

Geometria e Algebra Lineare

Carlo Petronio

Università di Pisa

Indice

1	Motivazioni	1
1.1	Forze applicate	1
1.2	Trasformazioni piane	4
1.2.1	Traslazioni	4
1.2.2	Omotetie	6
1.2.3	Rotazioni	7
1.3	Intersezione tra rette	10
1.4	Numeri di Fibonacci	11
2	Preliminari algebrici e logici	15
2.1	Numeri	15
2.1.1	Numeri naturali e interi	15
2.1.2	Numeri razionali e reali	16
2.1.3	Campi e gruppi commutativi	18
2.2	Polinomi	19
2.3	Dimostrazioni	20
2.3.1	Principio di induzione	20
2.3.2	Dimostrazioni per assurdo	21
3	Spazi vettoriali	23
3.1	Definizione	23
3.1.1	Conseguenze della definizione	24
3.1.2	Esempi	25
3.2	Sottospazi vettoriali	27
3.2.1	Esempi	28
4	Dimensione	31
4.1	Sottospazio generato e combinazioni lineari	31
4.2	Indipendenza lineare e basi	34

4.3	Dimensione	37
4.4	Estrazione e completamento	40
4.5	Formula di Grassmann	43
5	Applicazioni lineari	47
5.1	Definizione, esempi, nucleo, immagine	47
5.2	Somme dirette e proiezioni	51
5.3	Matrici e applicazioni associate	54
5.4	Spazio delle applicazioni lineari e matrice di un'applicazione	57
5.5	Cambiamenti di base e inversione	63
6	Sistemi lineari, rango, determinante	67
6.1	Rango e soluzioni di un sistema	68
6.2	Determinante	70
6.3	Matrice inversa e rango	77
7	Sottospazi affini	81
7.1	Equazioni cartesiane e parametriche	81
7.2	Rette nel piano, rette e piani nello spazio	84
7.3	Somma di sottospazi	87
8	Numeri complessi	91
8.1	Il campo complesso	91
8.2	Il piano complesso	93
8.3	Forma polare ed esponenziale complessa	95
8.4	Equazioni polinomiali	97
8.5	Spazi vettoriali complessi	99
9	Prodotti scalari e hermitiani	103
9.1	Prodotti scalari	105
9.2	Norma e ortogonalità	107
9.3	Decomposizioni ortogonali e proiezioni	111
9.3.1	Rudimenti di calcolo differenziale in più variabili	112
9.3.2	Proiezioni ortogonali	115
9.3.3	Rette e piani nello spazio	116
9.4	Matrici simmetriche e ortogonali	118
9.5	Il caso complesso	120

10 Autovalori e diagonalizzazione	123
10.1 Matrici coniugate e diagonalizzabilità	125
10.2 Il teorema spettrale	129
10.2.1 Teorema spettrale e matrice hessiana	132
10.2.2 Versione complessa del teorema spettrale	133
10.2.3 Forma canonica di matrici antisimmetriche e ortogonali	134
11 Coniche metriche e affini	137
11.1 Modelli metrici delle coniche	137
11.1.1 Piano euclideo, isometrie e trasformazioni affini	137
11.1.2 Equazioni metriche di circonferenza, ellisse, parabola e iperbole	139
11.2 Matrice di una conica e classificazione affine delle coniche non degeneri	141
12 Spazi proiettivi	147
12.1 Relazioni di equivalenza e insiemi quoziente	147
12.2 Definizione di spazio proiettivo	149
12.2.1 La retta proiettiva reale	150
12.2.2 Il piano proiettivo reale	151
12.2.3 Superfici ottenute incollando lati di una figura piana . .	153
12.2.4 Spazio proiettivo come completamento all'infinito dello spazio affine	153
12.2.5 La retta proiettiva complessa	155
12.3 Coordinate omogenee e sottospazi proiettivi	155
12.4 Cambiamenti di coordinate affini e proiettivi	157
13 Coniche e quadriche affini e proiettive	159
13.1 Completamento proiettivo e punti all'infinito	159
13.2 Classificazione delle coniche proiettive e affini	161
13.3 Modelli affini delle quadriche	164
13.3.1 Insieme vuoto, ellissoide, paraboloidi ellittico	164
13.3.2 Iperboloidi a una e a due falde	164
13.3.3 Paraboloidi iperbolico	165
13.3.4 Punti all'infinito	166

13.4	Classificazione affine delle quadriche non degeneri	167
14	Curve nel piano e nello spazio	171
14.1	Curve parametrizzate e regolari	171
14.1.1	Cambiamento di parametro	173
14.1.2	Lunghezza	174
14.1.3	Integrale su una curva di una funzione	175
14.1.4	Parametro d'arco	176
14.1.5	Teoremi di Jordan e Dini	177
14.2	Curvatura di una curva piana	180
14.2.1	Circonferenza osculatrice e curvatura	180
14.2.2	Segno della curvatura e suo calcolo	182
14.3	Curvatura e torsione di una curva nello spazio	184
14.3.1	Calcolo di curvatura e torsione	186
15	Integrazione su curve	189
15.1	1-forme e loro integrali	189
15.2	Forme esatte	191
15.2.1	Forme chiuse	193
15.3	Formula di Gauss-Green	195
15.3.1	Bordo di un insieme	196
15.3.2	Insiemi semplicemente connessi	198
16	Isometrie del piano	201
16.1	Il gruppo delle isometrie	201
16.1.1	Il piano euclideo e il gruppo delle isometrie	201
16.1.2	Tipi di isometria	202
16.1.3	Composizioni di isometrie	202
16.1.4	Il gruppo delle isometrie del piano	208
16.2	Descrizione analitica delle isometrie	209
16.2.1	Riflessioni e glissoriflessioni	210
16.2.2	Composizione di isometrie descritte analiticamente	213
17	Pavimentazioni del piano	215
17.1	Pavimentazioni regolari	215
17.2	Restrizioni sui sottogruppi delle pavimentazioni	218
17.2.1	Restrizioni sugli elementi con punti fissi	219
17.2.2	Restrizioni sulle traslazioni	223

17.3 I 17 sottogruppi delle pavimentazioni	224
17.3.1 Equivalenza di sottogruppi di pavimentazioni	224
17.3.2 Quoziente del piano rispetto a un sottogruppo	226
17.3.3 La lista	226

Capitolo 1

Motivazioni

Lo scopo fondamentale di questo corso è l'introduzione di tecniche di manipolazione di tipo *algebrico* per oggetti di natura non algebrica (come sarebbero, ad esempio, numeri, polinomi, espressioni), bensì *geometrica* (come punti, rette, piani, curve, superfici). Svilupperemo un linguaggio astratto e dei metodi che si prestano a trattare in modo unificato (e a risolvere!) problemi apparentemente molto diversi tra loro. Questi problemi hanno tutti importanti motivazioni nel “mondo reale” (ad esempio in questioni provenienti dall'ingegneria), ma noi non avremo tempo di insistere su di esse: lasciamo ai corsi successivi l'illustrazione delle applicazioni della teoria qui descritta.

In questo capitolo anticipiamo alcuni dei temi di cui ci occuperemo. Menzioniamo anche che i Capitoli 14 e 15 costituiscono una parziale eccezione rispetto a quanto appena enunciato, essendo basati su un approccio *analitico* piuttosto che *algebrico* alla geometria.

1.1 Forze applicate

Nella fisica del punto materiale rappresentiamo una forza F applicata a un punto p come una freccia (*vettore*) che parte da p (si veda Fig. 1.1). La freccia indica direzione e verso di F , e la sua lunghezza è proporzionale all'intensità di F . Se due forze sono applicate allo stesso p , l'effetto sul moto di p è quello di una singola forza, detta *risultante* e costruita con la regola del parallelogrammo, come illustrato in Fig. 1.2. Valgono le proprietà:

- $\text{ris}(F_1, F_2) = \text{ris}(F_2, F_1)$;
- $\text{ris}\left(F_1, \text{ris}(F_2, F_3)\right) = \text{ris}\left(\text{ris}(F_1, F_2), F_3\right)$.

