSERVAZIONE 2.2. Si noti che *ogni punto* della superficie deve essere contenuto nell'immagine di almeno una parametrizzazione locale; in generale, ci vorrà più di una parametrizzazione per ricoprire tutta la superficie (vedi il successivo Esempio 2.6).

ESEMPIO 2.5. Consideriamo il **piano** passante per il punto $P_0(x_0, y_0, z_0)$ e parallelo ai vettori (linearmente indipendenti) $\mathbf{a} = a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}$ e $\mathbf{b} = b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}$.

In questo caso, una sola parametrizzazione locale (D, φ) , dove $D = \mathbb{R}^2$ e

$$\varphi(u, v) = (x_0 + a_1 u + b_1 v, y_0 + a_2 u + b_2 v, z_0 + a_3 u + b_3 v),$$

è sufficiente per ricoprire tutto il piano (controllare che φ verifica le condizioni della definizione precedente).

ESEMPIO 2.6. Sia $S^2 = \{(x, y, z) \in \mathbb{R}^3 / x^2 + y^2 + z^2 = 1\}$ la **sfera** di centro l'origine e raggio 1.

Non è possibile ricoprire la sfera, che è compatta (vedi la Definizione 3.11 di [5]), con un'unica parametrizzazione locale. Infatti, non esiste alcun

omeomorfismo che manda un aperto di \mathbb{R}^2 in un compatto. In questa sezione, costruiremo tre diversi insiemi di parametrizzazioni locali che ricoprono la sfera, aventi ciascuno un numero decrescente di elementi.

Indicati con

$$U_+ = \{(x, y, z) \in \mathbb{R}^3 / z = \sqrt{1 - x^2 - y^2}, z \neq 0\},$$

$$U_- = \{(x, y, z) \in \mathbb{R}^3 / z = -\sqrt{1 - x^2 - y^2}, z \neq 0\},$$

gli “emisferi” nord e sud, rispettivamente, si controlla facilmente che le funzioni

$$\varphi_{\pm}: D \longrightarrow U_{\pm}, \quad (u, v) \longmapsto \varphi_{\pm}(u, v) = (u, v, \pm\sqrt{1 - u^2 - v^2}),$$

dove $D = \{(u, v) \in \mathbb{R}^2 / u^2 + v^2 < 1\}$, verificano le condizioni della Definizione 2.6 (si noti che φ_{\pm}^{-1} sono le proiezioni ortogonali degli emisferi sul piano Oxy). Quindi (D, φ_+) e (D, φ_-) sono due parametrizzazioni locali che ricoprono tutta la sfera, tranne l’“equatore”. Per ricoprire i punti mancanti, occorrono altre quattro parametrizzazioni locali (D, φ'_\pm) , (D, φ''_\pm) , dove:

$$\varphi'_\pm(u, v) = (\pm\sqrt{1 - u^2 - v^2}, u, v), \quad \varphi''_\pm(u, v) = (u, \pm\sqrt{1 - u^2 - v^2}, v).$$

Tutto ciò che è stato detto per la sfera di raggio 1, si estende (con ovvie modifiche: quali?) alla sfera $S^2(r)$ di raggio r qualsiasi.

Vogliamo ora introdurre il concetto di punto regolare di una superficie; vedremo successivamente che un punto è regolare se e solo se la superficie ha piano tangente in tale punto.

Data una superficie parametrizzata M , consideriamo una parametrizzazione locale (D, φ) e fissiamo un punto $P_0 = \varphi(u_0, v_0) \in \varphi(D)$. Le **curve coordinate** passanti per P_0 sono le curve

$$\alpha_1(u) = \varphi(u, v_0), \quad \alpha_2(v) = \varphi(u_0, v),$$

che si ottengono fissando una delle due variabili. I campi vettoriali tangenti alle curve coordinate sono:

$$\varphi_u(u, v_0) \stackrel{\text{def.}}{=} \alpha'_1(u) = \left(\frac{\partial x}{\partial u} \Big|_{(u, v_0)}, \frac{\partial y}{\partial u} \Big|_{(u, v_0)}, \frac{\partial z}{\partial u} \Big|_{(u, v_0)} \right),$$

$$\varphi_v(u_0, v) \stackrel{\text{def.}}{=} \alpha'_2(v) = \left(\frac{\partial x}{\partial v} \Big|_{(u_0, v)}, \frac{\partial y}{\partial v} \Big|_{(u_0, v)}, \frac{\partial z}{\partial v} \Big|_{(u_0, v)} \right).$$

DEFINIZIONE 2.7. *Un punto P_0 di una superficie M è **regolare** se esiste una parametrizzazione locale (D, φ) tale che: $P_0 = \varphi(u_0, v_0) \in \varphi(D)$ ed i vettori $\varphi_u(u_0, v_0)$ e $\varphi_v(u_0, v_0)$ sono linearmente indipendenti.*

*M è una **superficie regolare** se ogni suo punto è regolare. Un punto che non è regolare si dice **punto singolare**.*

OSSERVAZIONE 2.3. Si noti che un generico punto di $\varphi(D)$ è regolare se e solo se i vettori

$$\varphi_u(u, v) = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right), \quad \varphi_v(u, v) = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right)$$

verificano la condizione $\varphi_u(u, v) \times \varphi_v(u, v) \neq \mathbf{0}$. Ciò equivale a richiedere che la matrice (trasposta della Jacobiana di φ)

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{pmatrix}$$

abbia *rango 2 in ogni punto*. In tal caso, si dice che (D, φ) è una **parametrizzazione locale regolare**.

ESEMPIO 2.7. Il piano dell'Esempio 2.5 è una superficie regolare perchè i vettori $\varphi_u(u, v) = \mathbf{a}$ e $\varphi_v(u, v) = \mathbf{b}$ sono linearmente indipendenti per definizione.

ESEMPIO 2.8. Anche la sfera S^2 è una superficie regolare. Per giustificare questa affermazione, occorre verificare che le sei parametrizzazioni locali costruite nell'Esempio 2.6 sono regolari. Faremo la verifica solo per (D, φ_+) ; gli altri casi sono del tutto simili (cercare di capire che cosa cambia). Derivando parzialmente φ_+ rispetto ad u ed a v , si trova:

$$(\varphi_+)_u(u, v) = \left(1, 0, \frac{-u}{\sqrt{1-u^2-v^2}}\right), \quad (\varphi_+)_v(u, v) = \left(0, 1, \frac{-v}{\sqrt{1-u^2-v^2}}\right)$$

e quindi:

$$(\varphi_+)_u(u, v) \times (\varphi_+)_v(u, v) = \left(\frac{u}{\sqrt{1-u^2-v^2}}, \frac{v}{\sqrt{1-u^2-v^2}}, 1\right) \neq \mathbf{0}.$$

Concludiamo questa sezione con altre due parametrizzazioni della sfera, particolarmente importanti per le applicazioni.

ESEMPIO 2.9. Coordinate sferiche. Se P è un punto della sfera $S^2(r)$, indichiamo con P' la proiezione ortogonale di P sul piano Oxy . Detti: u l'angolo fra i vettori \mathbf{i} e $\overrightarrow{OP'}$, v l'angolo fra i vettori \overrightarrow{OP} e $\overrightarrow{OP'}$, le coordinate cartesiane di P sono: $(r \cos v \cos u, r \cos v \sin u, r \sin v)$. La parametrizzazione locale associata alle **coordinate sferiche** (r, u, v) è

$$\varphi(u, v) = (r \cos v \cos u, r \cos v \sin u, r \sin v), \quad (u, v) \in (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Si verifichi che è regolare. La parametrizzazione non comprende i punti per cui $u = 0$ (una semicirconferenza) e $v = \pm \frac{\pi}{2}$ (i poli). Per ricoprire i punti mancanti, occorrono altre tre parametrizzazioni che si ricavano dalla precedente con ovvie modifiche.

ESEMPIO 2.10. Proiezioni stereografiche. Vogliamo ora ricoprire la sfera con il numero minimo possibile di parametrizzazioni. Per semplicità considereremo solo la sfera S^2 di raggio uno.

Sia $P'(x', y', 0)$ un generico punto del piano Oxy . Tracciata la retta che unisce P' con il "polo nord" $N(0, 0, 1)$, cerchiamo le coordinate del punto P (diverso dal polo nord) in cui tale retta incontra la sfera. Dato che $\overrightarrow{NP'} = x' \mathbf{i} + y' \mathbf{j} - \mathbf{k}$, le equazioni parametriche della retta NP' sono:

$$\begin{cases} x = x' t \\ y = y' t \\ z = 1 - t \end{cases}$$

Sostituendo tali espressioni nell'equazione della sfera e risolvendo rispetto a t , si trovano le coordinate di P :

$$\left(\frac{2x'}{x'^2 + y'^2 + 1}, \frac{2y'}{x'^2 + y'^2 + 1}, \frac{x'^2 + y'^2 - 1}{x'^2 + y'^2 + 1} \right).$$

La funzione che manda P' in P è una biiezione tra il piano e la sfera privata del polo nord. La funzione inversa prende il nome di **proiezione stereografica dal polo nord** della sfera sul piano. Le considerazioni precedenti permettono di costruire una parametrizzazione (D, φ_N) , dove $D = \mathbb{R}^2$ e

$$\varphi_N(u, v) = \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{u^2 + v^2 - 1}{u^2 + v^2 + 1} \right),$$

che ricopre tutta la sfera tranne il polo nord (verificare che è regolare!). Per includere il punto mancante, si ricorre alla proiezione stereografica dal “polo sud” $S(0, 0, -1)$ che individua la parametrizzazione (D, φ_S) , dove $D = \mathbb{R}^2$ e

$$\varphi_S(u, v) = \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{1 - u^2 - v^2}{1 + u^2 + v^2} \right).$$

3. Esempi di superfici regolari

ESEMPIO 2.11. *Grafici di funzioni di due variabili.*

Data una funzione differenziabile di due variabili $z = f(x, y)$, definita su un dominio (aperto) D del piano, il suo grafico è l'insieme dei punti

$$M = \{P \in \mathbb{R}^3 / P = (x, y, f(x, y)), (x, y) \in D\}.$$

Una parametrizzazione locale che ricopre interamente M è (D, φ) , dove:

$$\varphi(u, v) = (u, v, f(u, v)), \quad (u, v) \in D.$$

Poichè $\varphi_u(u, v) = (1, 0, \frac{\partial f}{\partial u})$ e $\varphi_v(u, v) = (0, 1, \frac{\partial f}{\partial v})$, i vettori $\varphi_u(u, v)$ e $\varphi_v(u, v)$ sono, in ogni punto, linearmente indipendenti e, quindi, M è una superficie regolare. (D, φ) viene detta **parametrizzazione di Monge**.

Queste osservazioni consentono di concludere immediatamente che, ad esempio, il paraboloido ellittico $z = x^2 + y^2$ ed il paraboloido iperbolico (detto anche “superficie a sella”) $z = x^2 - y^2$ sono superfici regolari.

ESEMPIO 2.12. *Superfici di livello.*

Sia $w = g(x, y, z)$ una funzione differenziabile di tre variabili il cui gradiente $\nabla g = \frac{\partial g}{\partial x} \mathbf{i} + \frac{\partial g}{\partial y} \mathbf{j} + \frac{\partial g}{\partial z} \mathbf{k}$ non si annulla in nessun punto del dominio. Fissata una costante reale c , si dice **superficie di livello** di g l'insieme di punti (se non è vuoto)

$$M_c = \{(x, y, z) \in \mathbb{R}^3 / g(x, y, z) = c\}.$$

Proviamo che M_c è una superficie regolare, mostrando che ogni suo punto è contenuto in una parametrizzazione regolare. Sia $P_0 = (x_0, y_0, z_0)$ un generico punto di M_c . Dato che $(\nabla g)(x_0, y_0, z_0) \neq \mathbf{0}$, almeno una derivata parziale non sarà nulla in P_0 (e quindi anche in un intorno di P_0 , a causa della continuità delle derivate parziali). Supponiamo, ad esempio, che

$$\left(\frac{\partial g}{\partial z} \right)_{(x_0, y_0, z_0)} \neq 0.$$

Si può allora applicare il teorema delle funzioni implicite: esiste una funzione $z = h(x, y)$, definita in un intorno di (x_0, y_0) , che risolve l'equazione $g(x, y, z) = c$, ossia tale che $h(x_0, y_0) = z_0$ e $g(x, y, h(x, y)) = c$. A questo punto è facile verificare che $\varphi(u, v) = (u, v, h(u, v))$ individua una parametrizzazione locale regolare e che $\varphi(x_0, y_0) = P_0$.

Esempi di superfici regolari del tipo precedente sono le *quadriche* proprie:

$$M = \{(x, y, z) \in \mathbb{R}^3 / \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}, \quad (\text{ellissoidi})$$

$$M' = \{(x, y, z) \in \mathbb{R}^3 / \frac{x^2}{a^2} \pm \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1\}, \quad (\text{iperboloidi})$$

$$M'' = \{(x, y, z) \in \mathbb{R}^3 / \frac{x^2}{a^2} \pm \frac{y^2}{b^2} - 2z = 0\}, \quad (\text{paraboloidi})$$

(a, b e c costanti reali strettamente positive).

ESEMPIO 2.13. *Superfici rigate.*

Supponiamo che $\alpha(u) = (x(u), y(u), z(u))$, $u \in I$, sia una curva regolare e che $\mathbf{w}(u) = w_1(u)\mathbf{i} + w_2(u)\mathbf{j} + w_3(u)\mathbf{k}$ sia un campo vettoriale lungo α che *non si annulla in nessun punto*.

La superficie M , formata dalle rette passanti per i punti $\alpha(u)$ della curva e parallele a $\mathbf{w}(u)$, è una **superficie rigata** avente α come **direttrice**. Le rette contenute in M sono le **generatrici** della rigata.

Una parametrizzazione di una superficie rigata è

$$\varphi(x, y) = (x(u) + v w_1(u), y(u) + v w_2(u), z(u) + v w_3(u)), \quad (u, v) \in I \times \mathbb{R}.$$

Si riconosce facilmente se una parametrizzazione individua una rigata: le equazioni corrispondenti dipendono linearmente da uno dei due parametri. In generale, una superficie rigata può avere punti singolari; sono i punti in cui $\varphi_u(u, v) \times \varphi_v(u, v) = \mathbf{0}$, dove: $\varphi_u(u, v) = \alpha'(u) + v \mathbf{w}'(u)$ e $\varphi_v(u, v) = \mathbf{w}(u)$.

Le principali famiglie di superfici rigate sono le seguenti.

– *Cilindri generalizzati.* Sono caratterizzati dal fatto che le direttrici sono parallele fra di loro e, quindi, $\mathbf{w}(u) = \mathbf{w}_0$ è un campo costante.

Ad esempio, $\varphi(u, v) = (2 \cos u + v, 3 \sin u, v)$ è la parametrizzazione locale del cilindro ellittico avente come direttrice l'ellisse $\alpha(u) = (2 \cos u, 3 \sin u, 0)$ e generatrici parallele al vettore $\mathbf{w}_0 = \mathbf{i} + \mathbf{k}$.

– *Coni generalizzati.* In questo caso, le generatrici passano tutte per un punto fisso, detto **vertice** del cono.

Ad esempio, il cono ordinario di vertice $V(0, 0, 1)$ ed avente come direttrice la circonferenza $\alpha(u) = (\cos u, \sin u, 0)$, è parametrizzato da

$$\varphi(u, v) = ((1 - v) \cos u, (1 - v) \sin u, v), \quad (u, v) \in (0, 2\pi) \times \mathbb{R}.$$

Verificare che V è un punto singolare del cono. Come al solito, in questa parametrizzazione mancano i punti di una generatrice: quella passante per $(1, 0, 0)$ e per V .

– *Superfici inviluppo delle tangenti.* Sono le superfici formate dalle rette tangenti ad una data curva (non piana). Se α è la direttrice, una parametrizzazione locale della rigata delle tangenti è data da: $\varphi(u, v) = \alpha(u) + v \alpha'(u)$. Si verifica direttamente che tutti i punti di α sono singolari.

– *Quadriche rigate.* Tra le varie quadriche considerate al termine dell'Esempio 2.12, due sono anche rigate: l'iperboloide ellittico (detto anche iperboloide ad una falda) e il paraboloido iperbolico, aventi equazioni, rispettivamente:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} - 2z = 0.$$

Verifichiamo, ad esempio, che l'iperboloide $x^2 + y^2 - z^2 = 1$ è rigato. Scritta l'equazione nella forma $(x - z)(x + z) = (1 - y)(1 + y)$, passando ai rapporti, si trova:

$$\frac{x - z}{1 - y} = \frac{1 + y}{x + z},$$

da cui, detto λ il comune valore costante dei rapporti, si ricava:

$$\begin{cases} x - z = \lambda(1 - y) \\ 1 + y = \lambda(x + z). \end{cases}$$

Al variare di $\lambda \in \mathbb{R}$, si ottiene un insieme di rette (detto **schiera di rette**) che, per costruzione, genera la superficie. Considerando un rapporto diverso dal precedente, si costruisce una nuova schiera di rette: per ogni punto dell'iperboloide passano due rette contenute nella superficie.

Verificare, per esercizio, che la superficie a sella $z = x^2 - y^2$ è rigata, determinando le due schiere di rette che la generano.

ESEMPIO 2.14. *Superfici di rotazione.*

Sia M la superficie che si ottiene ruotando una curva *piana* regolare α (curva profilo) attorno ad una retta a (l'asse di rotazione) complanare con α . Al fine di evitare possibili singolarità, si escludono le intersezioni della curva profilo con l'asse. Per trovare una parametrizzazione locale di M , è sempre possibile supporre che $\alpha(u) = (x(u), 0, z(u))$, $u \in I$, sia una curva del piano Oxz , che l'asse di rotazione coincida con l'asse z e che $x(u) > 0$. Un generico punto di M sta su una circonferenza passante per un punto $(x(u), 0, z(u))$ di α e di raggio $x(u)$. Pertanto, una parametrizzazione della superficie di rotazione sarà:

$$\varphi(u, v) = (x(u) \cos v, x(u) \sin v, z(u)), \quad (u, v) \in I \times (0, 2\pi).$$

Dato che $\|\varphi_u(u, v) \times \varphi_v(u, v)\|^2 = x(u)^2[x'(u)^2 + z'(u)^2]$, la parametrizzazione è regolare. Essa ricopre M tranne i punti della curva profilo. Le curve coordinate $\varphi(u_0, v)$ e $\varphi(u, v_0)$ sono dette, rispettivamente, **paralleli** e **meridiani**.

Esempi importanti di superfici di rotazione (oltre alle sfere) sono i **tori di rivoluzione**, generati dalla rotazione, attorno all'asse z , della circonferenza $\alpha(u) = (a + r \cos u, 0, r \sin u)$, di centro $(a, 0, 0)$ e raggio r , con $0 < r < a$. Una parametrizzazione locale del toro è (D, φ) , dove:

$$\varphi(u, v) = ((a + r \cos u) \cos v, (a + r \cos u) \sin v, r \sin u), \quad D = (0, 2\pi) \times (0, 2\pi).$$

Questa parametrizzazione esclude un meridiano e un parallelo; per ricoprire i punti mancanti occorrono altre tre parametrizzazioni (quali?).

4. Piani tangenti ad una superficie

Siano M una superficie parametrizzata ed $\alpha: I \rightarrow \mathbb{R}^3$ una curva tracciata sulla superficie, cioè tale che $\alpha(I) \subseteq M$. Consideriamo una parametrizzazione locale (D, φ) che “interseca la curva”, ossia: $\varphi(D) \cap \alpha(I) \neq \emptyset$. Dato che φ è biiettiva, ad ogni punto $\alpha(t)$ della curva che sta in $\varphi(D)$ corrisponde un unico punto $(u(t), v(t))$ di D tale che: $\alpha(t) = \varphi(u(t), v(t))$. In altri termini, ogni curva tracciata su di una superficie è (localmente) l'immagine di una curva del piano.

Fissato un punto P_0 di M , diremo che un vettore \mathbf{v} di \mathbb{R}^3 è **tangente alla superficie** in P_0 se esiste una curva α , definita su un intervallo I e contenuta in M , tale che:

$$\alpha(t_0) = P_0 \quad \text{e} \quad \alpha'(t_0) = \mathbf{v}, \quad \text{per un certo } t_0 \in I.$$

Si noti che tale curva non è unica; anzi ci sono infinite curve che verificano le condizioni precedenti.

PROPRIETÀ 2.3. *Dato un punto $P_0 \in M$, sia (D, φ) una parametrizzazione locale tale che $P_0 = \varphi(u_0, v_0)$. Ogni vettore \mathbf{v} tangente alla superficie in P_0 è combinazione lineare dei vettori $\varphi_u(u_0, v_0)$ e $\varphi_v(u_0, v_0)$.*

DIMOSTRAZIONE. Per definizione di vettore tangente, esiste una curva

$$\alpha(t) = \varphi(u(t), v(t)) = (x(u(t), v(t)), y(u(t), v(t)), z(u(t), v(t)))$$

tale che: $\alpha(t_0) = P_0$ e $\mathbf{v} = \alpha'(t_0) = (x'(t_0), y'(t_0), z'(t_0))$. Applicando la regola di derivazione delle funzioni composte, si ha:

$$x'(t_0) = \left. \frac{\partial x}{\partial u} \right|_{(u_0, v_0)} \left. \frac{du}{dt} \right|_{t_0} + \left. \frac{\partial x}{\partial v} \right|_{(u_0, v_0)} \left. \frac{dv}{dt} \right|_{t_0}, \dots$$

e quindi:

$$\mathbf{v} = \left. \frac{du}{dt} \right|_{t_0} \left(\left. \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right)_{(u_0, v_0)} + \left. \frac{dv}{dt} \right|_{t_0} \left(\left. \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right)_{(u_0, v_0)}.$$

Posto

$$a = \left. \frac{du}{dt} \right|_{t_0} \quad \text{e} \quad b = \left. \frac{dv}{dt} \right|_{t_0},$$

si vede che $\mathbf{v} = a\varphi_u(u_0, v_0) + b\varphi_v(u_0, v_0)$. □

La proprietà precedente ha due importanti conseguenze:

- l'insieme dei vettori tangenti ad una superficie in un suo punto è uno spazio vettoriale;
- se P_0 è un punto regolare, i vettori $\varphi_u(u_0, v_0)$ e $\varphi_v(u_0, v_0)$ sono linearmente indipendenti. In tal caso, l'insieme dei vettori tangenti in P_0 è uno spazio vettoriale di dimensione due, detto **piano tangente** alla superficie in P_0 , che si indica con $T_{P_0}M$.

DEFINIZIONE 2.8. *Un **campo vettoriale** X su una superficie parametrizzata M è una funzione che assegna ad ogni punto P di M un vettore $X(P)$ di \mathbb{R}^3 , avente estremo iniziale in P .*

*Se la superficie è regolare, un campo vettoriale X si dice **tangente** (oppure, **normale**) alla superficie se $X(P)$ appartiene al piano tangente $T_P M$ (oppure, è perpendicolare al piano tangente $T_P M$).*

OSSERVAZIONE 2.4. Se la superficie è regolare, per ogni parametrizzazione locale (D, φ) , i campi vettoriali $\varphi_u(u, v)$ e $\varphi_v(u, v)$ sono tangenti alla superficie, mentre $\varphi_u(u, v) \times \varphi_v(u, v)$ è normale.

Se M è la superficie di livello di una funzione g (vedi l'Esempio 2.12), il campo gradiente $\nabla g = \frac{\partial g}{\partial x} \mathbf{i} + \frac{\partial g}{\partial y} \mathbf{j} + \frac{\partial g}{\partial z} \mathbf{k}$ è normale alla superficie.

5. Superfici orientabili

Intuitivamente una superficie è orientabile se ha due facce distinte. Più precisamente, vale la

DEFINIZIONE 2.9. Una superficie parametrizzata M è **orientabile** se esiste un campo vettoriale normale e non nullo in ogni punto di M .

Si noti che:

- (1) La definizione precedente è equivalente a: M è orientabile se esiste un campo vettoriale *unitario* normale alla superficie in ogni suo punto.
- (2) Ogni superficie regolare M è *localmente* orientabile: se (D, φ) è una parametrizzazione locale, allora

$$N(u, v) = \frac{\varphi_u(u, v) \times \varphi_v(u, v)}{\|\varphi_u(u, v) \times \varphi_v(u, v)\|}$$

è un campo unitario e normale definito sull'aperto $\varphi(D)$.

- (3) Sono orientabili le seguenti superfici:
 - a) le superfici che possono essere ricoperte da un'unica parametrizzazione locale; rientrano in questa categoria le superfici che sono il grafico di una funzione di due variabili;
 - b) le superfici di livello delle funzioni di tre variabili: se $M_C = \{(x, y, z) \in \mathbb{R}^3 / g(x, y, z) = c\}$, allora

$$N(x, y, z) = \frac{\nabla g(x, y, z)}{\|\nabla g(x, y, z)\|}$$

è un campo unitario normale in ogni punto di M_c .

Per capire quando si può estendere un campo normale *locale* ad un campo normale *globale* è necessario vedere come cambia tale campo in un punto della superficie che sta nell'intersezione di due parametrizzazioni locali.

Dato un punto $P \in M$, siano (D, φ) e $(\bar{D}, \bar{\varphi})$ due parametrizzazioni locali tali che $P \in W = \varphi(D) \cap \bar{\varphi}(\bar{D})$.

Se $P = \varphi(u, v) = \bar{\varphi}(\bar{u}, \bar{v})$, il passaggio dalle coordinate locali (u, v) alle coordinate (\bar{u}, \bar{v}) è dato dalla funzione

$$(\bar{\varphi}^{-1} \circ \varphi)(u, v) = (\bar{u}(u, v), \bar{v}(u, v)), \quad (u, v) \in \varphi^{-1}(W).$$

Analogamente, la funzione

$$(\varphi^{-1} \circ \bar{\varphi})(\bar{u}, \bar{v}) = (u(\bar{u}, \bar{v}), v(\bar{u}, \bar{v})), \quad (\bar{u}, \bar{v}) \in \bar{\varphi}^{-1}(W),$$

individua il cambiamento dalle coordinate locali (\bar{u}, \bar{v}) alle coordinate (u, v) . In particolare, se $P \in W$, allora

$$P = \varphi(u, v) = (x(u, v), y(u, v), z(u, v)) = (x(u(\bar{u}, \bar{v}), v(\bar{u}, \bar{v})), \dots) = \bar{\varphi}(\bar{u}, \bar{v})$$

e quindi

$$\begin{aligned}\bar{\varphi}_{\bar{u}} &= \left(\frac{\partial x}{\partial \bar{u}}, \frac{\partial y}{\partial \bar{u}}, \frac{\partial z}{\partial \bar{u}} \right) = \left(\frac{\partial x}{\partial u} \frac{\partial u}{\partial \bar{u}} + \frac{\partial x}{\partial v} \frac{\partial v}{\partial \bar{u}}, \dots \right) \\ &= \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right) \frac{\partial u}{\partial \bar{u}} + \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right) \frac{\partial v}{\partial \bar{u}} \\ &= \frac{\partial u}{\partial \bar{u}} \varphi_u + \frac{\partial v}{\partial \bar{u}} \varphi_v.\end{aligned}$$

Nello stesso modo, si trova

$$\bar{\varphi}_{\bar{v}} = \frac{\partial u}{\partial \bar{v}} \varphi_u + \frac{\partial v}{\partial \bar{v}} \varphi_v$$

e quindi la matrice di passaggio dalla base (φ_u, φ_v) alla base $(\bar{\varphi}_{\bar{u}}, \bar{\varphi}_{\bar{v}})$ è la matrice Jacobiana

$$\frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} = \begin{pmatrix} \frac{\partial u}{\partial \bar{u}} & \frac{\partial u}{\partial \bar{v}} \\ \frac{\partial v}{\partial \bar{u}} & \frac{\partial v}{\partial \bar{v}} \end{pmatrix}.$$

Si noti che $\det \frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} \neq 0$.

Vogliamo ora esprimere la condizione di orientabilità di una superficie mediante il cambiamento di coordinate locali. Siano M una superficie orientabile e N un campo vettoriale unitario normale alla superficie. Fissata una parametrizzazione locale (D, φ) , scambiando eventualmente u con v , è sempre possibile fare in modo che il verso del campo N coincida con quello del campo

$$\frac{\varphi_u \times \varphi_v}{\|\varphi_u \times \varphi_v\|}.$$

Se $(\bar{D}, \bar{\varphi})$ è un'altra parametrizzazione locale tale che $p \in \bar{\varphi}(\bar{D})$, allora

$$\begin{aligned}\bar{\varphi}_{\bar{u}} \times \bar{\varphi}_{\bar{v}} &= \left(\frac{\partial u}{\partial \bar{u}} \varphi_u + \frac{\partial v}{\partial \bar{u}} \varphi_v \right) \times \left(\frac{\partial u}{\partial \bar{v}} \varphi_u + \frac{\partial v}{\partial \bar{v}} \varphi_v \right) \\ &= \left(\frac{\partial u}{\partial \bar{u}} \frac{\partial v}{\partial \bar{v}} - \frac{\partial v}{\partial \bar{u}} \frac{\partial u}{\partial \bar{v}} \right) \varphi_u \times \varphi_v \\ &= \det \frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} \varphi_u \times \varphi_v.\end{aligned}$$

Di conseguenza, il verso del versore

$$\frac{\bar{\varphi}_{\bar{u}} \times \bar{\varphi}_{\bar{v}}}{\|\bar{\varphi}_{\bar{u}} \times \bar{\varphi}_{\bar{v}}\|}.$$

coincide con quello di N su $W = \varphi(D) \cap \bar{\varphi}(\bar{D})$ se e solo se

$$\det \frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} > 0.$$

Vale quindi il

TEOREMA 2.4. *Una superficie M è orientabile se e solo se è possibile ricoprire M con un sistema di parametrizzazioni locali in modo tale che il cambiamento di coordinate locali avvenga mediante una matrice Jacobiana con determinante positivo.*

Sfrutteremo il teorema per provare che il nastro di Möbius non è orientabile. Per costruire tale superficie, consideriamo la circonferenza α del piano Oxy di centro l'origine e raggio 2, percorsa in senso *orario* a partire dal punto $C_0(0, 2, 0)$. Le sue equazioni sono:

$$\alpha(u) = (2 \operatorname{sen} u, 2 \cos u, 0), \quad u \in (0, 2\pi).$$

Sia A_0B_0 il segmento perpendicolare all'asse y , passante per C_0 e di lunghezza 2. Facciamolo ruotare lungo la circonferenza α in modo tale che il centro del segmento stia su α e, nel punto $\alpha(u)$, il segmento corrispondente formi l'angolo $\frac{u}{2}$ con la verticale. Il **nastro di Möbius** M è la superficie rigata avente α come curva direttrice ed avente per generatrici i segmenti (di lunghezza 2) paralleli al campo vettoriale $\mathbf{w}(u) = A - C$, dove A e $C = \alpha(u)$ sono, rispettivamente, un estremo del segmento ruotato ed il centro dello stesso segmento. Si noti che, dopo una rotazione completa, il segmento ritorna nella posizione iniziale con gli estremi scambiati. Una parametrizzazione locale del nastro di Möbius è data da

$$\varphi(u, v) = \alpha(u) + v(A - \alpha(u)), \quad (u, v) \in (0, 2\pi) \times (-1, 1).$$

Per scrivere le equazioni di M è necessario trovare coordinate dell'estremo A del segmento. Sia A' la proiezione ortogonale di A sul piano Oxy . Dato che l'angolo fra i segmenti OC e CA vale $\frac{\pi}{2} - \frac{u}{2}$, si ha:

$$\overline{A'C} = \overline{AC} \cos\left(\frac{\pi}{2} - \frac{u}{2}\right) = \operatorname{sen} \frac{u}{2}$$

e quindi

$$\overline{OA'} = \overline{OC} - \overline{A'C} = 2 - \operatorname{sen} \frac{u}{2},$$

da cui si ricava che $A \left((2 - \operatorname{sen} \frac{u}{2}) \operatorname{sen} u, (2 - \operatorname{sen} \frac{u}{2}) \cos u, \cos \frac{u}{2} \right)$ e

$$\mathbf{w}(u) = A - \alpha(u) = \left(-\operatorname{sen} \frac{u}{2} \operatorname{sen} u, -\operatorname{sen} \frac{u}{2} \cos u, \cos \frac{u}{2} \right).$$

In conclusione

$$\varphi(u, v) = \left((2 - v \operatorname{sen} \frac{u}{2}) \operatorname{sen} u, (2 - v \operatorname{sen} \frac{u}{2}) \cos u, v \cos \frac{u}{2} \right), \\ (u, v) \in D = (0, 2\pi) \times (-1, 1).$$

In questa parametrizzazione mancano i punti per cui $u = 0$, ossia i punti del segmento iniziale A_0B_0 . Costruiamo una nuova parametrizzazione $(\overline{D}, \overline{\varphi})$, simile alla precedente, avente come punto iniziale della circonferenza il punto $C_1(2, 0, 0)$ e come segmento iniziale il segmento passante per C_1 , che risulta essere inclinato di $\frac{\pi}{4}$ rispetto alla verticale. Le equazioni della circonferenza sono:

$$\alpha(\overline{u}) = (2 \cos \overline{u}, -2 \operatorname{sen} \overline{u}, 0), \quad \overline{u} \in (0, 2\pi).$$

Tenuto conto del fatto che adesso, nel punto $\alpha(\overline{u})$, il segmento che ruota è inclinato di $\frac{\pi}{4} + \frac{\overline{u}}{2}$ rispetto alla verticale, la nuove equazioni parametriche di

M sono:

$$\begin{cases} x = \left[2 - \bar{v} \operatorname{sen} \left(\frac{\pi}{4} + \frac{\bar{u}}{2} \right) \right] \cos \bar{u} \\ y = - \left[2 - \bar{v} \operatorname{sen} \left(\frac{\pi}{4} + \frac{\bar{u}}{2} \right) \right] \operatorname{sen} \bar{u}, & (\bar{u}, \bar{v}) \in \bar{D} = (0, 2\pi) \times (-1, 1). \\ z = \bar{v} \cos \left(\frac{\pi}{4} + \frac{\bar{u}}{2} \right) \end{cases}$$

Si noti che $\bar{\varphi}(\bar{D})$ non comprende il segmento passante per $C_1(2, 0, 0)$, corrispondente a $\bar{u} = 0$ (e a $u = \frac{\pi}{2}$). Di conseguenza, l'aperto $W = \varphi(D) \cap \bar{\varphi}(\bar{D})$ coincide con il nastro di Möbius privato di due segmenti. In altri termini, W è l'unione di due aperti *disgiunti*:

$$W_1 = \{\varphi(u, v) \in M / \frac{\pi}{2} < u < 2\pi\}, \quad W_2 = \{\varphi(u, v) \in M / 0 < u < \frac{\pi}{2}\}.$$

Se $P \in W_1$, allora $P = \varphi(u, v) = \bar{\varphi}(\bar{u}, \bar{v})$, dove

$$\begin{cases} \bar{u} = u - \frac{\pi}{2} \\ \bar{v} = v. \end{cases}$$

La matrice Jacobiana del cambiamento di coordinate locali su W_1 è

$$\frac{\partial(\bar{u}, \bar{v})}{\partial(v, v)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

e quindi $\det \frac{\partial(\bar{u}, \bar{v})}{\partial(v, v)} = 1$.

Se $P \in W_2$, i segmenti generatori hanno verso opposto a seconda del fatto che $P \in \varphi(D)$ oppure $P \in \bar{\varphi}(\bar{D})$. Dato che

$$\begin{cases} \bar{u} = u + \frac{3\pi}{2} \\ \bar{v} = -v, \end{cases}$$

la matrice Jacobiana del cambiamento di coordinate locali su W_2 è

$$\frac{\partial(\bar{u}, \bar{v})}{\partial(v, v)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

e $\det \frac{\partial(\bar{u}, \bar{v})}{\partial(v, v)} = -1$.

Ciò significa che un campo vettoriale unitario normale definito su $\varphi(D)$ non può essere esteso ad un campo su $\bar{\varphi}(\bar{D})$ in modo tale che il suo verso sia coerente su $W = \varphi(D) \cap \bar{\varphi}(\bar{D})$. In conclusione, il nastro di Möbius non è una superficie orientabile.

6. Metrica su una superficie: prima forma quadratica

In questa sezione verranno introdotti alcuni concetti di carattere metrico che consentiranno, tra l'altro, di introdurre una nozione intrinseca di distanza tra punti e di calcolare l'area di certe porzioni di superficie.

Data una superficie regolare M , fissiamo un punto P di M e consideriamo una parametrizzazione locale (D, φ) tale che $P = \varphi(u, v)$. Presi due vettori tangenti $\mathbf{u}, \mathbf{v} \in T_P M$, la Proprietà 2.3 dice che essi sono combinazione

lineare di $\varphi_u(u, v)$ e $\varphi_v(u, v)$ e, quindi:

$$\mathbf{u} = x_1\varphi_u(u, v) + x_2\varphi_v(u, v), \quad \mathbf{v} = y_1\varphi_u(u, v) + y_2\varphi_v(u, v),$$

per certi scalari x_1, x_2, y_1, y_2 . Allo scopo di semplificare le notazioni, nel seguito verrà spesso omessa la dipendenza dei vettori da u e v . Calcoliamo il prodotto scalare di \mathbf{u} e \mathbf{v} :

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (x_1\varphi_u + x_2\varphi_v) \cdot (y_1\varphi_u + y_2\varphi_v) \\ &= x_1y_1\varphi_u \cdot \varphi_u + (x_1y_2 + x_2y_1)\varphi_u \cdot \varphi_v + x_2y_2\varphi_v \cdot \varphi_v. \end{aligned}$$

Seguendo la tradizione, poniamo

$$E = \varphi_u \cdot \varphi_u = \|\varphi_u\|^2, \quad F = \varphi_u \cdot \varphi_v, \quad G = \varphi_v \cdot \varphi_v = \|\varphi_v\|^2$$

e riscriviamo il prodotto scalare nella forma seguente:

$$\mathbf{u} \cdot \mathbf{v} = E x_1 y_1 + F (x_1 y_2 + x_2 y_1) + G x_2 y_2.$$

Al variare di \mathbf{u} e \mathbf{v} nel piano tangente, l'espressione precedente individua una forma bilineare simmetrica su $T_P M$. Dato che tale forma bilineare è definita positiva, il determinante della matrice associata sarà strettamente positivo:

$$\det \begin{pmatrix} E & F \\ F & G \end{pmatrix} = EG - F^2 > 0.$$

La relativa forma quadratica

$$\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2 = E (x_1)^2 + 2F x_1 x_2 + G (x_2)^2$$

viene detta **prima forma quadratica fondamentale** della superficie.

Si noti che i **coefficienti** E , F e G della prima forma quadratica sono degli scalari che dipendono dalle coordinate del punto P ; al variare di P in $\varphi(D)$ diventano funzioni (differenziabili) di u e v .

OSSERVAZIONE 2.5. Dato un vettore tangente $\mathbf{v} \in T_P M$, consideriamo una curva $\alpha(t) = \varphi(u(t), v(t))$, passante per P e tangente a \mathbf{v} . Dalla dimostrazione della Proprietà 2.3 si ricava che

$$\mathbf{v} = \frac{du}{dt} \varphi_u + \frac{dv}{dt} \varphi_v$$

e, quindi

$$\|\mathbf{v}\|^2 = E \left(\frac{du}{dt} \right)^2 + 2F \frac{du}{dt} \frac{dv}{dt} + G \left(\frac{dv}{dt} \right)^2.$$

Passando ai differenziali, si ha l'espressione classica della prima forma quadratica:

$$ds^2 = E (du)^2 + 2F dudv + G (dv)^2.$$

In questa formula, ds^2 ha il significato di lunghezza di un arco "infinitesimo" di curva tracciato sulla superficie.

ESEMPIO 2.15. Consideriamo il piano passante per un punto P_0 e parallelo ai vettori \mathbf{a} e \mathbf{b} . Dato che tali vettori sono linearmente indipendenti, cambiando eventualmente base, possiamo sempre supporre che formino una base ortonormale. Usando la parametrizzazione locale dell'Esempio 2.5, si ha che $\varphi_u = \mathbf{a}$ e $\varphi_v = \mathbf{b}$. Pertanto i coefficienti della prima forma quadratica sono:

$$E = \mathbf{a} \cdot \mathbf{a} = 1, \quad F = \mathbf{a} \cdot \mathbf{b} = 0, \quad G = \mathbf{b} \cdot \mathbf{b} = 1.$$

Vediamo ora alcune importanti applicazioni della prima forma quadratica.

Angolo fra due curve su una superficie. Siano α_1 e α_2 due curve su M che si incontrano in un punto $P = \alpha_1(t_1) = \alpha_2(t_2)$. L'**angolo** θ fra le due curve è l'angolo formato dai rispettivi vettori tangenti in P e quindi

$$\cos \theta = \frac{\alpha'_1(t_1) \cdot \alpha'_2(t_2)}{\|\alpha'_1(t_1)\| \|\alpha'_2(t_2)\|}.$$

In particolare, se α_1 e α_2 sono curve coordinate, allora $\alpha'_1 = \varphi_u$ e $\alpha'_2 = \varphi_v$ e il loro angolo è dato da

$$\cos \theta = \frac{\varphi_u \cdot \varphi_v}{\|\varphi_u\| \|\varphi_v\|} = \frac{F}{\sqrt{E} \sqrt{G}}.$$

Da questa espressione si ottiene subito la seguente

PROPRIETÀ 2.5. *Le curve coordinate sono ortogonali se e solo se $F = 0$.*

ESEMPIO 2.16. Sia M il cilindro circolare retto ottenuto ruotando la retta $\alpha(u) = (1, 0, u)$ attorno all'asse z . Una parametrizzazione locale di M è (vedi l'Esempio 2.14):

$$\varphi(u, v) = (\cos v, \sin v, u), \quad (u, v) \in \mathbb{R} \times (0, 2\pi).$$

Dato che $\varphi_u(u, v) = (0, 0, 1)$ e $\varphi_v(u, v) = (-\sin v, \cos v, 0)$, i coefficienti della prima forma quadratica sono: $E = G = 1$ e $F = 0$. Notare che le linee coordinate (circonferenze e rette, in questo caso) sono ortogonali.

OSSERVAZIONE 2.6. Il fatto che la prima forma quadratica del cilindro coincida con quella del piano (Esempio 2.15) implica che queste due superfici sono localmente isometriche: ciò significa che, in un intorno sufficientemente piccolo di un punto, il cilindro ha le stesse proprietà metriche (angoli e distanze) del piano. D'altro canto, per ragioni di carattere topologico, il cilindro ed il piano non sono omeomorfi e, di conseguenza, non possono essere globalmente isometrici. Per una spiegazione più precisa di queste affermazioni, si veda il capitolo 4 di [4].

Distanza tra coppie di punti di una superficie. Sia $\alpha(t) = \varphi(u(t), v(t))$ un arco di curva che congiunge due punti $P_1 = \alpha(t_1)$ e $P_2 = \alpha(t_2)$ di M . Dato che (vedi l'Osservazione di pag. 36)

$$\alpha'(t) = \frac{du}{dt} \varphi_u + \frac{dv}{dt} \varphi_v,$$

la lunghezza dell'arco di curva è (Definizione 1.7):

$$\mathcal{L}(\alpha) = \int_{t_1}^{t_2} \|\alpha'(t)\| dt = \int_{t_1}^{t_2} \sqrt{E \left(\frac{du}{dt}\right)^2 + 2F \frac{du}{dt} \frac{dv}{dt} + G \left(\frac{dv}{dt}\right)^2} dt.$$

Supponiamo ora che la superficie sia *connessa* (Definizione 2.4). La **distanza** di due punti P_1 e P_2 di M si definisce come l'estremo inferiore delle lunghezze degli archi di curva che congiungono i due punti. Più precisamente, si prova che la funzione:

$$d(P_1, P_2) = \inf_{\gamma} \mathcal{L}(\gamma),$$

dove γ è una qualsiasi curva che unisce P_1 a P_2 , verifica tutte le condizioni della Definizione 1.7 di [5] e che quindi (M, d) è uno spazio metrico. Inoltre la topologia metrica (Proprietà 1.15 di [5]) coincide con la topologia di M indotta dalla topologia standard di \mathbb{R}^3 .

L'ipotesi che la superficie sia connessa serve per garantire l'esistenza di una curva che passi per i due punti. Se M fosse sconnessa, non ci sarebbero curve che congiungono punti appartenenti a componenti connesse distinte.

Un arco di curva avente lunghezza minima tra due suoi punti si dice **geodetica**. Ad esempio, le geodetiche del piano sono i segmenti e quelle della sfera sono gli archi di cerchio massimo.

Area di una porzione di superficie. Data una parametrizzazione locale (D, φ) , l'**area** della porzione $\varphi(D)$ della superficie è

$$\mathcal{A} = \int_D \|\varphi_u(u, v) \times \varphi_v(u, v)\| \, dudv,$$

supposto che l'integrale sia convergente. Dato che $\|\varphi_u \times \varphi_v\|$ è l'area di un parallelogramma di lati φ_u e φ_v , una motivazione intuitiva dell'espressione precedente è che $\|\varphi_u \times \varphi_v\| \, dudv$ rappresenta l'area di una "porzione infinitesima" di superficie.

Tenuto conto che vale l'identità vettoriale $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ (la dimostrazione è semplice: basta scrivere i vettori come combinazione lineare della base standard e calcolare direttamente il primo ed il secondo membro), si ha:

$$\|\varphi_u \times \varphi_v\|^2 = \|\varphi_u\|^2 \|\varphi_v\|^2 - (\varphi_u \cdot \varphi_v)^2 = EG - F^2 \quad (*)$$

e quindi la formula per il calcolo dell'area diventa:

$$\mathcal{A} = \int_D \sqrt{EG - F^2} \, dudv.$$

ESEMPIO 2.17. Calcoliamo l'area della sfera $S^2(r)$ di raggio r . La parametrizzazione più conveniente è quella data dalle coordinate sferiche, introdotte nell'Esempio 2.9 di pag. 27. Con semplici calcoli si ha che

$$E = \varphi_u \cdot \varphi_u = r^2 \cos^2 v, \quad F = \varphi_u \cdot \varphi_v = 0, \quad G = \varphi_v \cdot \varphi_v = r^2.$$

L'area della sfera vale

$$\mathcal{A} = \int_D r^2 \cos v \, dudv = r^2 \int_0^{2\pi} \left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos v \, dv \right) du = 4\pi r^2.$$

Si noti che, anche se $\varphi(D)$ non ricopre tutta la sfera, la parte mancante (una semicirconferenza più i poli) è un insieme di area nulla e quindi non modifica il valore dell'integrale.

ESEMPIO 2.18. Troviamo l'area del toro di rivoluzione usando la parametrizzazione dell'Esempio 2.14 di pag. 30. In questo caso

$$E = r^2, \quad F = 0, \quad G = (a + r \cos u)^2$$

e quindi

$$\mathcal{A} = \int_D r(a + \cos u) \, dudv = r \int_0^{2\pi} \left(\int_0^{2\pi} (a + r \cos u) \, du \right) dv = 4ar\pi^2.$$

7. Operatore forma: seconda forma quadratica

In questa sezione, allo scopo di rendere precisa l'idea intuitiva che una superficie si "incurva" nello spazio, verrà introdotta la nozione di curvatura Gaussiana. L'idea geometrica è quella di considerare la variazione del piano tangente, individuato da un versore normale, quando il punto di contatto con la superficie descrive una curva.

Fissato un punto P di una superficie regolare M , consideriamo una parametrizzazione locale (D, φ) tale che $P = \varphi(u_0, v_0)$. Sia $\mathbf{v} \in T_P M$ un vettore tangente alla superficie in P e supponiamo che $\mathbf{v} = a \varphi_u(u_0, v_0) + b \varphi_v(u_0, v_0)$. Inoltre, sia $\alpha(t) = \varphi(u(t), v(t))$ una *qualsiasi* curva passante per P e tangente a \mathbf{v} , ossia tale che:

$$u(t_0) = u_0, \quad v(t_0) = v_0 \quad \text{e} \quad \left. \frac{du}{dt} \right|_{t_0} = a, \quad \left. \frac{dv}{dt} \right|_{t_0} = b. \quad (*)$$

Il campo vettoriale, detto **campo normale**,

$$N(u, v) = \frac{\varphi_u(u, v) \times \varphi_v(u, v)}{\|\varphi_u(u, v) \times \varphi_v(u, v)\|}, \quad (u, v) \in D,$$

ha norma unitaria ed è, in ogni punto, perpendicolare al piano tangente. Come al solito, indicheremo con N_u e N_v i campi vettoriali aventi come componenti le derivate parziali delle componenti di N , rispetto ad u e v . Dato che N ha norma costante, N_u e N_v sono *tangenti* alla superficie in ogni punto del loro dominio (il Lemma 1.4 vale anche per le superfici, con ovvie modifiche: quali?).