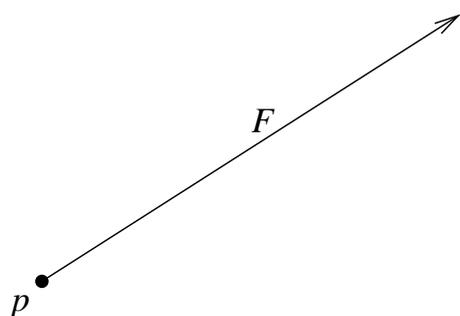


Figura 1.1: Una forza F applicata a un punto p

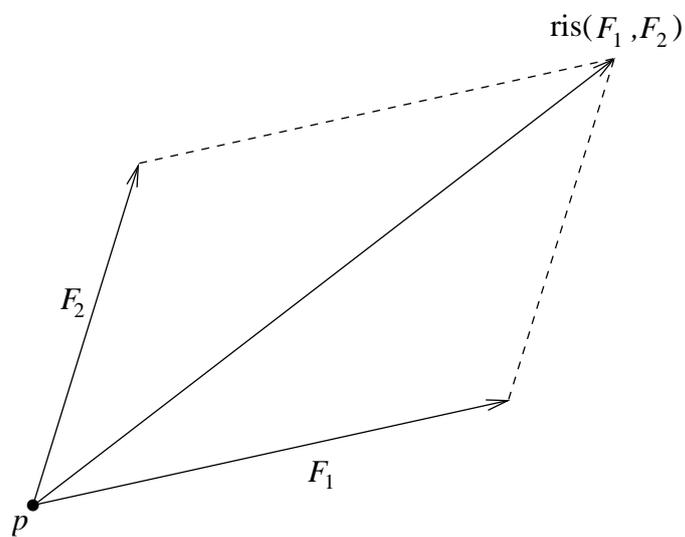


Figura 1.2: Risultante di due forze F_1 e F_2

(La prima è facile; verificare per esercizio la seconda). Inoltre se chiamiamo 0 la forza di intensità nulla applicata a p (della quale non contano direzione e verso, naturalmente) abbiamo anche:

- $\text{ris}(0, F) = F$;
- $\text{ris}(F, \widehat{F}) = 0$ se \widehat{F} è la forza che ha direzione e intensità uguali a quelle di F , ma verso opposto.

Esaminando le quattro proprietà elencate ci si accorge che sono perfettamente analoghe a quelle dell'operazione “+” (addizione) tra numeri:

- $x + y = y + x$;
- $x + (y + z) = (x + y) + z$;
- $0 + x = x$;
- $x + (-x) = 0$.

Facciamo allora la scommessa di indicare $\text{ris}(F_1, F_2)$ con $F_1 + F_2$, sperando che l'analogia con l'operazione “+” tra numeri dia frutti. Bisogna fare attenzione: scrivendo $F_1 + F_2$ stiamo dando un *nuovo significato* al simbolo “+”, cioè stiamo definendo una *nuova operazione*, poiché essa si applica a oggetti che non sono più numeri (sono forze). Lo stiamo facendo sulla base dell'analogia delle proprietà dell'operazione “risultante” tra forze con quelle dell'operazione “addizione” tra numeri.

Ci sono altre operazioni che possiamo fare con le forze. Ad esempio, data una forza F possiamo raddoppiarne, triplicarne o dimezzarne l'intensità, senza modificare direzione e verso, oppure possiamo cambiarne il verso. È naturale interpretare questo come la moltiplicazione della forza per un fattore 2 , 3 , $\frac{1}{2}$ e -1 , rispettivamente. Più in generale è chiaro che senso dare alla forza $\lambda \cdot F$ se λ è un numero. Osserviamo che, come sopra con il “+”, abbiamo introdotto una *nuova operazione* “ \cdot ” che ha come “input” una coppia di oggetti del tipo numero–forza e ha come “output” una forza. Naturalmente l'uso dello stesso simbolo impiegato per la moltiplicazione tra numeri ha senso se sono analoghe le proprietà algebriche. E di fatto si possono facilmente verificare le identità seguenti:

- $(\lambda_1 \cdot \lambda_2) \cdot F = \lambda_1 \cdot (\lambda_2 \cdot F)$

(si osservi che dei quattro “.” che compaiono in questa formula, il primo rappresenta la moltiplicazione tra numeri, mentre gli altri tre rappresentano la nuova operazione introdotta);

$$\bullet (\lambda_1 + \lambda_2) \cdot F = \lambda_1 \cdot F + \lambda_2 \cdot F$$

(in questo caso il primo “+” è quello tra numeri, mentre l’altro “+” e i tre “.” hanno il nuovo significato);

$$\bullet \lambda \cdot (F_1 + F_2) = \lambda \cdot F_1 + \lambda \cdot F_2;$$

$$\bullet \lambda \cdot 0 = 0 \cdot F = 0$$

(nell’ultima formula il primo e il terzo “0” rappresentano la forza di intensità nulla, mentre il secondo è il numero zero). Sulla base di queste proprietà sarà lecito manipolare espressioni contenenti le nuove operazioni introdotte così come si sarebbe fatto per le corrispondenti espressioni algebriche.

1.2 Trasformazioni piane

Vediamo ora come una struttura algebrica simile a quella appena scoperta per le forze applicate emerga in un ambito di natura apparentemente del tutto diversa, quello delle trasformazioni del piano cartesiano.

1.2.1 Traslazioni

Una classe di trasformazioni del piano cartesiano che molti dovrebbero avere incontrato è quella delle traslazioni. Una traslazione T è determinata da un vettore v avente certe componenti (x_0, y_0) , come in Fig. 1.3-sinistra. Essa agisce su un generico punto p come indicato in Fig. 1.3-centro. Esaminiamo ora come si esprimano le coordinate (x', y') del punto trasformato $T(p)$ in funzione delle coordinate (x, y) del punto iniziale p . La relazione, che si deduce facilmente da Fig. 1.3-destra, è la seguente:

$$(\spadesuit) \quad \begin{cases} x' = x + x_0 \\ y' = y + y_0. \end{cases}$$

Interpretiamo ora le coordinate di un punto o le componenti di un vettore, che sono una coppia di numeri, come un oggetto unico, che scriviamo come colonna (questa scelta si rivelerà molto utile, come vedremo già in questo primo capitolo) e che indichiamo con il simbolo stesso che denota il punto o il vettore:

$$p' = \begin{pmatrix} x' \\ y' \end{pmatrix}, \quad p = \begin{pmatrix} x \\ y \end{pmatrix}, \quad v = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

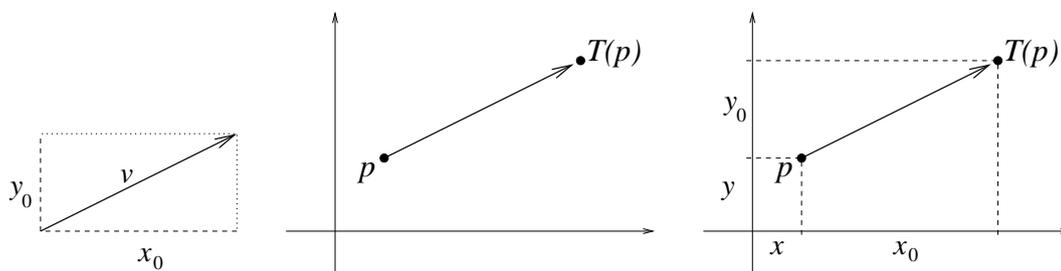


Figura 1.3: Un vettore nel piano, la traslazione a esso associata, e la corrispondente trasformazione delle coordinate

È allora naturale scrivere la relazione (\spadesuit) semplicemente come:

$$p' = p + v.$$

In altre parole stiamo definendo ancora un nuovo significato per il simbolo “+”. Ora abbiamo come “input” della nuova operazione due colonne di numeri $\begin{pmatrix} x \\ y \end{pmatrix}$ e $\begin{pmatrix} z \\ w \end{pmatrix}$, e come “output” una nuova colonna $\begin{pmatrix} x+z \\ y+w \end{pmatrix}$. Come sopra, è sensato denotare con “+” la nuova operazione poiché valgono le seguenti proprietà algebriche analoghe a quelle dell’addizione tra numeri:

- $0 + p = p$

dove 0 indica la colonna $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$;

- $p + (-p) = 0$

dove $-p$ indica la colonna $\begin{pmatrix} -x \\ -y \end{pmatrix}$ se $p = \begin{pmatrix} x \\ y \end{pmatrix}$;

- $p_1 + p_2 = p_2 + p_1$;

- $p_1 + (p_2 + p_3) = (p_1 + p_2) + p_3$.

In realtà possiamo ricondurre questa nuova operazione + a quella introdotta sopra tra forze. Se infatti interpretiamo p come la forza applicata all’origine delle coordinate e avente “testa” nel punto di coordinate p , abbiamo una completa coerenza, come illustrato in Fig. 1.4.

Per concludere con le traslazioni osserviamo anche che, interpretando sia i punti del piano cartesiano sia i vettori come colonne di due numeri, abbiamo

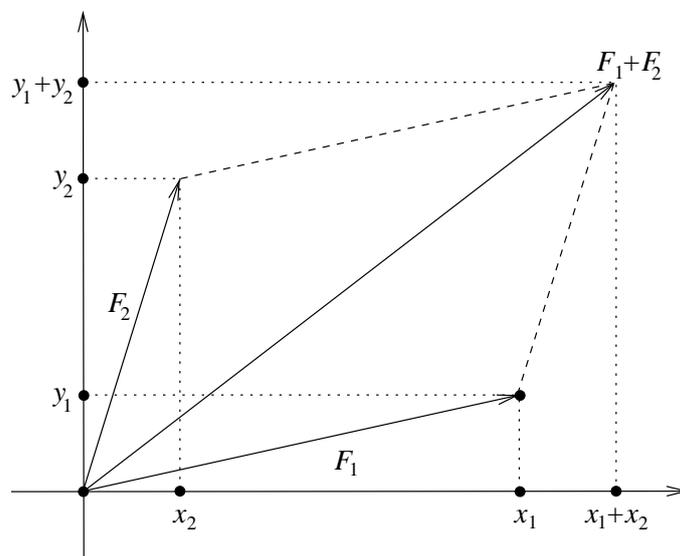
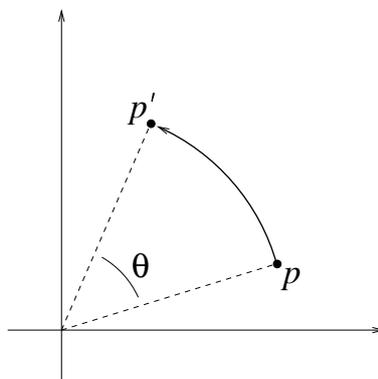


Figura 1.4: Risultante tra forze espressa in coordinate

di fatto abbandonato la distinzione tra questi tipi di oggetto. La coerenza di questo fatto si spiega con la costruzione appena vista: un punto p determina il vettore che ha “coda” nell’origine degli assi e “testa” in p . Viceversa un vettore v determina un punto se si trasla la “coda” di v nell’origine e se ne considera la “testa”.