Consideriamo la restrizione $N(t) = N(u(t), v(t))$ del campo normale ai punti della curva α e calcoliamo la sua derivata nel punto P : come risultato otterremo un vettore che esprime la variazione del piano tangente quando ci si avvicina a P lungo la curva α . Applicando, come al solito, la regola di derivazione delle funzioni composte, tenuto conto delle condizioni iniziali (*), otteniamo:

$$\begin{aligned} N'(t_0) &= \left. \frac{dN(u(t), v(t))}{dt} \right|_{t_0} = N_u(u_0, v_0) \left. \frac{du}{dt} \right|_{t_0} + N_v(u_0, v_0) \left. \frac{dv}{dt} \right|_{t_0} \\ &= a N_u(u_0, v_0) + b N_v(u_0, v_0). \end{aligned}$$

Il calcolo precedente permette di trarre alcune importanti conclusioni:

- il vettore $N'(t_0)$ dipende solo dal punto P e dal vettore \mathbf{v} e *non dalla curva* che interviene nel calcolo;
- essendo combinazione lineare di $N_u(u_0, v_0)$ e $N_v(u_0, v_0)$, $N'(t_0)$ è *tangente* alla superficie;
- $N'(t_0)$ dipende in modo *lineare* dalle componenti di \mathbf{v} .

Possiamo riassumere le affermazioni precedenti nella seguente

PROPRIETÀ 2.6. *La funzione*

$$S_P: T_P M \longrightarrow T_P M, \quad \mathbf{v} \longmapsto S_P(\mathbf{v}) = -N'(t_0),$$

è un endomorfismo del piano tangente, detto **operatore forma** in P .

OSSERVAZIONE 2.7. In molti testi, l'operatore forma prende il nome di **applicazione di Gauss–Weingarten**. Il segno meno ha il solo scopo di semplificare alcune formule successive.

PROPRIETÀ 2.7. S_P è un endomorfismo simmetrico di $T_P M$, ossia:

$$S_P(\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot S_P(\mathbf{w}), \quad \text{per ogni } \mathbf{v}, \mathbf{w} \in T_P M.$$

DIMOSTRAZIONE. Dati $\mathbf{v}, \mathbf{w} \in T_P M$, supponiamo che $\mathbf{v} = a_1\varphi_u + b_1\varphi_v$ e $\mathbf{w} = a_2\varphi_u + b_2\varphi_v$. Allora:

$$\begin{aligned} S_P(\mathbf{v}) \cdot \mathbf{w} &= -(a_1N_u + b_1N_v) \cdot (a_2\varphi_u + b_2\varphi_v) \\ &= -(a_1a_2N_u \cdot \varphi_u + a_1b_2N_u \cdot \varphi_v + a_2b_1N_v \cdot \varphi_u + b_1b_2N_v \cdot \varphi_v). \end{aligned}$$

Derivando l'identità $N \cdot \varphi_u = 0$ rispetto a v e $N \cdot \varphi_v = 0$ rispetto ad u , si ha:

$$N_v \cdot \varphi_u + N \cdot \varphi_{uv} = 0, \quad N_u \cdot \varphi_v + N \cdot \varphi_{vu} = 0.$$

Dato che $\varphi_{uv} = \varphi_{vu}$, otteniamo:

$$N_v \cdot \varphi_u = -N \cdot \varphi_{uv} = N_u \cdot \varphi_v$$

e quindi

$$S_P(\mathbf{v}) \cdot \mathbf{w} = -[a_1a_2N_u \cdot \varphi_u + (a_1b_2 + a_2b_1)N_u \cdot \varphi_v + b_1b_2N_v \cdot \varphi_v].$$

Essendo tale espressione simmetrica rispetto alle componenti di \mathbf{v} e \mathbf{w} , si ha la tesi. \square

Una conseguenza della proprietà precedente è che la forma bilineare:

$$(\mathbf{v}, \mathbf{w}) \in T_P M \times T_P M \longmapsto S_P(\mathbf{v}) \cdot \mathbf{w} \in \mathbb{R}$$

è simmetrica. La forma quadratica associata

$$\mathbf{v} = a\varphi_u + b\varphi_v \longmapsto S_P(\mathbf{v}) \cdot \mathbf{v} = -(a^2N_u \cdot \varphi_u + 2abN_u \cdot \varphi_v + b^2N_v \cdot \varphi_v)$$

si dice **seconda forma quadratica fondamentale** della superficie in P .

Seguendo ancora una volta la tradizione, indichiamo con

$$l = -N_u \cdot \varphi_u, \quad m = -N_u \cdot \varphi_v, \quad n = -N_v \cdot \varphi_v,$$

i coefficienti della seconda forma quadratica. Derivando l'identità $N \cdot \varphi_u = 0$ rispetto ad u e $N \cdot \varphi_v = 0$ rispetto a v , si ottengono le seguenti espressioni equivalenti dei coefficienti:

$$l = N \cdot \varphi_{uu}, \quad m = N \cdot \varphi_{uv}, \quad n = N \cdot \varphi_{vv},$$

Sia

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

la matrice dell'endomorfismo S_P rispetto alla base $(\varphi_u(u_0, v_0), \varphi_v(u_0, v_0))$ di $T_P M$. Dato che, in generale, tale base non è ortonormale, non è detto che la matrice sia simmetrica.

Dalla definizione stessa dell'operatore forma ricaviamo che $S_P(\varphi_u) = -N_u$ e quindi:

$$\begin{aligned} l &= -N_u \cdot \varphi_u = S_P(\varphi_u) \cdot \varphi_u = (a_{11}\varphi_u + a_{21}\varphi_v) \cdot \varphi_u \\ &= a_{11}\varphi_u \cdot \varphi_u + a_{21}\varphi_v \cdot \varphi_u = a_{11}E + a_{21}F. \end{aligned}$$

In modo analogo, si ottiene la relazione $m = a_{11}F + a_{21}G$. Risolvendo il sistema lineare (nelle incognite a_{11} e a_{21})

$$\begin{cases} Ea_{11} + Fa_{21} = l \\ Fa_{11} + Ga_{21} = m \end{cases}$$

si trova:

$$a_{11} = \frac{Gl - Fm}{EG - F^2}, \quad a_{21} = \frac{Em - Fl}{EG - F^2}.$$

La seconda colonna della matrice di S_P si ottiene in modo analogo a partire da $S_P(\varphi_v) = -N_v$:

$$a_{12} = \frac{Gm - Fn}{EG - F^2}, \quad a_{22} = \frac{En - Fm}{EG - F^2}.$$

ESEMPIO 2.19. Consideriamo il piano passante per un punto P_0 e parallelo ai vettori \mathbf{a} e \mathbf{b} che formano una base ortonormale (vedi l'Esempio 2.15). Sotto queste ipotesi, il campo vettoriale normale $N = \mathbf{a} \times \mathbf{b}$ è costante e, quindi, $N'(t) = \mathbf{0}$, in ogni punto e rispetto a qualsiasi curva. Di conseguenza l'operatore forma, in ogni punto del piano, è l'endomorfismo nullo.

ESEMPIO 2.20. Consideriamo la sfera $S^2(r)$ parametrizzata mediante le coordinate sferiche (Esempio 2.9).

Nell'Esempio 2.17 si era visto che $E = r^2 \cos^2 v$, $F = 0$, $G = r^2$. Con semplici calcoli si trova che

$$\begin{aligned} N(u, v) &= (\cos v \cos u, \cos v \sin u, \sin v) \\ \varphi_{uu}(u, v) &= (-r \cos v \cos u, -r \cos v \sin u, 0), \\ \varphi_{uv}(u, v) &= (r \sin v \sin u, -r \sin v \cos u, 0), \\ \varphi_{vv}(u, v) &= (-r \cos v \cos u, -r \cos v \sin u, -r \sin v), \end{aligned}$$

e quindi:

$$l = N \cdot \varphi_{uu} = -r \cos^2 v, \quad m = N \cdot \varphi_{uv} = 0, \quad n = N \cdot \varphi_{vv} = -r.$$

Inserendo tali dati nelle espressioni di pag. 41, si vede che gli elementi della matrice dell'operatore forma sono:

$$a_{11} = -\frac{1}{r}, \quad a_{12} = a_{21} = 0, \quad a_{22} = -\frac{1}{r}.$$

Ciò significa che $S_P = -1/r Id$, ossia l'operatore forma è, in ogni punto della sfera, proporzionale all'endomorfismo identico.

8. Curvature principali, Gaussiana e media

Abbiamo visto nella sezione precedente che l'operatore forma è un endomorfismo simmetrico (Proprietà 2.7). Di conseguenza, ha autovalori reali ed è diagonalizzabile. Indicati con $k_1(P)$ e $k_2(P)$ gli autovalori di S_P , è sempre possibile trovare una *base ortonormale* $\{\mathbf{e}_1, \mathbf{e}_2\}$ di $T_P M$ formata da autovettori. In tal caso,

$$S_P(\mathbf{e}_1) = k_1(P) \mathbf{e}_1, \quad S_P(\mathbf{e}_2) = k_2(P) \mathbf{e}_2,$$

e la matrice dell'endomorfismo S_P , rispetto alla base di autovettori, è diagonale:

$$\begin{pmatrix} k_1(P) & 0 \\ 0 & k_2(P) \end{pmatrix}.$$

DEFINIZIONE 2.10. 1) $k_1(P)$ e $k_2(P)$ sono le **curvature principali** di M in P , associate alle **direzioni principali di curvatura** \mathbf{e}_1 ed \mathbf{e}_2 .

2) La **curvatura Gaussiana** di M in P è:

$$K(P) \stackrel{\text{def.}}{=} k_1(P) k_2(P).$$

3) La **curvatura media** di M in P è:

$$H(P) \stackrel{\text{def.}}{=} k_1(P) + k_2(P).$$

OSSERVAZIONE 2.8. Le curvature Gaussiana e media sono, rispettivamente, il *determinante* e la *traccia* dell'endomorfismo S_P . Si noti che, cambiando parametrizzazione locale (per esempio, scambiando u e v), k_1 e k_2 possono cambiare segno. In tal caso, H si muta nel suo opposto mentre K resta invariata.

ESEMPIO 2.21. L'operatore forma del piano è l'endomorfismo nullo in ogni punto (Esempio 2.19). Quindi: $k_1 = k_2 = 0$ e $K = H = 0$.

Nel caso della sfera di raggio r , dall'Esempio 2.20 si deduce immediatamente che, in ogni punto,

$$k_1 = k_2 = -\frac{1}{r} \quad \text{e} \quad K = \frac{1}{r^2}, \quad H = -\frac{2}{r}.$$

Notare che in questi due esempi la curvatura Gaussiana è costante: nulla, nel primo caso, negativa nel secondo.

Considerata la matrice di S_P rispetto alla base $\{\varphi_u, \varphi_v\}$ (vedi pag. 40), il *polinomio caratteristico* dell'endomorfismo è

$$p(\lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}.$$

Tenuto conto che $K = a_{11}a_{22} - a_{12}a_{21}$ e $H = a_{11} + a_{22}$ (vedi l'Osservazione di 42), le curvature principali sono le radici dell'equazione caratteristica

$$\lambda^2 - H\lambda + K = 0.$$

Sostituendo agli a_{ij} le loro espressioni in funzione dei coefficienti delle due forme quadratiche fondamentali (vedi pag. 41), si ha la

PROPRIETÀ 2.8. *Le curvature Gaussiana e media di una superficie sono date da:*

$$K = \frac{ln - m^2}{EG - F^2}, \quad H = \frac{En - 2Fm + Gl}{EG - F^2}.$$

ESEMPIO 2.22. L'**elicoide** è la superficie generata dalle rette perpendicolari all'asse z e passanti per i punti dell'elica $\alpha(u) = (\cos u, \sin u, u)$, $u \in \mathbb{R}$. Detti $P'(0, 0, u)$ un generico punto dell'asse z e $P''(\cos u, \sin u, u)$ il corrispondente punto dell'elica, scritte le equazioni in forma parametrica della retta passante per P' e parallela al vettore $\overrightarrow{P'P''}$, si trova che una parametrizzazione locale dell'elicoide è:

$$\varphi(u, v) = (v \cos u, v \sin u, u), \quad (u, v) \in \mathbb{R} \times \mathbb{R}.$$

Calcoliamo le curvature di questa superficie. I coefficienti delle due forme quadratiche fondamentali sono:

$$E = 1 + v^2, \quad F = 0, \quad G = 1; \quad l = n = 0, \quad m = \frac{1}{\sqrt{1 + v^2}}.$$

Sostituendo tali espressioni nelle formule della Proprietà 2.8, si trovano le curvatures Gaussiana e media:

$$K = -\frac{1}{(1+v^2)^2}, \quad H = 0.$$

Le curvatures principali, in un punto generico dell'elicoide, sono le radici dell'equazione:

$$\lambda^2 - \frac{1}{(1+v^2)^2} = 0$$

e valgono:

$$k_1(P) = -\frac{1}{1+v^2}, \quad k_2(P) = \frac{1}{1+v^2}.$$

Fissato un punto $P_0 = \varphi(0,1)$ dell'elicoide, cerchiamo una base ortonormale di autovettori dell'operatore forma S_{P_0} . In primo luogo è necessario trovare la matrice di S_{P_0} rispetto alla base $\varphi_u(0,1) = (0,1,1)$ e $\varphi_v(0,1) = (1,0,0)$. Dopo aver calcolato il valore dei coefficienti delle due forme quadratiche in S_{P_0} , applicando le formule di pag. 41 si ha la matrice:

$$\begin{pmatrix} 0 & \frac{\sqrt{2}}{4} \\ \frac{\sqrt{2}}{2} & 0 \end{pmatrix}.$$

Occorre diagonalizzare tale matrice. Sappiamo già che i suoi autovalori (le curvatures principali in P_0) sono $k_1(P_0) = -1/2$ e $k_2(P_0) = 1/2$. Gli autovettori dell'endomorfismo S_{P_0} , relativi all'autovalore $k_1(P_0) = -1/2$, sono i vettori $\mathbf{v} = x_1\varphi_u(0,1) + x_2\varphi_v(0,1)$, dove (x_1, x_2) è una soluzione non nulla del sistema lineare omogeneo (ottenuto ponendo $\lambda = -1/2$ nella matrice usata per il polinomio caratteristico, vedi pag. 42):

$$\begin{cases} \frac{1}{2}x_1 + \frac{\sqrt{2}}{4}x_2 = 0 \\ \frac{\sqrt{2}}{2}x_1 + \frac{1}{2}x_2 = 0. \end{cases}$$

La soluzione generale del sistema è

$$\begin{cases} x_1 = t \\ x_2 = -\sqrt{2}t, \end{cases} \quad t \in \mathbb{R}.$$

Una base del relativo autospazio si ottiene scegliendo una soluzione particolare. Ad per esempio, per $t = 1$, si ha l'autovettore

$$\mathbf{v}_1 = \varphi_u(0,1) - \sqrt{2}\varphi_v(0,1) = (-\sqrt{2}, 1, 1).$$

Dato che siamo interessati ad una base ortonormale, occorre passare al versore:

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(-\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right).$$

In modo del tutto analogo, l'autospazio relativo all'autovalore $k_2(P_0) = 1/2$ si trova risolvendo il sistema:

$$\begin{cases} -\frac{1}{2}x_1 + \frac{\sqrt{2}}{4}x_2 = 0 \\ \frac{\sqrt{2}}{2}x_1 - \frac{1}{2}x_2 = 0 \end{cases}.$$

Una base è data dall'autovettore $\mathbf{v}_2 = \varphi_u(0, 1) + \sqrt{2}\varphi_v(0, 1) = (\sqrt{2}, 1, 1)$ o anche, normalizzando:

$$\mathbf{e}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2}\right).$$

Si noti che \mathbf{e}_1 e \mathbf{e}_2 sono perpendicolari (conseguenza di una proprietà valida per tutti gli endomorfismi simmetrici) e quindi individuano una base ortonormale di autovettori di S_{P_0} .

ESEMPIO 2.23. Consideriamo il cilindro circolare retto descritto dalla parametrizzazione locale dell'Esempio 2.16, in cui si era visto che $E = G = 1$ e $F = 0$. Dato che $N = \varphi_u \times \varphi_v = (-\cos v, -\sin v, 0)$ e

$$\varphi_{uu} = \varphi_{uv} = (0, 0, 0), \quad \varphi_{vv} = (-\cos v, -\sin v, 0),$$

si ha: $l = m = 0$ e $n = 1$.

Le curvatures Gaussiana e media sono: $K = 0$ e $H = 1$. Le curvatures principali, soluzioni dell'equazione $\lambda^2 - \lambda = 0$, valgono $k_1 = 0$, $k_2 = 1$. La matrice dell'operatore forma, in un generico punto del cilindro, rispetto alla base φ_u, φ_v è

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Poichè tale matrice è già in forma diagonale ed è anche ortogonale, si deduce che $\{\varphi_u, \varphi_v\}$ è una base ortonormale di autovettori in ogni punto del cilindro.

ESERCIZIO 2.24. Sia M la superficie che si ottiene ruotando la curva **tratrice** $\alpha(u) = (\sin u, 0, \cos u + \ln(\operatorname{tg} u/2))$, $u \in (0, \pi)$, attorno all'asse z . Trovare i punti singolari di M e verificare che la curvatura Gaussiana è costante e vale $K = -1$. Per questo motivo, M è detta **pseudo-sfera**.

9. Simboli di Christoffel e Teorema Egregium di Gauss

Dall'espressione della curvatura Gaussiana trovata nella Proprietà 2.8, sembra che K dipenda in modo essenziale dall'operatore forma e quindi dal modo in cui la superficie M è immersa nello spazio. Proveremo invece che K dipende solo dai coefficienti E , F e G della prima forma quadratica fondamentale e dalle loro derivate. In altri termini, la curvatura Gaussiana è intrinseca alla superficie.

Sia (D, φ) una parametrizzazione locale della superficie M . I campi vettoriali φ_u , φ_v e $N = \frac{\varphi_u \times \varphi_v}{\|\varphi_u \times \varphi_v\|}$ sono linearmente indipendenti in ogni punto di $\varphi(D)$. In generale, i campi derivati secondi non sono nè tangenti nè normali alla superficie e quindi saranno, in ogni punto, combinazione lineare dei vettori della base $(\varphi_u, \varphi_v, N)$. Poniamo

$$\begin{aligned} \varphi_{uu} &= \Gamma_{11}^1 \varphi_u + \Gamma_{11}^2 \varphi_v + L_1 N, & \varphi_{uv} &= \Gamma_{12}^1 \varphi_u + \Gamma_{12}^2 \varphi_v + L_2 N, \\ \varphi_{vu} &= \Gamma_{21}^1 \varphi_u + \Gamma_{21}^2 \varphi_v + L_3 N, & \varphi_{vv} &= \Gamma_{22}^1 \varphi_u + \Gamma_{22}^2 \varphi_v + L_4 N, \end{aligned}$$

dove gli scalari Γ_{ij}^k , detti **simboli di Christoffel**, e L_i sono da determinare. Si noti che:

- $\varphi_{u,v} = \varphi_{v,u}$ implica che $\Gamma_{12}^1 = \Gamma_{21}^1$, $\Gamma_{12}^2 = \Gamma_{21}^2$ (i simboli di Christoffel sono simmetrici rispetto agli indici inferiori) e $L_2 = L_3$.
- $l = \varphi_{uu} \cdot N = L_1$, $m = \varphi_{uv} \cdot N = L_2$, $n = \varphi_{vv} \cdot N = L_4$.

Inoltre, a partire dal fatto che

$$\varphi_{uu} \cdot \varphi_u = \Gamma_{11}^1 \varphi_u \cdot \varphi_u + \Gamma_{11}^2 \varphi_v \cdot \varphi_u = \Gamma_{11}^1 E + \Gamma_{11}^2 F,$$

e tenuto conto dell'identità

$$\varphi_{uu} \cdot \varphi_u = \frac{1}{2}(\varphi_u \cdot \varphi_u)_u = \frac{1}{2}E_u,$$

si ricava l'equazione

$$\Gamma_{11}^1 E + \Gamma_{11}^2 F = \frac{1}{2}E_u.$$

In modo analogo, dalla relazione

$$\varphi_{uu} \cdot \varphi_v = \Gamma_{11}^1 \varphi_u \cdot \varphi_v + \Gamma_{11}^2 \varphi_v \cdot \varphi_v = \Gamma_{11}^1 F + \Gamma_{11}^2 G,$$

assieme all'identità

$$\varphi_{uu} \cdot \varphi_v = (\varphi_u \cdot \varphi_v)_u - \varphi_u \cdot \varphi_{uv} = (\varphi_{uu} \cdot \varphi_v)_u - \frac{1}{2}(\varphi_u \cdot \varphi_u)_v = F_u - \frac{1}{2}E_v,$$

si ricava l'equazione

$$\Gamma_{11}^1 F + \Gamma_{11}^2 G = F_u - \frac{1}{2}E_v.$$

Risolvendo il sistema formato dalle due equazioni precedenti, si trova

$$\Gamma_{11}^1 = \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}, \quad \Gamma_{11}^2 = \frac{2EF_u - EE_v - FE_u}{2(EG - F^2)}.$$

Procedendo nello stesso modo, si possono esprimere i rimanenti simboli di Christoffel in funzione dei coefficienti della prima forma quadratica e delle loro derivate. Vale pertanto la

PROPRIETÀ 2.9. *I simboli di Christoffel di una superficie dipendono solo dalla metrica e non dalla sua immersione nello spazio.*

ESERCIZIO 2.25. Sia M la superficie generata dalla rotazione della curva piana $\alpha(u) = (x(u), 0, z(u))$ attorno all'asse z . Verificare che i simboli di Christoffel di M , rispetto alla parametrizzazione locale costruita nell'Esempio 2.14, sono:

$$\Gamma_{11}^1 = \frac{x'(u)x''(u) + z'(u)z''(u)}{x'(u)^2 + z'(u)^2}, \quad \Gamma_{11}^2 = 0, \quad \Gamma_{12}^1 = 0,$$

$$\Gamma_{12}^2 = \frac{x'(u)}{x(u)}, \quad \Gamma_{22}^1 = -\frac{x(u)x'(u)}{x'(u)^2 + z'(u)^2}, \quad \Gamma_{22}^2 = 0.$$

Vogliamo ora esprimere la curvatura Gaussiana in funzione dei simboli di Christoffel e delle loro derivate. A tal fine, notiamo che dall'identità $(\varphi_{uu})_v = (\varphi_{uv})_u$ si ottiene l'uguaglianza

$$\begin{aligned} (\Gamma_{11}^1)_v \varphi_u + \Gamma_{11}^1 \varphi_{uv} + (\Gamma_{11}^2)_v \varphi_v + \Gamma_{11}^2 \varphi_{vv} + l_v N + l N_v = \\ (\Gamma_{12}^1)_u \varphi_u + \Gamma_{12}^1 \varphi_{uu} + (\Gamma_{12}^2)_u \varphi_v + \Gamma_{12}^2 \varphi_{vu} + m_u N + m N_u. \end{aligned}$$

Sostituendo nell'equazione precedente le espressioni di φ_{uu} , φ_{uv} , φ_{vv} (vedi l'inizio di questa sezione) e quelle di N_u , N_v (vedi le formule di pag. 41), uguagliando le componenti di φ_v a primo e secondo membro si ottiene il

TEOREMA 2.10 (Teorema Egregium di Gauss). *La curvatura Gaussiana di una superficie è data da*

$$K = \frac{1}{E} [(\Gamma_{11}^2)_u + \Gamma_{11}^1 \Gamma_{12}^2 - (\Gamma_{12}^2)_v - \Gamma_{12}^1 \Gamma_{11}^2 + \Gamma_{11}^2 \Gamma_{22}^2 - \Gamma_{12}^1 \Gamma_{12}^2]$$

e quindi dipende solo dalla metrica.

10. Curvature normali

In questa sezione verrà presa in esame la relazione tra la curvatura Gaussiana e la curvatura delle curve tracciate su una superficie. Ciò consentirà di dare un significato geometrico alla seconda forma quadratica ed alle curvature principali.

Fissati un punto P della superficie M ed un versore $\mathbf{u} \in T_P M$, sia $\beta(s)$ una curva tracciata su M , parametrizzata mediante l'ascissa curvilinea, passante per P e tangente a \mathbf{u} . Quindi: $\beta(s_0) = P$ e $\beta'(s_0) = \mathbf{u}$, per un certo s_0 appartenente al dominio di β . Possiamo considerare due campi vettoriali definiti lungo β , entrambi di norma unitaria: $N(s)$, un campo ortogonale alla superficie e $N_\beta(s)$ il campo normale alla curva (vedi a pag. 8). Se $\theta(s)$ è l'angolo fra i versori $N(s)$ e $N_\beta(s)$, si ha che $\cos \theta(s) = N(s) \cdot N_\beta(s)$.

Dato che $\beta(s)$ è una curva su M , $\beta'(s)$ è un campo tangente alla superficie e quindi: $\beta'(s) \cdot N(s) = 0$. Derivando questa espressione rispetto ad s , si trova: $\beta''(s) \cdot N(s) + \beta'(s) \cdot N'(s) = 0$, da cui, tenuto conto della prima formula di Frenet (Proprietà 1.6), si ricava:

$$-\beta'(s) \cdot N'(s) = \beta''(s) \cdot N(s) = T'(s) \cdot N(s) = k(s) N_\beta(s) \cdot N(s) = k(s) \cos \theta(s),$$

dove $k(s)$ è la curvatura di β .

Calcoliamo ora la seconda forma quadratica in corrispondenza al versore \mathbf{u} . Tenuto conto delle condizioni iniziali e della definizione di S_P , abbiamo:

$$S_P(\mathbf{u}) \cdot \mathbf{u} = S_P(\beta'(s_0)) \cdot \beta'(s_0) = -N'(s_0) \cdot \beta'(s_0) = k(s_0) \cos \theta(s_0).$$

Si noti che, per definizione di operatore forma, il primo membro dipende solo da P e da \mathbf{u} e non dalla curva β . Di conseguenza, la seguente definizione è geometricamente motivata.

DEFINIZIONE 2.11. *La curvatura normale di una superficie M nella direzione di un versore $\mathbf{u} \in T_P M$ è*

$$k_n(\mathbf{u}) \stackrel{\text{def}}{=} k(s_0) \cos \theta(s_0).$$

Dato che $k_n(\mathbf{u})$ non dipende dalla curva che interviene nel calcolo, possiamo ricorrere a delle curve speciali. Indicato con π il piano ortogonale alla superficie passante per P e parallelo ai vettori $N(s_0)$ e \mathbf{u} , la **sezione normale** determinata da P ed \mathbf{u} è la curva $\gamma(s)$, parametrizzata con l'ascissa curvilinea, che si ottiene intersecando M con π . In tal caso, il versore normale a γ è contenuto nel piano π ed è quindi parallelo a $N(s_0)$. Di conseguenza, $N_\gamma(s_0) = \pm N(s_0)$ e $\cos \theta(s_0) = \pm 1$.

Detta k_N la curvatura della sezione normale, si ha che $k_n(\mathbf{u}) = \pm k_N(s_0)$, ossia: *la curvatura normale coincide (a meno del segno) con la curvatura della sezione normale*. Ricordato che $k_n(\mathbf{u}) = S_P(\mathbf{u}) \cdot \mathbf{u}$, l'osservazione precedente è un'interpretazione geometrica della seconda forma quadratica fondamentale. Un'altra importante conseguenza è il classico

TEOREMA 2.11 (di Meusnier). *Tutte le curve tracciate su una superficie M che hanno la stessa tangente in un punto P determinano la stessa curvatura normale in P .*

OSSERVAZIONE 2.9. Il segno della curvatura normale dipende dal vettore \mathbf{u} ; tuttavia quando è costante fornisce un'informazione di carattere geometrico. Se $k_n(\mathbf{u}) > 0$, per ogni $\mathbf{u} \in T_P M$, si ha che $N(s_0) = N_\gamma(s_0)$ per ogni sezione normale γ passante per P . Ciò significa che la superficie si “incurva” nella stessa direzione di $N(s_0)$. Ovviamente, se $k_n(\mathbf{u}) < 0$, per ogni $\mathbf{u} \in T_P M$, la superficie si “incurva” nella direzione opposta a quella $N(s_0)$.

ESEMPIO 2.26. In alcuni casi particolarmente semplici, è possibile trovare direttamente le curvature normali.

Se M è il cilindro con la parametrizzazione dell'Esempio 2.23, la sezione normale, in un punto qualsiasi, determinata dal versore \mathbf{k} è un meridiano, cioè una retta, e quindi $k_n(\mathbf{k}) = 0$. Invece la curvatura normale nella direzione di un versore \mathbf{u} tangente ad un parallelo vale 1, dato che la sezione normale è una circonferenza di raggio 1.

Sia M la superficie a sella $z = x^2 - y^2$. Calcoliamo le curvature normali nell'origine e nella direzione dei versori $\mathbf{u}_1 = \mathbf{i}$, $\mathbf{u}_2 = \mathbf{j}$ e $\mathbf{u}_3 = \mathbf{i} + \mathbf{j}$. Le sezioni normali (parabole, nei primi due casi, una retta nel terzo caso) hanno equazione:

$$\beta_1(t) = (t, 0, -t^2), \quad \beta_2(t) = (0, t, t^2), \quad \beta_3(t) = (t, t, 0).$$

Le curvature normali coincidono con le curvature di queste curve nell'origine: $k_n(\mathbf{u}_1) = -2$, $k_n(\mathbf{u}_2) = 2$, $k_n(\mathbf{u}_3) = 0$. Si noti che le curvature hanno segni diversi: ciò conferma che, nell'origine, la sella si incurva in modo diverso al variare della direzione.

In generale, non è facile calcolare la curvatura normale passando attraverso le sezioni normali. È molto più semplice applicare la formula di Eulero.

Sia $\{\mathbf{e}_1, \mathbf{e}_2\}$ una base ortonormale di autovettori di $S_P M$. Ogni versore $\mathbf{u} \in T_P M$ si può scrivere nella forma: $\mathbf{u} = \cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2$, dove θ è l'angolo fra \mathbf{u} e \mathbf{e}_1 . Vale la seguente **formula di Eulero**.

PROPRIETÀ 2.12.

$$k_n(\mathbf{u}) = k_1(P) \cos^2 \theta + k_2(P) \sin^2 \theta.$$

DIMOSTRAZIONE. Basta calcolare la curvatura normale, ricordando il suo significato geometrico (vedi pag. 46):

$$\begin{aligned} k_n(\mathbf{u}) &= S_P(\mathbf{u}) \cdot \mathbf{u} = S_P(\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2) \cdot (\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2) \\ &= [\cos \theta S_P(\mathbf{e}_1) + \sin \theta S_P(\mathbf{e}_2)] \cdot (\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2) \\ &= [\cos \theta k_1(P) \mathbf{e}_1 + \sin \theta k_2(P) \mathbf{e}_2] \cdot (\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2) \\ &= k_1(P) \cos^2 \theta + k_2(P) \sin^2 \theta. \end{aligned}$$

□

La derivata della formula di Eulero è $k'_n(\theta) = (k_2(P) - k_1(P)) \sin(2\theta)$. Ci sono due possibilità.

- $k_1(P) = k_2(P) = c$. Poichè $k'_n = 0$, $k_n(\mathbf{u})$ è costante, tutte le direzioni sono principali e l'operatore forma $S_P = cId$ è il prodotto di uno scalare per l'endomorfismo identico. In tal caso, P si dice **punto ombelicale**.
- $k_1(P) \neq k_2(P)$. Allora k'_n si annulla solo per $\theta = 0$ e per $\theta = \pi/2$, che corrispondono ad \mathbf{e}_1 ed \mathbf{e}_2 , rispettivamente. Ciò significa che le direzioni principali di curvatura sono quelle in cui la superficie ha la massima o la minima curvatura e che *le curvatures principali sono il valore massimo e minimo delle curvatures normali*.

ESEMPIO 2.27. Riprendiamo l'Esempio 2.22 di pag. 42 e calcoliamo la curvatura normale dell'elicoide nel punto $P_0 = \varphi(0, 1)$ e nella direzione del vettore $\mathbf{v} = \sqrt{2}\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$. Indicato con \mathbf{u} il versore di \mathbf{v} , si ha:

$$\cos \theta = \mathbf{u} \cdot \mathbf{e}_1 = \frac{\sqrt{10}}{10} (\sqrt{2}, 2, 2) \cdot \left(-\frac{\sqrt{2}}{2}, \frac{1}{2}, \frac{1}{2} \right) = \frac{\sqrt{10}}{10}.$$

Applicando la formula di Eulero, si ottiene:

$$k_n(\mathbf{u}) = -\frac{1}{2} \cdot \frac{1}{10} + \frac{1}{2} \cdot \frac{9}{10} = \frac{2}{5},$$

da cui si vede che, nella direzione del vettore \mathbf{v} , la superficie si incurva nella stessa direzione del versore normale $N(0, 1) = \sqrt{2}/2(0, 1, -1)$.

Il segno della curvatura Gaussiana consente di avere un'idea della forma di una superficie, nell'intorno di un suo punto.

- Se $K(P) > 0$, le curvatures principali hanno lo stesso segno. Dalla formula di Eulero si vede che k_n ha lo stesso segno in ogni direzione: tutte le sezioni normali si incurvano allo stesso modo ed il piano tangente in P non attraversa la superficie. Intuitivamente, in un intorno di P la superficie ha la stessa forma di un ellissoide. Per questo motivo, un punto P in cui $K(P) > 0$ si dice **ellittico**.
- Se $K(P) < 0$, le curvatures principali hanno segno opposto e quindi k_n cambia di segno, a seconda delle direzioni, e si annulla in due direzioni particolari dette **direzioni asintotiche**. Il piano tangente attraversa la superficie che, in un intorno di P , ha la stessa forma della superficie a sella vicino all'origine. In questo caso, P si dice **punto iperbolico**.
- Se invece $K(P) = 0$, ci sono due possibilità.
 - La prima è che solo una curvatura principale sia nulla: in tal caso, la curvatura normale è nulla in una direzione (detta **direzione asintotica**) ed ha segno costante in tutte le altre. Localmente la superficie assomiglia ad un cilindro e P si dice **punto parabolico**.
 - La seconda possibilità è che tutte e due le curvatures principali siano nulle: l'operatore forma è nullo e la superficie assomiglia localmente ad un piano. Per questo motivo, P si dice **punto planare**.

ESEMPIO 2.28. Classifichiamo i punti del toro di rivoluzione. Usando la parametrizzazione di pag. 30, si trova che la curvatura Gaussiana del toro è

$$K = \frac{\cos u}{r(a + r \cos u)}.$$

Quindi: se $u = \pi/2$ oppure $u = 3\pi/2$ i punti del toro sono parabolici; se $0 < u < \pi/2$ oppure $3\pi/2 < u < 2\pi$ (la parte “esterna”) i punti sono ellittici; se $\pi/2 < u < 3\pi/2$ (la parte “interna”) i punti sono iperbolici.

Un’ulteriore motivazione della terminologia precedente si ha considerando il seguente insieme di vettori tangenti:

$$\mathcal{D}_P = \{\mathbf{v} \in T_P M / S_P(\mathbf{v}) \cdot \mathbf{v} = \pm 1\}.$$

Posto $\mathbf{v} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2$, dove $\{\mathbf{e}_1, \mathbf{e}_2\}$ è una base ortonormale di autovettori dell’operatore forma, si vede subito che le equazioni di \mathcal{D}_P sono:

$$k_1(P)x_1^2 + k_2(P)x_2^2 = \pm 1.$$

Si tratta di una coppia di coniche (eventualmente a punti immaginari) che prende il nome di **indicatrice di Dupin**. Esaminiamo le varie possibilità in relazione al segno delle curvature principali.

- Se P è un punto ellittico, k_1 e k_2 hanno lo stesso segno e \mathcal{D}_P è un’ellisse reale (più un’ellisse a punti immaginari). L’intersezione della superficie con un piano parallelo (e “vicino”) al piano tangente è proprio un’ellisse.
- Se P è un punto iperbolico, k_1 e k_2 hanno segno opposto e \mathcal{D}_P è una coppia di iperboli aventi gli stessi asintoti, le cui direzioni coincidono con le direzioni asintotiche. In questo caso, l’intersezione della superficie con certi piani paralleli al piano tangente è proprio una coppia di iperboli.
- Se P è un punto parabolico, \mathcal{D}_P coincide con una coppia di rette parallele, ossia una conica degenera di tipo parabolico. Infine, se P è planare, \mathcal{D}_P non ha punti reali.

Terminiamo questa sezione caratterizzando i piani e le sfere come le uniche superfici i cui punti sono tutti ombelicali.

TEOREMA 2.13. *Sia M una superficie regolare connessa. Allora:*

- (1) M è parte di un piano se e solo se $k_1(P) = k_2(P) = 0$, per ogni $P \in M$.
- (2) M è parte di una sfera se e solo se $k_1(P) = k_2(P) \neq 0$, per ogni $P \in M$.

DIMOSTRAZIONE. Sappiamo già che tutti i punti del piano e della sfera sono ombelicali (vedi l’Esempio 2.21). Resta da provare il viceversa: se una superficie è formata solo da punti ombelicali allora è parte di un piano o di una sfera.

La dimostrazione si divide in due parti. La prima, di carattere *locale*, consiste nel far vedere che, fissata una parametrizzazione (D, φ) , le curvature principali, non solo coincidono punto per punto, ma hanno lo stesso valore in ogni punto di $\varphi(D)$. Successivamente si prova che, se il valore (costante) delle curvature principali è zero (oppure, diverso da zero), $\varphi(D)$ è parte di un piano (o di una sfera).

Dato che una parametrizzazione locale può non ricoprire interamente la superficie, non è escluso il caso in cui M sia l’unione di un pezzo di piano e di un pezzo di sfera. Per eliminare questa possibilità è necessario far intervenire l’ipotesi topologica: nella seconda parte della dimostrazione, di

carattere *globale*, si prova che, se la superficie è connessa, allora o è un piano o è una sfera, ma non entrambi.

Data una parametrizzazione locale (D, φ) tale che $P = \varphi(u, v)$, iniziamo col far vedere che la funzione $\lambda(P) = k_1(P) = k_2(P)$ è costante su $\varphi(D)$. L'ipotesi che ogni punto sia ombelicale implica che l'operatore forma sia il prodotto di uno scalare per l'identità ed ogni vettore (non nullo) sia un autovettore. Di conseguenza: $S_P(\varphi_u) = \lambda(P) \varphi_u$ e $S_P(\varphi_v) = \lambda(P) \varphi_v$. Ricordato che $S_P(\varphi_u) = -N_u$ e $S_P(\varphi_v) = -N_v$ (vedi pag. 39), si trova:

$$N_u = -\lambda(P) \varphi_u, \quad N_v = -\lambda(P) \varphi_v. \quad (*)$$

Derivando la prima uguaglianza rispetto a v e la seconda rispetto ad u (notare che $\lambda(P)$ è una funzione di u e v), otteniamo

$$N_{uv} = -\lambda_v \varphi_u - \lambda \varphi_{uv}, \quad N_{vu} = -\lambda_u \varphi_v - \lambda \varphi_{vu},$$

da cui si ricava $\lambda_v \varphi_u - \lambda_u \varphi_v = \mathbf{0}$. Dato che la superficie è regolare, φ_u e φ_v sono linearmente indipendenti e, quindi, $\lambda_u = \lambda_v = 0$, ossia: λ è costante su $\varphi(D)$. Ci sono due possibilità.

- $\lambda(P) = 0$, per ogni $P \in \varphi(D)$. Da (*), si ricava che $N_u = N_v = \mathbf{0}$ e, quindi, $N(u, v) = N_0$ è un campo costante. Per provare che $P = \varphi(u, v)$ sta in un piano, fissiamo un punto $P_0 = \varphi(u_0, v_0)$ e verifichiamo che i vettori $\varphi(u, v) - P_0$ e N_0 sono perpendicolari. A tal fine, basta notare che la funzione $(\varphi(u, v) - P_0) \cdot N_0$ è costante (e, quindi, nulla) dato che

$$\begin{aligned} [(\varphi(u, v) - P_0) \cdot N_0]_u &= \varphi_u(u, v) \cdot N_0 = 0, \\ [(\varphi(u, v) - P_0) \cdot N_0]_v &= \varphi_v(u, v) \cdot N_0 = 0. \end{aligned}$$

- $\lambda(P) = \lambda_0 \neq 0$, per ogni $P \in \varphi(D)$. Verifichiamo che P sta su una sfera di raggio $r = 1/|\lambda_0|$. Consideriamo la funzione (a valori in \mathbb{R}^3) $\gamma(u, v) = \varphi(u, v) + 1/\lambda_0 N(u, v)$. Tenuto conto delle condizioni (*), si vede subito che le sue derivate parziali sono nulle. Quindi, $\gamma(u, v) = C$ è un punto fisso tale che:

$$\|\varphi(u, v) - C\| = \left\| \frac{1}{\lambda_0} N(u, v) \right\| = r.$$

In conclusione, $\varphi(D)$ è parte di una sfera di centro C e raggio r .

Passiamo alla seconda parte della dimostrazione. Fissiamo un punto P_0 di M e consideriamo un *generico* punto P della superficie. Faremo vedere che se P_0 sta su un piano (oppure una sfera), anche P sta sullo stesso piano (o sulla stessa sfera).

Dato che M è connessa, è possibile costruire una curva α , definita sull'intervallo $[0, 1]$, tale che: $\alpha([0, 1]) \subseteq M$, $\alpha(0) = P_0$ e $\alpha(1) = P$. Fissato un numero reale $t \in [0, 1]$, esiste una parametrizzazione locale (D_t, φ_t) per cui $\alpha(t)$ sta nell'aperto $U_t = \varphi_t(D_t)$. Poichè α è continua, $\alpha^{-1}(U_t)$ è un aperto di $[0, 1]$. Al variare di t , questi insiemi formano un ricoprimento aperto dell'intervallo $[0, 1]$. Essendo tale intervallo compatto (vedi la Definizione 3.11 e la Proprietà 3.24 di [5]), è possibile ricoprire $[0, 1]$ con un numero finito di aperti $\alpha^{-1}(U_{t_1}), \dots, \alpha^{-1}(U_{t_k})$ che verificano le seguenti condizioni:

$$0 \in \alpha^{-1}(U_{t_1}), \quad 1 \in \alpha^{-1}(U_{t_k}), \quad \alpha^{-1}(U_{t_i}) \cap \alpha^{-1}(U_{t_{i+1}}) \neq \emptyset, \quad i = 1, \dots, k-1.$$

Di conseguenza, gli aperti U_{t_1}, \dots, U_{t_k} formano un ricoprimento finito di $\alpha([0, 1])$ per cui:

$$P_0 \in U_{t_1}, \quad P \in U_{t_k}, \quad U_{t_i} \cap U_{t_{i+1}} \neq \emptyset, \quad i = 1, \dots, k-1.$$

Nella prima parte della dimostrazione si è visto che ogni aperto U_t è un pezzo di piano o di sfera. Dato che due aperti successivi U_{t_i} e $U_{t_{i+1}}$ hanno intersezione non vuota, non è possibile che siano di tipo diverso. Pertanto l'ultimo aperto della catena U_{t_k} deve essere dello stesso tipo del primo U_{t_1} . Ripetendo lo stesso ragionamento per ogni punto della superficie, si ha la tesi. \square

11. Esercizi svolti sulle superfici

ESERCIZIO 2.29. Sia M la superficie generata dalla rotazione della curva $\alpha(u) = (u, 0, \ln u)$ attorno all'asse z .

- Costruire una parametrizzazione locale della superficie M e trovare gli eventuali punti singolari.
- Trovare la curvatura Gaussiana e media in un punto qualsiasi di M .
- Calcolare la curvatura normale nel punto $P_0(1, 0, 0)$ e nella direzione del vettore tangente alla curva $\gamma(t) = \varphi(t, 1 - t^2)$.

SOLUZIONE. a) Seguendo l'Esempio 2.14, una parametrizzazione locale della superficie è data da:

$$\varphi(u, v) = (u \cos v, u \sin v, \ln u), \quad (u, v) \in (0, +\infty) \times (0, 2\pi).$$

Per trovare i punti critici, calcoliamo

$$\varphi_u(u, v) = (\cos v, \sin v, \frac{1}{u}), \quad \varphi_v(u, v) = (-u \sin v, u \cos v, 0)$$

e cerchiamo per quali valori dei parametri si annulla il prodotto vettoriale $\varphi_u(u, v) \times \varphi_v(u, v) = -\cos v \mathbf{i} - \sin v \mathbf{j} + u \mathbf{k}$. A tal fine, basta vedere quando si annulla il *modulo* del prodotto vettoriale. Dato che $\|\varphi_u(u, v) \times \varphi_v(u, v)\| = \sqrt{1 + u^2}$, non ci sono punti singolari.

b) Si applicano le formule di pag. 42. I coefficienti della prima forma quadratica sono:

$$E = \varphi_u \times \varphi_u = \frac{1 + u^2}{u^2}, \quad F = \varphi_u \times \varphi_v = 0, \quad G = \varphi_v \times \varphi_v = u^2.$$

Notare che $EG - F^2 = 1 + u^2 = \|\varphi_u \times \varphi_v\|^2$, in accordo con il risultato precedente. Per calcolare i coefficienti della seconda forma fondamentale, occorre trovare le derivate seconde ed un campo unitario normale alla superficie:

$$\begin{aligned} \varphi_{uu} &= (0, 0, -\frac{1}{u^2}), & \varphi_{uv} &= (-\sin v, \cos v, 0), & \varphi_{vv} &= (-u \cos v, -u \sin v, 0), \\ N &= \frac{\varphi_u \times \varphi_v}{\|\varphi_u \times \varphi_v\|} = \frac{1}{\sqrt{1 + u^2}} (-\cos v, -\sin v, u). \end{aligned}$$

Dato che

$$l = N \cdot \varphi_{uu} = -\frac{1}{u\sqrt{1 + u^2}}, \quad m = N \cdot \varphi_{uv} = 0, \quad n = N \cdot \varphi_{vv} = \frac{u}{\sqrt{1 + u^2}},$$

le curvatures Gaussiana e media valgono, rispettivamente:

$$K = -\frac{1}{(1+u^2)^2}, \quad H = \frac{1}{u\sqrt{(1+u^2)^3}}.$$

c) Per poter calcolare la curvatura normale mediante la formula di Eulero (Proprietà 2.12), è necessario trovare la matrice dell'operatore forma nel punto P_0 e costruire una base ortonormale di $T_{P_0}M$ formata da autovettori. Per semplificare i calcoli, conviene esprimere i coefficienti delle due forme quadratiche in $P_0 = \varphi(1, 0)$:

$$E = 2, \quad F = 0, \quad G = 1; \quad l = -\frac{\sqrt{2}}{2}, \quad m = 0, \quad n = \frac{\sqrt{2}}{2}.$$

Dalle formule di pag. 41 si trova che la matrice di S_{P_0} , rispetto alla base $\{\varphi_u(1, 0), \varphi_v(1, 0)\}$ è:

$$\begin{pmatrix} -\frac{\sqrt{2}}{4} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Poichè tale matrice è già diagonale, si ha immediatamente che le curvatures principali in P_0 (gli autovalori) sono $k_1(P_0) = -\sqrt{2}/4$ e $k_2(P_0) = \sqrt{2}/2$, mentre $\varphi_u(1, 0) = (1, 0, 1)$ e $\varphi_v(1, 0) = (0, 1, 0)$ è una base ortogonale di autovettori. Per avere una base ortonormale, non resta che normalizzare la base precedente: $\mathbf{e}_1 = \frac{\sqrt{2}}{2}(\mathbf{i} + \mathbf{k})$, $\mathbf{e}_2 = \mathbf{j}$. Dalla dimostrazione della Proprietà 2.3 si vede che il vettore tangente alla curva γ in $P_0 = \gamma(1)$ è:

$$\gamma'(1) = \frac{du}{dt}\Big|_{t=1} \varphi_u(1, 0) + \frac{dv}{dt}\Big|_{t=1} \varphi_v(1, 0) = \varphi_u(1, 0) - 2\varphi_v(1, 0) = (1, -2, 1).$$

Il corrispondente versore è $\mathbf{u} = \sqrt{6}/6(1, -2, 1)$. La curvatura normale nella direzione del versore \mathbf{u} si ottiene dalla formula di Eulero, tenuto conto che $\cos \theta = \mathbf{u} \cdot \mathbf{e}_1 = 1/\sqrt{3}$:

$$k_n(\mathbf{u}) = -\frac{\sqrt{2}}{4} \cdot \frac{1}{3} + \frac{\sqrt{2}}{2} \cdot \frac{2}{3} = \frac{\sqrt{2}}{4}.$$

ESERCIZIO 2.30. Sia M la superficie descritta dalla parametrizzazione locale $\varphi(u, v) = (\cos u - v \sin u, \sin u + v \cos u, v)$, $(u, v) \in \mathbb{R}^2$.

- Trovare la curvatura Gaussiana e media in un punto qualsiasi di M . Di che superficie si tratta?
- Determinare le curvatures principali e le direzioni principali di curvatura nel punto $P_0 = \varphi(0, 0)$.
- Calcolare la curvatura della sezione normale determinata dal punto P_0 e dal versore $\mathbf{u} = 1/2(-\mathbf{j} + \sqrt{3}\mathbf{k})$.

SOLUZIONE. a) Dato che $\varphi_u(u, v) = (-\sin u - v \cos u, \cos u - v \sin u, 0)$, e $\varphi_v(u, v) = (-\sin u, \cos u, 1)$, i coefficienti della prima forma quadratica sono: $E = 1 + v^2$, $F = 1$, $G = 2$. Inoltre, essendo

$$\begin{aligned} \varphi_{uu} &= (-\cos u + v \sin u, -\sin u - v \cos u, 0), & \varphi_{uv} &= (-\cos u, -\sin u, 0), \\ \varphi_{vv} &= (0, 0, 0), & N &= \frac{1}{\sqrt{1+2v^2}}(\cos u - v \sin u, \sin u + v \cos u, -v), \end{aligned}$$

i coefficienti della seconda forma quadratica sono:

$$l = -\frac{1+v^2}{\sqrt{1+2v^2}}, \quad m = -\frac{1}{\sqrt{1+2v^2}}, \quad n = 0,$$

e le curvatures Gaussiana e media valgono:

$$K = -\frac{1}{(1+2v^2)^2}, \quad H = -\frac{2v^2}{\sqrt{(1+2v^2)^3}}.$$

Le funzioni che individuano la parametrizzazione locale sono lineari rispetto alla variabile v : ciò significa che M è una superficie rigata (Esempio 2.13). Più precisamente, M è la rigata avente come direttrice la circonferenza $\alpha(u) = (\cos u, \sin u, 0)$. Le rette generatrici, essendo parallele ai vettori $\mathbf{w}(u) = (-\sin u, \cos u, 1) = \alpha'(u) + \mathbf{k}$, sono tutte inclinate di 45° rispetto al piano che contiene la circonferenza. Si può allora concludere che la superficie è un iperboloide ad una falda.

b) Procedendo come nell'esercizio precedente, si trova che la matrice dell'operatore forma, rispetto alla base formata dai vettori $\varphi_u(0,0) = (0, 1, 0)$ e $\varphi_v(0,0) = (0, 1, 1)$, è:

$$\begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Questa volta la matrice non è più diagonale ed occorre trovare gli autovalori e gli autovettori. Dal polinomio caratteristico della matrice $p(\lambda) = \lambda^2 - 1$, si deduce che le curvatures principali in P_0 sono: $k_1(P_0) = -1$ e $k_2(P_0) = 1$. Per trovare le direzioni principali di curvatura si proceda come nell'Esercizio 2.22 di pag. 42. Risulta che una base ortonormale di autovettori è formata dai vettori $\mathbf{e}_1 = \mathbf{j}$ ed $\mathbf{e}_2 = -\mathbf{k}$.

c) Applicando la formula di Eulero, si trova che: $k_n(\mathbf{u}) = 1/2$.

ESERCIZIO 2.31. Sia M la superficie descritta dalla parametrizzazione locale $\varphi(u, v) = (u, e^u \cos v, e^u \sin v)$, $(u, v) \in (-\infty, +\infty) \times (0, 2\pi)$.

- Trovare le curvatures Gaussiana e media in un punto qualsiasi di M . Di che superficie si tratta?
- Calcolare la curvatura della sezione normale determinata dal punto $P_0 = \varphi(0,0)$ e dal versore $\mathbf{u} = \sqrt{3}/3(\mathbf{i} + \mathbf{j} - \mathbf{k})$.
- Calcolare l'area della parte di superficie descritta dai punti $\varphi(u, v)$ con $(u, v) \in D_a = (a, 0) \times (0, 2\pi)$, dove $a < 0$. Stabilire se l'area della parte di superficie descritta dai punti $\varphi(u, v)$ con $u < 0$ è limitata.

SOLUZIONE. a) Dato che

$$\begin{aligned} \varphi_u(u, v) &= (1, e^u \cos v, e^u \sin v), & \varphi_v(u, v) &= (0, -e^u \sin v, e^u \cos v), \\ \varphi_{uu}(u, v) &= (0, e^u \cos v, e^u \sin v), & \varphi_{uv}(u, v) &= (0, -e^u \sin v, e^u \cos v), \\ \varphi_{vv}(u, v) &= (0, -e^u \cos v, -e^u \sin v), & N &= \frac{1}{\sqrt{1+e^{2u}}}(e^u, -\cos v, -\sin v), \end{aligned}$$

i coefficienti delle due forme quadratiche fondamentali sono:

$$E = 1 + e^{2u}, \quad F = 0, \quad G = e^{2u}; \quad l = -\frac{e^u}{\sqrt{1+e^{2u}}}, \quad m = 0, \quad n = \frac{e^u}{\sqrt{1+e^{2u}}}.$$

Le curvatures Gaussiana e media valgono:

$$K = -\frac{1}{(1+e^{2u})^2}, \quad H = \frac{1}{e^u \sqrt{(1+e^{2u})^3}}.$$

M è la superficie generata dalla rotazione della curva $\alpha(u) = (u, e^u, 0)$, contenuta nel piano Oxy , attorno all'asse x .

b) La matrice dell'operatore forma in P_0 è:

$$\begin{pmatrix} -\frac{\sqrt{2}}{4} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Tale matrice è diagonale e, quindi, una base ortonormale di autovettori è:

$$\mathbf{e}_1 = \frac{\varphi_u(0,0)}{\|\varphi_u(0,0)\|} = \frac{\sqrt{2}}{2} (1, 1, 0), \quad \mathbf{e}_2 = \frac{\varphi_v(0,0)}{\|\varphi_v(0,0)\|} = (0, 0, 1).$$

Dalla formula di Eulero si ricava che: $k_n(\mathbf{u}) = 3\sqrt{2}/8$.

c) Applicando la formula di pag.38, si vede che l'area della regione $\varphi(D_a)$ è:

$$\begin{aligned} \mathcal{A}_a &= \int_{D_a} \sqrt{EG - F^2} \, dudv = \int_0^{2\pi} \left(\int_a^0 e^u \sqrt{1+e^{2u}} \, du \right) dv \\ &= 2\pi \int_a^0 e^u \sqrt{1+e^{2u}} \, du = 2\pi \int_{e^a}^1 \sqrt{1+w^2} \, dw, \quad \text{dove } w = e^u. \end{aligned}$$

Dato che $\int \sqrt{1+w^2} \, dw = 1/2 [\ln(w + \sqrt{1+w^2}) + w\sqrt{1+w^2}]$ (vedi il corso di Analisi Matematica II), l'area vale:

$$\mathcal{A}_a = \pi [\ln(1 + \sqrt{2}) + \sqrt{2} - \ln(e^a + \sqrt{1+e^{2a}}) - e^a \sqrt{1+e^{2a}}].$$

Poichè il

$$\lim_{a \rightarrow -\infty} \mathcal{A}_a = \pi [\ln(1 + \sqrt{2}) + \sqrt{2}]$$

è finito, l'area della parte di superficie formata dai punti $\varphi(u, v)$, con $u < 0$, è finita.

ESERCIZIO 2.32. Sia M la superficie descritta dalla parametrizzazione locale

$$\varphi(u, v) = \left(u, \frac{v^2}{2}, \frac{u^2}{2} - \frac{v^3}{3} \right), \quad (u, v) \in D.$$

- a) Dopo aver determinato D in modo tale che la superficie sia regolare, calcolare le curvatures Gaussiana e media in un generico punto di M . Classificare i punti di M .
- b) Trovare la curvatura della sezione normale determinata dal punto $P_0 = \varphi(1, 1)$ e dal versore $\mathbf{u} = \sqrt{2}/2 (\mathbf{i} + \mathbf{j})$.

SOLUZIONE. a) Dato che: $\varphi_u(u, v) = (1, 0, u)$, $\varphi_v(u, v) = (0, v, -v^2)$, $\varphi_u(u, v) \times \varphi_v(u, v) = (-uv, v^2, v)$, gli eventuali punti singolari si trovano annullando $\|\varphi_u(u, v) \times \varphi_v(u, v)\| = |v| \sqrt{u^2 + v^2 + 1}$. Pertanto, la superficie è regolare se $D = \{(u, v) \in \mathbb{R}^2 / v \neq 0\}$.

I coefficienti delle due forme quadratiche sono:

$$E = 1 + u^2, \quad F = -uv^2, \quad G = v^4 + v^2,$$

$$l = \frac{v}{|v|\sqrt{u^2 + v^2 + 1}}, \quad m = 0, \quad n = -\frac{v^2}{|v|\sqrt{u^2 + v^2 + 1}}.$$

Le curvatures Gaussiana e media valgono:

$$K = -\frac{1}{v(u^2 + v^2 + 1)^2}, \quad H = \frac{v^3 - u^2 + v - 1}{|v|\sqrt{(u^2 + v^2 + 1)^3}}.$$

Esaminando il segno di K , si deduce che i punti di M per cui $v < 0$ sono ellittici, quelli per cui $v > 0$ sono iperbolici.

b) La matrice dell'operatore forma nel punto $P_0 = \varphi(1, 1)$, rispetto alla base formata dai vettori $\varphi_u(1, 1) = (1, 0, 1)$ e $\varphi_v(1, 1) = (0, 1, -1)$ è:

$$\begin{pmatrix} \frac{2\sqrt{3}}{9} & -\frac{\sqrt{3}}{9} \\ \frac{\sqrt{3}}{9} & -\frac{2\sqrt{3}}{9} \end{pmatrix}.$$

Annullando il polinomio caratteristico $p(\lambda) = \lambda^2 - 1/9$, si trovano le curvatures principali: $k_1(P_0) = -1/3$ e $k_2(P_0) = 1/3$. Gli autovettori relativi all'autovalore $-1/3$ si trovano risolvendo il sistema lineare:

$$\begin{cases} (2\sqrt{3} + 3)x_1 - \sqrt{3}x_2 = 0 \\ \sqrt{3}x_1 + (-2\sqrt{3} + 3)x_2 = 0. \end{cases}$$

Considerando la soluzione non nulla $x_1 = 1$, $x_2 = 2 + \sqrt{3}$, si ha una base del relativo autospazio $\mathbf{v}_1 = \varphi_u(1, 1) + (2 + \sqrt{3})\varphi_v(1, 1) = (1, 2 + \sqrt{3}, -1 - \sqrt{3})$. Dopo aver calcolato

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{6(2 + \sqrt{3})}} (1, 2 + \sqrt{3}, -1 - \sqrt{3}), \quad \cos^2 \theta = (\mathbf{u} \cdot \mathbf{e}_1)^2 = \frac{1}{2},$$

si può applicare la formula di Eulero e si trova: $k_n(\mathbf{u}) = 0$.

Gruppi di matrici

1. Strutture algebriche

In questa sezione vengono introdotte alcune importanti strutture algebriche che saranno cruciali nello studio dei gruppi di Lie.

DEFINIZIONE 3.1. *Un insieme G dotato di un'operazione*

$$G \times G \longrightarrow G, \quad (x, y) \longmapsto x \circ y$$

per cui valgono le seguenti proprietà:

- (1) *associativa: $(x \circ y) \circ z = x \circ (y \circ z)$, $x, y, z \in G$;*
- (2) *esiste l'elemento neutro: $\exists e \in G, \forall x \in G / x \circ e = e \circ x = x$;*
- (3) *esiste l'inverso di ogni elemento: $\forall x \in G, \exists x^{-1} \in G / x \circ x^{-1} = x^{-1} \circ x = e$;*

prende il nome di **gruppo**.

TEOREMA 3.1. *Sia (G, \circ) un gruppo. Allora:*

- (1) *l'elemento neutro e di G è unico,*
- (2) *per ogni elemento x di G il suo inverso x^{-1} è unico,*
- (3) *si ha*

$$(x \circ y)^{-1} = y^{-1} \circ x^{-1}, \quad x, y \in G.$$

ESERCIZIO 3.1. Dimostrare il teorema precedente.

DEFINIZIONE 3.2. *Se l'operazione di un gruppo (G, \circ) è anche commutativa (ossia $x \circ y = y \circ x$, $x, y \in G$) il gruppo si dice **abeliano**.*

ESEMPIO 3.2. Uno spazio vettoriale V è un gruppo abeliano rispetto all'operazione di somma.

ESERCIZIO 3.3. Verificare che i seguenti esempi sono gruppi e determinare i loro elementi neutri e gli inversi di ogni loro elemento:

- 1) $(\mathbb{R} - \{0\}, \cdot)$.
- 2) $(\mathbb{Z}, +)$.
- 3) $(\{0\}, +)$.
- 4) $(\{1\}, \cdot)$.

ESEMPIO 3.4. Si considerino gli insiemi di n elementi delle radici n -esime dell'unità dati da:

$$G_n = \left\{ e^{2k\pi i/n} = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right), \quad k = 0, 1, 2, \dots, n-1 \right\}.$$

Tali insiemi, con l'operazione di prodotto di numeri complessi, sono esempi di gruppi con un numero finito di elementi. Ad esempio $G_4 = \{1, i, -1, -i\}$.

Si osserva che, rappresentando sul piano di Argand-Gauss i gruppi G_n , si ottengono evidenti simmetrie. Infatti gli elementi di G_n sono i vertici di un poligono regolare con n lati e centro nell'origine del riferimento.

Le seguenti strutture algebriche verranno utilizzate nelle sezioni seguenti.

DEFINIZIONE 3.3. *Un insieme \mathcal{A} su cui sono definite due operazioni: la somma “+” e il prodotto “·”, prende il nome di **anello** se:*

- (1) $(\mathcal{A}, +)$ è un gruppo abeliano,
- (2) il prodotto è associativo, cioè

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c, \quad a, b, c \in \mathcal{A},$$

- (3) valgono le proprietà distributive del prodotto rispetto alla somma

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c, \quad a, b, c \in \mathcal{A}.$$

Se per il prodotto esiste l'elemento neutro, l'anello si dice **dotato di unità**. Se il prodotto è commutativo, l'anello si dice **commutativo**.

ESEMPIO 3.5. L'insieme dei numeri relativi, $(\mathbb{Z}, +, \cdot)$, con le consuete operazioni di somma e di prodotto è un esempio di anello commutativo con unità. Se si considera l'insieme dei numeri pari $2\mathbb{Z}$ con le consuete operazioni di somma e di prodotto si ottiene un esempio di anello commutativo ma privo di unità. L'insieme delle matrici quadrate di ordine n , con $n \geq 2$, ad elementi reali, con le solite operazioni di somma e prodotto tra matrici, è un anello con unità non commutativo.

DEFINIZIONE 3.4. *Un insieme \mathbb{K} , su cui sono definite due operazioni: la somma “+” e il prodotto “·”, prende il nome di **campo** se:*

- (1) $(\mathbb{K}, +, \cdot)$ è un anello commutativo con unità,
- (2) $(\mathbb{K} - \{0\}, \cdot)$ è un gruppo commutativo.

ESEMPIO 3.6. \mathbb{R}, \mathbb{C} sono esempi di campi, rispetto alle consuete definizioni di somma e di prodotto.

DEFINIZIONE 3.5. *Uno spazio vettoriale V su cui si definisce un prodotto che è bilineare rispetto alla somma e al prodotto per scalari prende il nome di **algebra**. Se per il prodotto esiste l'elemento neutro, l'algebra si dice **dotata di unità**. Se il prodotto è commutativo, l'algebra si dice **commutativa**.*

ESEMPIO 3.7. Lo spazio vettoriale reale $\mathbb{R}^{n,n}$ della matrici quadrate ad elementi reali è un esempio di algebra non commutativa con unità. Lo spazio vettoriale (per esempio a destra) dei quaternioni è un esempio di algebra con unità.

2. I gruppi lineari generali reali e complessi

I gruppi di Lie di matrici, studiati nel successivo Capitolo 5, sono tutti riconducibili a sottogruppi dei gruppi lineari generali, reali o complessi, definiti in questa sezione.

ESEMPIO 3.8. L'insieme

$$GL(n, \mathbb{R}) = \{A \in \mathbb{R}^{n,n} / \exists A^{-1}\} = \{A \in \mathbb{R}^{n,n} / \det(A) \neq 0\}$$

è un gruppo (non abeliano) rispetto al prodotto di matrici. L'elemento neutro è la matrice unità $I \in \mathbb{R}^{n,n}$. $GL(n, \mathbb{R})$ prende il nome di **gruppo lineare generale**.

Si osservi che $(\mathbb{R}^{n,n}, +)$ è un gruppo abeliano rispetto alla somma di matrici, ma non rispetto al prodotto di matrici (se $\det(A) = 0$, A non è invertibile).

Si osservi inoltre che

$$GL(1, \mathbb{R}) = \{a \in \mathbb{R} / \exists a^{-1}\} = \mathbb{R} - \{0\} = \mathbb{R}^*.$$

DEFINIZIONE 3.6. Sia (G, \circ) un gruppo e $H \subseteq G$. H si dice **sottogruppo** di G se (H, \circ) è un gruppo rispetto alla restrizione dell'operazione \circ di G ad H , ciò equivale a richiedere che

- (1) $\forall x, y \in H \implies x \circ y \in H$;
- (2) $\forall x \in H \implies x^{-1} \in H$.

ESERCIZIO 3.9. Dimostrare che la definizione di sottogruppo H di G appena enunciata equivale a

$$\forall x, y \in H \implies x \circ y^{-1} \in H.$$

Osservare, inoltre, che l'elemento neutro $e \in G$ necessariamente appartiene ad H .

ESEMPIO 3.10. Ogni sottospazio vettoriale \mathcal{W} di uno spazio vettoriale V è un sottogruppo di $(V, +)$.

ESEMPIO 3.11. Il seguente sottoinsieme di $GL(n, \mathbb{R})$

$$SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) / \det(A) = 1\}$$

è un sottogruppo di $GL(n, \mathbb{R})$ e prende il nome di **gruppo lineare speciale reale**. La verifica è lasciata per esercizio. Si osservi che $SL(1, \mathbb{R}) = \{1\}$.

ESEMPIO 3.12. Il seguente sottoinsieme

$$He = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} / a, b, c \in \mathbb{R} \right\}$$

è un importante esempio di sottogruppo di $GL(3, \mathbb{R})$ che prende il nome di **gruppo di Heisenberg**. Si lascia per esercizio la verifica che He è un esempio di sottogruppo di $GL(3, \mathbb{R})$.

DEFINIZIONE 3.7. Siano (G_1, \circ) e (G_2, \circ) due gruppi. Una funzione

$$f: G_1 \longrightarrow G_2$$

si dice **omomorfismo** se

$$f(x \circ y) = f(x) \circ f(y), \quad x, y \in G_1.$$

Sono evidenti le nozioni di **endomorfismo**, **isomorfismo**.

TEOREMA 3.2. Sia $f: G_1 \longrightarrow G_2$ un omomorfismo di gruppi e siano e_1, e_2 i due elementi neutri di G_1 e di G_2 , rispettivamente. Allora

$$f(e_1) = e_2,$$

inoltre

$$f(x^{-1}) = f(x)^{-1}, \quad x \in G_1.$$

ESERCIZIO 3.13. 1) Dimostrare il teorema precedente.

2) Sia (G, \circ) un gruppo. La funzione

$$f: G \longrightarrow G, \quad x \longmapsto x^{-1}$$

è un endomorfismo di G ?

3) Verificare che la seguente funzione

$$f: \mathbb{R} \longrightarrow \mathbb{R}^*, \quad a \longmapsto e^a$$

è un omomorfismo tra i gruppi $(\mathbb{R}, +)$ e (\mathbb{R}^*, \cdot) .

L'ultimo esempio si estende in modo naturale ad una funzione

$$e^\bullet: \mathbb{R}^{n,n} \longrightarrow GL(n, \mathbb{R}), \quad A \longmapsto e^A.$$

Si veda il Capitolo 5 per la definizione e le principali proprietà di e^A .

Si consideri il campo dei numeri complessi \mathbb{C} . $(\mathbb{C}, +)$ e $(\mathbb{C} - \{0\}, \cdot)$ sono esempi di gruppi abeliani. Si ricorda che \mathbb{C} è uno spazio vettoriale su \mathbb{C} di dimensione 1 ma è anche uno spazio vettoriale su \mathbb{R} di dimensione 2. In questo caso una sua base è ad esempio $(1, i)$. Analogamente al caso reale, l'insieme delle matrici quadrate di ordine n ad elementi complessi $\mathbb{C}^{n,n}$ è un esempio di gruppo abeliano rispetto alla somma.

ESEMPIO 3.14. Il seguente insieme

$$GL(n, \mathbb{C}) = \{A \in \mathbb{C}^{n,n} / \exists A^{-1}\} = \{A \in \mathbb{C}^{n,n} / \det(A) \neq 0\}$$

è un gruppo rispetto al prodotto di matrici. $GL(n, \mathbb{C})$ prende il nome di **gruppo lineare generale complesso**.

Scopo delle prossime considerazioni è dimostrare il seguente teorema. I dettagli dei calcoli man mano indicati sono lasciati per esercizio.

TEOREMA 3.3. $GL(n, \mathbb{C})$ è isomorfo ad un sottogruppo proprio (strettamente incluso) di $GL(2n, \mathbb{R})$.

La dimostrazione è conseguenza delle seguenti considerazioni.

Si inizia con il considerare la funzione

$$f: \mathbb{C} \longrightarrow \mathbb{R}^2, \quad z = a + ib \longmapsto (a, b).$$

Si prova che f è un'applicazione \mathbb{R} -lineare, ossia:

- 1) $f(z_1 + z_2) = f(z_1) + f(z_2)$, $z_1, z_2 \in \mathbb{C}$;
- 2) $f(\lambda z) = \lambda f(z)$, $\lambda \in \mathbb{R}$, $z \in \mathbb{C}$.

Si vuole definire una funzione $\alpha: \mathbb{C}^{n,n} \longrightarrow \mathbb{R}^{2n,2n}$ che estenda la funzione precedente. A questo scopo si osservi che, nel caso particolare di $\mathbb{C}^{2,2}$, si ha

$$\begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} = \begin{pmatrix} a_1 + ib_1 & a_2 + ib_2 \\ a_3 + ib_3 & a_4 + ib_4 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} + i \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}.$$

Quindi ogni matrice $Z \in \mathbb{C}^{n,n}$ si può scrivere come

$$Z = A + iB$$

con $A, B \in \mathbb{R}^{n,n}$. Segue che $\mathbb{C}^{n,n}$ è uno spazio vettoriale reale di dimensione doppia della dimensione di $\mathbb{R}^{n,n}$ ossia

$$\dim(\mathbb{C}^{n,n}) = 2n^2.$$

Il prodotto di due matrici $Z_1 = A_1 + iB_1$, $Z_2 = A_2 + iB_2 \in \mathbb{C}^{n,n}$ è dato da

$$Z_1 Z_2 = (A_1 A_2 - B_1 B_2) + i(A_1 B_2 + B_1 A_2).$$

Si consideri la funzione $\alpha: \mathbb{C}^{n,n} \rightarrow \mathbb{R}^{2n,2n}$ definita da

$$\alpha(Z) = \alpha(A + iB) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}, \quad Z = A + iB \in \mathbb{C}^{n,n}. \quad (3.1)$$

Si verifica che

- 1) α è \mathbb{R} -lineare;
- 2) α non è suriettiva, infatti $\dim(\text{im } \alpha) = 2n^2 < (2n)^2$,
- 3) α è iniettiva, infatti $\ker \alpha = \{O\}$, (O è la matrice nulla di $\mathbb{C}^{n,n}$).
- 4) $\alpha(Z_1 Z_2) = \alpha(Z_1) \alpha(Z_2)$, $Z_1, Z_2 \in \mathbb{C}^{n,n}$,

DEFINIZIONE 3.8. L'immagine mediante la funzione α , definita in (3.1), della matrice $Z \in \mathbb{C}^{n,n}$ prende il nome di **rappresentazione reale di Z** .

OSSERVAZIONE 3.1. Si consideri un numero complesso z di modulo 1, ossia

$$z = e^{i\theta} = \cos \theta + i \sin \theta.$$

È chiaro che

$$\alpha(z) = R[\theta] = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

La rappresentazione reale permette di stabilire un isomorfismo tra i numeri complessi di modulo 1 e le rotazioni $R[\theta]$ di angolo θ del piano. Questa osservazione (che sarà chiarita meglio nel Capitolo 4) giustifica l'introduzione del concetto di *rappresentazione reale* di una matrice ad elementi complessi.

Data la funzione

$$J: \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto iz, \quad (3.2)$$

si verifica che J è un'applicazione \mathbb{C} -lineare. Considerando \mathbb{C} come spazio vettoriale su \mathbb{R} , si vede immediatamente che J è la rotazione anti-oraria di angolo $\pi/2$. La matrice associata a J , rispetto alla base $(1, i)$ è

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Si osservi che $J^2 = -I$ con $I \in \mathbb{R}^{2,2}$ matrice unità; quindi $J^{-1} = -J$ e $J = R[\frac{\pi}{2}]$.

In modo analogo si definisce la funzione

$$J: \mathbb{C}^n \rightarrow \mathbb{C}^n, \quad (z_1, z_2, \dots, z_n) \mapsto (iz_1, iz_2, \dots, iz_n).$$

Si verifica che J così definita è una applicazione \mathbb{C} -lineare la cui matrice associata rispetto alla base canonica di \mathbb{C}^n (considerato come spazio vettoriale

su \mathbb{C}) è

$$J = \begin{pmatrix} i & 0 & \dots & 0 \\ 0 & i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & i \end{pmatrix} = O + iI, \quad (3.3)$$

dove $O \in \mathbb{R}^{n,n}$ è la matrice nulla, mentre $I \in \mathbb{R}^{n,n}$ è la matrice unità. La rappresentazione reale di (3.3) non è altro che la matrice

$$J = \begin{pmatrix} O & -I \\ I & O \end{pmatrix} \in \mathbb{R}^{2n,2n}.$$

Per dimostrare il Teorema 3.3 bisogna far vedere che

$$\alpha(GL(n, \mathbb{C})) = \{P \in GL(2n, \mathbb{R}) / PJ = JP\}. \quad (3.4)$$

A tal fine occorre provare i seguenti fatti:

- $\{P \in GL(2n, \mathbb{R}) / PJ = JP\}$ è un sottogruppo di $GL(2n, \mathbb{R})$. La verifica segue facilmente dalla Definizione 3.6.
- $\alpha(\mathbb{C}^{n,n}) = \{P \in \mathbb{R}^{2n,2n} / PJ = JP\}$. A questo scopo innanzi tutto si dimostra che la matrice $\alpha(Z)$ verifica la relazione $PJ = JP$. Viceversa, considerata una generica matrice $P \in \mathbb{R}^{2n,2n}$ del tipo

$$P = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}, \quad X_i \in \mathbb{R}^{n,n}, \quad i = 1, \dots, 4,$$

imponendo che $PJ = JP$ segue $X_2 = -X_3$, $X_1 = X_4$, da cui la tesi.

- se $Z = A + iB \in GL(n, \mathbb{C})$, ossia se Z è invertibile, allora anche $\alpha(Z)$ è invertibile. Ma se Z è invertibile, esistono $X, Y \in \mathbb{R}^{n,n}$ tali che

$$(A + iB)(X + iY) = I$$

da cui segue

$$(AX - BY) + i(AX + BY) = I$$

ossia $AX - BY = I$ e $AX + BY = O$. È un semplice calcolo verificare che

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix}.$$

OSSERVAZIONE 3.2. Si vuole legare l'invertibilità di $Z = A + iB \in GL(n, \mathbb{C})$ al suo determinante.

Si osservi che $z = a + ib \in \mathbb{C} - \{0\}$ equivale a $|z| = \sqrt{a^2 + b^2} \neq 0$. Passando alla rappresentazione reale di z si ha

$$\alpha(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

e

$$\det \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a^2 + b^2,$$

ossia esiste z^{-1} se e solo se $\det \alpha(z) \neq 0$.

Questo fatto vale in generale per $Z \in \mathbb{C}^{n,n}$. Infatti si ha:

$$\det \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = |\det(A + iB)|^2 = |\det(Z)|^2, \quad (3.5)$$

Tale formula lega il $\det Z$ con il $\det \alpha(z)$. Rimane il problema (non banale) della dimostrazione della formula (3.5) per cui si rimanda alle note [13].

ESEMPIO 3.15. Il seguente sottoinsieme

$$SL(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) / \det(A) = 1\}$$

è un sottogruppo di $GL(n, \mathbb{C})$ e prende il nome di **gruppo lineare speciale complesso**. Si osservi che $SL(1, \mathbb{C}) = \{1\}$.

Se $Z \in \mathbb{C}^{m,n}$, si indichi con \bar{Z} la matrice **coniugata** di Z , ossia la matrice che ha come elementi i coniugati degli elementi di Z . Se $A \in \mathbb{R}^{m,n}$, cioè se A è reale, ovviamente $\bar{A} = A$. Per la coniugata di una matrice valgono ovvie proprietà che sono facile conseguenza dell'operazione di coniugio sui numeri complessi, e precisamente:

- (1) $\overline{Z_1 + Z_2} = \bar{Z}_1 + \bar{Z}_2$, $Z_1, Z_2 \in \mathbb{C}^{m,n}$;
- (2) $\overline{\lambda Z} = \bar{\lambda} \bar{Z}$, $\lambda \in \mathbb{C}, Z \in \mathbb{C}^{m,n}$;
- (3) $\overline{Z_1 Z_2} = \bar{Z}_1 \bar{Z}_2$, $Z_1 \in \mathbb{C}^{m,n}, Z_2 \in \mathbb{C}^{n,k}$;
- (4) ${}^t \bar{Z} = \overline{{}^t Z}$, $Z \in \mathbb{C}^{m,n}$;
- (5) $\det(\bar{Z}) = \overline{\det(Z)}$, $Z \in \mathbb{C}^{n,n}$.

ESEMPIO 3.16. Si può procedere ora al calcolo della matrice inversa di $Z \in \mathbb{C}^{n,n}$, nel caso in cui Z sia invertibile.

Sia $z = a + ib \in \mathbb{C}$, $z \neq 0$. È noto che

$$z^{-1} = \frac{a - ib}{a^2 + b^2} = \frac{\bar{z}}{z\bar{z}} = \bar{z}(z\bar{z})^{-1}.$$

Una formula analoga vale per una matrice invertibile $Z = A + iB \in \mathbb{C}^{n,n}$

$$Z^{-1} = \bar{Z}(Z\bar{Z})^{-1}.$$

Per il primo fattore si ottiene $\bar{Z} = A - iB$; per il secondo fattore si ha

$$Z\bar{Z} = (A + iB)(A - iB) = A^2 + B^2 - i(AB - BA).$$

Se si definisce **bracket o commutatore** l'operazione

$$[\ , \]: \mathbb{R}^{n,n} \times \mathbb{R}^{n,n} \longrightarrow \mathbb{R}^{n,n}, \quad (A, B) \longmapsto [A, B] = AB - BA,$$

si ottiene che

$$Z^{-1} = (A - iB)(A^2 + B^2 - i[A, B])^{-1}.$$

Nel caso particolare di matrici A, B che commutano si ha la formula analoga al caso di $z^{-1} \in \mathbb{C}$, ossia

$$Z^{-1} = (A - iB)(A^2 + B^2)^{-1}.$$

L'operazione $[\ , \]$ riveste una grande importanza nello studio delle **Algebre di Lie** (cfr. Capitolo 5).

3. I quaternioni

Il campo dei numeri reali \mathbb{R} è anche uno spazio vettoriale su sé stesso di dimensione 1 e base standard (1). Il campo dei numeri complessi \mathbb{C} è uno spazio vettoriale su \mathbb{R} di dimensione 2 e base standard $(1, i)$. Si pone, in modo evidente il problema di tentare di costruire un campo su \mathbb{R} di dimensione 3, per esempio con base $(1, i, j)$ con $j^2 = -1$.

Poniamo

$$\mathbb{R}^3 = \mathcal{L}(1, i, j) = \{a + ib + jc / a, b, c \in \mathbb{R}\}.$$

Se fosse un campo allora esisterebbero tre numeri reali x, y, z tali che

$$ij = x + iy + jz$$

ma allora

$$-j = (ii)j = i(ij) = i(x + iy + jz) = ix - y + ijz$$

ossia

$$-j = ix - y + (x + iy + jz)z = ix - y + xz + iy + jz^2$$

da cui si ha

$$-1 = z^2$$

che è palesemente assurdo.

Il problema è “quasi” risolto da Sir William Hamilton nel 1843 nel modo seguente. Si definisce l'insieme

$$\mathbb{H} = \{a + ib + jc + kd / a, b, c, d \in \mathbb{R}\}$$

su cui si introduce la somma, con l'usuale definizione di \mathbb{R}^4 e il prodotto che si ottiene considerando i, j, k come tre unità immaginarie tali che

$$i^2 = -1, \quad j^2 = -1, \quad k^2 = -1,$$

e “confondendo” i, j, k con i versori $\mathbf{i}, \mathbf{j}, \mathbf{k}$ dello spazio vettoriale dei vettori ordinari V_3 e il prodotto vettoriale tra essi, ponendo quindi

$$ij = k, \quad ik = -j, \quad ji = -k, \quad jk = i, \quad ki = j, \quad kj = -i.$$

Viene ad essere così definita su \mathbb{H} una struttura di campo non commutativo rispetto al prodotto, ossia un **corpo**.

Gli elementi di \mathbb{H} prendono il nome di **quaternioni** e \mathbb{H} è il corpo dei quaternioni.

Un quaternione $q = a + ib + jc + kd$ ha una **parte reale** a e una **parte immaginaria** $ib + jc + kd$.

ESERCIZIO 3.17. Dimostrare le seguenti proprietà su \mathbb{H} :

- 1) proprietà associativa del prodotto;
- 2) proprietà distributive del prodotto rispetto alla somma;
- 3) l'elemento neutro del prodotto è il numero 1;
- 4) se $q = a + ib + jc + kd \in \mathbb{H}$ e $q \neq 0$ allora

$$q^{-1} = \frac{a - ib - jc - kd}{a^2 + b^2 + c^2 + d^2}.$$

DEFINIZIONE 3.9. Dato il quaternione $q = a + ib + jc + kd$, si definisce **coniugato** di q il quaternione \bar{q} dato da

$$\bar{q} = a - (ib + jc + kd).$$

TEOREMA 3.4. Valgono le seguenti proprietà:

- (1) se $q = a + ib + jc + kd \in \mathbb{H}$, allora $q\bar{q} = \bar{q}q = a^2 + b^2 + c^2 + d^2$ è un numero reale non negativo;
- (2) l'inverso di un quaternione q non nullo è

$$q^{-1} = \left(\frac{1}{q\bar{q}} \right) \bar{q}; \quad (3.6)$$

- (3) $\overline{pq} = \bar{q}\bar{p}$, per ogni $p, q \in \mathbb{H}$ (attenzione!!);
- (4) $\overline{\bar{q}} = q$, per ogni $q \in \mathbb{H}$;
- (5) detto $|q| = \sqrt{q\bar{q}}$ il **modulo** del quaternione q , si ha: $|qp| = |q||p|$, per ogni $q, p \in \mathbb{H}$. Di conseguenza, se $q \neq 0$,

$$q^{-1} = \frac{\bar{q}}{|q|^2}.$$

ESERCIZIO 3.18. Dimostrare il teorema precedente.

OSSERVAZIONE 3.3. Si può usare il Quaternions Package di *Mathematica* per controllare le varie proprietà.

È ovvio che \mathbb{H} è uno spazio vettoriale reale di dimensione 4. Vediamo se \mathbb{H} sia anche uno spazio vettoriale complesso di dimensione 2.

Poiché

$$q = a + ib + jc + kd = (a + ib) + j(c - id) = u + jv, \quad (3.7)$$

la funzione

$$\beta: \mathbb{H} \longrightarrow \mathbb{C}^2, \quad q \longmapsto (u, v) \quad (3.8)$$

è una biiezione. β è sicuramente lineare rispetto alla somma ed è ovviamente \mathbb{R} -lineare. Posponiamo il problema della linearità rispetto a \mathbb{C} alle seguenti considerazioni. Si ricordi che \mathbb{H} **non** è un campo, ossia il prodotto in \mathbb{H} **non** è commutativo.

Se $q = u + jv$ si verifica facilmente che

$$\bar{q} = \bar{u} - jv.$$

Ci si propone di calcolare $q\bar{q}$. Si ha

$$q\bar{q} = (u + jv)(\bar{u} - jv) = u\bar{u} - u jv + jv\bar{u} - jv jv. \quad (3.9)$$

Si presenta il problema di calcolare uj , ossia

$$uj = (a + ib)j = aj + ibj = aj + ijb = ja + kb = j(a - ib)$$

vale a dire

$$uj = j\bar{u}. \quad (3.10)$$

Proseguendo il conto di (3.9) si ha

$$q\bar{q} = u\bar{u} - j\bar{u}v + jv\bar{u} - j(j\bar{v})v = u\bar{u} + v\bar{v}.$$

In conclusione

$$q\bar{q} = |u|^2 + |v|^2 = |q|^2.$$

Sostituendo in (3.6) segue

$$q^{-1} = \frac{\bar{u} - jv}{|u|^2 + |v|^2}.$$

Si osservi che la dicitura: V spazio vettoriale su \mathbb{H} è ambigua, è infatti necessario prestare particolare attenzione alla definizione di prodotto di un vettore di V per un elemento di \mathbb{H} . Si presentano, quindi, due possibilità

$$1. \quad \mathbb{H} \times V \longrightarrow V, \quad (\lambda, \mathbf{x}) \longmapsto \lambda \mathbf{x},$$

con questa definizione, dopo aver verificato le proprietà richieste, si perviene alla definizione di V spazio vettoriale su \mathbb{H} **a sinistra**.

Invece, mediante la definizione

$$2. \quad \mathbb{H} \times V \longrightarrow V, \quad (\lambda, \mathbf{x}) \longmapsto \mathbf{x} \lambda,$$

si perviene alla definizione di V spazio vettoriale su \mathbb{H} **a destra**.

Si osservi che se V è uno spazio vettoriale su \mathbb{H} a sinistra, allora la definizione di applicazione lineare rispetta la stessa convenzione.

Per capire meglio le conseguenze, per semplicità di esposizione, consideriamo \mathbb{H}^n spazio vettoriale su \mathbb{H} a sinistra. Sia

$$f: \mathbb{H}^n \longrightarrow \mathbb{H}^n$$

un'applicazione lineare, con matrice associata (rispetto una base fissata) $M \in \mathbb{H}^{n,n}$. Si sta costruendo la matrice M ponendo **sulle righe** le componenti delle immagini dei vettori di una base di \mathbb{H}^n . Sia $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{H}^n$ allora $f(\mathbf{x})$ in notazione matriciale è

$$(x_1, x_2, \dots, x_n)M,$$

in quanto l'usuale prodotto

$$M \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

non potrebbe rispettare le regole di linearità $M(\lambda \mathbf{x}) \neq \lambda M(\mathbf{x})$, $\lambda \in \mathbb{H}$. Ricordiamo che è anche necessario distinguere se \mathbb{H} è uno spazio vettoriale su \mathbb{C} a destra o a sinistra. Infatti, in generale $zq \neq qz$, con $z \in \mathbb{C}$ e $q \in \mathbb{H}$.

Sulla base di queste considerazioni, data la funzione β definita in (3.8) e \mathbb{H} spazio vettoriale su \mathbb{C} con prodotto a sinistra, controlliamo se β sia o meno lineare rispetto a \mathbb{C} . Si tratta di verificare che

$$\beta(\lambda q) = \lambda \beta(q), \quad \lambda \in \mathbb{C}.$$

Dalla relazione (3.10) si ottiene

$$\beta(\lambda q) = (\lambda u, \bar{\lambda} v),$$

invece

$$\lambda \beta(q) = (\lambda u, \lambda v)$$

non verificando quindi l'affermazione richiesta.

Se, invece, si considera \mathbb{H} come spazio vettoriale su \mathbb{C} a destra, l'uguaglianza è verificata.

OSSERVAZIONE 3.4. A maggior chiarimento, si osservi che la formula (3.7) dipende dalla scelta effettuata nella scrittura di q . Infatti, ponendo

$$q = a + bi + cj + dk = (a + bi) + (c + di)j = z_1 + z_2j = z_1 + j\bar{z}_2,$$

con $z_1, z_2 \in \mathbb{C}$. Con questa scelta, tutte le formule che seguono dovranno essere opportunamente modificate.

Un altro problema che si presenta in modo evidente è la definizione di determinante di una matrice ad elementi in \mathbb{H} . Si consideri per esempio la matrice

$$A = \begin{pmatrix} i & j \\ i & j \end{pmatrix}$$

di evidente rango 1. Il suo determinante però potrebbe essere calcolato nel modo seguente

$$\det(A) = ij - ji = 2ij = 2k \neq 0.$$

È perciò anche necessario formalizzare in modo privo di ambiguità il calcolo del determinante di una matrice ad elementi in \mathbb{H} che non può essere la naturale estensione della definizione di determinante di una matrice ad elementi in \mathbb{R} o in \mathbb{C} .

4. Gruppo lineare quaternionico

Si inizia con la ovvia definizione di **gruppo lineare generale quaternionico** ossia

$$GL(n, \mathbb{H}) = \{A \in \mathbb{H}^{n,n} / \exists A^{-1}\}.$$

Si vuole dimostrare il seguente teorema (analogo al Teorema 3.3).

TEOREMA 3.5. $GL(n, \mathbb{H})$ è isomorfo ad un sottogruppo proprio (strettamente incluso) di $GL(2n, \mathbb{C})$.

Questo teorema giustifica la scelta di definire la funzione β come in (3.8) anche se β non è un'applicazione lineare su \mathbb{C} .

Per la dimostrazione del Teorema 3.5 si procede come nel caso reale definendo la naturale estensione ad $\mathbb{H}^{n,n}$ della funzione α introdotta in (3.1), iniziando dal caso particolare di \mathbb{H} . Si definisce quindi

$$\alpha : \mathbb{H} \longrightarrow \mathbb{C}^{2,2}, \quad q = u + jv \longmapsto \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix}. \quad (3.11)$$

ESERCIZIO 3.19. Dimostrare che per la funzione α appena definita valgono le seguenti proprietà

1. $\alpha(q_1 + q_2) = \alpha(q_1) + \alpha(q_2)$, $q_1, q_2 \in \mathbb{H}$,
2. $\alpha(\lambda q) = \lambda \alpha(q)$, $\lambda \in \mathbb{R}, q \in \mathbb{H}$,
3. $\alpha(q_1 q_2) = \alpha(q_1) \alpha(q_2)$, $q_1, q_2 \in \mathbb{H}$,
4. α è iniettiva,
5. α non è suriettiva.

OSSERVAZIONE 3.5. Una definizione alternativa di α , ma assolutamente ad essa equivalente per i risultati ottenuti, è la seguente

$$\alpha: \mathbb{H} \longrightarrow \mathbb{C}^{2,2}, \quad q = u + jv \longmapsto \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}.$$

Usando questa definizione si ha immediatamente il collegamento con la formula (4.20).

OSSERVAZIONE 3.6. La scelta di $q = u + jv$, che comporta la definizione di

$$\beta: \mathbb{H} \longrightarrow \mathbb{C}^2, \quad q \rightarrow (u, v),$$

ha come conseguenza il fatto che β sia un isomorfismo di spazi vettoriali su \mathbb{C} a destra e giustifica la definizione di

$$\alpha: \mathbb{H} \longrightarrow \mathbb{C}^{2,2}, \quad q = u + jv \rightarrow \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix}.$$

Infatti, fissato $q = u + jv$, il seguente diagramma commuta:

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{\beta} & \mathbb{C}^2 \\ \gamma \downarrow & & \downarrow \alpha(q) \\ \mathbb{H} & \xrightarrow{\beta} & \mathbb{C}^2 \end{array}$$

dove

$$\gamma: \mathbb{H} \longrightarrow \mathbb{H}, \quad q' \rightarrow qq',$$

è il prodotto a sinistra. La verifica è lasciata per esercizio.

Sia $Q \in \mathbb{H}^{n,n}$. Q determina in modo naturale due matrici $U, V \in \mathbb{C}^{n,n}$ tali che

$$Q = U + jV.$$

La funzione α definita in (3.11) si estende quindi a

$$\alpha: \mathbb{H}^{n,n} \longrightarrow \mathbb{C}^{2n,2n}, \quad Q = U + jV \longmapsto \begin{pmatrix} U & -\bar{V} \\ V & \bar{U} \end{pmatrix}.$$

Per la funzione α appena definita valgono le estensioni evidenti delle proprietà enunciate nell'esercizio 3.19, la cui dimostrazione è lasciata per esercizio. L'immagine, mediante α , della matrice $Q \in \mathbb{H}^{n,n}$ prende il nome di **rappresentazione complessa della matrice Q** .

Il Teorema 3.5 segue dalla seguente uguaglianza

$$\alpha(GL(n, \mathbb{H})) = \{Z \in GL(2n, \mathbb{C}) / ZJ = J\bar{Z}\},$$

la cui dimostrazione è simile a quella del caso reale ed è lasciata per esercizio.

DEFINIZIONE 3.10. Si definisce **determinante** di una matrice $Q \in \mathbb{H}^{n,n}$ il determinante della matrice $\alpha(Q)$, rappresentazione complessa della matrice Q .