1.2.2 Omotetie

Consideriamo ora un’altra operazione elementare sui punti del piano cartesiano, vale a dire l’omotetia di ragione $\lambda \neq 0$ centrata nell’origine O . Geometricamente tale trasformazione mantiene ogni punto p sulla retta che contiene O e p , ma modifica la distanza tra O e p di un fattore $|\lambda|$, riflettendo inoltre rispetto all’origine il punto risultante se $\lambda < 0$. Se $\begin{pmatrix} x \\ y \end{pmatrix}$ sono le coordinate di p , è facile vedere che le coordinate del suo trasformato sono $\begin{pmatrix} \lambda \cdot x \\ \lambda \cdot y \end{pmatrix}$. È ora naturale denotare tale colonna di coordinate come $\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix}$, introducendo dunque una nuova operazione “ \cdot ” che ha come “input” un numero e una colonna con

Figura 1.5: La rotazione di angolo ϑ

due coordinate e come “output” un’altra colonna con due coordinate. Come al solito questa scelta è giustificata dalle proprietà algebriche dell’operazione. È anzi facile vedere che essa corrisponde esattamente alla moltiplicazione tra un numero e una forza, introdotta sopra, interpretando un punto p come il vettore con “coda” nell’origine e “testa” in p .

1.2.3 Rotazioni

Consideriamo ora una nuova classe di trasformazioni del piano cartesiano, quella delle rotazioni intorno all’origine degli assi. Una rotazione è determinata da un angolo ϑ , e agisce come suggerito in Fig. 1.5.

Determiniamo come si ricavano le coordinate $\begin{pmatrix} x' \\ y' \end{pmatrix}$ del punto ruotato p' in funzione delle coordinate $\begin{pmatrix} x \\ y \end{pmatrix}$ del punto iniziale p . Si tratta di un esercizio di trigonometria, svolto in Fig. 1.6: sono qui mostrati il triangolo avente vertici in O , in p e nella proiezione di p sull’asse delle ascisse, il ruotato di tale triangolo, e le lunghezze di vari segmenti, dalle quali la conclusione segue facilmente. Si trovano dunque le relazioni seguenti:

$$(\heartsuit) \quad \begin{cases} x' = \cos \vartheta \cdot x - \sin \vartheta \cdot y \\ y' = \sin \vartheta \cdot x + \cos \vartheta \cdot y. \end{cases}$$

Avendo interpretato le coordinate come un unico oggetto $\begin{pmatrix} x \\ y \end{pmatrix}$, è naturale

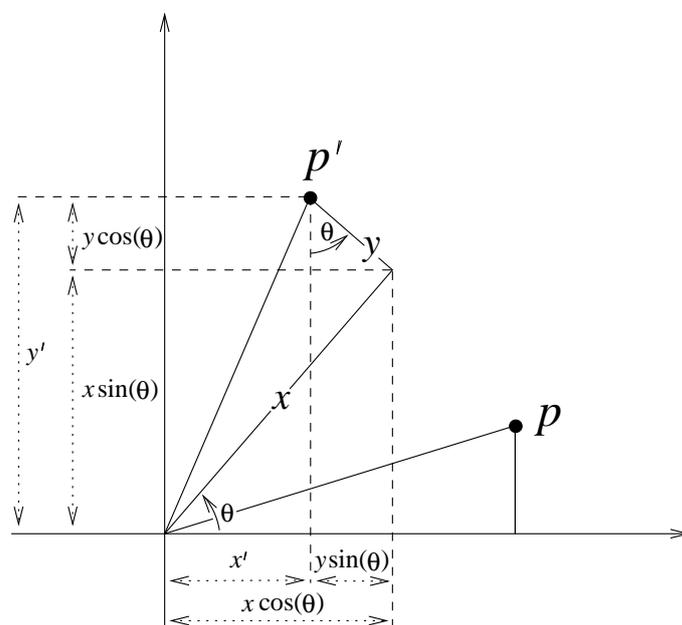


Figura 1.6: Coordinate di un punto ruotato di angolo ϑ

tentare di riscrivere le equazioni (\heartsuit) come un'unica equazione

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R_\vartheta \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

dove tanto l'oggetto R_ϑ che il significato di “ \cdot ” sono da inventare. Per R_ϑ inventiamo la tabella

$$R_\vartheta = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}$$

che rispecchia le posizioni dei coefficienti $\cos \vartheta$ e $\pm \sin \vartheta$ nelle equazioni (\heartsuit). A questo punto inventiamo la regola di moltiplicazione tra una tabella con due righe e due colonne e una colonna di altezza due in modo che risulti quello che desideriamo:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \cdot x + b \cdot y \\ c \cdot x + d \cdot y \end{pmatrix}.$$

Ancora una volta, ha senso usare il simbolo “ \cdot ” per l'operazione introdotta in virtù delle sue proprietà algebriche, che vedremo oltre.

Illustriamo ora un primo elemento di utilità delle nuove operazioni (che altrimenti potrebbero sembrare pure invenzioni formali, prive di scopo). Ricordiamo che la tabella R_ϑ l'abbiamo associata alla rotazione di angolo ϑ , nel senso che la rotazione di un punto di angolo ϑ si esegue moltiplicando (a sinistra) le coordinate del punto per R_ϑ . Se ora consideriamo due angoli ϑ e φ , possiamo considerare due modi di “moltiplicare” tra loro le corrispondenti R_ϑ e R_φ . Da un lato, geometricamente, abbiamo le rotazioni di angoli ϑ e φ , che possiamo comporre, cioè eseguire una dopo l'altra, ottenendo la rotazione di angolo $\vartheta + \varphi$. Dall'altro lato, possiamo facilmente inventare una ragionevole regola di moltiplicazione tra tabelle con due righe e due colonne. Date $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e $\begin{pmatrix} x & z \\ y & w \end{pmatrix}$, interpretiamo la seconda tabella come la scrittura, una di fianco all'altra, delle due colonne $\begin{pmatrix} x \\ y \end{pmatrix}$ e $\begin{pmatrix} z \\ w \end{pmatrix}$, e decidiamo che moltiplicando a sinistra per $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ si ottengono, una di fianco all'altra, le due colonne $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$ e $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} z \\ w \end{pmatrix}$, ovvero:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x & z \\ y & w \end{pmatrix} = \begin{pmatrix} a \cdot x + b \cdot y & a \cdot z + b \cdot w \\ c \cdot x + d \cdot y & c \cdot z + d \cdot w \end{pmatrix}.$$

Con questa regola abbiamo allora:

$$\begin{aligned} R_{\vartheta} \cdot R_{\varphi} &= \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \cdot \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \\ &= \begin{pmatrix} \cos \vartheta \cdot \cos \varphi - \sin \vartheta \cdot \sin \varphi & -\cos \vartheta \cdot \sin \varphi - \sin \vartheta \cdot \cos \varphi \\ \cos \vartheta \cdot \sin \varphi + \sin \vartheta \cdot \cos \varphi & \cos \vartheta \cdot \cos \varphi - \sin \vartheta \cdot \sin \varphi \end{pmatrix} \end{aligned}$$

mentre avevamo osservato che la composizione di R_{ϑ} ed R_{φ} conduceva a:

$$R_{\vartheta+\varphi} = \begin{pmatrix} \cos(\vartheta + \varphi) & -\sin(\vartheta + \varphi) \\ \sin(\vartheta + \varphi) & \cos(\vartheta + \varphi) \end{pmatrix}.$$

Le regole trigonometriche di calcolo per $\cos(\vartheta + \varphi)$ e $\sin(\vartheta + \varphi)$ comportano ora che $R_{\vartheta} \cdot R_{\varphi} = R_{\vartheta+\varphi}$. Dunque abbiamo il fatto notevole che la regola di moltiplicazione algebrica tra tabelle, inventata senza particolare motivazione geometrica, in realtà traduce esattamente la nozione geometrica di composizione delle due rotazioni. Alternativamente, dando per assodata la corrispondenza tra moltiplicazione e composizione, la relazione $R_{\vartheta} \cdot R_{\varphi} = R_{\vartheta+\varphi}$ ci consente di ricavare in modo facile le regole trigonometriche di calcolo per $\cos(\vartheta + \varphi)$ e $\sin(\vartheta + \varphi)$: basta ricordare come si associa R_{ϑ} a ϑ e la moltiplicazione tra tabelle.