Da questa definizione segue che se $q = u + jv = a + ib + jc + kd \in \mathbb{H}$ allora

$$\det(q) = \det \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix} = |u|^2 + |v|^2 = a^2 + b^2 + c^2 + d^2 \in \mathbb{R}. \quad (3.12)$$

Se $Q \in GL(n, \mathbb{H})$ allora $\det(Q) = \det(\alpha(Q))$, per definizione

$$\alpha(Q)J = J\overline{\alpha(Q)}$$

da cui

$$\det(\alpha(Q)) \det(J) = \det(J) \overline{\det(\alpha(Q))},$$

quindi

$$\det(\alpha(Q)) \in \mathbb{R}.$$

Ha quindi senso la seguente definizione di **gruppo lineare speciale quaternionico**:

$$SL(n, \mathbb{H}) = \{Q \in GL(n, \mathbb{H}) / \det(Q) = 1\}.$$

Si osservi che, per $n = 1$, da (3.12) segue subito l'importante uguaglianza

$$S^3(1) = \{(a, b, c, d) \in \mathbb{R}^4 / a^2 + b^2 + c^2 + d^2 = 1\} = SL(1, \mathbb{H}).$$

5. Gli ottetti di Cayley

Una naturale generalizzazione dell'algebra dei quaternioni \mathbb{H} è costituita dall'algebra degli **ottetti di Cayley**, che si definiscono estendendo l'algebra dei quaternioni. Si consideri la seguente rappresentazione di estendendo \mathbb{H}

$$\mathbb{H} = \mathbb{C} + \mathbb{C}j$$

in cui ogni quaternionione si vede come coppia di numeri complessi

$$q = a + bj = (a, b), \quad a, b \in \mathbb{C}$$

con il prodotto

$$(a, b)(c, d) = (a+bj)(c+dj) = (ac - b\bar{d}) + (ad + b\bar{c})j = (ac - \bar{b}d, ad + b\bar{c}). \quad (3.13)$$

Vengono così definiti otto elementi, detti unità, nel modo seguente:

$$\begin{aligned} e_0 &= (1, 0) & e_4 &= (0, 1) \\ e_1 &= (i, 0) & e_5 &= (0, i) \\ e_2 &= (j, 0) & e_6 &= (0, j) \\ e_3 &= (k, 0) & e_7 &= (0, k), \end{aligned}$$

che verificano le seguenti uguaglianze:

$$e_1^2 = e_2^2 = e_3^2 = e_4^2 = e_5^2 = e_6^2 = e_7^2 = -e_0.$$

I prodotti tra le unità e_h sono elencati nella tabella seguente e sono dati dalla regola (3.13), in cui, per esempio

$$e_3e_4 = (k, 0)(0, 1) = (0, k) = e_7.$$

Si osservi che non vale la proprietà associativa, infatti si ha

$$(e_h e_k) e_p = -e_h (e_k e_p), \quad h, k, p = 1, \dots, 7, \quad h \neq k \neq p.$$

Si definisce, quindi, l'algebra \mathbb{O} degli ottetti di Cayley come:

$$\mathbb{O} = \{a_0 e_0 + a_1 e_1 + a_2 e_2 + a_3 e_3 + a_4 e_4 + a_5 e_5 + a_6 e_6 + a_7 e_7 / a_h \in \mathbb{R}\}$$

con le naturali definizioni di somma e di prodotto per numeri reali. È chiaro che \mathbb{O} è uno spazio vettoriale su \mathbb{R} di dimensione 8 ed è un'algebra non associativa, non commutativa, ma con unità.

| | e_0 | e_1 | e_2 | e_3 | e_4 | e_5 | e_6 | e_7 |
|-------|-------|--------|--------|--------|--------|--------|--------|--------|
| e_0 | e_0 | e_1 | e_2 | e_3 | e_4 | e_5 | e_6 | e_7 |
| e_1 | e_1 | $-e_0$ | e_3 | $-e_2$ | e_5 | $-e_4$ | $-e_7$ | e_6 |
| e_2 | e_2 | $-e_3$ | $-e_0$ | e_1 | e_6 | e_7 | $-e_4$ | $-e_5$ |
| e_3 | e_3 | e_2 | $-e_1$ | $-e_0$ | e_7 | $-e_6$ | e_5 | $-e_4$ |
| e_4 | e_4 | $-e_5$ | $-e_6$ | $-e_7$ | $-e_0$ | e_1 | e_2 | e_3 |
| e_5 | e_5 | e_4 | $-e_7$ | e_6 | $-e_1$ | $-e_0$ | $-e_3$ | e_2 |
| e_6 | e_6 | e_7 | e_4 | $-e_5$ | $-e_2$ | e_3 | $-e_0$ | $-e_1$ |
| e_7 | e_7 | $-e_6$ | e_5 | e_4 | $-e_3$ | $-e_2$ | e_1 | $-e_0$ |

È evidente che il sottospazio vettoriale $\mathcal{L}(e_0, e_1)$ è isomorfo a \mathbb{C} , mentre $\mathcal{L}(e_0, e_1, e_2, e_3)$ è isomorfo a \mathbb{H} .

Si dimostra che vale la proprietà associativa con tre elementi, di cui almeno due uguali tra di loro:

$$(xx)y = x(xy), \quad \forall x, y \in \mathbb{O}.$$

Di grande interesse è il gruppo delle matrici $A \in GL(8, \mathbb{R})$ che mantengono il prodotto di elementi di \mathbb{O} , ossia tali che

$$A(xy) = A(x)A(y), \quad x, y \in \mathbb{O}.$$

Dato che $A(e_0) = e_0$, la matrice A deve avere la forma seguente:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix}, \quad A' \in GL(7, \mathbb{R}).$$

Il sottogruppo di $GL(7, \mathbb{R})$ delle matrici A' così definite coincide con il gruppo eccezionale G_2 . Per ulteriori informazioni sugli ottetti di Cayley si veda [2].

Forme bilineari e prodotti scalari

1. Gruppi di matrici invarianti per forme bilineari

Sia

$$B: V \times V \longrightarrow \mathbb{R}$$

una **forma bilineare non degenera** definita su V , spazio vettoriale reale di $\dim(V) = n$.

Ci si propone di studiare l'insieme delle matrici $A \in GL(n, \mathbb{R})$ per cui

$$B(AX, AY) = B(X, Y), \quad X, Y \in V. \quad (4.1)$$

Con la definizione precedente si confonde il vettore $\mathbf{x} \in V$ con la matrice colonna $X \in \mathbb{R}^{n,1}$ delle sue componenti rispetto ad una base prefissata in V . Si richiede, sostanzialmente, di studiare le matrici dei cambiamenti di base in V che mantengono invariato il valore di $B(\mathbf{x}, \mathbf{y})$ quando sono fissati i vettori \mathbf{x} e \mathbf{y} . (Infatti si chiede che $B(X, Y) = B(X', Y')$, con X, X' componenti di \mathbf{x} rispetto a due basi diverse in V e Y, Y' componenti di \mathbf{y} rispetto a due basi diverse in V , ma è noto che se $A \in GL(n, \mathbb{R})$ è la matrice del cambiamento di base, allora $X = AX'$.)

È un facile **esercizio** controllare che l'insieme delle matrici $A \in GL(n, \mathbb{R})$ che verificano la relazione (4.1) è un sottogruppo di $GL(n, \mathbb{R})$.

Si considerano i casi seguenti:

- 1) B forma bilineare simmetrica definita positiva, ossia un prodotto scalare.
- 2) B forma bilineare simmetrica definita negativa. Questo caso, in realtà coincide con il precedente, a meno di un cambiamento di segno.
- 3) B forma bilineare simmetrica di segnatura $(k, n - k)$.
- 4) B forma bilineare non degenera antisimmetrica.

Il primo caso si studia nei corsi elementari di Algebra Lineare, si ottiene, infatti, il **gruppo ortogonale**

$$O(n) = \{P \in GL(n, \mathbb{R}) / {}^t P P = I\}.$$

Si ricorda il seguente teorema, oggetto del primo corso di Algebra Lineare

TEOREMA 4.1. *Valgono le seguenti proprietà:*

- (1) *il prodotto di due matrici ortogonali è una matrice ortogonale;*
- (2) *la matrice identità I è ortogonale;*
- (3) *l'inversa P^{-1} di una matrice ortogonale P è ortogonale;*
- (4) *la trasposta ${}^t P$ di una matrice ortogonale P è ortogonale;*

(5) una matrice $P \in \mathbb{R}^{n,n}$ è ortogonale se e solo se le righe e le colonne di P sono le componenti di una base ortonormale in \mathbb{R}^n , rispetto al prodotto scalare standard;

(6) il determinante di una matrice ortogonale P vale ± 1 .

OSSERVAZIONE 4.1. Non vale il viceversa del punto (6) del teorema precedente. Ad esempio, se si considera la matrice:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

si ha che il determinante di A è 1 ma A non è ortogonale.

Si definisce, quindi, in modo naturale l'insieme

$$SO(n) = \{A \in O(n) / \det(A) = 1\} = O(n) \cap SL(n, \mathbb{R}) \quad (4.2)$$

che è un gruppo rispetto al prodotto di matrici. $SO(n)$ è detto **gruppo ortogonale speciale**.

2. Isometrie, similitudini

Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione finita. La definizione seguente estende (in modo naturale) a dimensioni superiori il concetto elementare di isometria o *movimento euclideo* nel piano e nello spazio.

DEFINIZIONE 4.1. Un endomorfismo f di uno spazio vettoriale euclideo (V, \cdot) prende il nome di isometria o trasformazione ortogonale se:

$$f(\mathbf{x}) \cdot f(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in V. \quad (4.3)$$

OSSERVAZIONE 4.2. È facile verificare che non è necessario assumere nella definizione di isometria che essa sia un'applicazione lineare, infatti la linearità di f è una conseguenza di (4.3).

Il teorema che segue afferma che la definizione di isometria impone che necessariamente essa sia un'isomorfismo.

TEOREMA 4.2. Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione n . Se f è un'isometria di (V, \cdot) , allora f è un automorfismo di V .

DIMOSTRAZIONE. È sufficiente dimostrare che $\ker f = \{\mathbf{o}\}$. Se \mathbf{x} è un vettore di $\ker f$ si ha $f(\mathbf{x}) = \mathbf{o}$, d'altra parte essendo f un'isometria segue:

$$f(\mathbf{x}) \cdot f(\mathbf{x}) = \|\mathbf{x}\|^2 = \|\mathbf{o}\|^2 = 0,$$

quindi $\mathbf{x} = \mathbf{o}$. □

Si può generalizzare la Definizione 4.1 precedente al caso di isomorfismi tra due spazi vettoriali euclidei in questo modo: dati due spazi vettoriali euclidei V e W , con la stessa dimensione, un isomorfismo $f : V \rightarrow W$ si dice isometria se “non cambia” il prodotto scalare.

ESEMPIO 4.1. Ogni rotazione $R[\theta]$ (in senso antiorario) di angolo θ del piano vettoriale V_2 è un'isometria. Inoltre, come è già stato affermato, la

matrice associata a $R[\theta]$ (rispetto ad una base ortonormale (\mathbf{i}, \mathbf{j}) di V_2) è la matrice ortogonale:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

ESEMPIO 4.2. L'identità $\text{id}: V \rightarrow V$ è un'isometria rispetto ad ogni prodotto scalare definito su V .

ESEMPIO 4.3. L'applicazione $-\text{id}: V \rightarrow V$ definita da $-\text{id}(\mathbf{x}) = -\mathbf{x}$, $\mathbf{x} \in V$, è un'isometria rispetto ad ogni prodotto scalare definito su V .

Alcune tra le principali proprietà delle isometrie sono riassunte nel seguente teorema.

TEOREMA 4.3. *Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione n .*

- (1) *Un endomorfismo f di V è un'isometria se e solo se non cambia la norma dei vettori:*

$$\|f(\mathbf{x})\| = \|\mathbf{x}\|, \quad \mathbf{x} \in V. \quad (4.4)$$

- (2) *Se f è un'isometria di V , allora la misura dell'angolo individuato dai vettori \mathbf{x} e \mathbf{y} di V coincide con la misura dell'angolo individuato dai vettori $f(\mathbf{x})$ e $f(\mathbf{y})$, per ogni $\mathbf{x}, \mathbf{y} \in V$.*

- (3) *La composizione di due isometrie è un'isometria.*

- (4) *L'inversa di un'isometria è un'isometria.*

- (5) *Un endomorfismo f di V è un'isometria di V se e solo se le immagini dei vettori di una base ortonormale di V formano una base ortonormale di V .*

- (6) *Un endomorfismo f di V è un'isometria di V se e solo se la matrice associata ad f , rispetto ad una base ortonormale di V , è una matrice ortogonale.*

DIMOSTRAZIONE. (1) Se f è un'isometria, la relazione (4.4) segue ponendo $\mathbf{y} = \mathbf{x}$ in (4.3). Viceversa, si dimostra che se vale (4.4) allora f è un'isometria. Dalla formula

$$\|(\mathbf{x} + \mathbf{y})\|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2. \quad (4.5)$$

si ottiene:

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2} (\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2)$$

e quindi

$$f(\mathbf{x}) \cdot f(\mathbf{y}) = \frac{1}{2} (\|f(\mathbf{x}) + f(\mathbf{y})\|^2 - \|f(\mathbf{x})\|^2 - \|f(\mathbf{y})\|^2).$$

Pertanto per la linearità di f e da (4.4) segue la relazione (4.3).

- (2) Se f è un'isometria, da (4.3) e dal punto (1) segue:

$$\|\mathbf{x}\| \|\mathbf{y}\| \cos(\widehat{\mathbf{x}\mathbf{y}}) = \|\mathbf{x}\| \|\mathbf{y}\| \cos(\widehat{f(\mathbf{x})f(\mathbf{y})}),$$

per ogni $\mathbf{x}, \mathbf{y} \in V$, dove $\widehat{\mathbf{x}\mathbf{y}}$ e $\widehat{f(\mathbf{x})f(\mathbf{y})}$ indicano, rispettivamente, l'angolo tra i vettori \mathbf{x} e \mathbf{y} ed i vettori $f(\mathbf{x})$ e $f(\mathbf{y})$ (cfr. la Definizione 4.2 riportata di seguito). Pertanto:

$$\cos(\widehat{\mathbf{x}\mathbf{y}}) = \cos(\widehat{f(\mathbf{x})f(\mathbf{y})}). \quad (4.6)$$

(3) Se f e g sono isometrie, allora:

$$\|(g \circ f)(\mathbf{x})\| = \|g(f(\mathbf{x}))\| = \|f(\mathbf{x})\| = \|\mathbf{x}\|,$$

per ogni $\mathbf{x} \in V$.

(4) Sia f un'isometria. Si ha: $\|f(f^{-1}(\mathbf{x}))\| = \|\text{id}(\mathbf{x})\|$, dove id è l'identità di V , ma

$$\|f(f^{-1}(\mathbf{x}))\| = \|f^{-1}(\mathbf{x})\| = \|\mathbf{x}\|,$$

per ogni $\mathbf{x} \in V$, da cui la tesi.

(5) Se f è un'isometria e $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ una base ortonormale di V , allora $\mathcal{B}' = (f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n))$ è una base ortonormale perché f mantiene immutata sia la norma dei vettori sia i loro prodotti scalari. Viceversa, siano $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ e $\mathcal{B}' = (f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n))$ due basi ortonormali di V . Dato:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n$$

in V , allora:

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + \dots + x_n^2,$$

d'altra parte, per la linearità di f :

$$f(\mathbf{x}) = x_1f(\mathbf{e}_1) + x_2f(\mathbf{e}_2) + \dots + x_nf(\mathbf{e}_n),$$

quindi:

$$\|f(\mathbf{x})\|^2 = x_1^2 + x_2^2 + \dots + x_n^2,$$

da cui $\|f(\mathbf{x})\| = \|\mathbf{x}\|$. Si osservi che il calcolo della norma dei vettori \mathbf{x} e $f(\mathbf{x})$ ha assunto l'espressione suddetta in quanto riferito a due basi ortonormali. Quest'ultima affermazione segue dal Teorema 4.4 ricordato alla fine della dimostrazione.

(6) Sia A la matrice associata ad f , rispetto ad una base ortonormale \mathcal{B} di V , e siano $Y = AX$ le equazioni di f rispetto a \mathcal{B} . Supponendo per ogni $\mathbf{x} \in V$, $\|f(\mathbf{x})\|^2 = \|\mathbf{x}\|^2$ e ricordando che $\|\mathbf{x}\|^2 = {}^tXX$ si ha:

$${}^tXX = {}^tX'X' = {}^t(AX)(AX) = {}^tX{}^tAAX,$$

da cui segue la tesi. Infatti

$${}^tX(I - {}^tAA)X = 0 \implies {}^tAA = I.$$

Si osservi che $\|\mathbf{x}\|^2 = {}^tXX$ se e solo se X è la matrice colonna delle componenti di \mathbf{x} rispetto ad una base ortonormale.

□

DEFINIZIONE 4.2. Sia (V, \cdot) uno spazio vettoriale euclideo. Dati due vettori $\mathbf{x}, \mathbf{y} \in V$ non nulli, si definisce angolo tra i due vettori \mathbf{x} e \mathbf{y} l'angolo $\theta \in [0, \pi]$ tale che:

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

TEOREMA 4.4. Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione n .

(1) Se $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ è una base ortonormale e:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n, \quad \mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \dots + y_n\mathbf{e}_n$$

sono due vettori qualsiasi di V , allora il loro prodotto scalare

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \dots + x_ny_n = {}^tXY, \quad (4.7)$$

dove tXY denota il prodotto della matrice riga e della matrice colonna formate rispettivamente dalle componenti di \mathbf{x} e \mathbf{y} , rispetto alla base \mathcal{B} .

(2) Se $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ è una base di V tale che, per ogni coppia di vettori:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n, \quad \mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \dots + y_n\mathbf{e}_n$$

di V si ha:

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \dots + x_ny_n = {}^tXY, \quad (4.8)$$

dove tXY denota il prodotto della matrice riga e della matrice colonna formate rispettivamente dalle componenti di \mathbf{x} e \mathbf{y} , rispetto alla base \mathcal{B} , allora \mathcal{B} è una base ortonormale.

OSSERVAZIONE 4.3. Dai punti (3) e (4) del teorema precedente segue che l'insieme delle isometrie di uno spazio vettoriale euclideo (V, \cdot) è un gruppo rispetto alla composizione di funzioni. Inoltre, fissata una base ortonormale \mathcal{B} nello spazio vettoriale euclideo (V, \cdot) , per la proprietà 6. si può stabilire, in modo naturale, un isomorfismo tra l'insieme delle isometrie di (V, \cdot) ed il gruppo ortogonale $O(n)$ associando ad ogni isometria f di V la matrice ortogonale $A = M^{\mathcal{B}, \mathcal{B}}(f)$.

Si osservi che se un automorfismo di uno spazio vettoriale euclideo (V, \cdot) mantiene invariati gli angoli tra i vettori di V allora non è necessariamente un'isometria. Un esempio elementare è dato dall'automorfismo:

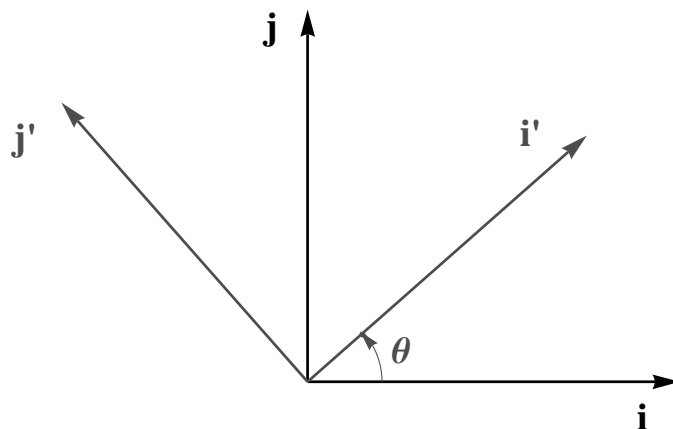
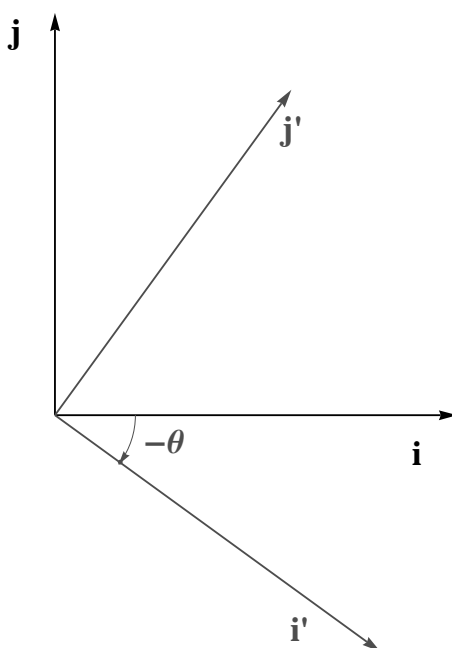
$$f : V \longrightarrow V, \quad \mathbf{x} \longmapsto 2\mathbf{x}, \quad (4.9)$$

ossia dalla funzione 2id .

Si consideri ora il caso particolare del cambiamento di base tra basi ortonormali in un piano vettoriale V_2 con lo scopo di determinare gli elementi della matrice ortogonale $P = (p_{ij}) \in \mathbb{R}^{2,2}$ che lo regola. Date due basi ortonormali $\mathcal{B} = (\mathbf{i}, \mathbf{j})$ e $\mathcal{B}' = (\mathbf{i}', \mathbf{j}')$ la seconda base si può ottenere dalla prima in uno dei modi rappresentati nelle Figure 2.1, 2.2, 2.3. Si osservi che le prime due figure sono in realtà dello stesso tipo e corrispondono ad una rotazione che la base \mathcal{B} compie per sovrapporsi alla base \mathcal{B}' , invece nella terza figura la base \mathcal{B} deve effettuare un movimento di riflessione (non interno al piano vettoriale V_2) per sovrapporsi alla base \mathcal{B}' .

Nel primo caso (cfr. Fig. 1) si ha:

$$\begin{cases} \mathbf{i}' = \cos \theta \mathbf{i} + \sin \theta \mathbf{j} \\ \mathbf{j}' = \cos \left(\theta + \frac{\pi}{2} \right) \mathbf{i} + \sin \left(\theta + \frac{\pi}{2} \right) \mathbf{j} = -\sin \theta \mathbf{i} + \cos \theta \mathbf{j}, \end{cases}$$

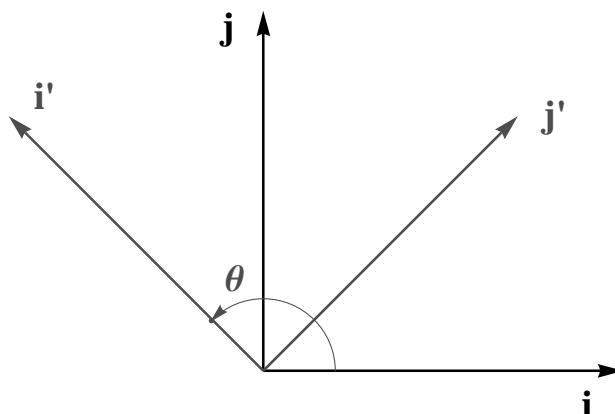
FIGURA 1. Cambiamento di basi ortonormali in V_2 , primo casoFIGURA 2. Cambiamento di basi ortonormali in V_2 , secondo caso

quindi la matrice del cambiamento di base da \mathcal{B} a \mathcal{B}' è:

$$P = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (4.10)$$

Si osservi che $\det(P) = 1$. In questo contesto, per mettere in evidenza la dipendenza della matrice P dall'angolo θ , è preferibile usare la notazione:

$$P = R[\theta].$$

FIGURA 3. Cambiamento di basi ortonormali in V_2 , terzo caso

Nel secondo caso (cfr. Fig. 2), con un procedimento analogo a quello appena descritto, si ottiene:

$$R[-\theta] = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (4.11)$$

anche in questo caso $\det(R[-\theta]) = 1$. Si osservi che la matrice (4.11) corrisponde alla matrice (4.10) in cui al posto di θ si considera l'angolo $-\theta$.

Nel terzo caso (cfr. Fig. 3) si ha:

$$P = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}. \quad (4.12)$$

invece, in questo caso $\det(P) = -1$.

Si osservi, inoltre, che

$$R[\theta]^2 = R[2\theta]$$

$$R[\theta + \varphi] = R[\theta]R[\varphi]$$

$$R[-\theta] = R[\theta]^{-1}.$$

ESEMPIO 4.4. Gli elementi di $O(2)$, ovvero le matrici ortogonali di ordine 2, sono necessariamente o di tipo (4.10) o di tipo (4.12). Pertanto gli endomorfismi del piano vettoriale V_2 con matrice associata, rispetto ad una base ortonormale $\mathcal{B} = (\mathbf{i}, \mathbf{j})$, di tipo (4.10) o di tipo (4.12) sono isometrie di V_2 . Nel caso di matrici di tipo (4.10) si ottengono le rotazioni $R[\theta]$ già considerate nell'Esempio 4.1. Se si considera invece un endomorfismo r_θ del piano vettoriale V_2 con matrice associata, rispetto ad una base ortonormale $\mathcal{B} = (\mathbf{i}, \mathbf{j})$, di tipo (4.12), ossia:

$$A_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix},$$

si ha un'isometria di V_2 tale che $r_\theta \circ r_\theta$ coincide con l'identità di V_2 , ovvero per cui $r_\theta^{-1} = r_\theta$.

ESERCIZIO 4.5. Si consideri su \mathbb{R}^2 la struttura euclidea determinata dal prodotto scalare

$$(x_1, x_2) \cdot (x_2, y_2) = x_1 y_1 + 4x_2 y_2.$$

Verificare che l'automorfismo di \mathbb{R}^2 la cui matrice associata, rispetto alla base canonica $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ di \mathbb{R}^2 , è la matrice:

$$A = \begin{pmatrix} \frac{\sqrt{3}}{2} & 1 \\ -\frac{1}{4} & \frac{\sqrt{3}}{2} \end{pmatrix},$$

è un'isometria di \mathbb{R}^2 . Per quale motivo A non è una matrice ortogonale?

SOLUZIONE. Le equazioni di f , rispetto alla base \mathcal{B} , sono:

$$\begin{cases} x'_1 = \frac{\sqrt{3}}{2}x_1 + x_2 \\ x'_2 = -\frac{1}{4}x_1 + \frac{\sqrt{3}}{2}x_2, \end{cases}$$

con $(x'_1, x'_2) = f((x_1, x_2))$. Per provare che f è un'isometria di (\mathbb{R}^2, \cdot) è sufficiente verificare che:

$$\|f(\mathbf{x})\|^2 = (x'_1)^2 + 4(x'_2)^2 = x_1^2 + 4x_2^2 = \|\mathbf{x}\|^2,$$

con $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$. La matrice A non è un elemento di $O(2)$ perché non è associata ad una base ortonormale (rispetto al prodotto scalare introdotto), infatti $\|\mathbf{e}_2\| = 2$.

ESERCIZIO 4.6. Nello spazio vettoriale euclideo (V, \cdot) di dimensione 4, riferito alla base ortonormale $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$, sono dati i vettori:

$$\mathbf{a} = 2\mathbf{e}_1 - \mathbf{e}_3 + 2\mathbf{e}_4, \quad \mathbf{b} = \mathbf{e}_3 - \mathbf{e}_4.$$

(1) Detto \mathbf{c} il versore di \mathbf{a} , verificare che la funzione

$$f: V \rightarrow V, \quad \mathbf{x} \mapsto \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{c})\mathbf{c}$$

è un'isometria di V .

(2) Calcolare $\|f^{-1}(\mathbf{b})\|$ ed il coseno dell'angolo θ tra i vettori $f^{-1}(\mathbf{a})$ e $f^{-1}(\mathbf{b})$.

SOLUZIONE. (1) La funzione f è un endomorfismo in quanto:

$$\begin{aligned} f(\lambda\mathbf{x} + \mu\mathbf{y}) &= (\lambda\mathbf{x} + \mu\mathbf{y}) - 2((\lambda\mathbf{x} + \mu\mathbf{y}) \cdot \mathbf{c})\mathbf{c} \\ &= \lambda\mathbf{x} - 2(\lambda\mathbf{x} \cdot \mathbf{c})\mathbf{c} + \mu\mathbf{y} - 2(\mu\mathbf{y} \cdot \mathbf{c})\mathbf{c} \\ &= \lambda f(\mathbf{x}) + \mu f(\mathbf{y}), \end{aligned}$$

per ogni $\lambda, \mu \in \mathbb{R}$ e per ogni $\mathbf{x}, \mathbf{y} \in V$. Inoltre:

$$\begin{aligned} \|f(\mathbf{x})\|^2 &= f(\mathbf{x}) \cdot f(\mathbf{x}) = (\mathbf{x} - 2(\mathbf{x} \cdot \mathbf{c})\mathbf{c}) \cdot (\mathbf{x} - 2(\mathbf{x} \cdot \mathbf{c})\mathbf{c}) \\ &= \mathbf{x} \cdot \mathbf{x} - 4(\mathbf{x} \cdot \mathbf{c})(\mathbf{x} \cdot \mathbf{c}) + 4(\mathbf{x} \cdot \mathbf{c})^2(\mathbf{c} \cdot \mathbf{c}) = \|\mathbf{x}\|^2. \end{aligned}$$

(2) Essendo f un'isometria, anche f^{-1} è un'isometria. Quindi:

$$\begin{aligned} \|f^{-1}(\mathbf{b})\| &= \|\mathbf{b}\| = \sqrt{2}, \\ \cos \theta &= \cos(\widehat{f^{-1}(\mathbf{a})f^{-1}(\mathbf{b})}) = \cos(\widehat{\mathbf{a}\mathbf{b}}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} = -\frac{1}{\sqrt{2}}. \end{aligned}$$

Si lascia per esercizio la determinazione della matrice A associata a f rispetto alla base \mathcal{B} e la verifica del fatto che A sia una matrice ortogonale.

OSSERVAZIONE 4.4. Siano (V, \cdot) uno spazio vettoriale euclideo di dimensione n e \mathcal{W} un iperpiano vettoriale di V . La funzione:

$$g : V \longrightarrow V, \quad \mathbf{x} \longmapsto \begin{cases} \mathbf{x}, & \mathbf{x} \in \mathcal{W} \\ -\mathbf{x}, & \mathbf{x} \in \mathcal{W}^\perp, \end{cases}$$

che coincide con l'identità su \mathcal{W} e che associa ad ogni vettore del complemento ortogonale \mathcal{W}^\perp di \mathcal{W} il proprio opposto, è un endomorfismo di V . Infatti, si può verificare che:

$$\frac{1}{2}(\mathbf{x} + g(\mathbf{x}))$$

coincide con la proiezione ortogonale di \mathbf{x} su \mathcal{W} . L'endomorfismo g prende il nome di *simmetria ortogonale* o *riflessione* rispetto all'iperpiano vettoriale \mathcal{W} . È un esercizio verificare che g è un'isometria dello spazio vettoriale euclideo (V, \cdot) e che $g \circ g$ coincide con l'identità di V . Infine, se \mathbf{c} è un versore ortogonale a \mathcal{W} , allora:

$$g(\mathbf{x}) = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{c})\mathbf{c}.$$

Di conseguenza l'isometria f dell'esercizio precedente non è altro che la simmetria ortogonale rispetto all'iperpiano vettoriale $\mathcal{L}(\mathbf{a})^\perp$ di equazione:

$$2x_1 - x_3 + 2x_4 = 0.$$

Infine si può verificare che l'endomorfismo r_θ del piano vettoriale V_2 , considerato nell'Esempio 4.4, è una simmetria ortogonale rispetto ad una retta vettoriale. Infatti, si ha:

$$\begin{aligned} A_\theta &= \begin{pmatrix} 1 - 2\sin^2 \frac{\theta}{2} & 2\cos \frac{\theta}{2} \sin \frac{\theta}{2} \\ 2\cos \frac{\theta}{2} \sin \frac{\theta}{2} & 1 - 2\cos^2 \frac{\theta}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 2 \begin{pmatrix} \sin \frac{\theta}{2} \\ -\cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \sin \frac{\theta}{2} & -\cos \frac{\theta}{2} \end{pmatrix}. \end{aligned}$$

Automorfismi di uno spazio vettoriale euclideo che mantengano invariata la misura degli angoli tra coppie di vettori e le loro immagini, ma in generale non le norme dei vettori, sono dati dalle similitudini, di cui l'automorfismo (4.9) ne è un esempio. Più precisamente si può enunciare la seguente definizione.

DEFINIZIONE 4.3. Sia $f : V \longrightarrow V$ un automorfismo di uno spazio vettoriale euclideo (V, \cdot) , f prende il nome di *similitudine* di rapporto μ se:

$$\|f(\mathbf{x})\| = \mu\|\mathbf{x}\|, \quad \mathbf{x} \in V,$$

da cui si deduce che μ deve essere un numero reale positivo non nullo.

OSSERVAZIONE 4.5. Ogni isometria è una similitudine di rapporto 1. Mentre l'automorfismo definito da (4.9) è una similitudine di rapporto 2.

Il teorema che segue, la cui dimostrazione è lasciata per esercizio, riassume alcune tra le principali proprietà delle similitudini.

TEOREMA 4.5. *Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione n .*

- (1) *Se f è una similitudine di V , allora la misura dell'angolo individuato dai vettori \mathbf{x} e \mathbf{y} di V coincide con la misura dell'angolo individuato dai vettori $f(\mathbf{x})$ e $f(\mathbf{y})$, per ogni $\mathbf{x}, \mathbf{y} \in V$.*
- (2) *La matrice associata ad una similitudine di rapporto μ , rispetto ad una base ortonormale di V , è data dal prodotto μA con $A \in O(n)$.*

ESERCIZIO 4.7. Si determinino esplicitamente le condizioni che devono essere verificate da un automorfismo f di uno spazio euclideo V imponendo che la misura dell'angolo individuato dai vettori \mathbf{x} e \mathbf{y} di V coincide con la misura dell'angolo individuato dai vettori $f(\mathbf{x})$ e $f(\mathbf{y})$, per ogni $\mathbf{x}, \mathbf{y} \in V$.

3. Autovalori delle isometrie, similitudini

Gli autovalori delle isometrie e delle similitudini di uno spazio vettoriale euclideo (V, \cdot) , considerate nel Paragrafo 2, non possono assumere valori qualsiasi. Infatti si può provare il seguente teorema.

TEOREMA 4.6. *Sia (V, \cdot) uno spazio vettoriale euclideo,*

- (1) *gli autovalori di un'isometria f di (V, \cdot) sono 1 e -1 ;*
- (2) *gli autovalori di una similitudine di rapporto μ di (V, \cdot) sono μ e $-\mu$.*

DIMOSTRAZIONE. (1) Sia λ un autovalore di un'isometria $f : V \rightarrow V$, quindi $f(\mathbf{x}) = \lambda\mathbf{x}$. Pertanto:

$$\|f(\mathbf{x})\|^2 = \|\lambda\mathbf{x}\|^2 = \lambda^2\|\mathbf{x}\|^2 = \|\mathbf{x}\|^2,$$

da cui la tesi.

(2) La dimostrazione di (2) è analoga ed è lasciata per esercizio. \square

OSSERVAZIONE 4.6. Si osservi che se un'automorfismo di uno spazio vettoriale euclideo (V, \cdot) ha autovalori pari a 1 e -1 non è detto che sia un'isometria di (V, \cdot) . Per esempio, si consideri l'automorfismo di \mathbb{R}^2 definito dalla matrice:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

rispetto alla base canonica \mathcal{B} di \mathbb{R}^2 . Se si considera su \mathbb{R}^2 la struttura di spazio vettoriale euclideo determinata dal prodotto scalare standard, si ha per il Teorema 4.3 che f non è un'isometria di \mathbb{R}^2 (perché?).

OSSERVAZIONE 4.7. Si osservi che il Teorema 4.6 afferma che se un'isometria ammette autovalori, allora questi sono necessariamente 1 o -1 . Infatti le rotazioni di angolo θ nel piano (con $\theta \neq 0, \pi$) **non** ammettono autovalori, invece, nel caso di \mathbb{R}^3 ogni isometria ammette sempre o l'autovalore 1 o l'autovalore -1 .

ESERCIZIO 4.8. Si dimostra che una matrice di $SO(n)$, con n dispari ammette **sempre** l'autovalore $+1$. Invece una matrice ortogonale A con $\det(A) = -1$ ammette **sempre, qualunque sia il suo ordine**, autovalore

−1. Si pensi per esempio all'Osservazione 4.4 in cui chiaramente \mathcal{W} è l'autospazio relativo all'autovalore +1 e \mathcal{W}^\perp l'autospazio relativo all'autovalore −1.

Se si calcolano invece gli autovalori della matrice (4.12) si ottengono autovalori ± 1 , quindi ogni riflessione nel piano è diagonalizzabile. Ciò **non** significa che ogni matrice $A \in O(n)$ con $\det(A) = -1$ è diagonalizzabile! Si pensi per esempio alla matrice

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Procediamo ora con la dimostrazione del fatto che se $A \in O(n)$ e $\det(A) = -1$ allora $\det(A + I) = 0$, con I matrice unità di ordine n , ossia A ammette autovalore −1:

$$\begin{aligned} \det(A + I) &= (+1) \det(A + I) \\ &= -\det({}^t A) \det(A + I) \\ &= -\det({}^t A A + {}^t A) \\ &= -\det(I + {}^t A) \\ &= -\det({}^t(I + {}^t A)) \\ &= -\det(A + I). \end{aligned}$$

Sia, ora, A una matrice ortogonale, di ordine $2n + 1$, tale che $\det(A) = 1$, si vuole dimostrare che $\det(A - I) = 0$. Si ha:

$$\begin{aligned} \det(A - I) &= \det({}^t A) \det(A - I) \\ &= \det({}^t A A - {}^t A) \\ &= \det(I - {}^t A) \\ &= \det({}^t(I - {}^t A)) \\ &= \det(-(A - I)) \\ &= (-1)^{2n+1} \det(A - I) \\ &= -\det(A - I). \end{aligned}$$

4. Forme bilineari simmetriche non degeneri indefinite

Affrontiamo ora lo studio del gruppo delle matrici $A \in GL(n, \mathbb{R})$ per cui

$$B(AX, AY) = B(X, Y),$$

con B forma bilineare simmetrica reale non degenera di segnatura $(k, n - k)$ definita su uno spazio vettoriale reale V di dimensione n . Il Teorema 4.7, riportato in fondo al paragrafo, afferma che è sempre possibile determinare una base di V mediante la quale la forma bilineare B assume la seguente

espressione

$$B(X, Y) = {}^tX \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix} Y,$$

dove I_j indica la matrice unità di ordine j . Si conviene di indicare con il simbolo $I_{k, n-k}$ le matrici del tipo seguente

$$I_{k, n-k} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix}.$$

Di conseguenza le matrici A cercate verificano la relazione

$${}^tAI_{k, n-k}A = I_{k, n-k},$$

definendo il seguente sottogruppo di $GL(n, \mathbb{R})$

$$O(k, n-k) = \{A \in GL(n, \mathbb{R}) / {}^tAI_{k, n-k}A = I_{k, n-k}\}.$$

È chiaro che $O(n, 0) = O(n)$ invece $O(3, 1)$ è il **gruppo di Lorentz**. Procediamo ora con il calcolo del determinante di una matrice $A \in O(k, n-k)$, si ha

$$\det({}^tAI_{k, n-k}A) = \det(I_{k, n-k})$$

da cui segue

$$\det({}^tA) \det(A) = 1$$

ossia

$$\det(A) = \pm 1.$$

È così possibile definire, anche in questo caso, il gruppo

$$SO(k, n-k) = O(k, n-k) \cap SL(n, \mathbb{R}).$$

ESERCIZIO 4.9. Verificare che gli elementi del gruppo $O(1, 1)$ possono essere solo di questo tipo:

$$A_1 = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}, \quad A_2 = \begin{pmatrix} -\cosh t & \sinh t \\ \sinh t & -\cosh t \end{pmatrix}$$

$$A_3 = \begin{pmatrix} \cosh t & -\sinh t \\ \sinh t & -\cosh t \end{pmatrix}, \quad A_4 = \begin{pmatrix} -\cosh t & -\sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

SOLUZIONE. Si consideri una matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(1, 1).$$

Dalla definizione segue

$$\begin{cases} a^2 - c^2 = 1 \\ ab - cd = 0 \\ b^2 - d^2 = -1. \end{cases}$$

Da $a^2 = 1 + c^2$ segue che $a = \pm \cosh t$, analogamente $d = \pm \cosh t$ da cui la tesi.

TEOREMA 4.7. *Una forma quadratica Q definita su uno spazio vettoriale V , di dimensione n , ammette sempre una forma normale. Se $Q(\mathbf{x}) = {}^tXAX$, con A matrice associata a Q rispetto ad una base \mathcal{B} di V , il numero dei coefficienti pari a 1, che compare nella forma normale, è uguale al numero di autovalori positivi della matrice A , contati con la relativa molteplicità, il numero dei coefficienti pari a -1 è uguale al numero di autovalori negativi di A , contati con la relativa molteplicità, il numero dei coefficienti pari a 0 coincide con la molteplicità dell'autovalore 0 di A .*

5. Forme simplettiche reali

Si ricorda che lo spazio vettoriale $\mathbb{R}^{n,n}$ si decompone nel modo seguente

$$\mathbb{R}^{n,n} = \mathcal{S}(\mathbb{R}^{n,n}) \oplus \mathcal{A}(\mathbb{R}^{n,n}),$$

con $\mathcal{S}(\mathbb{R}^{n,n})$ sottospazio vettoriale delle matrici simmetriche e $\mathcal{A}(\mathbb{R}^{n,n})$ sottospazio vettoriale delle matrici antisimmetriche.

Nei paragrafi precedenti e nel corso di base di Algebra Lineare ci si è occupati dello studio delle forme bilineari simmetriche non degeneri, quindi delle matrici invertibili di $\mathcal{S}(\mathbb{R}^{n,n})$. In questo paragrafo ci si propone di studiare in modo più approfondito le forme bilineari associate alle matrici antisimmetriche e precisamente le forme bilineari $B : V \times V \rightarrow \mathbb{R}$ (con V spazio vettoriale reale di dimensione n) per cui

$$B(\mathbf{x}, \mathbf{y}) = -B(\mathbf{y}, \mathbf{x}) \quad \mathbf{x}, \mathbf{y} \in V.$$

Si osservi che se B è una forma bilineare antisimmetrica, allora

$$B(\mathbf{x}, \mathbf{x}) = 0, \quad \mathbf{x} \in V.$$

Fissata una base in V è facile verificare che la matrice associata ad una forma bilineare antisimmetrica è una matrice antisimmetrica.

Sia A una matrice antisimmetrica, ossia ${}^tA = -A$. Quindi

$$\det(A) = (-1)^n \det(A)$$

segue che se n è un numero dispari, allora $\det(A) = 0$. Di conseguenza le forme bilineari antisimmetriche non degeneri **possono solo essere definite su spazi vettoriali di dimensione pari.**

DEFINIZIONE 4.4. *Una forma bilineare antisimmetrica non degenera B definita su uno spazio vettoriale reale V di dimensione $2n$ prende il nome di **forma simplettica.***

ESEMPIO 4.10. La forma bilineare su \mathbb{R}^2

$$B(\mathbf{x}, \mathbf{y}) = x_2y_1 - x_1y_2, \quad \mathbf{x} = (x_1, x_2), \quad \mathbf{y} = (y_1, y_2)$$

è una forma simplettica. Si osservi che la sua matrice associata (rispetto alla base canonica di \mathbb{R}^2) è

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Anche la forma bilineare

$$B(\mathbf{x}, \mathbf{y}) = x_1y_2 - x_2y_1$$

è un esempio di forma simplettica su \mathbb{R}^2 . La sua matrice associata è

$$-J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Il seguente teorema stabilisce un risultato fondamentale per le forme simplettiche.

TEOREMA 4.8. *Sia $B : V \times V \rightarrow \mathbb{R}$ una forma simplettica con V spazio vettoriale reale di dimensione $2n$. È sempre possibile determinare una base*

$$\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n)$$

di V rispetto alla quale la forma simplettica B assume la seguente espressione

$$B(X, Y) = {}^tXJY,$$

*(con la usuale confusione di vettore con matrice colonna delle componenti). Una tale base prende il nome di **base simplettica**. Se $\mathbf{x} = (x_1, \dots, x_{2n})$, $\mathbf{y} = (y_1, \dots, y_{2n})$ rispetto a tale base, la forma simplettica B si scrive come*

$$B(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^n (x_{n+k}y_k - x_ky_{n+k}).$$

DIMOSTRAZIONE. La dimostrazione procede per induzione su n . Per spiegare i vari passi si inizia con il caso $\dim(V) = 2$.

Sia $B : V \times V \rightarrow \mathbb{R}$ ($\dim(V) = 2$) una forma simplettica. Si consideri una base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ di V . Poiché B è non degenere, esiste un vettore \mathbf{f} di V per cui $B(\mathbf{e}_1, \mathbf{f}) \neq 0$, (in caso contrario $\mathbf{e}_1 \in \ker B$.) Si pone

$$\mathbf{f}_1 = -\frac{1}{B(\mathbf{e}_1, \mathbf{f})} \mathbf{f},$$

da cui si deduce $B(\mathbf{e}_1, \mathbf{f}_1) = -1$ ossia $\mathcal{C} = (\mathbf{e}_1, \mathbf{f}_1)$ è la base simplettica cercata.

Nel caso $\dim(V) = 2n$ si inizia allo stesso modo, individuando 2 vettori linearmente indipendenti $\mathbf{e}_1, \mathbf{f}_1$ per cui $B(\mathbf{e}_1, \mathbf{f}_1) = -1$. Si procede considerando il sottospazio vettoriale $\mathcal{L}(\mathbf{e}_1, \mathbf{f}_1)^{\perp B}$ ortogonale, rispetto a B , di $\mathcal{L}(\mathbf{e}_1, \mathbf{f}_1)$, non prima di aver dimostrato i seguenti Lemmi. \square

LEMMA 4.9. *Sia B una forma simplettica definita su uno spazio vettoriale V di dimensione $2n$ e sia \mathcal{W} un sottospazio vettoriale di V allora*

$$\dim(\mathcal{W}^{\perp B}) = 2n - \dim(\mathcal{W}),$$

dove $\mathcal{W}^{\perp B}$ indica il sottospazio vettoriale ortogonale a \mathcal{W} rispetto a B .

OSSERVAZIONE 4.8. Il teorema precedente **non** afferma che \mathcal{W} e $\mathcal{W}^{\perp B}$ sono in somma diretta.

DIMOSTRAZIONE. Sia $A \in \mathbb{R}^{2n, 2n}$ la matrice associata a B . Essendo B non degenere $\det(A) \neq 0$, inoltre $B(\mathbf{x}, \mathbf{y}) = B(X, Y) = {}^tXAY$. D'altra parte

$$\mathcal{W}^{\perp B} = \{\mathbf{x} \in V / B(\mathbf{x}, \mathbf{y}) = 0, \forall \mathbf{y} \in \mathcal{W}\}.$$

La relazione ${}^tX(AY) = 0$ si può leggere come il prodotto scalare tra i vettori X e AY rispetto ad una opportuna base ortonormale di V . La matrice A (di rango massimo, ossia con determinante non nullo) si può interpretare come la matrice associata all'isomorfismo $f : V \rightarrow V$ (rispetto alla stessa base ortonormale considerata). Essendo f un isomorfismo si ha

$$\dim(f(\mathcal{W})) = \dim(\mathcal{W}),$$

quindi

$$\dim(f(\mathcal{W}))^\perp = \dim(\mathcal{W})^\perp.$$

Per teorema sul complemento ortogonale di un sottospazio vettoriale (rispetto al prodotto scalare), segue la tesi. Ossia

$$\dim(\mathcal{W}^{\perp_B}) = \dim(f(\mathcal{W})^\perp) = \dim(\mathcal{W}^\perp) = 2n - (\dim \mathcal{W})$$

dove il primo simbolo " \perp_B " indica l'ortogonale rispetto a B mentre il secondo simbolo " \perp " indica l'ortogonale rispetto al prodotto scalare introdotto. \square

OSSERVAZIONE 4.9. Si osservi che il Teorema appena dimostrato è valido per ogni forma bilineare non degenere.

OSSERVAZIONE 4.10. In generale, la somma di \mathcal{W} e di \mathcal{W}^{\perp_B} (rispetto alla forma simplettica B) non è detto che sia diretta, per esempio se $\mathcal{W} = \mathcal{L}(\mathbf{x})$ allora $\mathcal{W} \subset \mathcal{W}^{\perp_B}$.

Se si considera una base simplettica $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n, \mathbf{f}_1, \dots, \mathbf{f}_n)$ si vede facilmente che posto $\mathcal{W} = \mathcal{L}(\mathbf{e}_1, \dots, \mathbf{e}_n)$, oppure $\mathcal{W} = \mathcal{L}(\mathbf{f}_1, \dots, \mathbf{f}_n)$ allora $\mathcal{W}^{\perp_B} = \mathcal{W}$.

LEMMA 4.10. *Sia \mathcal{W} un sottospazio vettoriale di V tale che la restrizione della forma simplettica B a \mathcal{W} sia ancora non degenere, allora*

$$\mathcal{W} \oplus \mathcal{W}^{\perp_B} = V.$$

DIMOSTRAZIONE. Se $\mathbf{x} \in \mathcal{W} \cap \mathcal{W}^{\perp_B}$ allora $\mathbf{x} \in \mathcal{W}$ e $\mathbf{x} \in \mathcal{W}^{\perp_B}$, quindi

$$B(\mathbf{x}, \mathbf{y}) = 0, \quad \forall \mathbf{y} \in \mathcal{W},$$

ciò significa che \mathbf{x} appartiene al nucleo della forma simplettica ottenuta dalla restrizione di B a \mathcal{W} . Quindi $\mathbf{x} = \mathbf{o}$. \square

I due Lemmi sopra citati garantiscono l'esistenza di una base simplettica.

ESEMPIO 4.11. Sia B la forma simplettica su \mathbb{R}^4 associata alla matrice

$$A = \begin{pmatrix} 0 & 4 & 2 & 3 \\ -4 & 0 & -1 & 5 \\ -2 & 1 & 0 & -2 \\ -3 & -5 & 2 & 0 \end{pmatrix}$$

rispetto alla base standard $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ di \mathbb{R}^4 . Si vuole determinare una base simplettica (rispetto a B) di \mathbb{R}^4 . Si osserva che $\det(A) = 441$ quindi B è non degenere. Si pone

$$\mathbf{v}_1 = \mathbf{e}_1.$$

Poiché $B(\mathbf{e}_1, \mathbf{v}_3) = 2$ si pone

$$\mathbf{f}_1 = -\frac{1}{B(\mathbf{e}_1, \mathbf{v}_3)} \mathbf{v}_3 = -\frac{1}{2} \mathbf{v}_3.$$

Si considera ora $\mathcal{W} = \mathcal{L}(\mathbf{e}_1, \mathbf{f}_1)$ e si calcola $\mathcal{W}^{\perp B}$ di equazioni

$$\begin{cases} 2x_1 - x_2 + 2x_4 = 0 \\ 4x_2 + 2x_3 + 3x_4 = 0, \end{cases}$$

e con base $((1, 2, -4, 0), (-2, 0, -3, 2))$. Si pone

$$\mathbf{e}_2 = (1, 2, -4, 0).$$

Inoltre si ha $B(\mathbf{e}_2, (-2, 0, -3, 2)) = 42$. Si pone infine

$$\mathbf{f}_2 = -\frac{1}{42}(-2, 0, -3, 2).$$

La base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{f}_1, \mathbf{f}_2)$ è la base simplettica richiesta.

Sia B una forma simplettica, le matrici $A \in \mathbb{R}^{2n, 2n}$ che lasciano invariante tale forma simplettica, cioè

$$B(A\mathbf{X}, A\mathbf{Y}) = B(\mathbf{X}, \mathbf{Y}),$$

sono dunque le matrici A per cui

$${}^t A J A = J.$$

Si definisce così il gruppo **simplettico reale**

$$Sp(2n, \mathbb{R}) = \{A \in GL(2n, \mathbb{R}) / {}^t A J A = J\},$$

i cui elementi prendono il nome di **matrici simplettiche**.

ESEMPIO 4.12. Si dimostra facilmente che $Sp(2, \mathbb{R}) = SL(2, \mathbb{R})$ (**esercizio!**)

Si procede ora con il calcolo del determinante di una matrice simplettica. Immediatamente si ottiene

$$\det({}^t A J A) = \det({}^t A) \det(J) \det(A) = \det(A)^2 \det(J) = \det(J),$$

da cui segue subito

$$\det(A) = \pm 1.$$

In realtà vale il seguente teorema di dimostrazione non banale, per approfondimenti si veda per esempio [9].

TEOREMA 4.11. *Le matrici simplettiche hanno necessariamente determinante pari a 1.*

6. Matrici hermitiane e anti-hermitiane

Si introducono ora due sottoinsiemi dello spazio vettoriale delle matrici quadrate complesse $\mathbb{C}^{n, n}$, che sono l'analogo, nel caso di $\mathbb{R}^{n, n}$, del sottospazio vettoriale delle matrici simmetriche $\mathcal{S}(\mathbb{R}^{n, n})$ e di quello delle matrici anti-simmetriche $\mathcal{A}(\mathbb{R}^{n, n})$. Precisamente, si definiscono l'insieme delle *matrici hermitiane*:

$$\mathcal{H}(\mathbb{C}^{n, n}) = \{A \in \mathbb{C}^{n, n} / {}^t A = \overline{A}\}$$

e l'insieme delle *matrici anti-hermitiane*:

$$\mathcal{AH}(\mathbb{C}^{n, n}) = \{A \in \mathbb{C}^{n, n} / {}^t A = -\overline{A}\}.$$

Chiaramente una matrice reale simmetrica $A \in \mathcal{S}(\mathbb{R}^{n,n})$ è hermitiana ed una matrice reale antisimmetrica $A \in \mathcal{A}(\mathbb{R}^{n,n})$ è anti-hermitiana. Ad esempio la matrice:

$$\begin{pmatrix} 2 & 1+i \\ 1-i & -3 \end{pmatrix}$$

è hermitiana. Si osservi che gli elementi sulla sua diagonale principale sono reali. Questa proprietà non è casuale, infatti gli elementi sulla diagonale principale di una matrice hermitiana A sono necessariamente reali, in quanto devono coincidere, per definizione, con i propri coniugati. Si può verificare che la somma di due matrici hermitiane è hermitiana e l'inversa di una matrice hermitiana invertibile è hermitiana. Analoghe proprietà valgono per le matrici anti-hermitane. Come per le matrici reali simmetriche, la matrice prodotto AB di due matrici hermitiane A e B è hermitiana solo se A e B commutano, cioè se e solo se $AB = BA$.

ESERCIZIO 4.13. Si verifichi che l'insieme delle matrici hermitiane $\mathcal{H}(\mathbb{C}^{n,n})$, con le operazioni di somma e prodotto per un numero reale, è uno spazio vettoriale su \mathbb{R} di dimensione n^2 . Inoltre, si provi invece che $\mathcal{H}(\mathbb{C}^{n,n})$, con le operazioni di somma e prodotto per un numero complesso, non è un sottospazio vettoriale complesso di $\mathbb{C}^{n,n}$.

SOLUZIONE Siano A e B due matrici hermitiane, ossia ${}^tA = \overline{A}$ e ${}^tB = \overline{B}$, si tratta di dimostrare che la matrice $A + B$ è hermitiana, infatti:

$${}^t(A + B) = {}^tA + {}^tB = \overline{A} + \overline{B} = \overline{(A + B)}.$$

Sia $\lambda \in \mathbb{R}$, si deve dimostrare che λA è una matrice hermitiana se $A \in \mathcal{H}(\mathbb{C}^{n,n})$, infatti:

$${}^t(\lambda A) = \lambda {}^tA = \lambda \overline{A} = \overline{(\lambda A)},$$

in quanto $\lambda = \overline{\lambda}$.

È evidente che questa proprietà è falsa se $\lambda \in \mathbb{C}$, quindi $\mathcal{H}(\mathbb{C}^{n,n})$ è uno spazio vettoriale *reale* ma *non* è un sottospazio vettoriale complesso di $\mathbb{C}^{n,n}$ inteso come spazio vettoriale complesso.

Come già accennato in precedenza, una matrice hermitiana A è del tipo:

$$A = \begin{pmatrix} a_{11} & a_{12} + ib_{12} & \dots & a_{1n} + ib_{1n} \\ a_{12} - ib_{12} & a_{22} & \dots & a_{2n} + ib_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} - ib_{1n} & a_{2n} - ib_{2n} & \dots & a_{nn} \end{pmatrix},$$

dove a_{hk}, b_{hk} , $h, k = 1, 2, \dots, n$ sono numeri reali. È quindi un esercizio dimostrare che una base di $\mathcal{H}(\mathbb{C}^{n,n})$.

$$\begin{aligned}
& \left(\begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \dots & \dots & 0 \\ 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix}, \right. \\
& \left. \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}, \right. \\
& \left. \begin{pmatrix} 0 & i & \dots & \dots & 0 \\ -i & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & \dots & i \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ -i & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & i \\ 0 & \dots & \dots & -i & 0 \end{pmatrix} \right).
\end{aligned}$$

Da questo segue che:

$$\dim(\mathcal{H}(\mathbb{C}^{n,n})) = \frac{n(n+1)}{2} + \frac{n(n-1)}{2} = n^2.$$

ESERCIZIO 4.14. Si verifichi che l'insieme delle matrici anti-hermitiane $\mathcal{AH}(\mathbb{C}^{n,n})$, con le operazioni di somma e prodotto per un numero reale, è uno spazio vettoriale su \mathbb{R} di dimensione n^2 . Inoltre, si provi invece che $\mathcal{AH}(\mathbb{C}^{n,n})$, con le operazioni di somma e prodotto per un numero complesso, non è un sottospazio vettoriale complesso di $\mathbb{C}^{n,n}$.

SOLUZIONE In modo analogo all'esercizio precedente si dimostra che $\mathcal{AH}(\mathbb{C}^{n,n})$ è un sottospazio vettoriale reale di $\mathbb{C}^{n,n}$ ma non è un sottospazio vettoriale complesso.

È facile verificare, a partire dalla definizione, che una matrice anti-hermitiana A è del tipo:

$$A = \begin{pmatrix} ib_{11} & a_{12} + ib_{12} & \dots & a_{1n} + ib_{1n} \\ -a_{12} + ib_{12} & ib_{22} & \dots & a_{2n} + ib_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} + ib_{1n} & -a_{2n} + ib_{2n} & \dots & ib_{nn} \end{pmatrix},$$

dove a_{hk}, b_{hk} , $h, k = 1, 2, \dots, n$ sono numeri reali. È quindi un esercizio dimostrare che una base di $\mathcal{AH}(\mathbb{C}^{n,n})$ è

$$\begin{aligned}
& \left(\left(\begin{pmatrix} 0 & 1 & \dots & \dots & 0 \\ -1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ -1 & 0 & \dots & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & -1 & 0 \end{pmatrix} \right) \\
& \left(\begin{pmatrix} i & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & i & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 0 \\ 0 & \dots & \dots & 0 & i \end{pmatrix}, \right. \\
& \left. \begin{pmatrix} 0 & i & \dots & \dots & 0 \\ i & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & \dots & i \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ i & 0 & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & i \\ 0 & \dots & \dots & i & 0 \end{pmatrix} \right).
\end{aligned}$$

Da questo segue che $\dim(\mathcal{AH}(\mathbb{C}^{n,n})) = n^2$. Si osservi, inoltre, che una matrice $A \in \mathbb{C}^{n,n}$ è hermitiana se e solo se iA è anti-hermitiana.

È valido, in campo complesso, il seguente teorema, analogo a quello dimostrato nel caso delle matrici quadrate ad elementi reali.

TEOREMA 4.12. *Lo spazio vettoriale reale $\mathbb{C}^{n,n}$ si decompone nel modo seguente:*

$$\mathbb{C}^{n,n} = \mathcal{H}(\mathbb{C}^{n,n}) \oplus \mathcal{AH}(\mathbb{C}^{n,n}).$$

DIMOSTRAZIONE. Si procede come nella dimostrazione del teorema di decomposizione di $\mathbb{R}^{n,n}$ nella somma diretta dei sottospazi vettoriali delle matrici simmetriche e antisimmetriche, tenendo conto che ogni matrice $A \in \mathbb{C}^{n,n}$ si decompone come:

$$A = \frac{1}{2}(A + \bar{t}A) + \frac{1}{2}(A - \bar{t}A)$$

e che $(A + \bar{t}A) \in \mathcal{H}(\mathbb{C}^{n,n})$ e $(A - \bar{t}A) \in \mathcal{AH}(\mathbb{C}^{n,n})$. Inoltre, è facile dimostrare che $\mathcal{H}(\mathbb{C}^{n,n}) \cap \mathcal{AH}(\mathbb{C}^{n,n}) = \{O\}$ con O matrice nulla di $\mathbb{C}^{n,n}$. \square

OSSERVAZIONE 4.11. In letteratura è spesso usata la notazione:

$$A^* = \bar{t}A,$$

pertanto una matrice A è hermitiana se e solo se $A = A^*$ ed è anti-hermitiana se e solo se $A = -A^*$.

Per le matrici A^* valgono le seguenti proprietà la cui dimostrazione è lasciata per esercizio.

- (1) $(A^*)^* = A$,
- (2) $(A + B)^* = A^* + B^*$,
- (3) $(\lambda A)^* = \bar{\lambda} A^*$,
- (4) $(AB)^* = B^* A^*$,

per ogni $A, B \in \mathbb{C}^{n,n}$ e per ogni $\lambda \in \mathbb{C}$.

ESERCIZIO 4.15. Dimostrare che la circonferenza $S^1(1)$ del piano è un gruppo isomorfo a $SO(2)$ tramite l'omomorfismo di gruppi

$$e^{i\theta} = \cos \theta + i \sin \theta \mapsto \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

7. Spazi vettoriali hermitiani

L'operazione di prodotto scalare introdotta su uno spazio vettoriale reale può essere estesa mediante la seguente definizione agli spazi vettoriali complessi.

DEFINIZIONE 4.5. Sia V uno spazio vettoriale su \mathbb{C} . Si definisce prodotto hermitiano su V la funzione:

$$\cdot: V \times V \longrightarrow \mathbb{C}, \quad (\mathbf{x}, \mathbf{y}) \longmapsto \mathbf{x} \cdot \mathbf{y},$$

per cui valgano le seguenti proprietà:

- (1) $\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}} \quad \mathbf{x}, \mathbf{y} \in V$;
- (2) $(\mathbf{x}_1 + \mathbf{x}_2) \cdot \mathbf{y} = \mathbf{x}_1 \cdot \mathbf{y} + \mathbf{x}_2 \cdot \mathbf{y}, \quad \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$;
- (3) $(\lambda \mathbf{x}) \cdot \mathbf{y} = \lambda(\mathbf{x} \cdot \mathbf{y}), \quad \lambda \in \mathbb{C}, \mathbf{x}, \mathbf{y} \in V$;
- (4) $\mathbf{x} \cdot \mathbf{x} \geq 0, \mathbf{x} \in V \quad \text{e} \quad \mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{o},$

dove $\overline{\mathbf{y} \cdot \mathbf{x}}$ indica il complesso coniugato di $\mathbf{y} \cdot \mathbf{x}$.

Uno spazio vettoriale complesso V su cui è definito un prodotto hermitiano “ \cdot ” prende il nome di *spazio vettoriale hermitiano* o di *spazio vettoriale euclideo complesso* e si indica, in generale, con la scrittura (V, \cdot) .

OSSERVAZIONE 4.12. (1) Come conseguenza delle Proprietà (1) e (3) della Definizione 4.5 si ha:

$$\mathbf{x} \cdot (\lambda \mathbf{y}) = \overline{(\lambda \mathbf{y}) \cdot \mathbf{x}} = \overline{\lambda(\mathbf{y} \cdot \mathbf{x})} = \bar{\lambda} \overline{(\mathbf{y} \cdot \mathbf{x})} = \bar{\lambda}(\mathbf{x} \cdot \mathbf{y}), \quad \lambda \in \mathbb{C}, \mathbf{x}, \mathbf{y} \in V.$$

Inoltre, dalle Proprietà 1., 2. e 3 della Definizione 4.5 si ottiene:

$$\mathbf{x} \cdot (\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2) = \bar{\lambda}_1(\mathbf{x} \cdot \mathbf{y}_1) + \bar{\lambda}_2(\mathbf{x} \cdot \mathbf{y}_2), \quad \lambda_1, \lambda_2 \in \mathbb{C}, \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in V. \quad (4.13)$$

- (2) La Proprietà (4) della Definizione 4.5 ha senso in quanto $\mathbf{x} \cdot \mathbf{x} = \overline{\mathbf{x} \cdot \mathbf{x}}$ e pertanto $\mathbf{x} \cdot \mathbf{x}$ è un numero reale.

Nel caso degli spazi vettoriali hermitiani valgono esempi e teoremi analoghi a quelli già dimostrati per gli spazi vettoriali euclidei.

ESEMPIO 4.16. Sullo spazio vettoriale complesso \mathbb{C}^n si può definire un prodotto hermitiano che, se ristretto a \mathbb{R}^n , coincide con il prodotto scalare standard. Precisamente si pone:

$$(x_1, x_2, \dots, x_n) \cdot (y_1, y_2, \dots, y_n) = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n = \sum_{i=1}^n x_i \bar{y}_i, \quad (4.14)$$

per ogni $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in \mathbb{C}^n$. Con la notazione matriciale:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

la (4.14) si scrive come:

$$X \cdot Y = {}^t X \bar{Y},$$

dove \bar{Y} indica la matrice complessa coniugata di Y . È un esercizio dimostrare che (4.14) è un prodotto hermitiano che, spesso, prende il nome di *prodotto hermitiano standard su \mathbb{C}^n* , che viene così dotato della struttura di spazio vettoriale hermitiano. È ancora un esercizio dimostrare che:

$$X \cdot Y = {}^t \bar{X} Y$$

è un altro prodotto hermitiano su \mathbb{C}^n . I due prodotti hermitiani sono l'uno il coniugato dell'altro ed entrambi coincidono con il prodotto scalare standard di \mathbb{R}^n , quando li si riferisce a numeri reali.

Ogni spazio vettoriale complesso, di dimensione finita, ammette almeno un prodotto hermitiano, vale infatti il seguente teorema.

TEOREMA 4.13. *Sia V uno spazio vettoriale complesso di dimensione n . Data una base $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ di V , la funzione:*

$$\cdot : V \times V \longrightarrow \mathbb{C},$$

definita da:

$$\mathbf{x} \cdot \mathbf{y} = {}^t X \bar{Y},$$

dove:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

e (x_1, x_2, \dots, x_n) e (y_1, y_2, \dots, y_n) sono le componenti rispettivamente di \mathbf{x} e \mathbf{y} rispetto alla base \mathcal{B} , è un prodotto hermitiano su V .

La dimostrazione è un esercizio.

La proprietà (4) della Definizione 4.5 permette, come per gli spazi vettoriali euclidei, di definire la norma di un vettore anche nel caso di uno spazio vettoriale hermitiano (V, \cdot) . Infatti si definisce *norma* di un vettore \mathbf{x} di V il numero reale non negativo dato da:

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

ed è lasciata al Lettore la verifica della proprietà:

$$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|, \quad \mathbf{x} \in V, \lambda \in \mathbb{C},$$

dove $|\lambda|$ indica il modulo del numero complesso λ .

ESEMPIO 4.17. Nello spazio vettoriale hermitiano (\mathbb{C}^n, \cdot) con il prodotto hermitiano standard (4.14), la norma di \mathbf{x} è data da:

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i \bar{x}_i} = \sqrt{\sum_{i=1}^n |x_i|^2},$$

dove $|x_i|$ indica il modulo del numero complesso x_i , $i = 1, 2, \dots, n$.

Continuano a valere, anche nel caso degli spazi vettoriali hermitiani, la disuguaglianza di Cauchy–Schwarz e la disuguaglianza triangolare, anche se la loro dimostrazione è più laboriosa rispetto a quella del caso euclideo. Precisamente, vale il teorema seguente.

TEOREMA 4.14. *Su uno spazio vettoriale hermitiano (V, \cdot) , la funzione:*

$$\|\cdot\|: V \longrightarrow \mathbb{R}, \quad \mathbf{x} \longmapsto \|\mathbf{x}\|$$

verifica le seguenti proprietà:

- (1) *Teorema di Pitagora:* $\mathbf{x} \cdot \mathbf{y} = 0 \implies \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.
- (2) *Disuguaglianza di Cauchy–Schwarz:* $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$, $\mathbf{x}, \mathbf{y} \in V$, dove $|\mathbf{x} \cdot \mathbf{y}|$ indica il modulo del numero complesso $\mathbf{x} \cdot \mathbf{y}$.
- (3) *Disuguaglianza triangolare (o di Minkowski):* $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, $\mathbf{x}, \mathbf{y} \in V$.

DIMOSTRAZIONE. (1) Segue da:

$$\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + \mathbf{x} \cdot \mathbf{y} + \overline{\mathbf{x} \cdot \mathbf{y}} + \|\mathbf{y}\|^2.$$

- (2) La disuguaglianza di Cauchy–Schwarz è banalmente verificata se uno dei due vettori coincide con il vettore nullo. Si supponga, quindi, che \mathbf{x} e \mathbf{y} siano entrambi diversi dal vettore nullo. Si inizia la dimostrazione nel caso particolare $\|\mathbf{y}\| = 1$. Si ha:

$$\begin{aligned} 0 &\leq \|\mathbf{x} - (\mathbf{x} \cdot \mathbf{y}) \mathbf{y}\|^2 = (\mathbf{x} - (\mathbf{x} \cdot \mathbf{y}) \mathbf{y}) \cdot (\mathbf{x} - (\mathbf{x} \cdot \mathbf{y}) \mathbf{y}) \\ &= \|\mathbf{x}\|^2 - \overline{(\mathbf{x} \cdot \mathbf{y})}(\mathbf{x} \cdot \mathbf{y}) - (\mathbf{x} \cdot \mathbf{y})(\mathbf{y} \cdot \mathbf{x}) \\ &\quad + (\mathbf{x} \cdot \mathbf{y})\overline{(\mathbf{x} \cdot \mathbf{y})}\|\mathbf{y}\|^2 \\ &= \|\mathbf{x}\|^2 - 2|\mathbf{x} \cdot \mathbf{y}|^2 + |\mathbf{x} \cdot \mathbf{y}|^2\|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - |\mathbf{x} \cdot \mathbf{y}|^2. \end{aligned}$$

Quindi $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|$ se $\|\mathbf{y}\| = 1$. Se $\mathbf{y} \neq \mathbf{o}$, si può considerare il versore di \mathbf{y} . Applicando allora la precedente disuguaglianza si ottiene:

$$\left| \mathbf{x} \cdot \frac{\mathbf{y}}{\|\mathbf{y}\|} \right| \leq \|\mathbf{x}\|,$$

da cui la tesi.

3. La dimostrazione è analoga a quella vista nel caso di uno spazio vettoriale euclideo se si tiene conto che:

$$\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{x} = \mathbf{x} \cdot \mathbf{y} + \overline{\mathbf{x} \cdot \mathbf{y}} \leq 2|\mathbf{x} \cdot \mathbf{y}|. \quad (4.15)$$

Infatti, dato un numero complesso $z = a + ib$, con $a, b \in \mathbb{R}$ e $a = \operatorname{Re}(z)$, dove $\operatorname{Re}(z)$ indica la parte reale di z , si ha:

$$|z|^2 = z\bar{z} = a^2 + b^2 \geq a^2 = (\operatorname{Re}(z))^2.$$

D'altra parte $2\operatorname{Re}(z) = z + \bar{z}$ da cui:

$$2\operatorname{Re}(z) = z + \bar{z} \leq 2|z|,$$

ossia la (4.15). □

OSSERVAZIONE 4.13. Si osservi che, a differenza del caso reale, l'equazione $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ non esprime una condizione necessaria e sufficiente per l'ortogonalità di due vettori.

Considerati due vettori \mathbf{x} e \mathbf{y} non nulli di uno spazio vettoriale hermitiano (V, \cdot) , dalla disuguaglianza di Cauchy–Schwarz segue che:

$$\frac{|\mathbf{x} \cdot \mathbf{y}|}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1,$$

ma ciò significa solo che il modulo del numero complesso $(\mathbf{x} \cdot \mathbf{y})/(\|\mathbf{x}\| \|\mathbf{y}\|)$ è minore o uguale a 1. Pertanto non è possibile nel caso degli spazi vettoriali hermitiani introdurre il concetto di angolo tra due vettori nello stesso modo in cui è stato definito per gli spazi vettoriali euclidei. Nonostante ciò, sugli spazi vettoriali hermitiani, come nel caso reale, si può introdurre il concetto di ortogonalità. Più precisamente, due vettori \mathbf{x} e \mathbf{y} di uno spazio vettoriale hermitiano (V, \cdot) si dicono *ortogonali* se $\mathbf{x} \cdot \mathbf{y} = 0$. Di conseguenza, è valida anche in questo caso la definizione di base ortogonale e ortonormale e continua a valere anche per uno spazio vettoriale hermitiano il Lemma 4.15 riportato di seguito. Una base ortonormale su uno spazio vettoriale hermitiano è in generale chiamata *base unitaria*.

LEMMA 4.15. *Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione n . Un insieme \mathcal{I} di k vettori di V :*

$$\mathcal{I} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$$

tale che:

- (1) $k \leq n$;
- (2) $\mathbf{v}_i \neq \mathbf{o}$, $i = 1, 2, \dots, k$;
- (3) $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, $i \neq j$, $i, j = 1, 2, \dots, k$

è un insieme libero. Se $k = n$ allora \mathcal{I} è una base ortogonale di V .

ESEMPIO 4.18. La base canonica $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ di \mathbb{C}^n è una base unitaria dello spazio vettoriale hermitiano (\mathbb{C}^n, \cdot) con il prodotto hermitiano standard (4.14).

Inoltre, analogamente a quanto dimostrato per uno spazio vettoriale euclideo, in uno spazio vettoriale hermitiano (V, \cdot) utilizzando il processo di ortonormalizzazione di Gram–Schmidt (che vale anche nel caso complesso) si può costruire una base unitaria a partire da una base di V . La dimostrazione è analoga a quella vista per il Teorema 4.16 ricordato di seguito, facendo attenzione al fatto che, a differenza del caso reale, il prodotto hermitiano non è simmetrico, ossia *non* vale la relazione $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$, $\forall \mathbf{x}, \mathbf{y} \in V$, ma vale la proprietà 1. della Definizione 4.5.

TEOREMA 4.16. *Sia (V, \cdot) uno spazio vettoriale euclideo di dimensione n e sia:*

$$\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$$

una sua base. A partire da \mathcal{B} è possibile costruire una base ortonormale:

$$\mathcal{B}' = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$$

di V tale che:

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k), \quad k = 1, 2, \dots, n.$$

OSSERVAZIONE 4.14. La disuguaglianza di Cauchy–Schwarz può essere dimostrata in molti modi diversi. Si propone di seguito una seconda versione della dimostrazione precedente, che, anche se più lunga, è interessante per la sua interpretazione geometrica. Infatti, sorprendentemente, anche se in uno spazio vettoriale hermitiano non è definito l'angolo tra due vettori, il prodotto hermitiano permette di introdurre la nozione di proiezione ortogonale di un vettore su un altro vettore. Sia \mathbf{x} un vettore non nullo, si definisce il vettore:

$$\mathbf{w} = \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y},$$

che, nel caso di uno spazio vettoriale euclideo, rappresenterebbe un vettore ortogonale a \mathbf{y} in quanto \mathbf{w} è la differenza di \mathbf{x} con la sua proiezione ortogonale su \mathbf{y} . Anche in questo caso si dimostra che $\mathbf{w} \cdot \mathbf{y} = 0$, infatti:

$$\mathbf{w} \cdot \mathbf{y} = \left(\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y} \right) \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y} \cdot \mathbf{y} = 0.$$

Dal fatto che $\|\mathbf{w}\|^2 \geq 0$ si ha:

$$0 \leq \mathbf{w} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{x} = \left(\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y} \right) \cdot \mathbf{x} = \frac{\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - |\mathbf{y} \cdot \mathbf{x}|^2}{\|\mathbf{y}\|^2},$$

ossia la tesi.

È noto che le matrici ortogonali esprimono il cambiamento di base tra coppie di basi ortonormali di uno spazio vettoriale euclideo. Un risultato analogo è valido in uno spazio vettoriale hermitiano, con la differenza che la matrice P del cambiamento di base deve verificare la relazione:

$$\overline{tP} P = I.$$

Si può allora introdurre l'insieme delle *matrici unitarie* di $\mathbb{C}^{n,n}$ definito da:

$$U(n) = \{P \in \mathbb{C}^{n,n} / \overline{tP} P = I\},$$

che è l'analogo, nel caso complesso, dell'insieme $O(n)$ delle matrici ortogonali.

Si osservi che dalla relazione $\overline{tP} P = I$ si ottiene $tP \overline{P} = I$ e $P^{-1} = \overline{tP}$. Inoltre, una matrice unitaria ad elementi reali è ovviamente ortogonale ed una matrice ortogonale, considerata come matrice complessa, è unitaria. Per le matrici unitarie valgono le seguenti proprietà.

- TEOREMA 4.17. (1) Il prodotto di due matrici unitarie è una matrice unitaria.
- (2) La matrice identità I è unitaria.
- (3) L'inversa P^{-1} di una matrice unitaria P è unitaria.
- (4) La trasposta tP di una matrice unitaria P è unitaria.
- (5) Una matrice $P \in \mathbb{C}^{n,n}$ è unitaria se e solo se le righe e le colonne di P sono le componenti di una base ortonormale di \mathbb{C}^n , rispetto al prodotto hermitiano standard.
- (6) Il determinante di una matrice unitaria P è un numero complesso di modulo 1.
- (7) Gli autovalori di una trasformazione unitaria hanno modulo pari a 1.

La dimostrazione, analoga a quella vista per il Teorema 4.1, è lasciata al Lettore per esercizio.

OSSERVAZIONE 4.15. Segue dalle proprietà (1), (2), (3) del Teorema 4.17 che l'insieme delle matrici unitarie $U(n)$ è un gruppo rispetto al prodotto di matrici. $U(n)$ prende il nome di **gruppo unitario**.

ESEMPIO 4.19. Le matrici unitarie di ordine 1 sono facilmente individuabili, infatti si ha:

$$U(1) = \{z \in \mathbb{C} / z\bar{z} = 1\},$$

si tratta, quindi, dei numeri complessi di modulo 1. ed usando l'identificazione tra i numeri complessi ed il punti del piano data da:

$$z = x + iy = (x, y), \quad z \in \mathbb{C},$$

segue che gli elementi di $U(1)$ costituiscono la circonferenza nel piano di centro l'origine e raggio 1. Si ottiene di conseguenza l'identificazione tra $U(1)$ e $SO(2)$.

ESEMPIO 4.20. Si intende ora mettere in relazione la sfera $S^3(1)$ con il gruppo di matrici:

$$SU(2) = \{P \in \mathbb{C}^{2,2} / {}^tP\bar{P} = I, \det(P) = 1\}$$

e dimostrare che i due insiemi in realtà coincidono. In altri termini anche il gruppo $SU(2)$ è identificabile ad una sfera opportuna. Inoltre, per ciò che è stato visto nel Capitolo 3, si ottiene di conseguenza l'identificazione tra $SU(2)$ e $SL(1, \mathbb{H})$.

Sia:

$$P = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}, \quad z_1, z_2, z_3, z_4 \in \mathbb{C},$$

una generica matrice appartenente ad $SU(2)$. La condizione:

$${}^tP\bar{P} = \begin{pmatrix} z_1 & z_3 \\ z_2 & z_4 \end{pmatrix} \begin{pmatrix} \bar{z}_1 & \bar{z}_2 \\ \bar{z}_3 & \bar{z}_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

si traduce nel sistema:

$$\begin{cases} z_1\bar{z}_1 + z_3\bar{z}_3 = 1 \\ z_1\bar{z}_2 + z_3\bar{z}_4 = 0 \\ z_2\bar{z}_2 + z_4\bar{z}_4 = 1. \end{cases} \quad (4.16)$$

Supponendo che tutti gli elementi della matrice P non siano nulli (si lasciano per esercizio i vari casi particolari), dalla seconda equazione di (4.16) si ricava:

$$z_1 = -\frac{z_3\bar{z}_4}{\bar{z}_2} \quad (4.17)$$

e:

$$z_3 = -\frac{z_1\bar{z}_2}{\bar{z}_4}. \quad (4.18)$$

Sostituendo (4.17) nella condizione ulteriore:

$$\det(P) = z_1z_4 - z_2z_3 = 1 \quad (4.19)$$

segue:

$$z_3 = -\bar{z}_2.$$

Sostituendo (4.18) in (4.19) si ha anche:

$$z_1 = \overline{z_4}.$$

Di conseguenza si ottiene che le generiche matrici P di $SU(2)$ devono essere del tipo:

$$P = \begin{pmatrix} z_1 & z_2 \\ -\overline{z_2} & \overline{z_1} \end{pmatrix}, \quad (4.20)$$

con $\det(P) = z_1 \overline{z_1} + z_2 \overline{z_2} = 1$. Ponendo:

$$z_1 = x + iy, \quad z_2 = z + it, \quad x, y, z, t \in \mathbb{R},$$

e sostituendo nella condizione $\det(P) = 1$ si ottiene:

$$z_1 \overline{z_1} + z_2 \overline{z_2} = x^2 + y^2 + z^2 + t^2 = 1,$$

pervenendo così all'identificazione di $SU(2)$ con la sfera $S^3(1)$.

Siano (V, \cdot) uno spazio vettoriale hermitiano di dimensione n e \mathcal{W} un suo sottospazio vettoriale di dimensione $k \leq n$. Come nel caso degli spazi vettoriali euclidei si può definire il complemento ortogonale \mathcal{W}^\perp di \mathcal{W} , rispetto al prodotto scalare hermitiano, per cui valgono le stesse proprietà del caso reale.

Si conclude l'argomento osservando che la norma di un vettore individua il prodotto scalare o hermitiano che la definisce. Infatti è un facile esercizio dimostrare che in uno spazio vettoriale euclideo reale V vale la formula:

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2} (\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2), \quad \mathbf{x}, \mathbf{y} \in V.$$

Più complicata la relazione analoga che vale nel caso di uno spazio vettoriale hermitiano:

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= \frac{1}{2} (\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2) \\ &+ \frac{i}{2} (\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2), \quad \mathbf{x}, \mathbf{y} \in V. \end{aligned} \quad (4.21)$$

OSSERVAZIONE 4.16. Il gruppo unitario $U(n)$ è quindi formato da tutte le matrici P che mantengono inalterato il valore del prodotto hermitiano su coppie di vettori, al variare della base nello spazio vettoriale hermitiano V , in altri termini, se

$$H : V \times V \longrightarrow \mathbb{C}$$

indica un prodotto hermitiano su V , scritto, rispetto ad una base unitaria, nel modo standard, allora le matrici $A \in GL(n, \mathbb{C})$ per cui

$$H(A\mathbf{X}, A\mathbf{Y}) = H(\mathbf{X}, \mathbf{Y}),$$

coincidono con le matrici unitarie.

8. Trasformazioni unitarie

Analogamente al caso reale, un endomorfismo f di uno spazio vettoriale hermitiano (V, \cdot) si dice una *trasformazione unitaria* o *operatore unitario* o *isometria complessa* se non cambia il prodotto hermitiano, cioè se vale la relazione (4.3). Come nel caso delle isometrie si può dimostrare il teorema che segue.

TEOREMA 4.18. *Sia (V, \cdot) uno spazio vettoriale hermitiano di dimensione n .*

- (1) *Un endomorfismo $f : V \rightarrow V$ è una trasformazione unitaria se e solo se:*

$$\|f(\mathbf{x})\| = \|\mathbf{x}\|, \quad \mathbf{x} \in V.$$

- (2) *La composizione di due trasformazioni unitarie è una trasformazione unitaria.*
 (3) *L'inversa di una trasformazione unitaria è una trasformazione unitaria.*
 (4) *Un endomorfismo f di V è una trasformazione unitaria di V se e solo se le immagini dei vettori di una base unitaria di V formano una base unitaria di V .*
 (5) *Un endomorfismo f di V è una trasformazione unitaria di V se e solo se la matrice associata ad f , rispetto ad una base unitaria di V , è una matrice unitaria.*

La dimostrazione del Teorema 4.18 è analoga a quella del Teorema 4.3 tenendo però conto che in questo caso vale la relazione (4.21).

OSSERVAZIONE 4.17. Come per le isometrie, dai punti 2. e 3. del Teorema 4.18 segue che l'insieme delle trasformazioni unitarie di uno spazio vettoriale hermitiano (V, \cdot) è un gruppo rispetto alla composizione di funzioni. Inoltre, fissata una base ortonormale \mathcal{B} nello spazio vettoriale hermitiano (V, \cdot) per il punto 5. dello stesso teorema si può stabilire, in modo naturale, un isomorfismo tra l'insieme delle trasformazioni unitarie di (V, \cdot) ed il gruppo unitario $U(n)$ delle matrici unitarie (cfr. Oss. 4.15) associando ad ogni trasformazione unitaria f di V la matrice unitaria $A = M^{\mathcal{B}, \mathcal{B}}(f)$.

9. Il teorema spettrale

Si ricordi che un endomorfismo f di uno spazio hermitiano (V, \cdot) si dice autoaggiunto se

$$f(\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot f(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in V.$$

TEOREMA 4.19. (1) *Gli autovalori di un endomorfismo autoaggiunto f di uno spazio vettoriale hermitiano (V, \cdot) sono reali.*

- (2) *Autovettori di un endomorfismo autoaggiunto f di uno spazio vettoriale hermitiano (V, \cdot) relativi ad autovalori diversi sono ortogonali.*
 (3) *Sia f un endomorfismo autoaggiunto di uno spazio vettoriale hermitiano (V, \cdot) di dimensione n , allora f è diagonalizzabile. Inoltre, esiste una base unitaria di V formata da autovettori di f .*
 (4) *Ogni matrice hermitiana è diagonalizzabile mediante una matrice unitaria.*

(5) Ogni endomorfismo f di uno spazio vettoriale complesso è autoaggiunto rispetto ad un opportuno prodotto hermitiano.

DIMOSTRAZIONE. Per dimostrare (1) si può osservare che se λ è un autovalore dell'endomorfismo autoaggiunto f e \mathbf{x} è un autovettore relativo a λ , si ha:

$$f(\mathbf{x}) \cdot \mathbf{x} = \lambda \mathbf{x} \cdot \mathbf{x}. \quad (4.22)$$

Poiché f è autoaggiunto risulta anche:

$$f(\mathbf{x}) \cdot \mathbf{x} = \mathbf{x} \cdot f(\mathbf{x}) = \mathbf{x} \cdot (\lambda \mathbf{x}) = \bar{\lambda} \mathbf{x} \cdot \mathbf{x}, \quad (4.23)$$

dove $\bar{\lambda}$ indica il complesso coniugato di λ . Essendo \mathbf{x} è un autovettore, segue $\mathbf{x} \cdot \mathbf{x} \neq 0$ e pertanto, confrontando (4.22) con (4.23) si ottiene $\lambda = \bar{\lambda}$. Le dimostrazioni delle proprietà (2), (3), (4) e (5) sono analoghe al caso reale. \square

OSSERVAZIONE 4.18. Come conseguenza del Teorema 4.19 si ottiene la dimostrazione del Lemma 4.20, riportato di seguito, in quanto dalla proprietà 1. si ha che ogni radice del polinomio caratteristico di una matrice hermitiana (e quindi in particolare di una matrice simmetrica) è reale.

Ogni matrice hermitiana è pertanto simile ad una matrice diagonale reale tramite una matrice unitaria. In altri termini, per ogni matrice hermitiana A esistono una matrice unitaria P ed una diagonale reale D per cui

$$D = P^{-1}AP = {}^t\bar{P}AP.$$

LEMMA 4.20. *Tutte le radici del polinomio caratteristico di una matrice simmetrica A di ordine n sono reali, o in modo equivalente, tutte le radici del polinomio caratteristico di un endomorfismo autoaggiunto f di uno spazio vettoriale euclideo (V, \cdot) sono reali. In altri termini, se $\lambda_1, \lambda_2, \dots, \lambda_k$ sono tutti gli autovalori distinti di un endomorfismo autoaggiunto f di V con rispettive molteplicità $m_{\lambda_1}, m_{\lambda_2}, \dots, m_{\lambda_k}$, allora si ha l'uguaglianza:*

$$m_{\lambda_1} + m_{\lambda_2} + \dots + m_{\lambda_k} = n,$$

dove n è la dimensione di V .

A differenza di ciò che succede nel caso reale per ciò che riguarda le matrici simmetriche, si proverà che se una matrice $A \in \mathbb{C}^{n,n}$ è diagonalizzabile in campo complesso mediante una matrice unitaria essa non è necessariamente hermitiana, ma verifica solo la relazione ${}^t\bar{A}A = A{}^t\bar{A}$ (cfr. Teor. 4.21). Ad esempio si avrà che le matrici ortogonali reali A (quindi tali che ${}^tA = A^{-1}$) sono diagonalizzabili in campo complesso mediante una matrice unitaria (cfr. Esercizio 4.21).

In modo naturale è necessario quindi introdurre la seguente definizione.

DEFINIZIONE 4.6. *Una matrice quadrata complessa $A \in \mathbb{C}^{n,n}$ di ordine n si dice normale se:*

$${}^t\bar{A}A = A{}^t\bar{A}.$$

Ricordando la notazione $A^ = {}^t\bar{A}$ (cfr. Oss. 4.11) si ha che una matrice quadrata complessa $A \in \mathbb{C}^{n,n}$ di ordine n si dice normale se:*

$$A^*A = AA^*.$$

ESERCIZIO 4.21. Verificare che le matrici unitarie, le matrici reali ortogonali, le matrici hermitiane e le matrici simmetriche (reali) sono esempi di matrici normali.

SOLUZIONE. Una matrice unitaria A è normale, in quanto da ${}^t\bar{A} = A^{-1}$, segue:

$$A {}^t\bar{A} = {}^t\bar{A} A = I.$$

Per le matrici ortogonali (reali) si ha che ${}^t\bar{A} = {}^tA = A^{-1}$ e pertanto:

$$A {}^t\bar{A} = {}^t\bar{A} A = I.$$

Per le matrici hermitiane si ha ${}^t\bar{A} = A$ e quindi:

$$A {}^t\bar{A} = {}^t\bar{A} A = A^2.$$

Infine, una matrice simmetrica (reale) A è normale in quanto ${}^t\bar{A} = {}^tA = A$ e pertanto

$$A {}^t\bar{A} = {}^t\bar{A} A = A^2.$$

Il teorema spettrale, valido nel caso reale per le matrici simmetriche, in campo complesso non vale solo per le matrici hermitiane ma in generale per le matrici normali. Più precisamente si ha il seguente teorema.

TEOREMA 4.21. (**Teorema spettrale in campo complesso**)

- (1) Se A è una matrice normale, allora esistono una matrice unitaria P ed una matrice diagonale D tali che $P^{-1}AP = D$.
- (2) Se A è una matrice diagonalizzabile mediante una matrice unitaria, allora A è normale.

DIMOSTRAZIONE. La dimostrazione di (1) è più complicata del caso reale, per la sua lettura si rimanda ad esempio a [8].

La dimostrazione di (2) è simile a quella nel caso reale per le matrici simmetriche. Infatti dal fatto che esista una matrice P di ordine n unitaria tale che $P^{-1}AP = D$, si deduce

$${}^t\bar{A} = {}^t(\bar{P} \bar{D} \bar{P}^{-1}) = {}^t\bar{P}^{-1} \bar{D} {}^t\bar{P} = P \bar{D} {}^t\bar{P}$$

in quanto ${}^t\bar{D} = \bar{D}$ e ${}^t\bar{P}^{-1} = P$. Quindi:

$${}^t\bar{A}A = (P \bar{D} {}^t\bar{P})(PDP^{-1}) = P \bar{D}DP^{-1},$$

$$A {}^t\bar{A} = (PDP^{-1})({}^t\bar{P}^{-1} \bar{D} {}^t\bar{P}) = P \bar{D}DP^{-1},$$

ma $D\bar{D} = \bar{D}D$. □

OSSERVAZIONE 4.19. Si osservi che gli autovalori di una matrice normale sono in generale numeri complessi e quindi la matrice D non è necessariamente reale.

Tutte le matrici complesse normali possono quindi essere diagonalizzate in \mathbb{C} con una base unitaria di autovettori. Tuttavia, questo non si verifica nel caso reale. Infatti, una matrice normale reale può avere autovalori immaginari, e quindi non essere diagonalizzabile in campo reale (pur rimanendo diagonalizzabile in campo complesso). Ne è un esempio la matrice di una rotazione del piano vettoriale V_2 di angolo diverso da 0 e θ .

Nel caso di una trasformazione unitaria di uno spazio vettoriale hermitiano (cfr. Par. 7 e 2) valgono le seguenti proprietà, per la dimostrazione si rimanda ad esempio a [8].

TEOREMA 4.22. *Sia f una trasformazione unitaria di uno spazio vettoriale hermitiano (V, \cdot) .*

- (1) *Se W è un sottospazio vettoriale invariante per f , anche il suo complemento ortogonale W^\perp è invariante per f .*
- (2) *Se V ha dimensione finita, allora V ammette una base unitaria formata da autovettori di f .*

Si osservi che il fatto che esista una base unitaria di V formata da autovettori di f segue anche dalla proprietà già enunciata che le matrici unitarie sono diagonalizzabili mediante matrici unitarie (cfr. Teorema 4.21.)

OSSERVAZIONE 4.20. Si precisa che il Teorema Spettrale dimostrato nel caso degli endomorfismi autoaggiunti è **un caso speciale** del Teorema 4.21 dimostrato nel caso delle matrici normali. Per esempio, nel caso degli endomorfismi autoaggiunti si ha che gli autovalori sono necessariamente reali, da cui segue che gli autospazi sono ortogonali e così via, invece, nel caso delle matrici normali gli autovalori **non** sono necessariamente reali.

Si intende ora presentare un metodo alternativo, che fa uso del teorema spettrale, per determinare una base simplettica relativa ad una forma simplettica. Per fare ciò è necessario dimostrare i lemmi seguenti.

LEMMA 4.23. *Sia A una matrice antisimmetrica reale, allora A^2 è una matrice simmetrica reale.*

DIMOSTRAZIONE. È quasi ovvia, infatti si ha per ipotesi che ${}^tA = -A$, dunque

$$({}^tA)^2 = ({}^tA) ({}^tA) = (-A)(-A) = A^2,$$

da cui la tesi. □

LEMMA 4.24. *Sia A una qualsiasi matrice quadrata, gli autovalori della matrice A^2 sono necessariamente il quadrato degli autovalori della matrice A .*

DIMOSTRAZIONE. Se λ è autovalore di A allora $AX = \lambda X$ con X autovettore. Quindi

$$A^2X = A(AX) = A(\lambda X) = \lambda AX = \lambda^2 X,$$

da cui segue che se X è un autovettore di A allora lo è anche di A^2 . □

LEMMA 4.25. *Ogni matrice antisimmetrica reale è diagonalizzabile in campo complesso.*

DIMOSTRAZIONE. La tesi segue dal fatto che iA è una matrice hermitiana, infatti

$${}^t(iA) = i {}^tA = -iA = \overline{iA},$$

da cui la tesi. □

Si deve ricordare ora che ogni autovalore di una matrice simmetrica reale è reale e che una matrice simmetrica è sempre diagonalizzabile.

LEMMA 4.26. *Sia A una matrice antisimmetrica di $\mathbb{R}^{2n,2n}$, allora gli autovalori di A^2 sono strettamente negativi. Quindi gli autovalori di A^2 hanno tutti necessariamente molteplicità uguale ad un numero pari.*

DIMOSTRAZIONE. Se $A^2 \in \mathbb{R}^{2n,2n}$ ha autovalori $\lambda_1, \dots, \lambda_k$ esiste una base ortonormale

$$(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{2n}) \quad (4.24)$$

di \mathbb{R}^{2n} di autovettori. Con un abuso di notazione, indichiamo le matrici colonne delle componenti dei vettori della base (4.24) come gli stessi vettori. Inoltre sia “ \cdot ” il prodotto scalare di \mathbb{R}^{2n} rispetto al quale la base (4.24) è ortonormale. Dimostriamo che $\lambda_1 < 0$. Si ha

$$A^2 \mathbf{v}_1 = \lambda_1 \mathbf{v}_1,$$

ossia

$$(A^2 \mathbf{v}_1) \cdot \mathbf{v}_1 = \lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_1 = \lambda_1 \|\mathbf{v}_1\|^2. \quad (4.25)$$

Ma

$$\begin{aligned} (A^2 \mathbf{v}_1) \cdot \mathbf{v}_1 &= {}^t(A^2 \mathbf{v}_1) \mathbf{v}_1 = {}^t \mathbf{v}_1 {}^t A^2 \mathbf{v}_1 \\ &= {}^t \mathbf{v}_1 {}^t A {}^t A \mathbf{v}_1 = {}^t \mathbf{v}_1 {}^t A (-A) \mathbf{v}_1 = -{}^t(A \mathbf{v}_1)(A \mathbf{v}_1) = -A \|\mathbf{v}_1\|^2 \end{aligned} \quad (4.26)$$

Dall’uguaglianza di (4.25) con (4.26) segue la tesi. \square

COROLLARIO 4.27. *Sia A una matrice antisimmetrica reale, considerata in campo complesso, allora i suoi autovalori sono totalmente immaginari e si trovano a coppie del tipo $i\lambda, -i\lambda$. Di conseguenza una tale matrice è sempre diagonalizzabile in campo complesso.*

DIMOSTRAZIONE. Segue dai Lemmi 4.24 e 4.26. \square

COROLLARIO 4.28. *Sia A una matrice antisimmetrica reale di ordine $2n + 1$, gli autovalori della matrice simmetrica A^2 hanno tutti molteplicità un numero pari.*

DIMOSTRAZIONE. Dai lemmi precedenti si ha:

- 1) ogni autovalore di A^2 è il quadrato di un autovalore della matrice A ed è un numero negativo;
- 2) ogni autovalore di A^2 è riferito agli stessi autovettori di A ;
- 3) la matrice A è diagonalizzabile. I suoi autovalori sono coppie del tipo $i\lambda, -i\lambda$. Siano \mathbf{x}_1 e \mathbf{x}_2 due autovettori ad essi riferiti, essendo gli autovalori corrispondenti diversi essi sono, necessariamente, linearmente indipendenti.
- 4) I vettori \mathbf{x}_1 e \mathbf{x}_2 sono anche due autovettori riferiti all’autovalore λ^2 di A^2 .
- 5) L’autospazio relativo a λ^2 di A^2 ha almeno dimensione 2.
- 6) Essendo A^2 una matrice simmetrica, la molteplicità dei suoi autovalori coincide con la dimensione dei suoi autospazi.