Concludiamo osservando che, quando abbiamo parlato della composizione delle due rotazioni di angoli ϑ e φ , non abbiamo avuto bisogno di specificare quale delle due trasformazioni sia eseguita prima dell'altra, poiché in questo caso l'una scelta o l'altra conducono allo stesso risultato. Osserviamo che ciò non è affatto vero per funzioni qualsiasi: ad esempio, se $f(t) = t^2$ e $g(t) = 2t$, si ha $(f \circ g)(t) = 4t^2$ e $(g \circ f)(t) = 2t^2$.

1.3 Intersezione tra rette

Nel piano cartesiano con coordinate x e y , una retta è il luogo definito da un'equazione del tipo $ax + by = k$ con a, b, k numeri fissati, purché a e b non siano entrambi nulli (altrimenti viene tutto il piano per $k = 0$ e l'insieme vuoto \emptyset per $c \neq k$). Date due rette di equazioni $ax + by = k$ e $cx + dy = h$, l'operazione geometrica di ricerca della loro intersezione (se esiste) corrisponde algebricamente alla ricerca delle soluzioni del sistema

$$\begin{cases} ax + by = k \\ cx + dy = h. \end{cases}$$

Usando le grafie e le operazioni introdotte in precedenza possiamo ora scrivere questo sistema come una singola equazione

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} k \\ h \end{pmatrix}.$$

Naturalmente il solo fatto di avere scritto le equazioni in maniera più sintetica non aiuta di per sé a risolverle. Vedremo però che al linguaggio introdotto si accompagneranno strumenti molto efficienti per la soluzione di un'equazione come quella appena scritta.

1.4 Numeri di Fibonacci

Menzioniamo ora un argomento che non pare a prima vista avere alcuna relazione con le nozioni di natura geometrica (vettori e coordinate) di cui ci si è occupati finora, ma che invece sarà possibile trattare con il linguaggio generale che ci avviamo a introdurre.

Supponiamo di piantare oggi (all'inizio del mese numero 1) un alberello con un solo ramo, che per tutto il mese cresce senza gettare nuovi rami e poi si sviluppa secondo queste regole:

- nessun ramo muore mai;
- alla metà di ogni nuovo mese, tutti i rami che non siano “neonati” (ovvero che non siano spuntati il mese precedente) gettano un nuovo ramo.

Supponendo che a ogni mese ogni ramo cresca verso l'alto della stessa altezza, l'aspetto dell'albero al settimo mese sarà quello in Fig. 1.7. Ci chiediamo ora quale sia il numero di rami f_n dell'albero all'inizio del mese n -esimo (tale f_n si chiama n -esimo numero di Fibonacci). Dalla Fig. 1.7 si desume che i valori di f_n per $n = 1, \dots, 7$ sono 1, 1, 2, 3, 5, 8, 13. Mostriamo ora che vale la seguente formula:

$$(\diamond) \quad f_n = f_{n-1} + f_{n-2}, \quad \text{per ogni } n \geq 3.$$

Infatti all'inizio del mese n sopravvivono tutti i rami che esistevano all'inizio del mese $n-1$, che sono nel numero di f_{n-1} , e in più vi sono tutti i rami buttati alla metà del mese $n-1$ dai rami non “neonati”, cioè da quelli che esistevano all'inizio del mese $n-2$. Questi ultimi sono nel numero di f_{n-2} , onde la (\diamond) .

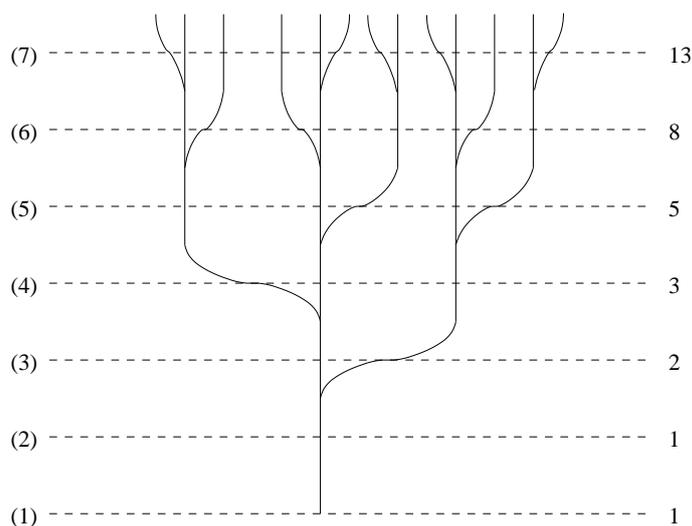


Figura 1.7: Un albero che cresce come la successione di Fibonacci

La (\diamond) è una formula cosiddetta “ricorsiva”, poiché consente di calcolare tutti i valori di f_n ripetendo ricorsivamente la stessa procedura a partire dai termini iniziali $f_1 = 1$ e $f_2 = 1$, ovvero:

$$\begin{aligned}
 f_1 &= 1, \\
 f_2 &= 1, \\
 f_3 &= f_2 + f_1 = 1 + 1 = 2, \\
 f_4 &= f_3 + f_2 = 2 + 1 = 3, \\
 f_5 &= f_4 + f_3 = 3 + 2 = 5 \\
 &\vdots
 \end{aligned}$$

In qualsiasi linguaggio di programmazione è molto semplice implementare questa procedura ricorsiva, ma il programma ottenuto è molto dispendioso in termini di tempo, poiché il calcolo di ciascun f_n richiede il calcolo di tutti gli f_m con $m < n$. Esiste tuttavia un metodo alternativo e molto efficiente di calcolo per f_n , basato sulla seguente (sorpriendente) identità:

$$(\clubsuit) \quad f_n = \frac{1}{\sqrt{5}} \cdot \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

(Già la sola verifica che il termine destro di questa relazione è intero per ogni n non è del tutto banale.) Vedremo che la formula (♣) si può non solo dimostrare facilmente, ma anche “spiegare” perfettamente, con il linguaggio che introdurremo a partire dal Capitolo 3.

Capitolo 2

Preliminari algebrici e logici

In questo capitolo richiamiamo le proprietà algebriche delle operazioni sui principali insiemi numerici e illustriamo, esemplificandole, due tecniche di dimostrazione.

2.1 Numeri

Questo paragrafo è dedicato agli insiemi dei numeri naturali, degli interi, dei razionali e dei reali.

2.1.1 Numeri naturali e interi

Useremo (senza definirlo) l'insieme dei numeri *naturali* $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ e il fatto che sui numeri naturali è definita un'operazione di addizione “+” binaria e interna (cioè che ha come “input” due naturali e come “output” un nuovo naturale), che soddisfa le seguenti proprietà:

1. $0 + n = n$ per ogni n ;
3. $n + (m + k) = (n + m) + k$ per ogni n, m, k ;
4. $n + m = m + n$ per ogni n, m .

Abbiamo intenzionalmente ommesso il numero 2 elencando le proprietà, per una ragione che sarà subito chiara. Se consideriamo, dato $n \in \mathbb{N}$, l'equazione

$$n + x = 0 \quad ,$$

osserviamo che essa non ha soluzioni $x \in \mathbb{N}$, ad eccezione che per $n = 0$, nel qual caso ha la soluzione $x = 0$. Per sopperire a questo inconveniente si

estende l'insieme \mathbb{N} dei numeri naturali a quello $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ dei numeri *interi*. Anche su questo insieme è definita un'operazione binaria interna di somma, che estende quella su \mathbb{N} , e che gode, oltre che delle proprietà (1), (3) e (4) elencate sopra, anche della seguente:

2. Per ogni $n \in \mathbb{Z}$ esiste $(-n) \in \mathbb{Z}$ tale che $n + (-n) = 0$.

Un utile esercizio: usando solo le proprietà elencate, provare che per ogni n l'equazione $n + x = 0$ ha un'unica soluzione $x \in \mathbb{Z}$, indicata appunto con $(-n)$.