\square

LEMMA 4.29. *Sia A una matrice antisimmetrica, se \mathbf{v}_1 è un autovettore di A^2 relativo all'autovalore λ_1 allora anche $A\mathbf{v}_1$ è un autovettore di A^2 relativo allo stesso autovalore λ_1 .*

DIMOSTRAZIONE. Tenendo conto che, per ipotesi $A^2\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ si ha

$$A^2(A\mathbf{v}_1) = A(A^2\mathbf{v}_1) = A(\lambda_1\mathbf{v}_1) = \lambda_1(A\mathbf{v}_1),$$

ossia la tesi. \square

LEMMA 4.30. *Nelle ipotesi del lemma precedente si ha che i vettori $\mathbf{v}_1, A\mathbf{v}_1$ sono linearmente indipendenti.*

DIMOSTRAZIONE. Se per assurdo esistesse $\mu \in \mathbb{R}$ tale che $A\mathbf{v}_1 = \mu\mathbf{v}_1$ allora

$$A(A\mathbf{v}_1) = \mu A\mathbf{v}_1 \rightarrow A^2\mathbf{v}_1 = \mu A\mathbf{v}_1 \rightarrow \lambda_1\mathbf{v}_1 = \mu^2\mathbf{v}_1.$$

Ma per il Lemma 4.26 $\lambda_1 < 0$ da cui l'assurdo. \square

Le seguenti affermazioni seguono dai lemmi precedenti e da dimostrazioni simili precedentemente incontrate. Innanzi tutto, se si considera l'automorfismo f associato alla matrice simplettica reale $A \in \mathbb{R}^{2n, 2n}$, si ha subito che il sottospazio vettoriale

$$\mathcal{L}(\mathbf{v}_1, A\mathbf{v}_1)$$

è invariante per f . Dal fatto che A^2 è diagonalizzabile mediante una base ortonormale di autovettori di cui \mathbf{v}_1 è il primo vettore si procede con il considerare un altro autovalore di A^2 , sia λ_2 e l'autovettore \mathbf{v}_2 ad esso relativo, di norma uno e ortogonale a \mathbf{v}_1 . Si dimostra che il sottospazio vettoriale

$$\mathcal{L}(\mathbf{v}_2, A\mathbf{v}_2)$$

è ortogonale al precedente, e così via, fino a determinare una base di \mathbb{R}^{2n} del tipo seguente

$$\mathcal{B} = (\mathbf{v}_1, A\mathbf{v}_1, \mathbf{v}_2, A\mathbf{v}_2, \dots, \mathbf{v}_n, A\mathbf{v}_n)$$

dove

$$(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$$

è una base ortonormale di \mathbb{R}^n .

Si tratta ora di determinare la matrice associata alla forma bilineare simplettica B (associata ad A) rispetto alla base \mathcal{B} .

Ci si limita al primo passo, relativo a $\mathcal{L}(\mathbf{v}_1, A\mathbf{v}_1)$. Si ha:

$$B(\mathbf{v}_1, A\mathbf{v}_1) = {}^t\mathbf{v}_1 A(A\mathbf{v}_1) = {}^t\mathbf{v}_1 \lambda_1 \mathbf{v}_1 = \lambda_1 \|\mathbf{v}_1\|.$$

In altri termini si ottiene al primo passo la matrice

$$\begin{pmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{pmatrix}$$

da cui, con un evidente cambiamento di base, si perviene alla matrice J e alla base simplettica richiesta.

ESERCIZIO 4.22. Utilizzando il metodo appena descritto e facendo uso del software *Mathematica*, si determini una base simplettica della forma bilineare simplettica associata alla matrice

$$A = \begin{pmatrix} 0 & -1 & 2 & 0 \\ 1 & 0 & -2 & 0 \\ -2 & 2 & 0 & 3 \\ 0 & 0 & -3 & 0 \end{pmatrix}.$$

10. Forme bilineari complesse

Si vogliono studiare ora le forme bilineari simmetriche $B : V \times V \rightarrow \mathbb{C}$ non degeneri e si intende dimostrare il seguente teorema.

TEOREMA 4.31. *Sia V uno spazio vettoriale complesso di dimensione n e sia B una forma bilineare simmetrica non degenera definita su V . È sempre possibile determinare una base di V per cui*

$$B(X, Y) = {}^tXY,$$

con la consueta confusione tra vettore e matrice colonna delle componenti.

DIMOSTRAZIONE. È necessario iniziare con una doverosa premessa che distingue le forme bilineari simmetriche definite sul campo dei numeri complessi \mathbb{C} da quelle definite su \mathbb{R} .

Supponiamo che per qualche motivo (che è l'oggetto della dimostrazione del teorema) la matrice simmetrica A associata a B rispetto ad una base di V si trasformi mediante un opportuno cambiamento di base ad una matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_2 & & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & \dots & \dots & \dots & \lambda_n \end{pmatrix}, \quad \lambda_k \in \mathbb{C}, \lambda_k \neq 0.$$

Ciò significa che se $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ è la base rispetto alla quale D è la matrice associata a B si ha che

$$B(\mathbf{e}_k, \mathbf{e}_k) = \lambda_k, \quad k = 1, \dots, n, \quad (4.27)$$

e i vettori della base sono ortogonali (rispetto a B) tra di loro. È così possibile, a partire da \mathcal{B} , determinare una nuova base $\mathcal{C} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n)$ per cui

$$B(\mathbf{f}_k, \mathbf{f}_k) = 1.$$

È sufficiente porre $\mathbf{f}_k = t_k \mathbf{e}_k$, sostituire nella relazione precedente e da (4.27) si ottiene

$$(t_k)^2 = \frac{1}{\lambda_k},$$

da cui la tesi. Si osservi che, nel caso complesso, non ha senso pensare ad un'eventuale concetto di segnatura, infatti se per esempio $B(\mathbf{e}_k, \mathbf{e}_k) = -1$ è sufficiente porre $\mathbf{f}_k = i\mathbf{e}_k$ e sostituendo si ottiene $i^2 B(\mathbf{e}_k, \mathbf{e}_k) = -1$ da cui segue $B(\mathbf{f}_k, \mathbf{f}_k) = 1$.

Sia nel caso reale sia nel caso complesso vale il seguente teorema, che insieme con l'osservazione precedente conclude la dimostrazione del Teorema 4.31. \square

TEOREMA 4.32. (Teorema di Gauss-Lagrange) *Sia B una forma bilineare simmetrica definita su uno spazio vettoriale V (reale o complesso) di dimensione n . È sempre possibile determinare una base di V mediante la quale la matrice associata a B sia diagonale.*

OSSERVAZIONE 4.21. Si osservi che la dimostrazione del Teorema di Gauss-Lagrange non può discendere dalla dimostrazione del Teorema Spettrale. Per esempio la matrice simmetrica complessa

$$A = \begin{pmatrix} 2i & 1 \\ 1 & 0 \end{pmatrix} \quad (4.28)$$

non è diagonalizzabile (ammette i come unico autovalore e non è diagonale), mentre può essere considerata come la matrice associata ad una forma bilineare simmetrica su \mathbb{C}^2 .

DIMOSTRAZIONE. (del Teorema di Gauss-Lagrange) Sia B una forma bilineare simmetrica non nulla. Ciò significa che esistono due vettori $\mathbf{x}, \mathbf{y} \in V$ tali che $B(\mathbf{x}, \mathbf{y}) \neq 0$. Da ciò segue che almeno uno tra i numeri

$$B(\mathbf{x}, \mathbf{x}), \quad B(\mathbf{y}, \mathbf{y}), \quad B(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y})$$

deve essere non nullo, come si deduce dalla formula

$$B(\mathbf{x}, \mathbf{y}) = \frac{1}{2} [B(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) - B(\mathbf{x}, \mathbf{x}) - B(\mathbf{y}, \mathbf{y})].$$

Sia quindi \mathbf{a} un vettore di V non isotropo, allora \mathbf{a} è il primo vettore della base \mathcal{B} cercata. La dimostrazione procede per induzione su $n = \dim(V)$.

Per spiegare meglio, costruiamo il secondo vettore di \mathcal{B} . Se si trova un vettore \mathbf{b} di V , non nullo e linearmente indipendente da \mathbf{a} , per cui $B(\mathbf{a}, \mathbf{b}) = 0$, allora \mathbf{b} è il secondo vettore della base. Altrimenti, preso \mathbf{b} per cui $B(\mathbf{a}, \mathbf{b}) \neq 0$, si cerca un vettore $\mathbf{c} = \mathbf{b} + \lambda \mathbf{a}$, con λ scalare, in modo tale che $B(\mathbf{a}, \mathbf{c}) = 0$. Si ottiene

$$\lambda = -\frac{B(\mathbf{a}, \mathbf{b})}{B(\mathbf{a}, \mathbf{a})},$$

quindi il vettore

$$\mathbf{c} = \mathbf{b} - \frac{B(\mathbf{a}, \mathbf{b})}{B(\mathbf{a}, \mathbf{a})} \mathbf{a}$$

è il secondo vettore della base \mathcal{B} .

Si procede, partendo dal primo vettore \mathbf{a} della base \mathcal{B} , calcolando $\mathcal{L}(\mathbf{a})^\perp$, ossia il sottospazio vettoriale ortogonale a $\mathcal{L}(\mathbf{a})$ rispetto a B . Essendo B una forma bilineare simmetrica qualsiasi è necessario dimostrare che

$$V = \mathcal{L}(\mathbf{a}) \oplus \mathcal{L}(\mathbf{a})^\perp \quad (4.29)$$

dopo di che il Teorema di Gauss-Lagrange è dimostrato, in quanto è sufficiente procedere con il metodo descritto su $\mathcal{L}(\mathbf{a})^\perp$.

Sia $\mathbf{x} \in \mathcal{L}(\mathbf{a}) \cap \mathcal{L}(\mathbf{a})^\perp$. Allora $\mathbf{x} = \lambda \mathbf{a}$ per qualche scalare λ in quanto $\mathbf{x} \in \mathcal{L}(\mathbf{a})$ da cui si ottiene $B(\lambda \mathbf{a}, \lambda \mathbf{a}) = \lambda^2 B(\mathbf{a}, \mathbf{a}) = 0$ che è assurdo.

Sia \mathbf{y} un generico vettore di V , allora

$$\mathbf{y} = \frac{B(\mathbf{a}, \mathbf{y})}{B(\mathbf{a}, \mathbf{a})} \mathbf{a} + \left(\mathbf{y} - \frac{B(\mathbf{a}, \mathbf{y})}{B(\mathbf{a}, \mathbf{a})} \mathbf{a} \right),$$

formula che completa la dimostrazione di (4.29). \square

ESEMPIO 4.23. Si consideri la matrice A definita in (4.28), si vuole determinare una base \mathcal{D} rispetto alla quale la forma bilineare simmetrica complessa associata ad A , rispetto alla base canonica $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ di \mathbb{C}^2 , sia associata alla matrice unit  (det(A) = -1 , quindi B   non degenere).

Si inizia con la determinazione di una base \mathcal{C} rispetto alla quale la matrice associata a B sia una matrice diagonale. Il primo vettore della base \mathcal{C} pu  essere \mathbf{e}_1 in quanto $B(\mathbf{e}_1, \mathbf{e}_1) = 2i$. Si determina il secondo vettore \mathbf{b} ponendo

$$\mathbf{b} = \mathbf{a}_2 + \lambda \mathbf{e}_1.$$

Da $B(\mathbf{e}_1, \mathbf{b}) = 0$ si ottiene

$$\lambda = \frac{1}{2}i.$$

Rispetto alla base $\mathcal{C} = (\mathbf{e}_1, \mathbf{b})$ la matrice associata a B  

$$D = \begin{pmatrix} 2i & 0 \\ 0 & \frac{1}{2}i \end{pmatrix}.$$

Si perviene alla base $\mathcal{D} = (\mathbf{f}_1, \mathbf{f}_2)$ richiesta ponendo $\mathbf{f}_1 = \mu_1 \mathbf{e}_1$ in modo tale che $B(\mathbf{f}_1, \mathbf{f}_1) = 1$ e $\mathbf{f}_2 = \mu_2 \mathbf{b}$ in modo tale che $B(\mathbf{f}_2, \mathbf{f}_2) = 1$ ossia

$$\mathbf{f}_1 = \left(\frac{1}{2} - i \frac{1}{2} \right) \mathbf{e}_1, \quad \mathbf{f}_2 = (1 - i) \mathbf{b}.$$

Si considerano quindi le matrici A per cui

$$B(AX, AY) = B(X, Y),$$

con B forma bilineare simmetrica complessa non degenere. Dal Teorema di Gauss-Lagrange segue la definizione del gruppo

$$O(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) / {}^tAA = I\}$$

che prende il nome di **gruppo ortogonale complesso**. Si vede facilmente che

$$A \in O(n, \mathbb{C}) \implies \det(A) = \pm 1.$$

Si definisce quindi

$$SO(n, \mathbb{C}) = O(n, \mathbb{C}) \cap SL(n, \mathbb{C}),$$

che prende il nome di **gruppo ortogonale speciale complesso**.

Si verifica anche che $O(1, \mathbb{C}) = O(1)$.

Sia ora $B : V \times V \longrightarrow \mathbb{C}$ una forma simplettica complessa. Nella dimostrazione del teorema di esistenza di una base simplettica (cfr. Teorema 4.8) si prescinde dal fatto che la forma simplettica sia reale o complessa, quindi vale lo stesso teorema anche nel caso complesso, permettendo quindi la definizione del gruppo

$$Sp(2n, \mathbb{C}) = \{A \in GL(2n, \mathbb{C}) / {}^tAJA = J\},$$

che è il **gruppo simplettico complesso**.

È facile verificare che ogni matrice simplettica complessa ha determinante pari a ± 1 . Nel modo analogo al caso reale si può dimostrare che in realtà $\det(A) = 1$.

ESERCIZIO 4.24. Verificare che

$$Sp(2, \mathbb{C}) = SL(2, \mathbb{C}).$$

Vale il seguente teorema che permette di concludere che anche il gruppo unitario $U(n)$ è in realtà rappresentabile come intersezione di due gruppi reali.

TEOREMA 4.33. *Si dimostra che*

$$U(n) \cong O(2n) \cap Sp(2n, \mathbb{R}),$$

dove il simbolo \cong indica che è possibile definire un isomorfismo di gruppi, tra i due gruppi considerati.

DIMOSTRAZIONE. Si consideri l'omomorfismo di gruppi

$$\alpha: GL(n, \mathbb{C}) \longrightarrow GL(2n, \mathbb{R}), \quad Z = A + iB \longmapsto \begin{pmatrix} A & -B \\ B & A \end{pmatrix}.$$

Si indichi ancora con α la restrizione di α a $U(n)$. La dimostrazione segue dai punti seguenti:

- 1) $\alpha(U(n)) \subseteq O(2n) \cap Sp(2n, \mathbb{R})$,
- 2) α è iniettiva,
- 3) α è suriettiva,
- 4) α è un omomorfismo di gruppi.

Si osserva subito che i punti 2) e 4) sono conseguenza delle analoghe proprietà di α definita su $\mathbb{C}^{n,n}$.

Imponendo che la matrice $Z = A + iB$ appartenga a $U(n)$ si ottengono le due condizioni seguenti:

$${}^tAA + {}^tBB = I$$

$${}^tAB \quad \text{deve essere simmetrica.}$$

Se si impone che $\alpha(U(n)) \subseteq O(2n)$ si ottengono esattamente le stesse condizioni. Ciò significa che

$$\alpha(A + iB) \in O(2n) \iff A + iB \in U(n),$$

ossia

$$\alpha(GL(n, \mathbb{C})) \cap O(2n) = \alpha(U(n)).$$

Anche imponendo che $\alpha(U(n)) \subseteq Sp(2n, \mathbb{R})$ si perviene alle stesse condizioni, ossia si ha

$$\alpha(GL(n, \mathbb{C})) \cap Sp(2n, \mathbb{R}) = \alpha(U(n)).$$

Da tutto ciò segue che il punto 1) è dimostrato.

Per il viceversa, si considera una matrice

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in GL(2n, \mathbb{R})$$

e, imponendo che appartenga a $O(2n) \cap Sp(2n, \mathbb{R})$, si trova innanzi tutto che deve essere della forma

$$M = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

ossia si ha che $M = \alpha(A - iB)$. Dalle considerazioni precedenti si ha la tesi. \square

OSSERVAZIONE 4.22. La dimostrazione precedente è formalmente corretta, anche se troppo lunga. Si può più informalmente ricavare lo stesso risultato anche dalle seguenti considerazioni (importanti).

Siano

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \quad \mathbf{y} = (y_1, y_2, \dots, y_n)$$

elementi di \mathbb{C}^n , allora, indicando con H il prodotto hermitiano (scritto rispetto ad una base unitaria) si ha

$$H(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n x_j \bar{y}_j.$$

Ponendo

$$x_j = (a_j + ib_j), \quad y_j = (c_j + id_j), \quad a_j, b_j, c_j, d_j \in \mathbb{R}$$

e sostituendo nell'espressione di H si ha

$$H(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n (a_j c_j + b_j d_j) + i \sum_{j=1}^n (b_j c_j - a_j d_j). \quad (4.30)$$

Ponendo

$$\mathbf{x} = (a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n), \quad \mathbf{y} = (c_1, c_2, \dots, c_n, d_1, d_2, \dots, d_n)$$

segue che la formula (4.30) si può scrivere come

$$H(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} + iB(\mathbf{x}, \mathbf{y}),$$

dove \cdot indica il prodotto scalare in \mathbb{R}^{2n} e B è una forma simplettica in \mathbb{R}^{2n} . Di conseguenza se una matrice A mantiene inalterato il prodotto hermitiano, deve necessariamente anche mantenere inalterati sia il prodotto scalare sia una forma simplettica, da cui, almeno intuitivamente, la tesi.

11. Prodotto scalare quaternionico

In questo paragrafo si intende introdurre un prodotto scalare anche nel caso quaternionico e lo si definirà, per iniziare, in \mathbb{H}^n . Si assume che \mathbb{H}^n sia uno spazio vettoriale su \mathbb{H} **a destra**, ciò significa che ha senso considerare le matrici colonna delle componenti dei vettori $\mathbf{q} \in \mathbb{H}^n$. Si pone

$$\mathbf{q} = (q_1, \dots, q_n), \quad \mathbf{q}' = (q'_1, \dots, q'_n), \quad \mathbf{q}, \mathbf{q}' \in \mathbb{H}^n,$$

Q e Q' indicheranno quindi le matrici colonna delle componenti di \mathbf{q} e di \mathbf{q}' . Come al solito, spesso si confonderanno le due notazioni.

Si definisce la funzione

$$B: \mathbb{H}^n \times \mathbb{H}^n \longrightarrow \mathbb{H}$$

ponendo

$$B(\mathbf{q}, \mathbf{q}') = \sum_{k=1}^n \bar{q}_k q'_k = {}^t \bar{Q} Q'. \quad (4.31)$$

Vale il seguente teorema.

TEOREMA 4.34. *Per la funzione B definita in (4.31) valgono le seguenti proprietà:*

- (1) B è una forma bilineare rispetto alla somma di vettori;
- (2) $B(\mathbf{q}, \mathbf{q}') = \overline{B(\mathbf{q}', \mathbf{q})}$, $\mathbf{q}, \mathbf{q}' \in \mathbb{H}^n$;
- (3) per ogni $\lambda \in \mathbb{H}$ e per ogni $\mathbf{q}, \mathbf{q}' \in \mathbb{H}^n$ si ha

$$B(\mathbf{q}\lambda, \mathbf{q}') = \bar{\lambda} B(\mathbf{q}, \mathbf{q}'), \quad B(\mathbf{q}, \mathbf{q}'\lambda) = B(\mathbf{q}, \mathbf{q}')\lambda;$$

- (4) $B(\mathbf{q}, \mathbf{q}) = \sum_{k=1}^n \bar{q}_k q_k = \sum_{k=1}^n q_k \bar{q}_k = \sum_{k=1}^n |q_k|^2 \geq 0$; inoltre $B(\mathbf{q}, \mathbf{q}) = 0$ se e solo se $\mathbf{q} = \mathbf{o}$.

La dimostrazione del teorema precedente è lasciata per esercizio.

Una forma bilineare B con le proprietà descritte nel teorema appena enunciato definisce un **prodotto scalare quaternionico**. Più precisamente si ha la seguente definizione.

DEFINIZIONE 4.7. *Sia V uno spazio vettoriale costruito su \mathbb{H} a destra, sia $B : V \times V \rightarrow \mathbb{H}$ una funzione che verifica le proprietà enunciate nel Teorema 4.34 allora B prende il nome di **prodotto scalare quaternionico** su V .*

Analogamente al caso reale e al caso hermitiano, si verifica facilmente che su ogni spazio vettoriale V di questo tipo è sempre possibile definire almeno un prodotto scalare quaternionico ed anche lasciata per esercizio la dimostrazione del teorema seguente.

TEOREMA 4.35. *Sia B un prodotto scalare quaternionico. È sempre possibile individuare una base di V rispetto alla quale B assume l'espressione (4.31).*

concetti appena introdotti permettono di definire la **norma** di $\mathbf{q} \in \mathbb{H}^n$ (o in V) nel modo ovvio:

$$\|\mathbf{q}\| = \sqrt{B(\mathbf{q}, \mathbf{q})}.$$

Si introduce, in modo naturale, il gruppo

$$Sp(n) = \{A \in GL(n, \mathbb{H}) / B(A\mathbf{q}, A\mathbf{q}') = B(\mathbf{q}, \mathbf{q}')\}$$

che prende il nome di **gruppo simplettico quaternionico**.

I due lemmi che seguono sono volti a studiare la relazione che definisce le matrici di $Sp(n)$.

LEMMA 4.36. *Siano $A, B \in \mathbb{H}^{n,n}$ allora*

$${}^t(AB) = {}^t \bar{B} {}^t \bar{A}$$

che equivale a

$${}^t(\overline{AB}) = {}^t \bar{B} {}^t \bar{A}.$$

DIMOSTRAZIONE. Ponendo $A = (a_{ij})$, $B = (b_{ij})$ con $a_{ij}, b_{ij} \in \mathbb{H}$, gli elementi di $AB = (c_{ij})$ sono

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

Per pervenire a BA è necessario commutare a_{ik} con b_{kj} ma, come è noto, il prodotto in \mathbb{H} non è commutativo. Per fare ciò si usa la proprietà

$$\overline{q q'} = \overline{q'} \overline{q}, \quad q, q' \in \mathbb{H}.$$

Precisamente si pone

$$\sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n \overline{\overline{a_{ik} b_{kj}}} = \sum_{k=1}^n \overline{b_{kj} \overline{a_{ik}}},$$

da cui la tesi. □

LEMMA 4.37. *Sia B un prodotto scalare quaternionico e $A \in \mathbb{H}^{n,n}$ allora*

$$B(A \mathbf{q}, \mathbf{q}') = B(\mathbf{q}, {}^t \overline{A} \mathbf{q}'), \quad \mathbf{q}, \mathbf{q}' \in \mathbb{H}^n.$$

DIMOSTRAZIONE. Si usa l'espressione (4.31) del prodotto scalare B e il risultato appena dimostrato per ottenere

$$B(A \mathbf{q}, \mathbf{q}') = {}^t(\overline{A \overline{Q}}) Q' = {}^t \overline{Q} {}^t \overline{A} Q' = {}^t \overline{Q} ({}^t \overline{A} Q'),$$

ossia la tesi. □

Siamo ora in grado di determinare la relazione che lega le matrici A di $Sp(n)$. Si ha

$$B(A \mathbf{q}, A \mathbf{q}') = B(\mathbf{q}, {}^t \overline{A} A \mathbf{q}') = B(\mathbf{q}, \mathbf{q}'),$$

quindi

$${}^t \overline{Q} ({}^t \overline{A} A Q') = {}^t \overline{Q} Q'.$$

Allora

$$Sp(n) = \{A \in Gl(n, \mathbb{H}) / {}^t \overline{A} A = I\},$$

dove I indica la matrice unità di ordine n . In altri termini, $Sp(n)$ ha la stessa definizione di $U(n)$.

ESEMPIO 4.25. Si osserva che

$$Sp(1) = SL(1, \mathbb{H}) \cong SU(2) \cong S^3(1).$$

Si dimostra il seguente teorema, analogo, nel caso complesso al Teorema 4.33.

TEOREMA 4.38. *È valida la seguente relazione*

$$Sp(n) \cong U(2n) \cap Sp(2n, \mathbb{C}).$$

DIMOSTRAZIONE. Si può procedere con lo stesso metodo descritto nella dimostrazione del Teorema 4.33, metodo formalmente corretto anche se, in questo caso, molto lungo.

Però più informalmente, si può scrivere l'espressione del prodotto scalare quaternionico $B(\mathbf{q}, \mathbf{q}')$, ponendo

$$\mathbf{q} = (q_1, q_2, \dots, q_n), \quad \mathbf{q}' = (q'_1, q'_2, \dots, q'_n), \quad q_k, q'_k \in \mathbb{H},$$

si ha

$$B(\mathbf{q}, \mathbf{q}') = \sum_{k=1}^n \bar{q}_k q'_k. \quad (4.32)$$

Usando l'identificazione di \mathbb{H} con \mathbb{C}^2 si pone

$$q_k = u_k + jv_k, \quad q'_k = u'_k + jv'_k, \quad u_k, v_k, u'_k, v'_k \in \mathbb{C},$$

proseguendo il calcolo di (4.32) si ha

$$B(\mathbf{q}, \mathbf{q}') = \sum_{k=1}^n (\bar{u}_k u'_k + \bar{v}_k v'_k) + j \sum_{k=1}^n (u_k v'_k - v_k u'_k),$$

ossia il prodotto scalare quaternionico di due vettori si può scrivere come la somma del loro prodotto hermitiano e del loro prodotto simplettico. Pertanto se $A \in Sp(n)$ allora

$$B(A\mathbf{q}, A\mathbf{q}') = B(\mathbf{q}, \mathbf{q}'),$$

e quindi la matrice A conserva entrambi questi prodotti, da cui la tesi. \square

Se si considera il calcolo del prodotto scalare quaternionico scritto identificando ogni quaternionone con \mathbb{R}^4 (anziché con \mathbb{C}^2) si ottiene un'importante caratterizzazione del prodotto scalare quaternionico.

Si considerino due quaternioni $q = a + ib + jc + kd$ e $q' = a' + ib' + jc' + kd'$, il loro prodotto scalare quaternionico $B(q, q')$ assume la seguente espressione

$$\begin{aligned} B(q, q') &= \bar{q}q' \\ &= (aa' + bb' + cc' + dd') + i(ab' - a'b - cd' + dc') \\ &\quad + j(ac' - ca' + bd' - db') + k(ad' - da' - bc' + cb'), \end{aligned} \quad (4.33)$$

esprimendo così B come somma del prodotto scalare standard di \mathbb{R}^4 e di tre forme simplettiche in \mathbb{R}^4 .

La stessa situazione si ripropone in generale con il calcolo di $B(\mathbf{q}, \mathbf{q}')$ con $\mathbf{q}, \mathbf{q}' \in \mathbb{H}^n$. Si ottiene

$$B(\mathbf{q}, \mathbf{q}') = \mathbf{q} \cdot \mathbf{q}' + iB_1(\mathbf{q}, \mathbf{q}') + jB_2(\mathbf{q}, \mathbf{q}') + kB_3(\mathbf{q}, \mathbf{q}')$$

dove \cdot indica il prodotto scalare standard in \mathbb{R}^{4n} e B_1, B_2, B_3 sono forme simplettiche diverse in \mathbb{R}^{4n} .

Per capire meglio la situazione descritta, scriviamo in notazione matriciale la formula (4.33) mettendo in evidenza le matrici associate alle forme simplettiche B_1, B_2, B_3 in questo caso particolare che facilmente si può estendere al caso generale (al solito si confondono i vettori con le matrici colonna delle loro componenti).

$$B(q, q') = {}^tQQ' + i {}^tQJ_1Q' + j {}^tQJ_2Q' + k {}^tQJ_3Q',$$

con

$$J_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

Si osservi che $J_2 = -J$ ed inoltre $J_1 J_2 = J_3$ (coerentemente con $ij = k$). Si ottiene così

$$Sp(n) \cong O(4n) \cap Sp_1(4n, \mathbb{R}) \cap Sp_2(4n, \mathbb{R}) \cap Sp_3(4n, \mathbb{R}),$$

dove

$$Sp_h(4n, \mathbb{R}) = \{A \in GL(4n, \mathbb{R}) / {}^t A J_h A = J_h\}, \quad h = 1, 2, 3.$$

È evidente che $Sp_2(4n, \mathbb{R}) = Sp(4n, \mathbb{R})$. Si tratta ora di capire il legame tra le matrici appartenenti ai diversi gruppi $Sp_h(4n, \mathbb{R})$. Appare subito chiaro che cambiamenti di base simplettiche lasciano invariati i gruppi introdotti. Sia $A \in Sp_1(4n, \mathbb{R})$, vale a dire

$${}^t A J_1 A = J_1,$$

sia $A' \in Sp_2(4n, \mathbb{R})$, vale a dire

$${}^t A' J_2 A' = J_2.$$

Si consideri il cambiamento di base $P \in GL(4n, \mathbb{R})$ nella forma bilineare B_1 che permette di passare da J_1 a J_2 , vale a dire

$${}^t P J_1 P = J_2.$$

Sostituendo nelle relazioni precedenti si ottiene

$$A = P A' P^{-1} \implies A' = P^{-1} A P,$$

ossia A e A' sono matrici simili (come era prevedibile), quindi la relazione tra i gruppi è del tipo

$$P^{-1} Sp_1(4n, \mathbb{R}) P = Sp_2(4n, \mathbb{R}),$$

di evidente definizione.

Si osservi che quanto descritto **non** è un caso particolare dei gruppi simplettici, ma si osserva in ciascuno dei vari gruppi di matrici studiati, ossia cambiamenti di base qualsiasi producono copie dei vari gruppi legati tra di loro dalla relazione appena trovata.

OSSERVAZIONE 4.23. Anche nel caso del prodotto scalare quaternionico vale il Teorema di Pitagora, ossia

$$B(\mathbf{q}, \mathbf{q}') = 0 \implies \|\mathbf{q} + \mathbf{q}'\|^2 = \|\mathbf{q}\|^2 + \|\mathbf{q}'\|^2,$$

con $\mathbf{q}, \mathbf{q}' \in V$ e V spazio vettoriale su \mathbb{H} a destra. La dimostrazione è un esercizio.

In modo analogo a quanto dimostrato nel caso hermitiano, possono essere adattate al caso del prodotto scalare quaternionico le dimostrazioni che permettono di ottenere, anche in questo caso le disuguaglianze di Cauchy-Schwarz e di Minkowski.

Più precisamente, dimostriamo ora la disuguaglianza di Cauchy-Schwarz data da

$$|B(\mathbf{q}, \mathbf{q}')| \leq \sqrt{B(\mathbf{q}, \mathbf{q})} \sqrt{B(\mathbf{q}', \mathbf{q}')},$$

con $\mathbf{q}, \mathbf{q}' \in V$, V spazio vettoriale su \mathbb{H} a destra e B prodotto scalare quaternionico.

Si considera il vettore $\mathbf{w} \in V$ dato da

$$\mathbf{w} = \mathbf{q} - \mathbf{q}' \frac{B(\mathbf{q}, \mathbf{q}')}{B(\mathbf{q}', \mathbf{q}')}$$

Si dimostra facilmente che $B(\mathbf{w}, \mathbf{q}') = 0$. Dalla relazione $B(\mathbf{w}, \mathbf{w}) \geq 0$, sostituendo il risultato ottenuto e procedendo come nella dimostrazione del caso hermitiano segue la tesi.

Mimando la stessa dimostrazione della disuguaglianza triangolare vista nel caso hermitiano, si ottiene la dimostrazione analoga della diguaglianza di Minkowski nel caso quaternionico.

Il seguente importantissimo teorema chiude l'identificazione dei gruppi di matrici con le sfere, affermando che non ci saranno altre identificazioni oltre quelle man mano indicate.

TEOREMA 4.39. (Teorema di Adams) *Le uniche sfere che hanno la struttura di gruppo di matrici sono*

$$S^0(1), \quad S^1(1), \quad S^3(1).$$

Si osservi che la sfera $S^0(1)$ è in realtà formata da soli due punti, quindi è un esempio di **gruppo finito** con due elementi.

Algebre e gruppi di Lie

1. Algebre di Lie

DEFINIZIONE 5.1. Un'algebra di Lie reale (complessa) è uno spazio vettoriale reale (complesso) \mathfrak{g} dotato di un'applicazione bilineare (detta **parentesi di Lie** o semplicemente "**bracket**")

$$[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}, \quad (\mathbf{x}, \mathbf{y}) \longmapsto [\mathbf{x}, \mathbf{y}],$$

che verifica le seguenti proprietà:

- (1) $[\mathbf{x}, \mathbf{y}] = -[\mathbf{y}, \mathbf{x}]$;
- (2) $[[\mathbf{x}, \mathbf{y}], \mathbf{z}] + [[\mathbf{y}, \mathbf{z}], \mathbf{x}] + [[\mathbf{z}, \mathbf{x}], \mathbf{y}] = 0$, per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathfrak{g}$.

L'espressione a sinistra di (2) si può scrivere come $\mathfrak{S}_{\mathbf{x}, \mathbf{y}, \mathbf{z}}[[\mathbf{x}, \mathbf{y}], \mathbf{z}]$ dove $\mathfrak{S}_{\mathbf{x}, \mathbf{y}, \mathbf{z}}$ denota la somma sulle permutazioni cicliche di $\mathbf{x}, \mathbf{y}, \mathbf{z}$. La formula (2) è detta **identità di Jacobi**.

ESEMPIO 5.1. (1) Sia V_3 lo spazio vettoriale reale ordinario. Il prodotto vettoriale \wedge conferisce a V_3 la struttura di algebra di Lie reale.

- (2) Un qualsiasi spazio vettoriale può sempre essere dotato della struttura di algebra di Lie **abeliana** definendo $[\mathbf{x}, \mathbf{y}] = \mathbf{0}$, per ogni coppia di vettori \mathbf{x}, \mathbf{y} in V .
- (3) Come già indicato, lo spazio vettoriale delle matrici $\mathbb{R}^{n,n}$ può essere dotato della struttura di algebra di Lie definendo

$$[A, B] = AB - BA, \quad A, B \in \mathbb{R}^{n,n}.$$

Per tale algebra di Lie si userà la notazione $\mathfrak{gl}(n, \mathbb{R})$.

- (4) In modo analogo, è evidente la definizione dell'algebra di Lie complessa $\mathfrak{gl}(n, \mathbb{C})$.

OSSERVAZIONE 5.1. Si osservi che sullo stesso spazio vettoriale è possibile definire diverse strutture di algebra di Lie, per esempio su $\mathbb{R}^{n,n}$ si può definire la struttura abeliana di algebra di Lie e la struttura introdotta al punto 3. degli esempi precedenti, la notazione $\mathfrak{gl}(n, \mathbb{R})$ è riferita solo a quest'ultima struttura. La definizione che segue spiega meglio questa osservazione introducendo il concetto di isomorfismo di algebre di Lie.

DEFINIZIONE 5.2. Un **omomorfismo di algebre di Lie** tra le due algebre di Lie \mathfrak{g} e \mathfrak{g}' , costruite sullo stesso campo, è una funzione

$$\varphi: \mathfrak{g} \longrightarrow \mathfrak{g}'$$

tale che

- (1) φ è un'applicazione lineare tra i due spazi vettoriali \mathfrak{g} e \mathfrak{g}' ;
- (2) vale la seguente relazione tra i bracket di \mathfrak{g} e \mathfrak{g}'

$$\varphi([\mathbf{x}, \mathbf{y}]) = [\varphi(\mathbf{x}), \varphi(\mathbf{y})], \quad \mathbf{x}, \mathbf{y} \in \mathfrak{g}.$$

Se φ è un isomorfismo di spazi vettoriali allora φ prende il nome di **isomorfismo di algebre di Lie**.

ESEMPIO 5.2. (1) Se $\dim(\mathfrak{g}) = 1$ allora l'unica struttura di algebra di Lie su \mathfrak{g} è quella abeliana, infatti se (\mathbf{e}_1) è una base di \mathfrak{g} allora

$$[\mathbf{x}, \mathbf{y}] = [\lambda \mathbf{e}_1, \mu \mathbf{e}_1] = \lambda \mu [\mathbf{e}_1, \mathbf{e}_1],$$

per ogni coppia di scalari λ, μ .

(2) Se $\dim(\mathfrak{g}) = 2$ allora si possono definire solo due strutture non isomorfe di algebra di Lie e precisamente:

- la struttura abeliana $[\mathbf{x}, \mathbf{y}] = \mathbf{o}$ per ogni coppia di vettori \mathbf{x}, \mathbf{y} di \mathfrak{g} .
- Sia $(\mathbf{e}_1, \mathbf{e}_2)$ una base di \mathfrak{g} , ponendo

$$[\mathbf{e}_1, \mathbf{e}_2] = \mathbf{e}_2$$

si definisce una struttura di algebra di Lie su \mathfrak{g} non abeliana.

È un **Esercizio** dimostrare che le altre possibili strutture non abeliane definibili su \mathfrak{g} sono tutte isomorfe a quella appena citata.

(3) La classificazione delle algebre di Lie di dimensione 3 è stata completata da J. Milnor in [10] in modo quasi elementare, sfruttando la nozione di prodotto vettoriale. Anche la classificazione in dimensione 4 è completata in lavori di ricerca meno accessibili alla loro trattazione nei corsi di laurea di base.

ESERCIZIO 5.3. Dimostrare che è sempre possibile definire un isomorfismo φ di algebre di Lie tra l'algebra di Lie \mathfrak{g} di dimensione 2 con base $\mathcal{B}(\mathbf{e}_1, \mathbf{e}_2)$ tale che $[\mathbf{e}_1, \mathbf{e}_2] = \mathbf{e}_2$ e se stessa considerata però con base $\mathcal{B}' = (\mathbf{e}'_1, \mathbf{e}'_2)$ e tale che $[\mathbf{e}'_1, \mathbf{e}'_2] = \lambda \mathbf{e}'_1 + \mu \mathbf{e}'_2$ con λ e μ non contemporaneamente nulli.

A questo scopo si indichi con

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

la matrice associata a φ rispetto alla base \mathcal{B} nel dominio e alla base \mathcal{B}' nel codominio. Esplicitando la condizione

$$\varphi([\mathbf{e}_1, \mathbf{e}_2]) = [\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2)],$$

si ottiene

$$\begin{cases} a_{12} = \lambda \det(A) \\ a_{22} = \mu \det(A). \end{cases}$$

Se sia λ sia μ non sono nulli, una possibile matrice A è

$$A = \begin{pmatrix} 1 & \lambda \\ \frac{1-\mu}{\mu} & \mu \end{pmatrix}.$$

È necessario introdurre anche la definizione (molto naturale) di sottoalgebra di Lie.

DEFINIZIONE 5.3. *Un sottospazio vettoriale \mathfrak{h} di un'algebra di Lie \mathfrak{g} si dice **sottoalgebra di Lie** di \mathfrak{g} se*

$$[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h},$$

ossia se

$$\forall \mathbf{x}, \mathbf{y} \in \mathfrak{h} \quad \Longrightarrow \quad [\mathbf{x}, \mathbf{y}] \in \mathfrak{h}.$$

2. Sottoalgebre di Lie di $\mathfrak{gl}(n, \mathbb{R})$

Il sottospazio vettoriale di $\mathbb{R}^{n,n}$ delle matrici triangolari superiori con diagonale nulla è un esempio di algebra di Lie che si indica, in generale con

$$\mathfrak{n} = \left\{ \left(\begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} / a_{ij} \in \mathbb{R} \right) \right\}. \quad (5.1)$$

Inoltre

$$\dim(\mathfrak{n}) = \frac{n(n-1)}{2}.$$

Le verifiche di quest'ultima affermazione e del fatto che \mathfrak{n} sia una sottoalgebra di Lie di $\mathfrak{gl}(n, \mathbb{R})$ sono lasciate per esercizio.

È molto importante il caso precedente per $n = 3$, che porta alla definizione dell'algebra di Lie

$$\mathfrak{h} = \left\{ \left(\begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} / x, y, z \in \mathbb{R} \right) \right\}$$

che prende il nome di **algebra di Lie di Heisenberg**. Ovviamente $\dim(\mathfrak{h}) = 3$. Se si fissa la base

$$\mathbf{e}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

si verifica facilmente che tutti i bracket sulla base sono nulli eccetto

$$[\mathbf{e}_1, \mathbf{e}_2] = \mathbf{e}_3.$$

Un altro esempio di sottoalgebra di Lie di $\mathfrak{gl}(n, \mathbb{R})$ è

$$\mathfrak{sl}(n, \mathbb{R}) = \{A \in \mathfrak{gl}(n, \mathbb{R}) / \operatorname{tr} A = 0\}.$$

Il fatto che $\mathfrak{sl}(n, \mathbb{R})$ sia un'algebra di Lie segue dalle proprietà della traccia di una matrice, infatti si ha

$$\operatorname{tr}([A, B]) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0.$$

È anche evidente che

$$\dim(\mathfrak{sl}(n, \mathbb{R})) = n^2 - 1.$$

Il sottospazio vettoriale delle matrici antisimmetriche è anche un esempio di sottoalgebra di Lie di $\mathfrak{gl}(n, \mathbb{R})$ e, considerando questa struttura, si indica nel modo seguente

$$\mathfrak{so}(n, \mathbb{R}) = \{A \in \mathfrak{gl}(n, \mathbb{R}) / A + {}^t A = 0\}.$$

La dimostrazione del fatto che $\mathfrak{so}(n, \mathbb{R})$ si ottiene calcolando mostrando che $[A, B]$ è una matrice antisimmetrica se A e B lo sono, si ha:

$$\begin{aligned} {}^t[A, B] &= {}^t(AB - BA) = {}^t(AB) - {}^t(BA) \\ &= {}^tB {}^tA - {}^tA {}^tB = BA - AB = -[A, B]. \end{aligned}$$

È noto che

$$\dim(\mathfrak{so}(n, \mathbb{R})) = \frac{n(n-1)}{2}.$$

ESERCIZIO 5.4. Dimostrare che non è possibile definire un isomorfismo di algebre di Lie tra $\mathfrak{so}(3, \mathbb{R})$ e \mathfrak{h} , pur essendo spazi vettoriali isomorfi.

ESERCIZIO 5.5. Verificare che il sottospazio vettoriale delle matrici simmetriche di $\mathbb{R}^{n,n}$ non è una sottoalgebra di Lie di $\mathfrak{gl}(n, \mathbb{R})$.

Un altro esempio di algebra di Lie è il seguente:

$$\mathfrak{so}(k, n-k) = \{A \in \mathfrak{gl}(n, \mathbb{R}) / {}^tA I_{k, n-k} + I_{k, n-k} A = 0\}.$$

Si lascia per esercizio la verifica che $\mathfrak{so}(k, n-k)$ è una sottoalgebra di Lie di $\mathfrak{gl}(n, \mathbb{R})$, invece si procede con il calcolo della sua dimensione.

Sia

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix},$$

una matrice di $\mathfrak{so}(k, n-k)$, di conseguenza, si pone $A_1 \in \mathbb{R}^{k,k}$, $A_2 \in \mathbb{R}^{k, n-k}$, $A_3 \in \mathbb{R}^{n-k, k}$, $A_4 \in \mathbb{R}^{n-k, n-k}$. Imponendo che A appartenga a $\mathfrak{so}(k, n-k)$ si ottengono le seguenti condizioni

- A_1 deve essere una matrice antisimmetrica,
- ${}^tA_2 = A_3$.
- A_4 deve essere una matrice antisimmetrica.

Di conseguenza si ha

$$\dim(\mathfrak{so}(k, n-k)) = \frac{k(k-1)}{2} + k(n-k) + \frac{(n-k)(n-k-1)}{2} = \frac{n(n-1)}{2}.$$

Nonostante $\dim(\mathfrak{so}(k, n-k)) = \dim(\mathfrak{so}(n))$ le due algebre di Lie non sono isomorfe ($k \neq n$). Inoltre non sono neppure isomorfe tra di loro le diverse algebre di Lie che si ottengono al variare di k , ad eccezione di $\mathfrak{so}(k, p) \cong \mathfrak{so}(p, k)$.

La seguente algebra di Lie

$$\mathfrak{sp}(2n, \mathbb{R}) = \{A \in \mathfrak{gl}(2n, \mathbb{R}) / {}^tAJ + JA = 0\}$$

prende il nome di **algebra di Lie simplettica reale**. È un esercizio controllare che effettivamente si tratta di una sottoalgebra di Lie di $\mathfrak{gl}(2n, \mathbb{R})$. Per il calcolo della sua dimensione, invece, si considera una matrice

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad A_1, A_2, A_3, A_4 \in \mathbb{R}^{n,n}$$

e si impone che A appartenga a $\mathfrak{sp}(2n, \mathbb{R})$. Si ottengono le seguenti condizioni

- A_3 deve essere una matrice simmetrica;
- ${}^tA_4 + A_1 = 0$;
- A_2 deve essere una matrice simmetrica.

Di conseguenza si ha

$$\dim(\mathfrak{sp}(2n, \mathbb{R})) = \frac{n(n+1)}{2} + n^2 + \frac{n(n+1)}{2} = n(2n+1).$$

ESERCIZIO 5.6. Verificare che:

- 1) $\mathfrak{gl}(1, \mathbb{R}) = \mathbb{R}$;
- 2) $\mathfrak{sl}(1, \mathbb{R}) = \{0\}$;
- 3) $\mathfrak{sp}(2, \mathbb{R}) = \mathfrak{sl}(2, \mathbb{R})$.