2.1.2 Numeri razionali e reali

Sull'insieme \mathbb{Z} degli interi, oltre alla somma abbiamo una seconda operazione binaria interna, indicata con “ \cdot ” e detta prodotto, che soddisfa le seguenti proprietà:

5. $1 \cdot n = n$ per ogni n ;
7. $n \cdot (m \cdot k) = (n \cdot m) \cdot k$ per ogni n, m, k ;
8. $n \cdot m = m \cdot n$ per ogni n, m .
9. $n \cdot (m + k) = n \cdot m + n \cdot k$ per ogni n, m, k ;

Di nuovo abbiamo intenzionalmente omesso la proprietà (6). La ragione è che, dato $n \in \mathbb{Z}$, l'equazione

$$n \cdot x = 1$$

non ha soluzioni x in \mathbb{Z} , a meno che n non sia $+1$ oppure -1 , nel qual caso la soluzione è n . Si passa dunque dall'insieme dei numeri interi a quello \mathbb{Q} dei numeri razionali, ovvero alle frazioni del tipo $\frac{p}{q}$ con $p, q \in \mathbb{Z}$ e $q \neq 0$. In realtà, poiché sappiamo che i numeri razionali si possono scrivere come frazioni in più modi, formalmente si dovrebbe definire \mathbb{Q} come l'insieme delle scritte del tipo $\frac{p}{q}$ come sopra, dove però $\frac{p}{q}$ è identificata a $\frac{p'}{q'}$ quando $p \cdot q' = p' \cdot q$ (il che, naturalmente, corrisponde al fatto che $\frac{p}{q}$ e $\frac{p'}{q'}$ hanno “lo stesso valore”). Analogamente, identifichiamo la frazione $\frac{n}{1}$ all'intero n , deducendo che \mathbb{Q} contiene \mathbb{Z} . Inoltre, le operazioni di somma e prodotto definite su \mathbb{Z} si estendono a \mathbb{Q} , e soddisfano su \mathbb{Q} le proprietà seguenti:

1. $0 + x = x$ per ogni x ;
2. Per ogni x esiste $(-x)$ tale che $x + (-x) = 0$;
3. $x + (y + z) = (x + y) + z$ per ogni x, y, z ;

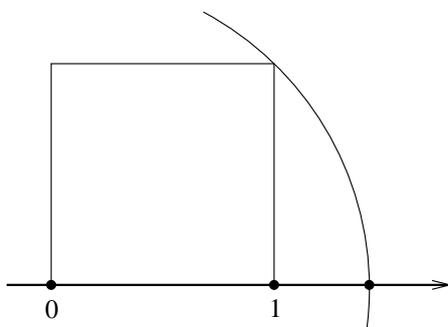


Figura 2.1: Un punto non razionale della retta

4. $x + y = y + x$ per ogni x, y ;
5. $1 \cdot x = x$ per ogni x ;
6. Per ogni $x \neq 0$ esiste x^{-1} tale che $x \cdot x^{-1} = 1$;
7. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ per ogni x, y, z ;
8. $x \cdot y = y \cdot x$ per ogni x, y ;
9. $x \cdot (y + z) = x \cdot y + x \cdot z$ per ogni x, y, z .

Nella proprietà (6) è necessario ipotizzare $x \neq 0$ perché dalle altre proprietà è possibile dedurre che $0 \cdot y = 0$ per ogni y .

Estendiamo ora ulteriormente l'insieme \mathbb{Q} dei razionali all'insieme, indicato con \mathbb{R} , dei numeri reali. La motivazione dell'estensione è questa volta di natura geometrico-analitica, piuttosto che algebrica come per i passaggi da \mathbb{N} a \mathbb{Z} e poi da \mathbb{Z} a \mathbb{Q} . Per spiegarla, consideriamo una retta e fissiamo su questa retta due punti distinti che facciamo corrispondere ai numeri razionali 0 e 1. Possiamo allora estendere questa corrispondenza, associando a ogni numero un punto. Ad esempio al numero $-\frac{3}{2}$ corrisponderà il punto che sta dalla parte opposta di 1 rispetto a 0 e che ha distanza da 0 una volta e mezza quella che 1 ha da 0. La funzione così definita da \mathbb{Q} alla retta risulta allora iniettiva (cioè numeri diversi danno punti diversi). Essa non è però surgettiva, ovvero ci sono punti che non corrispondono ad alcun numero razionale. Un esempio è il punto x costruito nella Fig. 2.1, la cui distanza da 0 è uguale alla lunghezza della diagonale del quadrato unitario. Per il Teorema di Pitagora, se x corrispondesse a un numero razionale esso dovrebbe soddisfare l'identità $x^2 = 2$, mentre è possibile

vedere (lo faremo più avanti in questo capitolo) che l'equazione $x^2 = 2$ non ha soluzioni in \mathbb{Q} .

Avendo scoperto che \mathbb{Q} rappresenta alcuni ma non tutti i punti della retta, definiamo l'insieme \mathbb{R} dei numeri reali semplicemente come l'insieme dei punti della retta, e usiamo (senza dimostrarlo), il fatto che le operazioni “+” e “.” definite su \mathbb{Q} si estendono a \mathbb{R} e continuano a soddisfare le proprietà (1)-(9) che abbiamo elencato sopra per \mathbb{Q} .

2.1.3 Campi e gruppi commutativi

Avendo incontrato due insiemi (cioè \mathbb{Q} e \mathbb{R}) che possiedono due operazioni soddisfacenti lo stesso lungo elenco di proprietà, introduciamo ora un singolo termine, quello di *campo*, per indicare un insieme siffatto. Nel Capitolo 8 incontreremo un altro esempio di campo, quello \mathbb{C} dei numeri complessi. Osservando inoltre che le proprietà (1)-(4) della definizione di campo riguardano la sola operazione “+”, mentre le (5)-(8) riguardano solo “.” e sono formalmente pressoché analoghe alle (1)-(4), introduciamo ora un'altra nuova terminologia. Chiamiamo *gruppo commutativo* un insieme sul quale è definita un'operazione (binaria e interna) tale che esiste un elemento neutro, ogni elemento ammette un opposto, e valgono le proprietà associativa e commutativa. Notiamo che, per essere del tutto precisi, un gruppo commutativo non è il solo insieme, bensì la terna insieme–operazione–elemento neutro. Dalla discussione sugli insiemi numerici fin qui svolta segue allora che possediamo già i seguenti esempi di gruppo commutativo:

- $(\mathbb{Z}, +, 0)$;
- $(\mathbb{Q}, +, 0)$;
- $(\mathbb{R}, +, 0)$;
- $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$, mentre, è importante osservarlo, non è un gruppo commutativo l'insieme $(\mathbb{Q}, \cdot, 1)$, poiché l'elemento 0 non ammette “opposto rispetto alla moltiplicazione” (cioè, non ammette inverso); è questa l'unica ma sostanziale differenza tra la proprietà (6) e la (2) tra quelle di campo;
- $(\mathbb{R} \setminus \{0\}, \cdot, 1)$.

2.2 Polinomi

Avremo spesso a che fare con i polinomi, ovvero con le scritture del tipo

$$a_0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_k \cdot x^k,$$

dove $k \in \mathbb{N}$, $a_0, \dots, a_k \in \mathbb{R}$, e x è un simbolo che non appartiene a \mathbb{R} (detto *indeterminata*). Sottolineiamo che un polinomio per noi è una espressione del tipo $a_0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_k \cdot x^k$, non la funzione polinomiale da \mathbb{R} a \mathbb{R} che eventualmente gli si può associare. Definiamo ora $\mathbb{R}[x]$ come l'insieme di tutti i polinomi, ovvero delle espressioni del tipo appena descritto, avendo però una piccola accortezza che ora illustriamo. Notiamo che formalmente l'espressione $\frac{1}{2} - 7x^2$ non è del tipo descritto, mentre $\frac{1}{2} + 0 \cdot x - 7x^2$ lo è. Inoltre, entrambe le espressioni $\frac{3}{4} + 2x$ e $\frac{3}{4} + 2x + 0 \cdot x^2$ sono del tipo descritto, e sono diverse. Tuttavia noi certamente vogliamo considerare $\frac{1}{2} - 7x^2$ un polinomio, e vogliamo considerare $\frac{3}{4} + 2x$ e $\frac{3}{4} + 2x + 0 \cdot x^2$ uguali tra loro. Adottiamo quindi la convenzione che nella scrittura di un polinomio è possibile introdurre o cancellare a piacimento monomi del tipo $0 \cdot x^h$.

Ad un polinomio $p(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_k \cdot x^k$ che non sia il polinomio nullo si associa un intero $\deg(p(x))$, detto *grado*, definito come

$$\deg(p(x)) = \max\{h : a_h \neq 0\}.$$

Al polinomio nullo si può convenzionalmente assegnare il grado $-\infty$, oppure -1 , oppure nessun grado.

Anche sull'insieme dei polinomi sono definite due operazioni naturali “+” e “·”. Per descriverle in termini generali conviene introdurre un'ulteriore convenzione. Osserviamo intanto che un'espressione

$$a_0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_k \cdot x^k$$

può essere descritta dalla sommatoria $\sum_{n=0}^k a_n \cdot x^n$. Inoltre, se $h \geq k$ e si pone

$a_n = 0$ per $n = k + 1, \dots, h$, la stessa espressione è uguale a $\sum_{n=0}^h a_n \cdot x^n$. Poiché nell'ultima sommatoria il valore di h è arbitrariamente grande, è ragionevole ora convenire che un polinomio è un'espressione del tipo $\sum_{n=0}^{\infty} a_n \cdot x^n$, dove i coefficienti a_n sono tutti nulli ad eccezione che per un numero finito di indici

n . Le formule per la somma ed il prodotto di polinomi sono allora date da:

$$\begin{aligned} \left(\sum_{n=0}^{\infty} a_n \cdot x^n \right) + \left(\sum_{n=0}^{\infty} b_n \cdot x^n \right) &= \sum_{n=0}^{\infty} (a_n + b_n) \cdot x^n, \\ \left(\sum_{n=0}^{\infty} a_n \cdot x^n \right) \cdot \left(\sum_{n=0}^{\infty} b_n \cdot x^n \right) &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k \cdot b_{n-k} \right) \cdot x^n. \end{aligned}$$

Si può verificare per esercizio che queste operazioni soddisfano le proprietà (1)-(5) e (7)-(9) della definizione di campo, ma non la (6).

2.3 Dimostrazioni

Illustriamo ed esemplifichiamo in questo paragrafo due tecniche dimostrative alle quali faremo spesso ricorso nel seguito.

2.3.1 Principio di induzione

Questo principio è utile quando si voglia dimostrare che una certa proposizione $\mathcal{P}(n)$, relativa a un numero naturale n , vale per ogni $n \in \mathbb{N}$. A questo fine è sufficiente dimostrare che:

- (passo base) vale la $\mathcal{P}(0)$, ovvero la proposizione $\mathcal{P}(n)$ per $n = 0$;
- (passo induttivo) se vale la proposizione $\mathcal{P}(n)$ per un generico $n \in \mathbb{N}$ (ipotesi induttiva), allora vale anche la $\mathcal{P}(n + 1)$.

È facile convincersi che questo davvero è sufficiente: vale infatti la $\mathcal{P}(0)$ per il passo base; applicando allora il passo induttivo per $n = 0$ si trova che vale $\mathcal{P}(1)$; di nuovo per il passo induttivo, ora con $n = 1$, si ha la $\mathcal{P}(2)$, e si procede deducendo che $\mathcal{P}(n)$ vale per ogni n . È importante osservare che nel passo induttivo la validità di $\mathcal{P}(n)$ è una *ipotesi*, ovvero non bisogna dimostrarla. Quello che bisogna fornire è un argomento che sia valido per ogni generico n e che consenta di far discendere la $\mathcal{P}(n + 1)$ dalla $\mathcal{P}(n)$. Un esempio di questo ragionamento è il seguente:

Proposizione 2.3.1. $\sum_{k=0}^n k = \frac{n(n+1)}{2} \quad \forall n \in \mathbb{N}.$

Dimostrazione. Chiamiamo $\mathcal{P}(n)$ l'affermazione che $\sum_{k=0}^n k$ vale $\frac{1}{2}n(n + 1)$, e dimostriamola per induzione su n . Il passo base dell'induzione, ovvero $\mathcal{P}(0)$, è

certamente valido, poiché i due numeri di cui bisogna verificare l'uguaglianza sono entrambi nulli.

Sia dunque $n \in \mathbb{N}$ generico, e *supponiamo* che $\mathcal{P}(n)$ sia vera, cioè che $\sum_{k=0}^n k$ valga $\frac{1}{2}n(n+1)$. Vogliamo dedurre che è vera la $\mathcal{P}(n+1)$, cioè che $\sum_{k=0}^{n+1} k$ vale $\frac{1}{2}(n+1)((n+1)+1)$, ovvero $\frac{1}{2}(n+1)(n+2)$. Infatti $\sum_{k=0}^{n+1} k$ vale $\left(\sum_{k=0}^n k\right) + (n+1)$, e per ipotesi induttiva il primo addendo di questa somma vale $\frac{1}{2}n(n+1)$. Si conclude ora semplicemente con un calcolo:

$$\frac{1}{2}n(n+1) + (n+1) = \left(\frac{n}{2} + 1\right)(n+2).$$

□

2.3.2 Dimostrazioni per assurdo

Un'altra tecnica che capita spesso di usare è la seguente. Volendo dimostrare, sotto opportune ipotesi, una proposizione \mathcal{P} , si suppone che \mathcal{P} sia falsa (*ipotesi dell'assurdo*) e, usando le ipotesi e fatti noti, si sviluppa un argomento che conduca alla negazione di una delle ipotesi o di un altro fatto noto (*contraddizione*). Dimostriamo ad esempio il fatto già annunciato sopra:

Proposizione 2.3.2. *L'equazione $x^2 = 2$ non ha soluzioni $x \in \mathbb{Q}$.*

Dimostrazione. Supponiamo per assurdo che una soluzione x esista. Utilizziamo ora il fatto noto che x , appartenendo a \mathbb{Q} , può essere scritto come una frazione $\frac{n}{m}$ dove n, m sono interi privi di fattori comuni (e $m \neq 0$). Per l'ipotesi dell'assurdo abbiamo allora $\left(\frac{n}{m}\right)^2 = 2$, dunque $n^2 = 2m^2$. Ora, se n fosse dispari, anche n^2 sarebbe dispari, mentre $2m^2$ è pari, perciò n è pari, ovvero $n = 2k$ con k intero. Sostituendo nell'identità $n^2 = 2m^2$ e semplificando un 2 troviamo allora $m^2 = 2k^2$. Per lo stesso argomento appena esposto, ne deduciamo che anche m è pari. Questo però è assurdo, perché n e m li avevamo supposti privi di fattori comuni, mentre avrebbero in comune il fattore 2. □

Capitolo 3

Spazi vettoriali

In questo capitolo introduciamo la classe di oggetti dei quali ci occuperemo principalmente fino al Capitolo 10, e forniamo numerosi esempi.

3.1 Definizione

L'introduzione della seguente nozione è giustificata dalle considerazioni svolte nel capitolo precedente:

Definizione 3.1.1. Uno *spazio vettoriale reale* (o *su* \mathbb{R}) è un insieme V (i cui elementi sono detti *vettori*) per il quale siano definite due operazioni “+” e “ \cdot ” che soddisfano certi assiomi che andiamo ora a elencare. L'operazione “+” è binaria e interna, ovvero è una funzione da $V \times V$ in V , che associa a una coppia di vettori (v, w) un terzo vettore $v + w$, detto *somma* di v e w . L'operazione “ \cdot ” è invece una funzione da $\mathbb{R} \times V$ in V , la quale dunque associa ad una coppia (λ, v) , costituita da un numero reale λ ed un vettore v , un nuovo vettore $\lambda \cdot v$, detto *prodotto* di v per lo *scalare* λ . Le proprietà che devono essere soddisfatte sono le seguenti:

1. Esiste $0 \in V$ tale che $0 + v = v$ per ogni v in V ;
2. Per ogni v in V esiste $(-v) \in V$ tale che $v + (-v) = 0$;
3. $v + (w + u) = (v + w) + u$ per ogni v, w, u in V ;
4. $v + w = w + v$ per ogni v, w in V ;
5. $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ per ogni λ in \mathbb{R} e v, w in V ;

6. $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$ per ogni λ, μ in \mathbb{R} e v in V ;
7. $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$ per ogni λ, μ in \mathbb{R} e v in V ;
8. $1 \cdot v = v$ e $0 \cdot v = 0$ per ogni v in V .

Notiamo che le proprietà (1)-(4) significano precisamente che V ha una operazione “+” di gruppo commutativo con elemento neutro 0. Sia la (5) sia la (6) sono leggi distributive, entrambe necessarie perché nel prodotto c’è una asimmetria dei due argomenti: quello sinistro è uno scalare, quello destro è un vettore. È un utile esercizio indagare nelle proprietà (5)-(8) la natura di ciascuno dei simboli “+”, “ \cdot ” e “0” (ovvero dire se il loro significato sia quello consueto in \mathbb{R} oppure quello relativo allo spazio vettoriale V).

Formalmente uno spazio vettoriale non è il solo insieme V , ma la quaterna $(V, +, \cdot, 0)$, ma nella pratica conviene riassumere nella sola scrittura V sia le operazioni di somma e prodotto per scalare relative a V sia l’elemento neutro della somma.

3.1.1 Conseguenze della definizione

Illustriamo ora come dagli assiomi della definizione di spazio vettoriale si possano far discendere altre naturali proprietà algebriche, analoghe a quelle note per le operazioni tra numeri. Innanzitutto le proprietà (3) e (7) rispettivamente ci consentono di dare un significato non ambiguo a espressioni del tipo:

$$v_1 + v_2 + \dots + v_n, \quad \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n \cdot v.$$

Inoltre, la (2) ci consente di “semplificare” le identità: se $v + w = v + u$ allora $(-v) + v + w = (-v) + v + u$, dunque $0 + w = 0 + u$, onde $w = u$ (qui, come anche nel seguito, usiamo la proprietà commutativa (4) senza neppure menzionarla). Dalla possibilità di semplificare identità del tipo $v + w = v + u$ segue anche subito l’unicità dell’opposto $(-v)$.

Dimostriamo ora che $(-1) \cdot v = (-v)$, specificando a ogni passaggio l’assioma usato:

$$\begin{aligned} (-1) \cdot v & \stackrel{(1)}{=} 0 + (-1) \cdot v \stackrel{(2)}{=} (-v) + v + (-1) \cdot v \\ & \stackrel{(8)}{=} (-v) + 1 \cdot v + (-1) \cdot v \stackrel{(6)}{=} (-v) + (1 - 1) \cdot v \\ & = (-v) + 0 \cdot v \stackrel{(8)}{=} (-v) + 0 \stackrel{(1)}{=} (-v). \end{aligned}$$

Possiamo inoltre provare che $\lambda \cdot 0 = 0$ per ogni λ , notando che

$$\lambda \cdot 0 \underset{(1)}{=} \lambda \cdot (0 + 0) \underset{(5)}{=} \lambda \cdot 0 + \lambda \cdot 0$$

e semplificando. Infine, verifichiamo che se $\lambda \neq 0$ e $\lambda \cdot v = 0$ allora $v = 0$:

$$0 = \lambda^{-1} \cdot 0 = \lambda^{-1} \cdot \lambda \cdot v = 1 \cdot v = v.$$

3.1.2 Esempi

Come già sottolineato, fanno parte della struttura di spazio vettoriale non solo l'insieme V , ma anche le operazioni “+” e “·” e l'elemento neutro 0 di “+”. Per descrivere un esempio dovremo dunque specificare tutti questi elementi.