3. Sottoalgebre di Lie di $\mathfrak{gl}(n, \mathbb{C})$

Si inizia la trattazione, analogamente al caso reale, considerando lo spazio vettoriale $\mathbb{C}^{n,n}$ delle matrici quadrate ad elementi complessi e la definizione di bracket data dal commutatore delle matrici. L'algebra di Lie complessa che viene così ad essere definita si indica $\mathfrak{gl}(n, \mathbb{C})$, è evidente che (su \mathbb{C})

$$\dim(\mathfrak{gl}(n, \mathbb{C})) = n^2.$$

Analogamente al caso reale si definisce l'algebra di Lie delle matrici complesse a traccia nulla, ossia

$$\mathfrak{sl}(n, \mathbb{C}) = \{A \in \mathfrak{gl}(n, \mathbb{C}) / \operatorname{tr}(A) = 0\}.$$

Si lascia per esercizio la verifica che si tratti di un'algebra di Lie, ovviamente si ha (su \mathbb{C})

$$\dim(\mathfrak{sl}(n, \mathbb{C})) = n^2 - 1.$$

Si considera ora il sottospazio vettoriale (reale) delle matrici anti-hermitiane. Si verifica facilmente che si tratta di un'algebra di Lie reale (a differenza del sottospazio delle matrici hermitiane) e si indica come

$$\mathfrak{u}(n) = \{A \in \mathfrak{gl}(n, \mathbb{C}) / {}^tA + \bar{A} = 0\}.$$

Si ricorda che (attenzione! su \mathbb{R})

$$\dim(\mathfrak{u}(n)) = n^2.$$

Inoltre le matrici anti-hermitiane non hanno traccia nulla, ha quindi senso considerare l'algebra di Lie

$$\mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C}) = \{A \in \mathfrak{u}(n) / \operatorname{tr}(A) = 0\},$$

quindi, di nuovo su \mathbb{R} ,

$$\dim(\mathfrak{su}(n)) = n^2 - 1.$$

Questo non deve essere confuso con l'algebra di Lie delle matrici complesse antisimmetriche definita da

$$\mathfrak{so}(n, \mathbb{C}) = \{A \in \mathfrak{gl}(n, \mathbb{C}) / {}^tA + A = 0\}$$

la cui dimensione (su \mathbb{C}) è

$$\dim(\mathfrak{so}(n, \mathbb{C})) = \frac{n(n-1)}{2}.$$

In modo totalmente analogo al caso reale, si definisce l'**algebra di Lie simplettica complessa** come

$$\mathfrak{sp}(2n, \mathbb{C}) = \{A \in \mathfrak{gl}(n, \mathbb{C}) / {}^tAJ + JA = 0\}$$

e anche in questo caso si ha (su \mathbb{C})

$$\dim(\mathfrak{sp}(2n, \mathbb{C})) = n(2n + 1).$$

ESEMPIO 5.7. Proviamo che l'algebra di Lie $\mathfrak{su}(2)$ è isomorfa a $\mathfrak{so}(3)$. Sia $\mathfrak{su}(2) = \mathcal{L}(E_1, E_2, E_3)$, dove

$$E_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Si ha

$$[E_1, E_2] = 2E_3, \quad [E_1, E_3] = -2E_2, \quad [E_2, E_3] = 2E_1.$$

Sia $\mathfrak{so}(3) = \mathcal{L}(F_1, F_2, F_3)$, dove

$$F_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad F_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Si ha

$$[F_1, F_2] = -F_3, \quad [F_1, F_3] = F_2, \quad [F_2, F_3] = -F_1.$$

Si verifichi che l'applicazione lineare $h: \mathfrak{su}(2) \rightarrow \mathfrak{so}(3)$ tale che

$$h(E_1) = 2F_1, \quad h(E_2) = 2F_2, \quad h(E_3) = -2F_3$$

è l'isomorfismo richiesto.

ESEMPIO 5.8. Si noti che se A è una matrice hermitiana, allora iA è una matrice antihermitiana e viceversa. Le matrici, dette **matrici di Pauli**,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

sono matrici hermitiane tali che

$$i\sigma_1 = E_3, \quad i\sigma_2 = E_2, \quad i\sigma_3 = E_1.$$

Lo spazio vettoriale $\mathcal{H}(\mathbb{C}^{2,2})$ delle matrici hermitiane di ordine 2 è reale, ha dimensione 4, ed ha come base $(I_2, \sigma_1, \sigma_2, \sigma_3)$, dove I_2 è la matrice unità di ordine 2. L'applicazione lineare definita da

$$1 \rightarrow I_2, \quad i \rightarrow -i\sigma_1, \quad j \rightarrow -i\sigma_2, \quad k \rightarrow -i\sigma_3,$$

individua un isomorfismo tra l'algebra \mathbb{H} dei quaternioni e l'algebra (rispetto al prodotto di matrici) generata da $\{I_2, i\sigma_1, i\sigma_2, i\sigma_3\}$.

4. Sottoalgebre di Lie di $\mathfrak{gl}(n, \mathbb{H})$

Analogamente al caso reale e al caso complesso si definisce l'algebra di Lie $\mathfrak{gl}(n, \mathbb{H})$ delle matrici $A \in \mathbb{H}^{n,n}$. Si intende con $\mathbb{H}^{n,n}$ lo spazio vettoriale a destra su \mathbb{H} delle matrici quadrate di ordine n ad elementi in \mathbb{H} e si considera come bracket il commutatore di due matrici. È chiaro che $\mathfrak{gl}(n, \mathbb{H})$ è un'algebra di Lie reale, in quanto $[A, B]$ è solo lineare rispetto ai numeri reali, quindi (su \mathbb{R})

$$\dim(\mathfrak{gl}(n, \mathbb{H})) = 4n^2.$$

Si introduce l'algebra di Lie **reale** definita nel modo seguente

$$\mathfrak{sp}(n, \mathbb{H}) = \{A \in \mathfrak{gl}(n, \mathbb{H}) / {}^t\bar{A} + A = 0\},$$

che è l'analogo di $\mathfrak{u}(n)$ nel caso complesso. Si osservi di nuovo che si tratta di un'algebra di Lie reale e ciò è dovuto alla presenza, nella sua definizione, del coniugato di una matrice. Per calcolarne la dimensione è necessario capire come siano gli elementi della generica matrice $A \in \mathfrak{sp}(n)$. Sulla diagonale si hanno elementi che verificano la condizione

$$q_{hh} + \overline{q_{hh}}, \quad h = 1, \dots, n,$$

ossia elementi totalmente immaginari, invece, in generale

$$q_{hl} + \overline{q_{hl}} = 0, \quad h, l = 1, \dots, n.$$

Quindi la dimensione (reale) di $\mathfrak{sp}(n)$ è

$$\dim(\mathfrak{sp}(n)) = 3n + 4 \frac{n(n-1)}{2} = n(2n+1).$$

Si osserva che, anche se $\dim(\mathfrak{sp}(n)) = \dim(\mathfrak{sp}(2n, \mathbb{R}))$ le due algebra di Lie non sono isomorfe, come stabilisce, anche già nel caso della dimensione 3, l'esercizio che segue.

ESERCIZIO 5.9. Verificare che non è possibile definire un isomorfismo di algebre di Lie tra $\mathfrak{sp}(1)$ e $\mathfrak{sp}(2, \mathbb{R})$.

SOLUZIONE. Per risolvere questo esercizio è necessario determinare una base per ciascuna delle due algebre di Lie e scrivere l'espressione dei bracket rispetto ai vettori della base. Innanzi tutto si osserva che

$$\mathfrak{sp}(1) = \mathcal{L}(i, j, k)$$

e

$$[i, j] = 2k, \quad [j, k] = 2i, \quad [i, k] = -2j.$$

Ponendo

$$\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

si ha

$$\mathfrak{sp}(2, \mathbb{R}) = \mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3),$$

e

$$[\mathbf{e}_1, \mathbf{e}_2] = 2\mathbf{e}_2, \quad [\mathbf{e}_1, \mathbf{e}_3] = -2\mathbf{e}_3, \quad [\mathbf{e}_2, \mathbf{e}_3] = \mathbf{e}_1.$$

Sia $\varphi: \mathfrak{sp}(1) \rightarrow \mathfrak{sp}(2, \mathbb{R})$ un isomorfismo con matrice associata $A = (a_{ij}) \in \mathbb{R}^{3,3}$ rispetto alle basi (i, j, k) nel dominio e $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ nel codominio. Si chiamino $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3$ i vettori riga della matrice A e si impongano le condizioni

$$\varphi([i, j]) = [\varphi(i), \varphi(j)], \quad \varphi([j, k]) = [\varphi(j), \varphi(k)], \quad \varphi([i, k]) = [\varphi(i), \varphi(k)].$$

Si ottiene

$$\mathbf{R}_1 = \mathbf{R}_2 \wedge \mathbf{R}_3, \quad \mathbf{R}_2 = \mathbf{R}_1 \wedge \mathbf{R}_2, \quad \mathbf{R}_3 = \mathbf{R}_3 \wedge \mathbf{R}_1,$$

che è chiaramente assurdo.

5. Esponenziale di una matrice

Sia A una matrice quadrata ad elementi in \mathbb{R} o in \mathbb{C} . Si vuole studiare lo sviluppo in serie di potenze

$$e^A = I + A + \frac{1}{2}A^2 + \dots + \frac{1}{n!}A^n + \dots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k. \quad (5.2)$$

È ovviamente indispensabile dimostrare che la serie appena introdotta converge ad una matrice. Ma prima di questa dimostrazione consideriamo alcuni esempi per capire meglio.

ESEMPIO 5.10. Sia

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

allora

$$e^D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2^2 & 0 \\ 0 & 3^2 \end{pmatrix} + \dots = \begin{pmatrix} e^2 & 0 \\ 0 & e^3 \end{pmatrix}.$$

L'esempio precedente porta quindi alla seguente proprietà la cui dimostrazione è un esercizio.

PROPRIETÀ 5.1. *Se D è una matrice diagonale (ad elementi reali o complessi) del tipo*

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \lambda_n \end{pmatrix},$$

allora

$$e^D = \begin{pmatrix} e^{\lambda_1} & 0 & \dots & \dots & 0 \\ 0 & e^{\lambda_2} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & e^{\lambda_n} \end{pmatrix}.$$

ESEMPIO 5.11. Sia

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix},$$

poiché

$$A^2 = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^3 = 0,$$

con 0 matrice nulla di ordine 3. Si ha

$$e^A = \begin{pmatrix} 1 & 1 & \frac{7}{2} \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'esempio appena descritto permette (in modo totalmente elementare) di dimostrare il seguente teorema.

TEOREMA 5.2. *La funzione esponenziale di una matrice definita sull'algebra di Lie di Heisenberg \mathfrak{h} ha come immagine il gruppo di Lie di Heisenberg H , ossia*

$$e^\bullet : \mathfrak{h} \longrightarrow H,$$

inoltre e^\bullet è una biiezione.

Il teorema precedente si può estendere all'algebra di Lie \mathfrak{n} introdotta in (5.1).

TEOREMA 5.3. *La serie di potenze (5.2) converge ad una matrice $X = (x_{ij})$ di ordine n ad elementi in \mathbb{R} (se $A \in \mathbb{R}^{n,n}$) o ad elementi in \mathbb{C} (se $A \in \mathbb{C}^{n,n}$).*

DIMOSTRAZIONE. Si tratta di dimostrare che il generico elemento di posto ij di (5.2) converge a x_{ij} . Si consideri allora tale elemento e lo si indichi come

$$(I)_{ij} + (A)_{ij} + \left(\frac{1}{2}A^2\right)_{ij} + \left(\frac{1}{3!}A^3\right)_{ij} + \dots \quad (5.3)$$

Si maggiaora questa serie indicando con m il più grande $|a_{ij}|$ della matrice A (si osservi che $|a_{ij}|$ denota il valore assoluto di a_{ij} se $A \in \mathbb{R}^{n,n}$ e $|a_{ij}|$ è il modulo di a_{ij} se $A \in \mathbb{C}^{n,n}$). È chiaro che per $(I)_{ij}$ si ha $m = 1$. Invece per $(A^2)_{ij}$ si ottiene nm^2 , per $(A^3)_{ij}$ si ha n^2m^3 e così via. Allora la serie (5.3) è dominata dalla serie

$$1 + \sum_{k=1}^{\infty} \frac{n^{k-1} m^k}{k!}.$$

Applicando a questa serie il criterio del rapporto, si ha che il rapporto tra l'elemento di posto $k+1$ e quello di posto k è

$$\frac{nm}{k+1},$$

quindi

$$\lim_{k \rightarrow \infty} \frac{nm}{k+1} = 0,$$

da cui si ottiene che la serie (5.3) converge assolutamente, quindi la matrice X esiste. \square

Si studiano ora le proprietà dell'esponenziale di una matrice e le loro conseguenze.

TEOREMA 5.4. *Sia A una matrice quadrata ad elementi reali o complessi. Allora*

(1) $e^0 = I$, con 0 matrice nulla e I matrice unità, dello stesso ordine;

$$(2) {}^t(e^A) = e^{tA}.$$

La dimostrazione è un esercizio. Ma segue subito dal teorema precedente il seguente corollario.

COROLLARIO 5.5. *A è una matrice simmetrica se e solo se e^A è una matrice simmetrica.*

TEOREMA 5.6. *Se A, B sono matrici quadrate di ordine n ad elementi reali o complessi tali che*

$$AB = BA$$

allora

$$e^{A+B} = e^A e^B.$$

DIMOSTRAZIONE. Si tratta di un calcolo che segue dalle serie

$$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \dots$$

$$e^B = I + B + \frac{1}{2}B^2 + \frac{1}{3!}B^3 + \dots$$

$$e^A e^B = I + B + \frac{1}{2}B^2 + \frac{1}{3!}B^3 + A + AB + \frac{1}{2}AB^2 + \frac{1}{2}A^2 + \dots$$

$$e^{A+B} = I + (A+B) + \frac{1}{2}(A+B)^2 + \dots$$

da cui la tesi (si osservi che è possibile condurre i calcoli in questo modo elementare in quanto le serie convergono assolutamente). \square

OSSERVAZIONE 5.2. L'esempio 5.13 mostrerà che l'ipotesi di commutatività delle matrici A, B è essenziale per la validità del teorema precedente.

COROLLARIO 5.7. *La matrice e^A è invertibile ed inoltre $e^{-A} = (e^A)^{-1}$.*

DIMOSTRAZIONE. Segue facilmente dal teorema precedente ponendo

$$e^0 = I = e^{A-A} = e^A e^{-A},$$

in quanto, ovviamente A e $-A$ sono matrici che commutano. \square

COROLLARIO 5.8. *Come evidente conseguenza del corollario precedente si ha nel caso reale*

$$e^\bullet : \mathfrak{gl}(n, \mathbb{R}) \longrightarrow GL(n, \mathbb{R})$$

e nel caso complesso

$$e^\bullet : \mathfrak{gl}(n, \mathbb{C}) \longrightarrow GL(n, \mathbb{C}).$$

OSSERVAZIONE 5.3. A differenza del caso dell'algebra di Lie di Heisenberg e del gruppo di Heisenberg, si vedrà con l'esempio 5.15 che la **prima** delle due applicazioni considerate nel corollario precedente **non** è suriettiva.

Il Lemma seguente permette di dimostrare che

$$e^\bullet : \mathfrak{sp}(2n, \mathbb{R}) \longrightarrow Sp(2n, \mathbb{R})$$

e (in modo analogo)

$$e^\bullet : \mathfrak{sp}(2n, \mathbb{C}) \longrightarrow Sp(2n, \mathbb{C}).$$

LEMMA 5.9. Sia $A \in \mathfrak{sp}(2n, \mathbb{R})$, oppure $A \in \mathfrak{sp}(2n, \mathbb{C})$, allora

$$e^{tA} J = J e^{-A}.$$

DIMOSTRAZIONE. Per ipotesi

$${}^tAJ = -JA,$$

quindi

$$\begin{aligned} e^{tA} J &= \left(I + {}^tA + \frac{1}{2}({}^tA)^2 + \frac{1}{3!}({}^tA)^3 + \dots \right) J \\ &= J + {}^tAJ + \frac{1}{2}({}^tA)({}^tA)J + \frac{1}{3!}({}^tA)({}^tA)({}^tA)J \dots \\ &= J - JA + \frac{1}{2}JA^2 - \frac{1}{3!}JA^3 + \dots \end{aligned}$$

ossia la tesi. □

TEOREMA 5.10. Se $A \in \mathfrak{sp}(2n, \mathbb{R})$ allora $e^A \in Sp(2n, \mathbb{R})$ e analogamente, se $A \in \mathfrak{sp}(2n, \mathbb{C})$ allora $e^A \in Sp(2n, \mathbb{C})$.

DIMOSTRAZIONE. Dal lemma appena dimostrato segue

$${}^t(e^A) J e^A = e^{tA} J e^A = J e^{-A} e^A = J,$$

ossia la tesi □

ESERCIZIO 5.12. Dimostrare (in modo analogo a quanto visto nel caso di $\mathfrak{sp}(2n, \mathbb{R})$) che

$$e^\bullet : \mathfrak{so}(k, n-k) \longrightarrow SO(k, n-k).$$

Il seguente teorema permette di calcolare l'esponenziale di una matrice diagonalizzabile.

TEOREMA 5.11. Siano A una matrice quadrata ad elementi reali o complessi, sia P una matrice invertibile dello stesso ordine della matrice A allora

$$e^{P^{-1}AP} = P^{-1}e^AP.$$

DIMOSTRAZIONE. È una conseguenza del calcolo diretto, infatti

$$\begin{aligned} e^{P^{-1}AP} &= I + (P^{-1}AP) + \frac{1}{2}(P^{-1}AP)^2 + \frac{1}{3!}(P^{-1}AP)^3 + \dots \\ &= I + P^{-1}AP + \frac{1}{2}(P^{-1}AP)(P^{-1}AP) \\ &\quad + \frac{1}{3!}(P^{-1}AP)(P^{-1}AP)(P^{-1}AP) + \dots \\ &= I + P^{-1}AP + \frac{1}{2}(P^{-1}A^2P) + \frac{1}{3!}(P^{-1}A^3P) + \dots \end{aligned}$$

da cui la tesi □

Se A è una matrice di ordine n diagonalizzabile, allora è ben noto che esistono una matrice diagonale D e una matrice invertibile P , dello stesso ordine, per cui $D = P^{-1}AP$, da cui segue

$$A = PDP^{-1},$$

di conseguenza

$$e^A = P e^D P^{-1},$$

essendo noto il calcolo di e^D si può così determinare e^A . I seguenti esempi chiariranno meglio quanto affermato.

ESEMPIO 5.13. Siano

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$$

verificare che $AB \neq BA$ e che $e^{A+B} \neq e^A e^B$. Si inizia con verificare che

$$AB = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Si ha $B^2 = 0$, con 0 matrice nulla di ordine 2, quindi

$$e^B = I + B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Per procedere al calcolo di e^A si verifica, innanzi tutto, che A è diagonalizzabile e poi si procede come descritto in precedenza.

Gli autovalori di A sono $\lambda_1 = 1$, $\lambda_2 = 0$. Gli autospazi ad essi relativi sono $V_{\lambda_1} = \mathcal{L}((1, 0))$, $V_{\lambda_2} = \mathcal{L}((1, -1))$. Si ottiene così

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix},$$

quindi

$$e^A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix}.$$

Da ciò segue

$$e^A e^B = \begin{pmatrix} e & -1 \\ 0 & 1 \end{pmatrix}.$$

Invece

$$A + B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

quindi

$$e^{A+B} = \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix}.$$

Si vuole ora enunciare un teorema fondamentale in questa trattazione la cui dimostrazione non sarà però completa.

TEOREMA 5.12. *Sia A una matrice quadrata di ordine n ad elementi reali o complessi, si ha*

$$\det(e^A) = e^{\text{tr}(A)}.$$

Alla dimostrazione è necessario anteporre alcune definizioni e teoremi.

DEFINIZIONE 5.4. Una matrice $N \in \mathbb{C}^{n,n}$ si dice **nilpotente** se esiste un intero positivo k tale che

$$N^k = 0,$$

con 0 matrice nulla di ordine n .

Si può dimostrare il seguente teorema

TEOREMA 5.13. *Le seguenti condizioni sono equivalenti:*

- (1) N è una matrice nilpotente di ordine n ,
- (2) il polinomio caratteristico di N è λ^n ;
- (3) $\text{tr}(N^h) = 0, \forall h > 0$;
- (4) è possibile determinare una matrice $P \in \mathbb{C}^{n,n}$ invertibile di ordine n tale che $N = P^{-1}SP$, con S matrice triangolare superiore con la diagonale formata da tutti zeri.

Se A è una matrice quadrata reale o complessa, non diagonalizzabile, è possibile (considerando A , se fosse reale, come matrice complessa) dimostrare il seguente teorema. In realtà il caso di A diagonalizzabile si può configurare come un caso particolare evidente di quanto si sta per affermare. Per la dimostrazione di questo teorema si rimanda a classici testi di Algebra Lineare, per esempio [12].

TEOREMA 5.14. (**Forma canonica di Jordan**) Sia $A \in \mathbb{C}^{n,n}$. Esiste una sola coppia di matrici S, N di $\mathbb{C}^{n,n}$ tali che

- (1) $A = S + N$;
- (2) $SN = NS$;
- (3) S è diagonalizzabile;
- (4) N è nilpotente.

ESEMPIO 5.14. La matrice

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

che, in generale, viene indicata come esempio di matrice non diagonalizzabile si scrive banalmente come

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Possiamo ora procedere con la dimostrazione del Teorema 5.12.

DIMOSTRAZIONE. La dimostrazione segue immediatamente dal Teorema 5.14 ma è più intuitivo dividere la sua dimostrazione nei casi seguenti.

(1) Se A è una matrice diagonale la dimostrazione è ovvia. Sia infatti

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n \end{pmatrix}$$

allora

$$e^{\text{tr}(A)} = e^{\lambda_1 + \lambda_2 + \dots + \lambda_n}.$$

Ma

$$e^A = \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & e^{\lambda_n} \end{pmatrix}$$

quindi

$$\det(e^A) = e^{\lambda_1} e^{\lambda_2} \dots e^{\lambda_n}.$$

(2) Se A è diagonalizzabile allora

$$\det(e^A) = \det(Pe^D P^{-1}) = \det(e^D) = e^{\text{tr}(D)},$$

ma una delle proprietà della traccia di una matrice afferma che

$$\text{tr}(D) = \text{tr}(P^{-1}AP) = \text{tr}(A).$$

(3) Se N è una matrice nilpotente, la tesi segue dal punto (4) del Teorema 5.13.

(4) Se A è una matrice qualsiasi, dal Teorema 5.14 si ha

$$\begin{aligned} \det(e^A) &= \det(e^{S+N}) = \det(e^S e^N) \\ &= \det(e^S) \det(e^N) = e^{\text{tr}(S)} e^{\text{tr}(N)} = e^{\text{tr}(S)}, \end{aligned}$$

da cui la tesi. □

Come conseguenza di questo teorema si ottiene che $e^\bullet : \mathfrak{so}(n) \longrightarrow SO(n)$. Inoltre (se si è in grado di determinare la forma canonica di Jordan di una matrice A) si ha anche il metodo per la determinazione esplicita di e^A .

OSSERVAZIONE 5.4. Dal Teorema 5.12 si ha che $\det(e^A) > 0$ e quindi l'immagine di $\mathfrak{gl}(n, \mathbb{R})$ mediante l'esponenziale è contenuta in $GL^+(n, \mathbb{R})$, sottogruppo di Lie di $GL(n, \mathbb{R})$ formato dalle matrici con determinante *positivo*. Tuttavia proveremo che l'esponenziale $e^\bullet : \mathfrak{gl}(n, \mathbb{R}) \longrightarrow GL^+(n, \mathbb{R})$ non è suriettivo.

ESEMPIO 5.15. Al fine di verificare la non suriettività dell'esponenziale, si consideri la matrice

$$X = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \in GL^+(2, \mathbb{R}).$$

Si vuole dimostrare che non esiste alcuna matrice $A \in \mathfrak{gl}(2, \mathbb{R})$ tale che $e^A = X$. Supponiamo, per assurdo, che una tale matrice esista. Allora

$$1 = \det(X) = e^{\text{tr}(A)} \implies \text{tr}(A) = 0.$$

Di conseguenza, tenendo conto che (nel caso della dimensione 2) la traccia di A è necessariamente la somma dei due autovalori λ_1, λ_2 di A (considerando $A \in \mathfrak{gl}(2, \mathbb{R})$) segue

$$\lambda_1 + \lambda_2 = 0.$$

Si distinguono, evidentemente, due casi

- (1) $\lambda_1 = \lambda_2 = 0$. La matrice A è necessariamente simile ad una matrice N del tipo seguente (non potendo essere la matrice nulla, da cui $e^A = I$)

$$N = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}.$$

Ossia esiste una matrice invertibile P di ordine 2 tale che $A = PNP^{-1}$. In questo caso

$$e^A = P \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} P^{-1},$$

da cui si ottiene che $\text{tr}(e^A) = \text{tr}(e^N) = 2$ che è palesemente assurdo, essendo $\text{tr}(X) = -2$.

- (2) $\lambda_1 = -\lambda_2 \neq 0$. Di conseguenza A è diagonalizzabile. Quindi anche $e^A = X$ è diagonalizzabile, ma ciò è palesemente assurdo.

OSSERVAZIONE 5.5. Si noti che il controesempio precedente non vale più nel caso complesso perchè se $e^{\text{tr}(A)} = 1$, con $A \in \mathfrak{gl}(n, \mathbb{C})$, si ha che $\text{tr}(A) = i2k\pi$, $k \in \mathbb{Z}$. Infatti vale il seguente Teorema.

TEOREMA 5.15. *L'applicazione esponenziale $e^\bullet: \mathfrak{gl}(n, \mathbb{C}) \rightarrow GL(n, \mathbb{C})$ è suriettiva.*

Sia B una forma bilineare non degenera (in generale: reale, complessa, quaternionica, simmetrica o simplettica) e sia e^{sM} , $s \in (-\epsilon, \epsilon) \subseteq \mathbb{R}$, $\epsilon > 0$, una curva passante per I e appartenente al gruppo di matrici che conservano B , cioè che verificano la condizione (4.1). Esplicitamente

$$e^{sM} = I + sM + \frac{1}{2}s^2M^2 + \dots$$

e quindi

$${}^t(e^{sM}) B (e^{sM}) = B.$$

Per trovare lo spazio tangente in I al gruppo di matrici considerato è necessario determinare il vettore tangente alla curva e^{sM} in I . Derivando la relazione precedente (rispetto ad s), si ha

$$e^{sM} {}^t M B e^{sM} + e^{sM} B e^{sM} M = 0,$$

da cui (per $s = 0$) segue

$${}^t M B + B M = 0.$$

In altri termini, usando questo tipo di calcolo e tenendo conto del Teorema 5.12, si dimostra facilmente che

- l'algebra di Lie di $SL(n, \mathbb{R})$ è $\mathfrak{sl}(n, \mathbb{R})$;
- l'algebra di Lie di $SL(n, \mathbb{C})$ è $\mathfrak{sl}(n, \mathbb{C})$;
- l'algebra di Lie di $SO(n)$ è $\mathfrak{so}(n)$;
- l'algebra di Lie di $SO(k, n-k)$ è $\mathfrak{so}(k, n-k)$;
- l'algebra di Lie di $U(n)$ è $\mathfrak{u}(n)$;
- l'algebra di Lie di $Sp(2n, \mathbb{R})$ è $\mathfrak{sp}(2n, \mathbb{R})$;
- l'algebra di Lie di $Sp(2n, \mathbb{C})$ è $\mathfrak{sp}(2n, \mathbb{C})$;
- l'algebra di Lie di $Sp(n)$ è $\mathfrak{sp}(n)$.

A titolo di esercizio determiniamo ora l'esponenziale di una matrice anti-simmetrica di ordine 2 sapendo che necessariamente si dovrà trovare una matrice di $SO(2)$.

ESEMPIO 5.16. Sia

$$D = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \quad a \in \mathbb{R},$$

si vuole calcolare e^A . Si inizia diagonalizzando A . I suoi autovalori sono $\lambda_1 = ia, \lambda_2 = -ia$. I relativi autospazi sono $V_{\lambda_1} = \mathcal{L}(1, -i)$, $V_{\lambda_2} = \mathcal{L}(1, i)$. Si ottiene

$$D = \begin{pmatrix} ia & 0 \\ 0 & -ia \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{1}{2} & -\frac{i}{2} \end{pmatrix}.$$

Di conseguenza

$$e^A = \begin{pmatrix} \frac{e^{ia} + e^{-ia}}{2} & \frac{ie^{ia} - e^{-ia}}{2} \\ \frac{-ie^{ia} + e^{-ia}}{2} & \frac{e^{ia} + e^{-ia}}{2} \end{pmatrix} = \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}.$$

Il teorema che seguente, di cui si omette la dimostrazione, intende (in un certo senso) estendere il Teorema 5.6 al caso di qualsiasi coppia di matrici A, B di ordine n ad elementi reali o complessi.

TEOREMA 5.16. (Teorema di Baker-Campbell-Hausdorff (BCH))
Sia G un gruppo di Lie di matrici e sia \mathfrak{g} la sua algebra di Lie. Il prodotto del gruppo G , in un intorno dell'elemento neutro I , è dato da:

$$e^{t\mathbf{x}} e^{t\mathbf{y}} = e^{\mathbf{z}}, \quad \mathbf{x}, \mathbf{y} \in \mathfrak{g}, \quad t \in \mathbb{R},$$

dove

$$\mathbf{z} = \sum_{n=0}^{+\infty} \mathbf{z}_n(\mathbf{x}, \mathbf{y}) t^n.$$

La serie così introdotta converge per valori di t sufficientemente piccoli ed i suoi coefficienti dipendono solo dai brackets iterati. Si ha:

$$\mathbf{z} = (\mathbf{x} + \mathbf{y})t + \frac{1}{2}[\mathbf{x}, \mathbf{y}]t^2 + \frac{1}{12}([\mathbf{x}, [\mathbf{x}, \mathbf{y}]] - [\mathbf{y}, [\mathbf{x}, \mathbf{y}]])t^3 \\ + \text{termini di ordine superiore}$$

La formula precedente prende il nome di **formula di BCH**.

Nell'esercizio che segue si intende spiegare la formula di BCH nel caso dell'algebra di Lie di Heisenberg.

ESERCIZIO 5.17. Nel caso dell'algebra di Lie di Heisenberg \mathfrak{h} la formula di BCH si riduce a

$$\mathbf{z} = \mathbf{x} + \mathbf{y} + \frac{1}{2}[\mathbf{x}, \mathbf{y}], \quad \mathbf{x}, \mathbf{y} \in \mathfrak{h}.$$

Si inizia considerando l'operazione

$$\mathbf{x} * \mathbf{y} = \mathbf{x} + \mathbf{y} + \frac{1}{2}[\mathbf{x}, \mathbf{y}], \quad \mathbf{x}, \mathbf{y} \in \mathfrak{h}.$$

Si dimostra che $(\mathfrak{h}, *)$ è un gruppo. Infatti vale la proprietà associativa, l'elemento neutro è 0 e l'opposto di \mathbf{x} è $-\mathbf{x}$. La verifica è lasciata per esercizio. Ponendo

$$\mathbf{x} = \begin{pmatrix} 0 & x_1 & x_3 \\ 0 & 0 & x_2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 0 & y_1 & y_3 \\ 0 & 0 & y_2 \\ 0 & 0 & 0 \end{pmatrix}$$

si ha

$$\mathbf{x} * \mathbf{y} = \begin{pmatrix} 0 & x_1 + y_1 & x_3 + y_3 + \frac{1}{2}(x_1y_2 - x_2y_1) \\ 0 & 0 & x_2 + y_2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Si calcoli ora

$$e^{\mathbf{x}*\mathbf{y}} = e^{\mathbf{x}+\mathbf{y}+\frac{1}{2}[\mathbf{x},\mathbf{y}]} = I + \mathbf{x} * \mathbf{y} + \frac{1}{2}(\mathbf{x} * \mathbf{y})^2,$$

si ha

$$e^{\mathbf{x}*\mathbf{y}} = \begin{pmatrix} 1 & x_1 + y_1 & x_3 + y_3 + x_1y_2 + \frac{1}{2}(x_1x_2 + y_1y_2) \\ 0 & 1 & x_2 + y_2 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.4)$$

D'altra parte si sa che $e^\bullet(\mathfrak{h}) = H$, è necessario quindi verificare che il prodotto definito mediante la formula di BCH in (5.4) coincide con l'usuale prodotto di matrici in H . A questo scopo si considerino $A, B \in H$ date da

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & b_{12} & b_{13} \\ 0 & 1 & b_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

si vede che ponendo

$$a_{12} = x_1, \quad b_{12} = y_1$$

$$a_{23} = x_2, \quad b_{23} = y_2$$

$$a_{13} = x_3 + \frac{1}{2}x_1x_2, \quad b_{13} = y_3 + \frac{1}{2}y_1y_2,$$

da cui si deduce che mediante la formula di BCH si ottiene la formula 5.4.

Varietà differenziabili

In questo capitolo verranno date le definizioni e le prime nozioni sulle varietà differenziabili, sui relativi spazi tangenti e sulle funzioni differenziabili. Tutti gli spazi topologici considerati saranno di Hausdorff e a base numerabile (vedi la Definizione 2.5).

1. Definizione di varietà differenziabile

Le varietà differenziabili sono state introdotte allo scopo di estendere alle dimensioni superiori la nozione di superficie parametrizzata: la loro definizione è in buona parte simile alla Definizione 2.6, con la significativa differenza dovuta al fatto che le varietà non sono necessariamente immerse in uno spazio euclideo. Iniziamo col definire le parametrizzazioni locali.

DEFINIZIONE 6.1. *Sia M uno spazio topologico. Una **parametrizzazione locale** è una coppia (U, φ) dove:*

- (1) U è un aperto di \mathbb{R}^n ,
- (2) $\varphi: U \subseteq \mathbb{R}^n \rightarrow \varphi(U) \subseteq M$ è un omeomorfismo ($\varphi(U)$ ha la topologia indotta da quella di M).

OSSERVAZIONE 6.1. Si noti che $\varphi(U)$ è un aperto di M . In molti testi, invece delle parametrizzazioni locali, si considerano le *carte locali* (U, φ) , dove U è un aperto di M e $\varphi: U \subseteq M \rightarrow \varphi(U) \subseteq \mathbb{R}^n$ è un omeomorfismo.

DEFINIZIONE 6.2. *Una **varietà differenziabile (reale)** di dimensione n è uno spazio topologico M di Hausdorff e a base numerabile dotato di un insieme di parametrizzazioni locali $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ che verifica le seguenti condizioni:*

- (1) *l'insieme $\{\varphi_\alpha(U_\alpha)\}_{\alpha \in A}$ è un ricoprimento di M , cioè*

$$\bigcup_{\alpha \in A} \varphi_\alpha(U_\alpha) = M;$$

- (2) *se $W_{\alpha\beta} \stackrel{\text{def.}}{=} \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta) \neq \emptyset$, le funzioni*

$$\varphi_\beta^{-1} \circ \varphi_\alpha: \varphi_\alpha^{-1}(W_{\alpha\beta}) \subseteq \mathbb{R}^n \rightarrow \varphi_\beta^{-1}(W_{\alpha\beta}) \subseteq \mathbb{R}^n,$$

$$\varphi_\alpha^{-1} \circ \varphi_\beta: \varphi_\beta^{-1}(W_{\alpha\beta}) \subseteq \mathbb{R}^n \rightarrow \varphi_\alpha^{-1}(W_{\alpha\beta}) \subseteq \mathbb{R}^n$$

sono differenziabili di classe C^∞ .

*In tal caso, si dice che l'insieme di parametrizzazioni locali $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ individua una **struttura differenziabile** su M .*

Confrontando la definizione di varietà differenziabile con quella di superficie parametrizzata, si vede che la condizione (2) è automaticamente

verificata nel caso delle superfici, essendo le funzioni φ_α differenziabili per definizione.

OSSERVAZIONE 6.2. Se, nella definizione precedente, ogni U_α è un sottoinsieme di \mathbb{C}^n e le funzioni $\varphi_\beta^{-1} \circ \varphi_\alpha$ e $\varphi_\alpha^{-1} \circ \varphi_\beta$ sono *olomorfe*, allora M è una *varietà complessa* di dimensione (complessa) n . Si noti che ogni varietà complessa di dimensione n è anche una varietà differenziabile reale di dimensione (reale) $2n$.

Sia M una varietà differenziabile di dimensione n . Data una parametrizzazione locale (U, φ) , se $p \in \varphi(U)$ allora $\varphi^{-1}(p) \in \mathbb{R}^n$ e quindi

$$\varphi^{-1}(p) = (x_1(p), \dots, x_n(p)).$$

Le funzioni

$$x_i: \varphi(U) \subseteq M \longrightarrow \mathbb{R}, \quad p \mapsto x_i(p), \quad i = 1, \dots, n$$

si dicono **coordinate locali** sull'aperto $\varphi(U)$.

Vediamo ora come si passa da un sistema di coordinate locali ad un altro. Se (U, φ) e (V, ψ) sono due parametrizzazioni locali tali che $p \in W \stackrel{\text{def.}}{=} \varphi(U) \cap \psi(V)$, si ha

$$\varphi^{-1}(p) = (x_1(p), \dots, x_n(p)), \quad \psi^{-1}(p) = (y_1(p), \dots, y_n(p)).$$

Quindi

$$(\psi^{-1} \circ \varphi)(x_1(p), \dots, x_n(p)) = (y_1(p), \dots, y_n(p)),$$

dove

$$y_i = f_i(x_1, \dots, x_n), \quad i = 1, \dots, n.$$

Le funzioni f_i sono differenziabili (punto (2) della Definizione 6.2). Dato che $\psi^{-1} \circ \varphi$ è biiettiva, la matrice Jacobiana

$$J(\psi^{-1} \circ \varphi) = \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} = \left(\frac{\partial y_i}{\partial x_j} \right)$$

ha determinante diverso da zero in ogni punto di $\varphi^{-1}(W)$.

2. Esempi di varietà differenziabili

ESEMPIO 6.1. Si consideri \mathbb{R}^n con la topologia standard. La coppia (\mathbb{R}^n, id) , dove id è la funzione identica, individua la struttura differenziabile standard su \mathbb{R}^n . Invece \mathbb{C}^n ha sia una struttura differenziabile complessa (di dimensione n) che una struttura differenziabile reale (di dimensione $2n$). Lo spazio quaternionico \mathbb{H}^n (identificato con \mathbb{R}^{4n}) ha una struttura differenziabile reale.

ESEMPIO 6.2. Sia V uno spazio vettoriale reale di dimensione n . Fissata una base $(\mathbf{v}_1, \dots, \mathbf{v}_n)$, si identifica V con \mathbb{R}^n mediante la biiezione

$$\varphi: \mathbb{R}^n \longrightarrow V, \quad (x_1, \dots, x_n) \mapsto x_1 \mathbf{v}_1 + \dots + x_n \mathbf{v}_n.$$

Ciò consente di definire una topologia su V : un sottoinsieme $A \subseteq V$ è aperto se e solo se $\varphi^{-1}(A)$ è un aperto di \mathbb{R}^n . Inoltre (\mathbb{R}^n, φ) è una struttura differenziabile su V . In particolare, $\mathbb{R}^{p,q}$, $\mathbb{C}^{p,q}$, $\mathbb{H}^{p,q}$ sono varietà differenziabili reali di dimensione, rispettivamente, pq , $2pq$ e $4pq$.

ESEMPIO 6.3. *Sottovarietà aperte.*

Siano M una varietà differenziabile di dimensione n e $S \subseteq M$ un sottoinsieme *aperto* non vuoto. Si considera su S la topologia indotta da quella di M (vedi la Definizione 2.3). Se $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ è una struttura differenziabile su M , l'insieme delle parametrizzazioni locali di S del tipo (V_α, ψ_α) , dove $V_\alpha = \varphi^{-1}(S \cap \varphi(U_\alpha))$ (se tale sottoinsieme non è vuoto) e ψ_α è la restrizione di φ a V_α , individua una struttura differenziabile su S . In tal modo S diventa una sottovarietà differenziabile aperta di dimensione n .

La struttura differenziabile di $GL(n, \mathbb{R})$ è un esempio della costruzione precedente. La funzione determinante

$$\det: \mathbb{R}^{n,n} \longrightarrow \mathbb{R}, \quad A \mapsto \det A,$$

è continua, dato che $\det A$ è un polinomio negli elementi di A . Quindi l'insieme $\det^{-1}(0) = \{A \in \mathbb{R}^{n,n} / \det A = 0\}$ è chiuso in $\mathbb{R}^{n,n}$. Il suo complementare $GL(n, \mathbb{R})$ è un sottoinsieme aperto di $\mathbb{R}^{n,n}$ e, di conseguenza, è una varietà differenziabile di dimensione n^2 . Nello stesso modo, si vede che $GL(n, \mathbb{C})$ è una varietà differenziabile (reale) di dimensione $2n^2$ e $GL(n, \mathbb{H})$ è una varietà differenziabile (reale) di dimensione $4n^2$.

ESEMPIO 6.4. *La sfera.*

La struttura differenziabile standard sulla sfera $S^n = \{(a_1, \dots, a_{n+1}) \in \mathbb{R}^{n+1} / a_1^2 + \dots + a_{n+1}^2 = 1\}$ si costruisce estendendo in modo naturale alle dimensioni superiori le proiezioni stereografiche dell'Esempio 2.10. Innanzitutto, si considera su S^n la topologia indotta dalla topologia standard di \mathbb{R}^{n+1} . Se $N = (0, \dots, 0, 1)$ e $S = ((0, \dots, 0, -1))$ sono, rispettivamente, i poli 'nord' e 'sud', le inverse delle proiezioni stereografiche sono date da:

$$\varphi: \mathbb{R}^n \longrightarrow S^n - \{N\}, \quad \mathbf{x} \mapsto \varphi(\mathbf{x}) = \left(\frac{2x_1}{1 + \|\mathbf{x}\|^2}, \dots, \frac{2x_n}{1 + \|\mathbf{x}\|^2}, \frac{\|\mathbf{x}\|^2 - 1}{\|\mathbf{x}\|^2 + 1} \right),$$

$$\psi: \mathbb{R}^n \longrightarrow S^n - \{S\}, \quad \mathbf{x} \mapsto \psi(\mathbf{y}) = \left(\frac{2y_1}{1 + \|\mathbf{y}\|^2}, \dots, \frac{2y_n}{1 + \|\mathbf{y}\|^2}, \frac{1 - \|\mathbf{y}\|^2}{1 + \|\mathbf{y}\|^2} \right),$$

$$\mathbf{x} = (x_1, \dots, x_n), \quad \|\mathbf{x}\|^2 = x_1^2 + \dots + x_n^2, \quad \mathbf{y} = (y_1, \dots, y_n), \quad \|\mathbf{y}\|^2 = y_1^2 + \dots + y_n^2.$$

Si controlla che il passaggio tra le due parametrizzazioni locali (\mathbb{R}^n, φ) e (\mathbb{R}^n, ψ) è dato dalla funzione (detta *inversione circolare*)

$$\psi^{-1} \circ \varphi: \varphi^{-1}(W) \longrightarrow \psi^{-1}(W), \quad W = \varphi(S^n) \cap \psi(S^n) = S^n - \{N, S\},$$

$$(\psi^{-1} \circ \varphi)(\mathbf{x}) = \left(\frac{x_1}{\|\mathbf{x}\|^2}, \dots, \frac{x_n}{\|\mathbf{x}\|^2} \right), \quad \mathbf{x} \in \varphi^{-1}(W) = \mathbb{R}^n - \{0\},$$

che è evidentemente differenziabile su $\varphi^{-1}(W)$. Una conclusione analoga vale per $(\psi^{-1} \circ \varphi)^{-1} = \varphi^{-1} \circ \psi$.

ESEMPIO 6.5. *Varietà differenziabili prodotto.*

Siano M_1 e M_2 due varietà differenziabili di dimensione, rispettivamente, n_1 e n_2 . Vogliamo costruire una struttura differenziabile sul prodotto cartesiano $M_1 \times M_2$. Innanzitutto, consideriamo la *topologia prodotto* su $M_1 \times M_2$. Una base di tale topologia è data dai sottoinsiemi del tipo $A_1 \times A_2$, dove A_1 è un aperto di M_1 e A_2 è un aperto di M_2 .

Se $\{(U_\alpha^1, \varphi_\alpha^1)\}_{\alpha \in A}$ è una struttura differenziabile su M_1 e $\{(U_\beta^2, \varphi_\beta^2)\}_{\beta \in B}$ è una struttura differenziabile su M_2 , la **struttura differenziabile prodotto**

su $M_1 \times M_2$ è data da $\{(U_\alpha^1 \times U_\beta^2, \varphi_{\alpha\beta})\}_{(\alpha,\beta) \in A \times B}$ dove le funzioni

$$\varphi_{\alpha\beta}: U_\alpha^1 \times U_\beta^2 \subseteq \mathbb{R}^{n_1+n_2} \longrightarrow M_1 \times M_2,$$

sono definite da:

$$\begin{aligned} \varphi_{\alpha\beta}(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) &= (\varphi_\alpha^1(x_1, \dots, x_{n_1}), \varphi_\beta^2(y_1, \dots, y_{n_2})), \\ (x_1, \dots, x_{n_1}) &\in U_\alpha^1, \quad (y_1, \dots, y_{n_2}) \in U_\beta^2. \end{aligned}$$

È evidente che $\dim(M_1 \times M_2) = n_1 + n_2$. Un esempio importante di varietà differenziabile prodotto è il *toro* di dimensione n : $T^n \stackrel{\text{def.}}{=} \underbrace{S^1 \times \dots \times S^1}_{n \text{ volte}}$.

3. Funzioni differenziabili tra varietà differenziabili

La struttura differenziabile permette di estendere il concetto di differenziabilità a funzioni aventi come dominio e immagine delle varietà che non sono necessariamente immerse in uno spazio euclideo.

DEFINIZIONE 6.3. *Siano M ed N due varietà differenziabili tali che $\dim M = n$ e $\dim N = k$. Una funzione $F: M \longrightarrow N$ è **differenziabile** se, per ogni punto $p \in M$, esistono due parametrizzazioni locali (U, φ) di M e (V, ψ) di N tali che:*

- (1) $p \in \varphi(U)$, $F(p) \in \psi(V)$ e $F(\varphi(U)) \subseteq \psi(V)$,
- (2) la funzione

$$\hat{F} \stackrel{\text{def.}}{=} \psi^{-1} \circ F \circ \varphi: U \subseteq \mathbb{R}^n \longrightarrow V \subseteq \mathbb{R}^k$$

è differenziabile.

\hat{F} si dice **espressione locale** della funzione F .

La definizione precedente non dipende dalla coppia di parametrizzazioni locali scelte. Siano (U', φ') e (V', ψ') altre due parametrizzazioni locali che verificano la condizione (1) della definizione precedente. Esse soddisfano anche alla condizione (2) dato che la funzione

$$\begin{aligned} \hat{F}' &= (\psi'^{-1})' \circ F \circ \varphi' = (\psi'^{-1})' \circ (\psi \circ \psi^{-1}) \circ F \circ (\varphi \circ \varphi^{-1}) \circ \varphi' \\ &= [(\psi'^{-1})' \circ \psi] \circ (\psi^{-1} \circ F \circ \varphi) \circ (\varphi^{-1} \circ \varphi') \end{aligned}$$

è differenziabile, essendo il prodotto di tre funzioni differenziabili.

DEFINIZIONE 6.4. *Una funzione differenziabile $F: M \longrightarrow N$ è un **diffeomorfismo** se*

- (1) F è un omeomorfismo,
- (2) F e F^{-1} sono differenziabili.

OSSERVAZIONE 6.3. Due varietà differenziabili diffeomorfe si considerano come un'unica varietà dato che il diffeomorfismo identifica sia le topologie che le strutture differenziabili. Inoltre, data una qualsiasi parametrizzazione locale (U, φ) di M , la funzione $\varphi: U \longrightarrow \varphi(U)$ è un diffeomorfismo tra $U \subseteq \mathbb{R}^n$ e $\varphi(U) \subseteq M$. Questa osservazione giustifica l'affermazione intuitiva che una varietà differenziabile si ottiene 'incollando' porzioni di uno spazio euclideo.

Una categoria molto importante di varietà differenziabili è costituita dai gruppi di Lie, vale a dire da quei gruppi che sono anche varietà in cui l'operazione prodotto è differenziabile. Più precisamente, si ha:

DEFINIZIONE 6.5. *Un gruppo G è un **gruppo di Lie** se*

- (1) G è una varietà differenziabile di dimensione n ;
- (2) le funzioni

$$\begin{aligned} G \times G &\longrightarrow G, & (a, b) &\mapsto ab, \\ G &\longrightarrow G, & a &\mapsto a^{-1}, \end{aligned}$$

sono differenziabili.

OSSERVAZIONE 6.4. Nella condizione (2), $G \times G$ ha la struttura di varietà differenziabile prodotto (vedi l'Esempio 6.5). Inoltre, essa può essere sostituita dalla condizione equivalente: la funzione

$$G \times G \longrightarrow G, \quad (a, b) \mapsto ab^{-1}$$

è differenziabile.

OSSERVAZIONE 6.5. Se G è un gruppo di Lie e $a \in G$ è un elemento fissato, le proprietà algebriche del gruppo sono caratterizzate dalle seguenti funzioni:

- la *traslazione sinistra* $L_a: G \longrightarrow G, \quad b \mapsto L_a(b) = ab,$
- la *traslazione destra* $R_a: G \longrightarrow G, \quad a \mapsto R_a(b) = ba,$
- il *coniugio* $I_a: G \longrightarrow G, \quad a \mapsto I_a(b) = aba^{-1}.$

Si verifica che queste tre funzioni sono *diffeomorfismi* di G .