Vettori colonna Cominciamo con lo spazio \mathbb{R}^n , ottenuto prendendo il prodotto cartesiano di n copie di \mathbb{R} , dove $n \geq 1$ è un intero qualsiasi. Per ragioni che saranno chiare nel Capitolo 5 (e sono state anticipate nel Capitolo 1), scriviamo un elemento x di \mathbb{R}^n come *vettore colonna*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

e chiamiamo x_i la i -esima *componente* di x . Definiamo le operazioni di spazio vettoriale “componente per componente”, ponendo cioè:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

$$\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Tutti gli assiomi si verificano immediatamente.

Matrici Dati interi $n, m \geq 1$ definiamo ora lo spazio $\mathcal{M}_{n \times m}(\mathbb{R})$ delle *matrici reali* $n \times m$, ovvero delle tabelle rettangolari di numeri reali aventi n righe e m colonne. Dunque $\mathcal{M}_{n \times m}(\mathbb{R})$ è l'insieme seguente:

$$\left\{ \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} : a_{ij} \in \mathbb{R} \text{ per } i = 1, \dots, n, j = 1, \dots, m \right\}.$$

Chiamiamo *posto* (i, j) della matrice quello sulla riga i -esima e colonna j -esima. Nella matrice generica scritta nella definizione, al posto (i, j) sta il coefficiente a_{ij} , e useremo sempre la convenzione di scrivere prima l'*indice di riga* (in questo caso i) e poi l'*indice di colonna* (in questo caso j). Utilizzeremo anche una singola lettera, generalmente maiuscola, per denotare un'intera matrice. Così ad esempio la matrice generica che compare nella definizione la scriveremo anche come

$$A = \left(a_{ij} \right)_{\substack{i=1, \dots, n \\ j=1, \dots, m}}.$$

Un'altra grafia utile è quella di indicare con $(A)_{ij}$ il coefficiente di posto (i, j) della matrice A (cosicché per la nostra A si ha $(A)_{ij} = a_{ij}$ semplicemente).

Introduciamo ora l'operazione di somma tra matrici usando "posto per posto" la somma consueta in \mathbb{R} :

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1m} + b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nm} + b_{nm} \end{pmatrix}.$$

Impiegando le grafie alternative introdotte possiamo anche scrivere la somma come segue:

$$\begin{aligned} \text{se } A = \left(a_{ij} \right)_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \quad \text{e} \quad B = \left(b_{ij} \right)_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \\ \text{allora } A + B = \left(a_{ij} + b_{ij} \right)_{\substack{i=1, \dots, n \\ j=1, \dots, m}}; \end{aligned}$$

oppure ancora, più facilmente: $(A + B)_{ij} = (A)_{ij} + (B)_{ij}$ per $i = 1, \dots, n$ e $j = 1, \dots, m$.

Definendo anche $(\lambda \cdot A)_{ij} = \lambda \cdot (A)_{ij}$ e $(0)_{ij} = 0$ per $i = 1, \dots, n$ e $j = 1, \dots, m$, è ora molto semplice osservare che resta definita su $\mathcal{M}_{n \times m}(\mathbb{R})$ una struttura di spazio vettoriale reale. Notiamo che $\mathcal{M}_{n \times 1}(\mathbb{R})$ si identifica naturalmente a \mathbb{R}^n .

Funzioni Sia ora S un insieme non vuoto, e consideriamo l'insieme $\mathcal{F}(S, \mathbb{R})$ di tutte le funzioni da S in \mathbb{R} . Vogliamo definire su $\mathcal{F}(S, \mathbb{R})$ una struttura di spazio vettoriale. Cominciamo con la somma. Abbiamo dunque due funzioni f e g da S in \mathbb{R} , e vogliamo definire una funzione $f + g$. Ricordiamo che una funzione $F : S \rightarrow \mathbb{R}$ è una legge che associa a ogni generico elemento s di S un determinato elemento $F(s)$ di \mathbb{R} , detto il valore di F nel punto s . Una funzione è dunque determinata precisamente quando se ne conoscano i valori su tutti

i punti. Volendo definire una funzione $f + g$, si tratta dunque di specificare quale sia il suo valore $(f + g)(s)$ su un generico elemento s di S . Lo facciamo ponendo $(f + g)(s) = f(s) + g(s)$. Analogamente, dati $\lambda \in \mathbb{R}$ e $f : S \rightarrow \mathbb{R}$, definiamo $\lambda \cdot f : S \rightarrow \mathbb{R}$ ponendo $(\lambda \cdot f)(s) = \lambda \cdot f(s)$ per ogni s in S . Definiamo infine $0 \in \mathcal{F}(S, \mathbb{R})$ ponendo $0(s) = 0$ per ogni s in S .

Verifichiamo in questo esempio che valgono le proprietà di spazio vettoriale, limitandoci alla (5). Vogliamo cioè verificare che, dati $\lambda \in \mathbb{R}$ e $f, g \in \mathcal{F}(S, \mathbb{R})$, si ha $\lambda \cdot (f + g) = \lambda \cdot f + \lambda \cdot g$. Dalla definizione di funzione appena ricordata segue che due funzioni sono uguali precisamente quando hanno lo stesso valore su ogni punto. Basta e bisogna dunque verificare che $(\lambda \cdot (f + g))(s) = (\lambda \cdot f + \lambda \cdot g)(s)$ per ogni s in S . Fissato un generico s , usando le definizioni delle operazioni, i due membri dell'identità da verificare si possono riscrivere come:

$$\begin{aligned}(\lambda \cdot (f + g))(s) &= \lambda \cdot ((f + g)(s)) = \lambda \cdot (f(s) + g(s)), \\(\lambda \cdot f + \lambda \cdot g)(s) &= (\lambda \cdot f)(s) + (\lambda \cdot g)(s) = \lambda \cdot f(s) + \lambda \cdot g(s).\end{aligned}$$

Si conclude allora usando la distributività delle operazioni su \mathbb{R} .

Osserviamo che, assegnato un qualsiasi spazio vettoriale V , si può definire una struttura di spazio vettoriale su $\mathcal{F}(S, V)$ procedendo in modo del tutto analogo a quanto fatto per $\mathcal{F}(S, \mathbb{R})$. Notiamo anche che se $S = \{1, \dots, n\}$ c'è una corrispondenza naturale, che rispetta le operazioni, tra $\mathcal{F}(S, \mathbb{R})$ e \mathbb{R}^n . Ad una funzione f si associa il vettore la cui componente i -esima è $f(i)$.

Polinomi Anche l'insieme $\mathbb{R}[x]$ dei polinomi reali nella indeterminata x è uno spazio vettoriale reale. Basta utilizzare come somma quella già definita tra polinomi, che ha il polinomio nullo come elemento neutro, e il prodotto per scalare

$$\lambda \cdot (a_0 + a_1 \cdot x + \dots + a_k \cdot x^k) = (\lambda \cdot a_0) + (\lambda \cdot a_1) \cdot x + \dots + (\lambda \cdot a_k) \cdot x^k.$$

Notiamo che abbiamo anche un'operazione di prodotto tra elementi di $\mathbb{R}[x]$, ma essa non fa parte della struttura di spazio vettoriale su $\mathbb{R}[x]$.

3.2 Sottospazi vettoriali

Un'ampia classe di esempi di spazio vettoriale si ottiene considerando quei sottoinsiemi di uno spazio vettoriale che lo sono a loro volta rispetto alle operazioni "ereditate" dallo spazio in cui sono contenuti. Più precisamente diamo la seguente:

Definizione 3.2.1. Dato uno spazio vettoriale V , diciamo che un sottoinsieme W di V è un *sottospazio vettoriale* di V se valgono le proprietà seguenti:

1. W contiene 0 ;
2. Se w_1, w_2 sono in W , allora $w_1 + w_2$ è in W ;
3. Se λ è in \mathbb{R} e w è in W , allora $\lambda \cdot w$ è in W .

Se W è un sottospazio di V , le proprietà contenute nella definizione sono sufficienti a garantire che un'operazione di somma ed una di prodotto per scalare sono definite su W . Poiché queste operazioni provengono da quelle di V , e queste soddisfacevano le proprietà (1)-(8) della definizione di spazio vettoriale, le stesse proprietà continuano a valere per le operazioni di W . Quindi in effetti W è a sua volta uno spazio vettoriale.

Osserviamo che, supponendo valida la proprietà (3), la (1) può essere sostituita dalla seguente più debole:

- 1'. W è non vuoto.

Tuttavia nei casi pratici, volendo provare che un assegnato W è un sottospazio, conviene sempre partire dalla verifica che W contiene 0 . Osserviamo anche che, supponendo valida la (1'), la congiunzione delle proprietà (2) e (3) è equivalente alla seguente:

- Se λ_1, λ_2 sono in \mathbb{R} e w_1, w_2 sono in W , allora $\lambda_1 \cdot w_1 + \lambda_2 \cdot w_2$ è in W .