ESEMPIO 6.6. Si è visto che il gruppo delle matrici invertibili $GL(n, \mathbb{R})$ è una varietà differenziabile di dimensione n^2 (vedi l'Esempio 6.3). Dato che il prodotto di due matrici e l'inversa di una matrice sono entrambi funzioni polinomiali degli elementi delle matrici considerate, la condizione (2) è verificata e $GL(n, \mathbb{R})$ è un gruppo di Lie.

Vale il seguente teorema, la cui dimostrazione non è elementare.

TEOREMA 6.1. *Se G è un gruppo di Lie e H è un sottogruppo **chiuso** di G , allora H ha una struttura di gruppo Lie indotta da quella di G .*

La conseguenza più importante di questo teorema è che i sottogruppi classici di matrici, essendo sottogruppi chiusi di $GL(n, \mathbb{R})$ o $GL(n, \mathbb{C})$, sono gruppi di Lie. In particolare, sono gruppi di Lie $SL(n, \mathbb{R}), SL(n, \mathbb{C}), O(n), O(k, n-k), SO(n), O(k, n-k), U(n), SU(n), Sp(n), Sp(n, \mathbb{R}), Sp(n, \mathbb{C})$.

Un caso importante di funzioni differenziabili è quello delle funzioni a valori reali. Più precisamente, data una varietà differenziabile M , una funzione $f: M \longrightarrow \mathbb{R}$ è differenziabile se, per ogni parametrizzazione locale (U, φ) , la sua espressione locale $\hat{f} = f \circ \varphi: U \longrightarrow \mathbb{R}$ è differenziabile (qui si considera \mathbb{R} dotato della struttura differenziabile dell'Esempio 6.1).

Si noti che, se $A \subseteq M$ è un *aperto*, ha senso considerare funzioni differenziabili 'locali', cioè definite su A .

Indichiamo con

$$\mathcal{C}^\infty(M) = \{f: M \longrightarrow \mathbb{R}, f \text{ differenziabile}\},$$

l'insieme delle funzioni differenziabili definite su una varietà differenziabile M ed a valori in \mathbb{R} . Si verifica facilmente che $\mathcal{C}^\infty(M)$ con le operazioni somma:

$$\forall f, g \in \mathcal{C}^\infty(M), \quad (f + g)(p) \stackrel{\text{def.}}{=} f(p) + g(p), \quad p \in M,$$

e prodotto per scalari:

$$\forall f \in \mathcal{C}^\infty(M), \lambda \in \mathbb{R}, \quad (\lambda f)(p) \stackrel{\text{def.}}{=} \lambda f(p), \quad p \in M,$$

è uno spazio vettoriale reale (di dimensione infinita). Inoltre $\mathcal{C}^\infty(M)$ con l'operazione prodotto:

$$\forall f, g \in \mathcal{C}^\infty(M), \quad (fg)(p) \stackrel{\text{def.}}{=} f(p)g(p), \quad p \in M,$$

ha la struttura di *algebra commutativa con unità* (vedi la Definizione 3.5).

4. Vettori tangenti e spazio vettoriale tangente

In questa sezione verrà introdotta la nozione di vettore tangente ad una varietà differenziabile e si proverà che l'insieme dei vettori tangenti in un punto forma uno spazio vettoriale che ha la stessa dimensione della varietà.

Dato che, generalmente, una varietà non è immersa in uno spazio euclideo, la definizione di vettore tangente deve fare a meno del solito aspetto geometrico e pertanto deve basarsi su considerazioni di carattere algebrico.

Iniziamo col richiamare la nozione di *derivata direzionale* in \mathbb{R}^n : essa servirà da motivazione per i concetti più astratti che verranno dati successivamente. Fissati un vettore $\mathbf{v} = (v_1, \dots, v_n)$ e un punto $p_0 = (x_1^0, \dots, x_n^0)$ di \mathbb{R}^n , la derivata direzionale di una funzione differenziabile $f: \mathbb{R}^n \rightarrow \mathbb{R}$, nel punto p_0 e nella direzione di \mathbf{v} è lo scalare

$$\mathbf{v}_{p_0}(f) \stackrel{\text{def.}}{=} \left. \frac{d}{dt} f(p_0 + t\mathbf{v}) \right|_{t=0},$$

dove $p_0 + t\mathbf{v} = (x_1^0 + tv_1, \dots, x_n^0 + tv_n)$ è la retta in \mathbb{R}^n passante per p_0 e parallela al vettore \mathbf{v} . Tenuto conto della regola di derivazione delle funzioni composte, si ha:

$$\mathbf{v}_{p_0}(f) = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)_{p_0} v_i = \sum_{i=1}^n \left(v_i \frac{\partial}{\partial x_i} \right)_{p_0} (f).$$

Si noti che la funzione (o, meglio, l'operatore)

$$\mathbf{v}_{p_0}: \mathcal{C}^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}, \quad f \mapsto \mathbf{v}_{p_0}(f)$$

- è *lineare*, cioè

$$\mathbf{v}_{p_0}(\lambda f + \mu g) = \lambda \mathbf{v}_{p_0}(f) + \mu \mathbf{v}_{p_0}(g), \quad \forall f, g \in \mathcal{C}^\infty(\mathbb{R}^n), \forall \lambda, \mu \in \mathbb{R};$$

- verifica la *regola di Leibnitz*

$$\mathbf{v}_{p_0}(fg) = \mathbf{v}_{p_0}(f)g(p_0) + f(p_0)\mathbf{v}_{p_0}(g), \quad \forall f, g \in \mathcal{C}^\infty(\mathbb{R}^n).$$

Un operatore soddisfa alle due proprietà precedenti viene detto *derivazione* dell'algebra $\mathcal{C}^\infty(\mathbb{R}^n)$. In conclusione: *ogni vettore di \mathbb{R}^n può essere visto come una derivazione dell'algebra delle funzioni differenziabili*. Questo punto di vista consente di estendere la nozione di vettore tangente ad una varietà differenziabile.

DEFINIZIONE 6.6. Sia M una varietà differenziabile di dimensione n . Fissato un punto $p \in M$, una **derivazione** dell'algebra $C^\infty(M)$ in p è una funzione

$$X_p: C^\infty(M) \longrightarrow \mathbb{R}, \quad f \mapsto X_p(f),$$

che verifica le seguenti condizioni:

(1) è lineare, cioè

$$X_p(\lambda f + \mu g) = \lambda X_p(f) + \mu X_p(g), \quad \forall f, g \in C^\infty(M), \forall \lambda, \mu \in \mathbb{R};$$

(2) vale la regola di Leibnitz

$$X_p(fg) = X_p(f)g(p_0) + f(p_0)X_p(g), \quad \forall f, g \in C^\infty(M).$$

PROPRIETÀ 6.2. L'insieme T_pM di tutte le derivazioni di $C^\infty(M)$ è uno spazio vettoriale, detto **spazio tangente** alla varietà M nel punto p .

DIMOSTRAZIONE. Date due derivazioni $X_p, Y_p \in T_pM$, la loro somma è definita da:

$$(X_p + Y_p)(f) \stackrel{\text{def.}}{=} X_p(f) + Y_p(f), \quad \forall f \in C^\infty(M).$$

Il prodotto di una derivazione X_p per uno scalare $\lambda \in \mathbb{R}$ è dato da

$$(\lambda X_p)(f) \stackrel{\text{def.}}{=} \lambda X_p(f), \quad \forall f \in C^\infty(M).$$

Occorre verificare (esercizio!) che $X_p + Y_p$ e λX_p sono derivazioni di $C^\infty(M)$ e che valgono gli assiomi che definiscono uno spazio vettoriale. Si noti che il vettore nullo coincide con la derivazione nulla: quella che associa ad ogni $f \in C^\infty(M)$ lo scalare zero. \square

Per provare che $\dim T_pM = \dim M$ occorre costruire esplicitamente una base di T_pM . Sia (U, φ) una parametrizzazione locale tale che $p \in \varphi(U)$ e siano x_1, \dots, x_n le relative funzioni coordinate locali. Consideriamo gli operatori

$$\left(\frac{\partial}{\partial x_i} \right)_p : C^\infty(M) \longrightarrow \mathbb{R}, \quad i = 1, \dots, n$$

definiti come segue:

$$\left(\frac{\partial f}{\partial x_i} \right)_p \stackrel{\text{def.}}{=} \left(\frac{\partial (f \circ \varphi)}{\partial x_i} \right)_{\varphi^{-1}(p)}.$$

Si verifica immediatamente che tali operatori sono derivazioni di $C^\infty(M)$ e quindi

$$\left(\frac{\partial}{\partial x_i} \right)_p \in T_pM, \quad i = 1, \dots, n.$$

Per dimostrare che individuano una base di T_pM sono necessari due lemmi preliminari.

LEMMA 6.3. Se $X_p \in T_pM$ e c è una costante reale (pensata come funzione costante su M), allora $X_p(c) = 0$.

DIMOSTRAZIONE. A causa della linearità di X_p si ha:

$$X_p(c) = X_p(c \cdot 1) = cX_p(1).$$

Inoltre, tenuto conto della regola di Leibnitz, si trova:

$$X_p(1) = X_p(1 \cdot 1) = X_p(1) \cdot 1 + 1 \cdot X_p(1) = 2X_p(1)$$

e quindi $X_p(1) = 0$. □

LEMMA 6.4. *Sia $f: M \rightarrow \mathbb{R}$ una funzione differenziabile. Fissato un punto $p \in M$, se (U, φ) è una parametrizzazione locale tale che $p \in \varphi(U)$, siano x_1, \dots, x_n le relative funzioni coordinate locali. Allora esistono n funzioni differenziabili*

$$g_i: \varphi(U) \subseteq M \rightarrow \mathbb{R}, \quad i = 1, \dots, n,$$

che verificano le seguenti condizioni:

(1) se $q \in \varphi(U)$ e $\varphi^{-1}(p) = (x_1^0, \dots, x_n^0)$, si ha

$$f(q) = f(p) + \sum_{i=1}^n (x_i(q) - x_i^0) g_i(q);$$

(2) inoltre

$$g_i(p) = \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\varphi^{-1}(p)}.$$

DIMOSTRAZIONE. Se $q \in \varphi(U)$, sia

$$\alpha(t) = \varphi^{-1}(p) + t[\varphi^{-1}(q) - \varphi^{-1}(p)]$$

la retta in \mathbb{R}^n passante per i punti $\varphi^{-1}(p)$ e $\varphi^{-1}(q) = (x_1(q), \dots, x_n(q))$. Le equazioni parametriche di α sono:

$$\alpha^i(t) = x_i^0 + t(x_i(q) - x_i^0), \quad i = 1, \dots, n.$$

Dato che $\alpha(0) = \varphi^{-1}(p)$ e $\alpha(1) = \varphi^{-1}(q)$, si ha: $p = \varphi[\alpha(0)]$, $q = \varphi[\alpha(1)]$ e quindi

$$f(q) - f(p) = (f \circ \varphi \circ \alpha)(1) - (f \circ \varphi \circ \alpha)(0) = \int_0^1 \frac{d(f \circ \varphi \circ \alpha)}{dt} dt.$$

Pertanto

$$\begin{aligned} f(q) &= f(p) + \int_0^1 \frac{d[(f \circ \varphi) \circ \alpha]}{dt} dt = \int_0^1 \sum_{i=1}^n \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\alpha(t)} \frac{d\alpha^i}{dt} dt \\ &= f(p) + \sum_{i=1}^n (x_i(q) - x_i^0) \int_0^1 \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\alpha(t)} dt \\ &= f(p) + \sum_{i=1}^n (x_i(q) - x_i^0) g_i(q), \end{aligned}$$

dove la funzione

$$g_i(q) \stackrel{\text{def.}}{=} \int_0^1 \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\alpha(t)} dt, \quad i = 1, \dots, n,$$

è differenziabile e dipende solamente da $q \in \varphi(U)$ (ricordare che $\alpha(t) = \varphi^{-1}(p) + t[\varphi^{-1}(q) - \varphi^{-1}(p)]$, p è fisso e t è la variabile d'integrazione).

Infine, si ha:

$$\begin{aligned} g_i(p) &= \int_0^1 \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\alpha(0)} dt = \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\varphi^{-1}(p)} \int_0^1 dt \\ &= \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\varphi^{-1}(p)}, \quad i = 1, \dots, n. \end{aligned}$$

□

Siamo ora in grado di provare il risultato più importante di questa sezione.

TEOREMA 6.5. *L'insieme di vettori tangenti $\left(\left(\frac{\partial}{\partial x_1} \right)_p, \dots, \left(\frac{\partial}{\partial x_n} \right)_p \right)$ è una base dello spazio vettoriale tangente $T_p M$ e quindi $\dim T_p M = \dim M$.*

DIMOSTRAZIONE. Iniziamo col verificare che i vettori $\left(\frac{\partial}{\partial x_1} \right)_p, \dots, \left(\frac{\partial}{\partial x_n} \right)_p$ sono generatori di $T_p M$. Sia X_p un generico vettore di $T_p M$. Tenuto conto dei Lemmi 6.3 e 6.4, se $f \in C^\infty(M)$, si ha:

$$\begin{aligned} X_p(f) &= X_p \left\{ f(p) + \sum_{i=1}^n (x_i(q) - x_i^0) g_i(q) \right\} \\ &= X_p(f(p)) + \sum_{i=1}^n X_p \left\{ (x_i(q) - x_i^0) g_i(q) \right\} \\ &= \sum_{i=1}^n \left\{ X_p(x_i(q) - x_i^0) g_i(p) + (x_i(p) - x_i^0) X_p(g_i(q)) \right\} \\ &= \sum_{i=1}^n X_p(x_i) \left(\frac{\partial(f \circ \varphi)}{\partial x_i} \right)_{\varphi^{-1}(p)} = \left[\sum_{i=1}^n X_p(x_i) \left(\frac{\partial}{\partial x_i} \right)_p \right] (f). \end{aligned}$$

Di conseguenza, il vettore

$$X_p = \sum_{i=1}^n X_p(x_i) \left(\frac{\partial}{\partial x_i} \right)_p$$

risulta essere combinazione lineare di $\left(\frac{\partial}{\partial x_1} \right)_p, \dots, \left(\frac{\partial}{\partial x_n} \right)_p$.

Verifichiamo che i vettori precedenti sono linearmente indipendenti. Supponiamo che

$$\sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i} \right)_p = O_p,$$

dove gli $a^i \in \mathbb{R}$ e O_p è il vettore nullo di $T_p M$, cioè la derivazione nulla di $C^\infty(M)$. Tenuto conto del fatto che

$$\left(\frac{\partial}{\partial x_i} \right)_p (x_j) = \left(\frac{\partial(x_j \circ \varphi)}{\partial x_i} \right)_{\varphi(p)} = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}, \quad i, j = 1, \dots, n,$$

si ha:

$$0 = O_p(x_j) = \sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i} \right)_p (x_j) = \sum_{i=1}^n a^i \delta_{ij} = a^j, \quad j = 1, \dots, n.$$

□

OSSERVAZIONE 6.6. Dal teorema precedente si ricava che ogni vettore tangente $X_p \in T_p M$ si può esprimere come combinazione lineare

$$X_p = \sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i} \right)_p, \quad a^i = X_p(x_i) \in \mathbb{R}.$$

Si noti che le componenti della combinazione lineare si ottengono applicando l'operatore X_p alle funzioni coordinate locali.

5. Differenziale di una funzione tra varietà differenziabili

In questa sezione la classica nozione di differenziale di una funzione reale di variabile reale viene estesa alle funzioni differenziabili tra varietà. Verrà data una definizione algebrica di differenziale; successivamente si vedrà che è equivalente ad una costruzione di carattere più geometrico.

Siano M ed N due varietà differenziabili di dimensione n e k , rispettivamente, e sia $F: M \rightarrow N$ una funzione differenziabile. Fissato un punto $p \in M$, vogliamo associare ad ogni vettore tangente $X_p \in T_p M$ un vettore $(F_*)_p(X_p)$, tangente nel punto $F(p)$ alla varietà N . A tale scopo, occorre specificare come $(F_*)_p(X_p)$ opera sugli elementi di $C^\infty(N)$, cioè sulle funzioni differenziabili definite su N e a valori reali. Poniamo pertanto

$$[(F_*)_p(X_p)](g) \stackrel{\text{def.}}{=} X_p(g \circ F), \quad \forall g \in C^\infty(N).$$

ESERCIZIO 6.7. Verificare che $(F_*)_p(X_p)$ è una derivazione di $C^\infty(N)$ e quindi $(F_*)_p(X_p) \in T_{F(p)}N$.

ESERCIZIO 6.8. Provare che la funzione

$$(F_*)_p: T_p M \rightarrow T_{F(p)}N, \quad X_p \mapsto (F_*)_p(X_p),$$

è lineare.

ESERCIZIO 6.9. Date tre varietà differenziabili M , N , P e due funzioni differenziabili

$$F: M \rightarrow N, \quad G: N \rightarrow P,$$

verificare che, per ogni punto $p \in M$, si ha:

$$(G \circ F_*)_p = (G_*)_{F(p)} \circ (F_*)_p.$$

DEFINIZIONE 6.7. L'applicazione lineare $(F_*)_p$ si dice **differenziale di F nel punto p** .

Vogliamo ora determinare l'espressione in coordinate locali del differenziale. Siano:

- (U, φ) una parametrizzazione locale tale che $p \in \varphi(U)$, con coordinate locali (x_1, \dots, x_n) ;
- (V, ψ) una parametrizzazione locale tale che $F(p) \in \psi(V)$ e $F(\varphi(U)) \subseteq V$, con coordinate locali (y_1, \dots, y_k) .

Se $\hat{F} = \psi^{-1} \circ F \circ \varphi: U \subseteq \mathbb{R}^n \longrightarrow V \subseteq \mathbb{R}^k$ è l'espressione locale di F , si ha:

$$\hat{F}(x_1, \dots, x_n) = (y_1, \dots, y_k) = (f_1(x_1, \dots, x_n), \dots, f_k(x_1, \dots, x_n))$$

e le funzioni

$$y_j = f_j(x_1, \dots, x_n), \quad j = 1, \dots, k,$$

sono differenziabili. Se

- $\mathcal{B} = \left(\left(\frac{\partial}{\partial x_1} \right)_p, \dots, \left(\frac{\partial}{\partial x_n} \right)_p \right)$ è una base di $T_p M$,
- $\mathcal{B}' = \left(\left(\frac{\partial}{\partial y_1} \right)_{F(p)}, \dots, \left(\frac{\partial}{\partial y_k} \right)_{F(p)} \right)$ è una base di $T_{F(p)} N$,

per trovare la matrice dell'applicazione lineare $(F_*)_p$, rispetto alle basi \mathcal{B} e \mathcal{B}' , occorre esprimere il vettore

$$(F_*)_p \left(\frac{\partial}{\partial x_i} \right)_p, \quad i = 1, \dots, n,$$

come combinazione lineare dei vettori di \mathcal{B}' . Le sue componenti sono date da (si ricordi l'Osservazione 6.6 di pag. 140):

$$\begin{aligned} \left[(F_*)_p \left(\frac{\partial}{\partial x_i} \right)_p \right] (y_j) &= \left(\frac{\partial}{\partial x_i} \right)_p (y_j \circ F) = \left(\frac{\partial (y_j \circ F \circ \varphi)}{\partial x_i} \right)_{\varphi^{-1}(p)} \\ &= \left(\frac{\partial f_j}{\partial x_i} \right)_{\varphi^{-1}(p)}, \end{aligned}$$

e quindi

$$\left[(F_*)_p \left(\frac{\partial}{\partial x_i} \right)_p \right] = \sum_{j=1}^k \left(\frac{\partial f_j}{\partial x_i} \right)_{\varphi^{-1}(p)} \left(\frac{\partial}{\partial y_j} \right)_{F(p)}.$$

Vale pertanto la

PROPRIETÀ 6.6. *La matrice del differenziale della funzione F , rispetto alle basi \mathcal{B} e \mathcal{B}' , è la matrice Jacobiana dell'espressione locale di F .*

OSSERVAZIONE 6.7. Dalla Proprietà precedente si deduce che, nel caso di una funzione differenziabile $F: \mathbb{R}^n \longrightarrow \mathbb{R}^k$, l'applicazione lineare F_* coincide con il differenziale 'classico'.

Vediamo ora di dare un'interpretazione geometrica dei vettori tangenti. Innanzitutto occorre la seguente

DEFINIZIONE 6.8. *Sia M una varietà differenziabile di dimensione n . Una **curva differenziabile** su M è una funzione differenziabile*

$$\alpha: (a, b) \subseteq \mathbb{R} \longrightarrow M, \quad t \mapsto \alpha(t)$$

(si pensa l'intervallo (a, b) come sottovarietà aperta di \mathbb{R}).

Sia (U, φ) una parametrizzazione locale di M , con coordinate locali (x_1, \dots, x_n) , tale che $\varphi(U) \cap \alpha(a, b) \neq \emptyset$. Allora, se $\alpha(t) \in \varphi(U) \cap \alpha(a, b)$, si ha:

$$\alpha(t) = (x_1(t), \dots, x_n(t))$$

dove

$$x_i(t) = (x_i \circ \alpha)(t), \quad i = 1, \dots, n.$$

Fissato un punto $t_0 \in (a, b)$ tale che $\alpha(t_0) \in \varphi(U) \cap \alpha(a, b)$, si ha:

- una base di $T_{t_0}(a, b)$ è data dall'operatore $\left(\frac{d}{dt}\right)_{t_0}$;
- una base di $T_{\alpha(t_0)}M$ è $\left(\left(\frac{\partial}{\partial x_1}\right)_{\alpha(t_0)} \cdots, \left(\frac{\partial}{\partial x_n}\right)_{\alpha(t_0)}\right)$;
- la matrice del differenziale $(\alpha_*)_{t_0}$, rispetto alla basi precedenti, è la matrice Jacobiana dell'espressione locale di α :

$$\left(\left(\frac{dx_1}{dt}\right)_{t_0} \cdots \left(\frac{dx_n}{dt}\right)_{t_0}\right);$$

- il vettore

$$\alpha'(t_0) \stackrel{\text{def.}}{=} (\alpha_*)_{t_0} \left(\frac{d}{dt}\right)_{t_0} = \sum_{i=1}^n \left(\frac{dx_i}{dt}\right)_{t_0} \left(\frac{\partial}{\partial x_i}\right)_{\alpha(t_0)} \in T_{\alpha(t_0)}M$$

prende il nome di *vettore tangente alla curva α nel punto $\alpha(t_0)$* .

Viceversa, vale la

PROPRIETÀ 6.7. *Dato un vettore tangente $X_p \in T_pM$, esistono infinite curve differenziabili $\alpha: (-\epsilon, \epsilon) \rightarrow M$ tali che:*

$$\alpha(0) = p, \quad \alpha'(0) = X_p.$$

In conclusione, ogni vettore tangente può essere visto come vettore tangente ad una curva differenziabile.

Applichiamo questo punto di vista per trovare lo spazio tangente nell'elemento neutro del gruppo di Lie $GL(n, \mathbb{R})$. Sia $I \in GL(n, \mathbb{R})$ la matrice unità di ordine n . Data una qualsiasi matrice $A \in \mathbb{R}^{n,n}$, consideriamo la curva (in $\mathbb{R}^{n,n}$)

$$\alpha(t) = I + tA, \quad t \in \mathbb{R}.$$

La funzione determinante è continua e, per il principio della conservazione del segno, si ha che $\det(I + tA) \neq 0$, se t è sufficientemente piccolo. Ciò significa che $\alpha(t) \in GL(n, \mathbb{R})$, per $|t| < \epsilon$. Dato che $\alpha(0) = I$ e $\alpha'(0) = A$, si conclude che $A \in T_I GL(n, \mathbb{R})$ e quindi $\mathbb{R}^{n,n} \subseteq T_I GL(n, \mathbb{R})$. Poichè $\dim \mathbb{R}^{n,n} = n^2 = \dim T_I GL(n, \mathbb{R})$, vale la

PROPRIETÀ 6.8. *Lo spazio tangente al gruppo di Lie $GL(n, \mathbb{R})$ nell'elemento neutro I coincide con $\mathbb{R}^{n,n}$.*

Sia ora $F: M \rightarrow N$ una funzione differenziabile. Fissato un punto $p \in M$, si consideri un vettore tangente $X_p \in T_pM$. Se $\alpha(t)$ è una curva differenziabile su M tale che $\alpha(0) = p$ e $\alpha'(0) = X_p$, allora $(F \circ \alpha)(t)$ è una curva differenziabile su N tale che $(F \circ \alpha)(0) = F(p)$ e $(F \circ \alpha)'(0) \in T_{F(p)}N$. Un'interpretazione geometrica del differenziale è data dalla seguente

PROPRIETÀ 6.9.

$$(F_*)_p(X_p) = (F \circ \alpha)'(0).$$

DIMOSTRAZIONE. Introdotta delle opportune parametrizzazioni locali (come all'inizio di questa sezione), si ha:

$$\begin{aligned} \varphi^{-1}(p) &= (x_1^0, \dots, x_n^0), & X_p &= \sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i}\right)_p, \\ \alpha(t) &= (x_1(t), \dots, x_n(t)), & t &\in (-\epsilon, \epsilon). \end{aligned}$$

Il fatto che $\alpha(0) = p$ e $\alpha'(0) = X_p$ si traduce nelle condizioni seguenti:

$$x_i(0) = x_i^0, \quad \left(\frac{dx_i(t)}{dt} \right)_0 = a^i, \quad i = 1, \dots, n.$$

L'espressione locale di F è:

$$\hat{F}(x_1, \dots, x_n) = (y_1, \dots, y_k) = (f_1(x_1, \dots, x_n), \dots, f_k(x_1, \dots, x_n)),$$

e quindi $(F \circ \alpha)(t) = (y_1(t), \dots, y_k(t))$, dove

$$y_j(t) = f_j(x_1(t), \dots, x_n(t)), \quad j = 1, \dots, k.$$

Allora

$$\begin{aligned} (F \circ \alpha)(t) &= \sum_{j=1}^k \left(\frac{dy_j(t)}{dt} \right)_0 \left(\frac{\partial}{\partial y_j} \right)_{F(p)} \\ &= \sum_{j=1}^k \left[\sum_{i=1}^n \left(\frac{\partial f_j}{\partial x_i} \right)_{\varphi^{-1}(p)} \left(\frac{dx_i(t)}{dt} \right)_0 \right] \left(\frac{\partial}{\partial y_j} \right)_{F(p)} \\ &= \sum_{j=1}^k \left[\sum_{i=1}^n \left(\frac{\partial f_j}{\partial x_i} \right)_{\varphi^{-1}(p)} a^i \right] \left(\frac{\partial}{\partial y_j} \right)_{F(p)} \\ &= \sum_{i=1}^n a^i \left[\sum_{j=1}^k \left(\frac{\partial f_j}{\partial x_i} \right)_{\varphi^{-1}(p)} \left(\frac{\partial}{\partial y_j} \right)_{F(p)} \right] = \sum_{i=1}^n a^i (F_*)_p \left(\frac{\partial}{\partial x_i} \right)_p \\ &= (F_*)_p \left[\sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i} \right)_p \right] = (F_*)_p(X_p). \end{aligned}$$

□

6. Fibrati tangenti e campi vettoriali

In questa sezione verrà introdotta la nozione di campo vettoriale differenziabile su una varietà. A tal scopo, occorre definire una struttura differenziabile sull'insieme di tutti i vettori tangenti.

Sia M una varietà differenziabile di dimensione n . Nel seguito, dati un punto $p \in M$ e un vettore tangente $X_p \in T_p M$, il vettore X_p verrà identificato con la coppia (p, X_p) . Consideriamo l'insieme

$$TM \stackrel{\text{def.}}{=} \bigcup_{p \in M} T_p M = \{(p, X_p), p \in M, X_p \in T_p M\}$$

formato da tutti i vettori tangenti. I sottoinsiemi di TM del tipo

$$\bigcup_{p \in A} T_p M = \{(p, X_p), p \in A, X_p \in T_p M\},$$

dove A è un aperto di M , individuano una base per la topologia di TM . Consideriamo una parametrizzazione locale (U, φ) di M . Se $p \in \varphi(U)$ e $X_p \in T_p M$, allora

$$X_p = \sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i} \right)_p, \quad a^i \in \mathbb{R},$$

e quindi possiamo identificare (p, X_p) con i $2n$ numeri reali

$$(x_1(p), \dots, x_n(p), a^1, \dots, a^n).$$

Questa osservazione è fondamentale per la costruzione delle parametrizzazioni locali su TM : dato che $U \times \mathbb{R}^n$ è un aperto di \mathbb{R}^{2n} , la coppia $(U \times \mathbb{R}^n, \psi)$, dove

$$\begin{aligned} \psi: U \times \mathbb{R}^n \subseteq \mathbb{R}^{2n} &\longrightarrow TM, \\ \psi(x_1, \dots, x_n, a^1, \dots, a^n) &= \left(p, \sum_{i=1}^n a^i \left(\frac{\partial}{\partial x_i} \right)_p \right), \quad p = \varphi(x_1, \dots, x_n), \end{aligned}$$

individua una parametrizzazione locale su TM . Al variare di (U, φ) si ottiene una struttura differenziabile su TM . La varietà differenziabile TM si dice **fibrato tangente** di M ed ha dimensione $2n$.

Un campo vettoriale su (un aperto di) M è una funzione che associa ad ogni punto un vettore tangente che dipende in modo differenziabile dal punto. Più precisamente, si ha:

DEFINIZIONE 6.9. Sia $A \subseteq M$ un aperto (considerato come sottovarietà aperta di M). Un **campo vettoriale differenziabile** su A è una funzione differenziabile

$$X: A \subseteq M \longrightarrow TM, \quad p \in A \mapsto X(p) = X_p \in T_p M.$$

Molto spesso, i campi vettoriali prendono il nome di *sezioni* (locali) del fibrato tangente.

ESEMPIO 6.10. Se (U, φ) è una parametrizzazione locale, i **campi coordinati locali** sono i campi vettoriali

$$\left(\frac{\partial}{\partial x_i} \right) : \varphi(U) \longrightarrow TM, \quad p \in \varphi(U) \mapsto \left(\frac{\partial}{\partial x_i} \right)_p, \quad i = 1, \dots, n.$$

OSSERVAZIONE 6.8. Se $X: M \longrightarrow TM$ è un campo vettoriale su M , data una parametrizzazione locale (U, φ) , X si può esprimere *localmente* come combinazione lineare dei campi coordinati locali:

$$X = \sum_{i=1}^n X^i(p) \frac{\partial}{\partial x_i}, \quad p \in \varphi(U),$$

dove, in questo caso, ogni $X^i(p)$ è una *funzione* differenziabile sull'aperto $\varphi(U)$.

Molto spesso nelle applicazioni i campi vettoriali vengono visti sotto un altro punto di vista. Dati un campo vettoriale X e una funzione differenziabile $f \in \mathcal{C}^\infty(M)$ su M , possiamo costruire una nuova funzione Xf nel modo seguente:

$$Xf: M \longrightarrow \mathbb{R}, \quad p \in M \mapsto (Xf)(p) \stackrel{\text{def.}}{=} X_p(f).$$

Si è visto che, data una parametrizzazione locale (U, φ) , si ha

$$X = \sum_{i=1}^n X^i(p) \frac{\partial}{\partial x_i}, \quad p \in \varphi(U),$$

e quindi, se $f \in C^\infty(M)$, la funzione

$$Xf = \sum_{i=1}^n X^i(p) \frac{\partial f}{\partial x_i},$$

risulta essere la somma di funzioni differenziabili. Ciò significa che anche Xf è differenziabile e quindi il campo vettoriale X individua un operatore su $C^\infty(M)$ (che si indica con la stessa lettera):

$$X: C^\infty(M) \longrightarrow C^\infty(M), \quad f \mapsto Xf.$$

Si verifica immediatamente che tale operatore

- è lineare: $X(\lambda f + \mu g) = \lambda(Xf) + \mu(Xg)$, $\forall f, g \in C^\infty(M)$, $\forall \lambda, \mu \in \mathbb{R}$;
- vale la regola di Leibnitz: $X(fg) = (Xf)g + f(Xg)$, $\forall f, g \in C^\infty(M)$.

Le proprietà precedenti provano che X è una *derivazione* dell'algebra $C^\infty(M)$. In conclusione: *dare un campo vettoriale differenziabile equivale a dare una derivazione dell'algebra delle funzioni differenziabili.*

A questo punto, consideriamo l'insieme $\mathfrak{X}(M)$ di tutti i campi vettoriali differenziabili su M . Si verifica immediatamente che $\mathfrak{X}(M)$ è uno spazio vettoriale reale (di dimensione infinita) con le operazioni:

- somma: se $X, Y \in \mathfrak{X}(M)$, il campo vettoriale $X + Y$ è definito da

$$(X + Y)(f) \stackrel{\text{def.}}{=} Xf + Yf, \quad \forall f \in C^\infty(M),$$

- prodotto per scalari: se $X \in \mathfrak{X}(M)$ e $\lambda \in \mathbb{R}$, il campo vettoriale λX è definito da

$$(\lambda X)(f) \stackrel{\text{def.}}{=} \lambda(Xf), \quad \forall f \in C^\infty(M).$$

Inoltre è possibile introdurre un prodotto tra campi vettoriali nel modo seguente: il prodotto (o 'bracket') dei campi X e Y è il campo vettoriale $[X, Y]$ definito da

$$[X, Y](f) \stackrel{\text{def.}}{=} X(Yf) - Y(Xf), \quad \forall f \in C^\infty(M).$$

Valgono le seguenti proprietà di verifica (quasi) immediata:

- anticommutativa:

$$[X, Y] = -[Y, X], \quad \forall X, Y \in \mathfrak{X}(M);$$

- linearità rispetto agli scalari:

$$[\lambda X + \mu Y, Z] = \lambda[X, Z] + \mu[Y, Z], \quad \forall X, Y, Z \in \mathfrak{X}(M), \lambda, \mu \in \mathbb{R};$$

-

$$[fX, gY] = fg[X, Y] - f(Xg)Y - g(Yf)X, \\ \forall X, Y \in \mathfrak{X}(M), \quad \forall f, g \in C^\infty(M);$$

- identità di Jacobi:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = O, \quad \forall X, Y, Z \in \mathfrak{X}(M),$$

(O è il campo vettoriale nullo: assegna ad ogni $p \in M$ il vettore nullo di $T_p M$).

Le proprietà precedenti si riassumono dicendo che $\mathfrak{X}(M)$ è un'algebra di Lie.

7. Algebra di Lie di un gruppo di Lie

In quest'ultima sezione si proverà che lo spazio tangente ad un gruppo di Lie G nell'*elemento neutro* ha la struttura di algebra di Lie ed è isomorfa ad una particolare sottoalgebra di Lie di $\mathfrak{X}(G)$.

Come si è visto nell'Osservazione 6.5, se $a \in G$ è un elemento fissato, le funzioni:

$$\begin{aligned} L_a: G &\longrightarrow G, & b &\longmapsto ab, \\ R_a: G &\longrightarrow G, & b &\longmapsto ba, \end{aligned}$$

sono diffeomorfismi analitici (ma non omomorfismi di gruppi), dette **traslazione sinistra** e **traslazione destra**, rispettivamente.

Si ricordi che se $F: M \longrightarrow N$ è una funzione differenziabile tra due varietà differenziabili, il differenziale di F nel punto $p \in M$ è un'applicazione lineare tra gli spazi tangenti $T_p M$ e $T_{F(p)} N$, indicata con $(F_*)_p$.

DEFINIZIONE 6.10. *Sia X un campo vettoriale su G . X si dice **invariante a sinistra (a destra)** se*

$$(L_a)_*|_b X_b = X_{ab}, \quad ((R_a)_*|_b X_b = X_{ba})$$

per ogni $a, b \in G$.

Verifichiamo che l'insieme dei campi vettoriali *invarianti a sinistra* ha la struttura di algebra di Lie ed è quindi una sottoalgebra di Lie dell'algebra $\mathfrak{X}(G)$, costituita *tutti* i campi vettoriali su G .

PROPRIETÀ 6.10. *Se X e Y sono due campi invarianti a sinistra (a destra), anche i campi $X + Y$, λX , con $\lambda \in \mathbb{R}$, e $[X, Y]$ sono invarianti a sinistra (a destra).*

DIMOSTRAZIONE. Le prime due affermazioni sono conseguenza del fatto che il differenziale è un'applicazione lineare. Per provare la terza, si osservi che, se F è un *diffeomorfismo* di G in G , dato un campo vettoriale X su G , si può definire un nuovo campo $F_* X$ nel modo seguente:

$$(F_* X)_a \stackrel{\text{def.}}{=} (F_*)_{F^{-1}(a)}(X_{F^{-1}(a)}), \quad \text{per ogni } a \in G.$$

Si prova che vale la relazione

$$F_*[X, Y] = [F_* X, F_* Y]$$

per ogni coppia di campi vettoriali X e Y su G (esercizio). La tesi segue dal fatto che un campo X è invariante a sinistra (a destra) se e solo se $(L_a)_* X = X$ (oppure $(R_a)_* X = X$), per ogni $a \in G$. Infatti, nel caso in cui X sia invariante a sinistra si ha:

$$[(L_a)_* X]_b = (L_a)_*_{(L_a)^{-1}(b)} X_{(L_a)^{-1}(b)} = (L_a)_*_{a^{-1}b} X_{a^{-1}b} = X_{a(a^{-1}b)} = X_b,$$

per ogni $a, b \in G$ (vale una dimostrazione analoga se X è invariante a destra). \square

DEFINIZIONE 6.11. *L'algebra formata dai campi invarianti a sinistra è detta **algebra di Lie del gruppo di Lie G** e si indica con \mathfrak{g} .*

OSSERVAZIONE 6.9. Naturalmente anche l'insieme dei campi invarianti a destra è un'algebra di Lie,

Il risultato principale di questa sezione è il seguente

TEOREMA 6.11. *Sia G un gruppo di Lie di dimensione n . La sua algebra di Lie \mathfrak{g} è isomorfa (come spazio vettoriale) a T_eG , spazio tangente a G nell'elemento neutro $e \in G$. Di conseguenza, anche \mathfrak{g} ha dimensione n .*

DIMOSTRAZIONE. Iniziamo col provare che, dato un vettore $\mathbf{x} \in T_eG$, esiste un *unico* campo vettoriale invariante a sinistra $X \in \mathfrak{g}$ tale che $X_e = \mathbf{x}$. Si definisce X nel modo seguente:

$$X_a = (L_a)_*|_e \mathbf{x}, \quad a \in G,$$

X è un campo invariante a sinistra poichè, tenuto conto dell'Esercizio 6.9, si ha:

$$X_{ab} = (L_{ab})_*|_e \mathbf{x} = (L_a)_*|_b (L_b)_*|_e \mathbf{x} = (L_a)_*|_b X_b, \quad a, b \in G.$$

Di conseguenza il campo X , univocamente determinato a partire dal vettore $\mathbf{x} \in T_eG$, sta in \mathfrak{g} . Non è difficile verificare che l'applicazione

$$\mathfrak{g} \longrightarrow T_eG, \quad X \longmapsto X_e,$$

è un isomorfismo di spazi vettoriali. \square

In base al Teorema precedente, si può introdurre su T_eG una struttura di algebra di Lie definendo, per ogni $\mathbf{x}, \mathbf{y} \in T_eG$,

$$[\mathbf{x}, \mathbf{y}] = [X, Y]_e,$$

dove X e Y sono gli unici campi invarianti a sinistra tali che $X_e = \mathbf{x}$, $Y_e = \mathbf{y}$. In tal modo, T_eG risulta isomorfo, *come algebra di Lie*, a \mathfrak{g} e quasi sempre le due algebre di Lie sono identificate.

Si è visto nella Proprietà 6.8 che lo spazio tangente al gruppo di Lie $GL(n, \mathbb{R})$ nell'elemento neutro I coincide (come spazio vettoriale) con $\mathbb{R}^{n,n}$. Vogliamo provare che $\mathbb{R}^{n,n}$, dotato dell'ordinaria struttura di algebra di Lie, è proprio l'algebra di Lie di $GL(n, \mathbb{R})$. Per ottenere questo risultato è necessario costruire esplicitamente un isomorfismo tra $\mathbb{R}^{n,n}$ e l'algebra di Lie dei campi invarianti a sinistra su $GL(n, \mathbb{R})$.

PROPRIETÀ 6.12. *L'algebra di Lie di $GL(n, \mathbb{R})$ è isomorfa a $\mathbb{R}^{n,n}$, dotato della solita struttura di algebra di Lie.*

DIMOSTRAZIONE. Iniziamo introducendo un sistema di coordinate globali su $GL(n, \mathbb{R})$ ponendo;

$$x_{ij}(a) = a_{ij}, \quad i, j = 1, \dots, n,$$

dove $a = (a_{ij}) \in GL(n, \mathbb{R})$. Ogni vettore tangente in I a $GL(n, \mathbb{R})$ si può scrivere nel modo seguente

$$\mathbf{x} = \sum_{i,j=1}^n A_{ij} \left(\frac{\partial}{\partial x_{ij}} \right)_I, \quad A_{ij} \in \mathbb{R}.$$

e, di conseguenza, il vettore \mathbf{x} si può identificare con la matrice $A = (A_{ij}) \in \mathbb{R}^{n,n}$.

Dato un vettore $\mathbf{x} \in T_I(GL(n, \mathbb{R}))$ vogliamo determinare l'unico campo invariante a sinistra X tale che $X_I = \mathbf{x}$. Si consideri la curva

$$\alpha(t) = I + tA, \quad t \in (-\epsilon, \epsilon) \subset \mathbb{R}.$$

A causa della continuità della funzione determinante, $\det(I + tA) \neq 0$, se t è sufficientemente piccolo, e quindi α è una curva in $GL(n, \mathbb{R})$ per la quale

$$\alpha(0) = I, \quad \alpha'(0) = \left(\frac{d\alpha}{dt} \right)_0 = \mathbf{x}.$$

Sia $a \in GL(n, \mathbb{R})$, dalla Proprietà 6.9 segue che

$$(L_a)_*|_I \mathbf{x} = \left(\frac{d\tilde{\alpha}}{dt} \right)_0,$$

dove

$$\tilde{\alpha}(t) = (L_a \circ \alpha)(t) = a + taA.$$

Dato che $\tilde{\alpha}(t)$ sta in $GL(n, \mathbb{R})$, le sue coordinate locali sono

$$\tilde{\alpha}_{ij}(t) = x_{ij}(\tilde{\alpha}(t)) = a_{ij} + t \sum_{k=1}^n a_{ik} A_{kj}$$

e quindi

$$\begin{aligned} (L_a)_*|_I \mathbf{x} &= \sum_{i,j=1}^n \left(\frac{d\tilde{\alpha}_{ij}}{dt} \right)_0 \left(\frac{\partial}{\partial x_{ij}} \right)_a = \sum_{i,j,k=1}^n a_{ik} A_{kj} \left(\frac{\partial}{\partial x_{ij}} \right)_a \\ &= \sum_{i,j,k=1}^n x_{ik}(a) A_{kj} \left(\frac{\partial}{\partial x_{ij}} \right)_a \end{aligned}$$

Dall'espressione precedente, si vede che l'unico campo invariante a sinistra X tale che $X_I = \mathbf{x}$ è dato da

$$X = \sum_{i,j,k=1}^n x_{ik} A_{kj} \frac{\partial}{\partial x_{ij}}.$$

Analogamente, considerato un secondo vettore $\mathbf{y} \in T_I(GL(n, \mathbb{R}))$ del tipo

$$\mathbf{y} = \sum_{i,j=1}^n B_{ij} \left(\frac{\partial}{\partial x_{ij}} \right)_I, \quad B_{ij} \in \mathbb{R},$$

l'unico campo invariante a sinistra Y tale che $Y_I = \mathbf{y}$ è

$$Y = \sum_{i,j,k=1}^n x_{ik} B_{kj} \frac{\partial}{\partial x_{ij}}.$$

Calcoliamo il bracket dei campi X e Y :

$$\begin{aligned}
[X, Y] &= \sum_{i,j,k=1}^n \sum_{l,r,s=1}^n \left[x_{ik} A_{kj} \frac{\partial}{\partial x_{ij}}, x_{lr} B_{rs} \frac{\partial}{\partial x_{ls}} \right] \\
&= \sum_{i,j,k,l,r,s=1}^n x_{ik} A_{kj} B_{rs} \frac{\partial x_{lr}}{\partial x_{ij}} \frac{\partial}{\partial x_{ls}} - \sum_{i,j,k,l,r,s=1}^n x_{lr} B_{rs} A_{kj} \frac{\partial x_{ik}}{\partial x_{ls}} \frac{\partial}{\partial x_{ij}} \\
&= \sum_{i,j,k,s=1}^n x_{ik} A_{kj} B_{js} \frac{\partial}{\partial x_{is}} - \sum_{j,l,r,s=1}^n x_{lr} B_{rs} A_{sj} \frac{\partial}{\partial x_{lj}} \\
&= \sum_{i,j,k,s=1}^n x_{ik} (A_{kj} B_{js} - B_{kj} A_{js}) \frac{\partial}{\partial x_{is}} = \sum_{i,k,s=1}^n x_{ik} [A, B]_{ks} \frac{\partial}{\partial x_{is}}.
\end{aligned}$$

Qui $[A, B]_{ks}$ indica l'elemento di posto (k, s) della matrice $AB - BA$, dove $A = (A_{ij})$ e $B = (B_{ij})$.

Prendendo il valore dei campi nell'elemento neutro I , si trova

$$[X, Y]_I = \sum_{i,s=1}^n [A, B]_{is} \left(\frac{\partial}{\partial x_{is}} \right)_I$$

e quindi $T_I(GL(n, \mathbb{R}))$ è isomorfo, come algebra di Lie, a $\mathbb{R}^{n,n}$. \square

OSSERVAZIONE 6.10. A questo punto siamo autorizzati ad usare la notazione $\mathfrak{gl}(n, \mathbb{R})$ per indicare $\mathbb{R}^{n,n}$, dotato della solita struttura di algebra di Lie. Vale un risultato analogo per i gruppi di Lie $GL(n, \mathbb{C})$ e $GL(n, \mathbb{K})$.

Gli altri gruppi di Lie di matrici sono sottogruppi dei vari gruppi lineari. Vale il seguente

TEOREMA 6.13. *Sia G un gruppo di Lie con algebra di Lie \mathfrak{g} . Se H è un sottogruppo di Lie di G , la sua algebra di Lie è data da:*

$$\mathfrak{h} = \{X \in \mathfrak{g} / \exp(tX) \in H, t \in \mathbb{R}\}.$$

Questo risultato permette di dimostrare, per esempio, che l'algebra di Lie di $SL(n, \mathbb{R})$ è $\mathfrak{sl}(n, \mathbb{R})$ e così via per tutti gli altri gruppi di matrici. In tal modo, risultano giustificate le notazioni scelte per le varie algebre di Lie di matrici.

Bibliografia

- [1] E. Abbena, A. Fino, G.M. Gianella, *Algebra Lineare e Geometria Analitica, Volume 1*, Aracne Editrice 2012.
- [2] J.H. Conway, D.H. Smith *On quaternions and octonions*, A.K. Peters, Natick MA 2003.
- [3] M.L. Curtis, *Matrix Groups*, Springer Verlag, New York 1984.
- [4] M. Do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice Hall, Englewood Cliffs, New Jersey 1976.
- [5] P.M. Gandini, *Appunti di Topologia*, pagina Moodle del corso.
- [6] A. Gray, E. Abbena, S. Salamon, *Modern Differential Geometry of Curves and Surfaces with Mathematica* Third edition, CRC Press, Boca Raton, FL, 2006.
- [7] B.C. Hall, *Lie Groups, Lie Algebras and Representations*, Graduate Texts in Mathematics 222, Springer, New York 2003
- [8] S. Lang, *Linear Algebra*, Addison Wesley, New York 1966.
- [9] D.S. Mackey, N. Mackey, *On the Determinant of Symplectic Matrices*, Numerical Analysis Report n. 422, Manchester Centre for Comp. Math. 2003.
- [10] Milnor J., *Curvature of left invariant metrics on Lie groups*, Adv. in Math. **21**(1976), 293–329.
- [11] B. O’Neil, *Elementary Differential Geometry*, Academic Press, New York 1997.
- [12] G. Strang, *Algebra Lineare e sue Applicazioni*, Liguori Editore 1985.
- [13] M. VanValkenburgh, *A Fun Linear Algebra Problem*, Berkeley University 2009.