3.2.1 Esempi

Elenchiamo ora numerosi esempi e non esempi di sottospazio vettoriale, suggerendo al lettore di verificare gli assiomi oppure individuare quali di essi siano violati. Solo alla fine dell'elenco specificheremo quali sottoinsiemi siano sottospazi e quali non lo siano. Cominciamo considerando i seguenti sottoinsiemi dello spazio vettoriale \mathbb{R}^n :

1. $\{x \in \mathbb{R}^n : x_1 = 0\}$;
2. $\{x \in \mathbb{R}^n : x_1 + x_2 + \dots + x_n = 0\}$;
3. $\{x \in \mathbb{R}^n : x_1 = 1\}$;
4. $\mathbb{Z}^n \subset \mathbb{R}^n$;
5. $\{x \in \mathbb{R}^n : x_1 \in \mathbb{Z}\}$;

6. $\{x \in \mathbb{R}^2 : x_2 = x_1^2\}$;
7. $\{x \in \mathbb{R}^2 : x_1 \cdot x_2 = 0\}$.

Passiamo quindi a sottoinsiemi di $\mathbb{R}[x]$, fissando un intero $d \geq 0$:

8. $\{p(x) \in \mathbb{R}[x] : \deg(p(x)) \leq d\}$, insieme denotato nel seguito con $\mathbb{R}_{\leq d}[x]$;
9. $\{p(x) \in \mathbb{R}[x] : \deg(p(x)) = d\}$;
10. $\{p(x) \in \mathbb{R}[x] : p(1) = 0\}$, dove $p(1)$ è il valore reale che si ottiene sostituendo il valore 1 alla indeterminata x ;
11. $\{p(x) \in \mathbb{R}[x] : p(0) = 1\}$,
12. $\{p(x) \in \mathbb{R}[x] : p(x) \text{ è multiplo di } x^3\}$;
13. $\{p(x) \in \mathbb{R}[x] : \text{la divisione di } p(x) \text{ per } x^3 \text{ ha resto } -1 + 2x^2\}$;
14. $\{p(x) \in \mathbb{R}[x] : x^2 \cdot p''(x) = 20p(x)\}$, dove $p''(x)$ rappresenta il polinomio derivata seconda di $p(x)$.

Sia adesso S un insieme non vuoto, ed s_0 un suo elemento fissato. Consideriamo in $\mathcal{F}(S, \mathbb{R})$ i sottoinsiemi:

15. $\{f \in \mathcal{F}(S, \mathbb{R}) : f(s_0) = 0\}$;
16. $\{f \in \mathcal{F}(S, \mathbb{R}) : f(s_0) = 1\}$;
17. $\{f \in \mathcal{F}(S, \mathbb{R}) : f(s) \geq 0 \forall s \in S\}$.

Concludiamo con i seguenti sottoinsiemi dello spazio delle funzioni reali sull'intervallo $[0, 1]$ di \mathbb{R} e dello spazio $\mathcal{F}(\mathbb{N}, \mathbb{R})$ delle successioni reali:

18. $\{f \in \mathcal{F}([0, 1], \mathbb{R}) : f \text{ è continua}\}$;
19. $\{a \in \mathcal{F}(\mathbb{N}, \mathbb{R}) : a_{n+2} = a_{n+1} + a_n \forall n \in \mathbb{N}\}$.

Sono esempi di sottospazi gli insiemi elencati ai punti 1, 2, 8, 10, 12, 14, 15, 18, 19. Tutti gli altri non lo sono.

Concludiamo il capitolo con un commento che si allaccia alle motivazioni anticipate nel Capitolo 1. Con la nozione di spazio vettoriale abbiamo sviluppato un linguaggio unificato per trattare oggetti di natura "ovviamente" geometrica (come i punti dello spazio \mathbb{R}^n) e oggetti più complicati (come polinomi o funzioni). Questo approccio è molto utile perché consente di applicare

ragionamenti e intuizioni di natura geometrica anche a questi oggetti più complicati, che possono avere origine anche da problemi dell'ingegneria. Un caso tipico, ad esempio, è quello in cui si considera come spazio vettoriale uno spazio di funzioni che rappresentano onde elettromagnetiche.

Capitolo 4

Dimensione

In questo capitolo introduciamo una nozione chiave che consente di stabilire quali spazi vettoriali sono tra loro “diversi” e quali sono “uguali” dal punto di vista delle operazioni di spazio vettoriale.

4.1 Sottospazio generato e combinazioni lineari

Fissiamo per tutto questo paragrafo uno spazio vettoriale reale V . Cominciamo con il seguente semplice:

Lemma 4.1.1. *Se W, Z sono sottospazi vettoriali di V , anche $W \cap Z$ lo è.*

Dimostrazione. A titolo di esempio dimostriamo la proprietà (2) della definizione, le altre sono ugualmente elementari. Siano $x, y \in W \cap Z$. Dobbiamo verificare che $x + y \in W \cap Z$. Poiché $x, y \in W$, si ha $x + y \in W$ per la proprietà (2) della definizione di sottospazio, che per ipotesi vale per W . Analogamente $x, y \in Z$, dunque $x + y \in Z$. Abbiamo provato che $x + y \in W$ e $x + y \in Z$, perciò $x + y \in W \cap Z$, come richiesto. \square

Osservazione 4.1.2. Se W e Z sono sottospazi vettoriali di V , l'insieme $W \cup Z$ può non esserlo, come si vede considerando $V = \mathbb{R}^2$, $W = \mathbb{R} \times \{0\}$ e $Z = \{0\} \times \mathbb{R}$. Si può anzi verificare per esercizio che $W \cup Z$ è un sottospazio precisamente quando $W \subset Z$ oppure $Z \subset W$.

Enunciamo ora, dando soltanto un'idea della dimostrazione formale, la seguente:

Proposizione 4.1.3. *Se S è un sottoinsieme di V esiste $W \subset V$ tale che:*

1. W è un sottospazio vettoriale di V ;
2. W contiene S ;
3. Se Z è un altro sottospazio vettoriale di V che contiene S , allora Z contiene W .

Il sottospazio W descritto da questa proposizione si può dunque descrivere come il più piccolo sottospazio vettoriale di V contenente S , ed è molto semplice verificare che esso è univocamente determinato da S . Infatti, se W e W' soddisfano le proprietà elencate, si deduce che $W \subset W'$ e $W' \subset W$, dunque $W = W'$. Il sottospazio W sarà denotato con $\text{Span}(S)$ e chiamato il *sottospazio generato da S* . Diremo anche che S genera $\text{Span}(S)$ oppure che S è un *insieme di generatori* per $\text{Span}(S)$.

Diamo dunque, come annunciato, l'idea di come si dimostri la Proposizione 4.1.3. Il punto è quello di usare il Lemma 4.1.1. Infatti:

- Esistono dei W che soddisfano le proprietà (1) e (2), infatti basta prendere $W = V$;
- Se W_1 e W_2 soddisfano (1) e (2), anche $W_1 \cap W_2$ le soddisfa, ed è più piccolo di entrambi (o eventualmente uguale).

Perciò si continua a intersecare, fino a che non si trova il W più piccolo possibile. Formalmente però, per dare significato a questa idea, bisogna definire $\text{Span}(S)$ come l'intersezione di *tutti* i sottospazi vettoriali di V che contengono S , osservando preliminarmente che la famiglia dei sottospazi che contengono S non è vuota, dato che contiene il sottospazio banale V stesso, ed estendendo il Lemma 4.1.1 all'intersezione di una famiglia qualsiasi di sottospazi. Poiché vedremo a breve una caratterizzazione alternativa molto concreta di $\text{Span}(S)$, non ci soffermiamo oltre sulla dimostrazione della Proposizione 4.1.3.

Osservazione 4.1.4.

- $\text{Span}(\emptyset) = \{0\}$;
- $\text{Span}(\text{Span}(S)) = \text{Span}(S)$ per ogni S ;
- S è un sottospazio vettoriale di V se e soltanto se $\text{Span}(S) = S$.

Definizione 4.1.5. Chiamiamo *combinazione lineare* una scrittura del tipo

$$(\spadesuit) \quad \lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n$$

dove $n \geq 1$ è un intero, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ e $v_1, \dots, v_n \in V$. I numeri λ_i sono detti *coefficienti* della combinazione lineare.

Osserviamo che la scrittura (\spadesuit) , grazie all'associatività della somma in V , determina in modo non ambiguo un vettore di V , e di fatto si usa il termine *combinazione lineare* tanto per l'espressione (\spadesuit) quanto per il vettore che essa determina.

Mostriamo ora che $\text{Span}(S)$ può essere caratterizzato come l'insieme di *tutte le possibili combinazioni lineari di elementi di S* :

Proposizione 4.1.6. *Se $S \neq \emptyset$ si ha*

$$\text{Span}(S) = \{\lambda_1 \cdot s_1 + \dots + \lambda_n \cdot s_n : n \geq 1, \lambda_1, \dots, \lambda_n \in \mathbb{R}, s_1, \dots, s_n \in S\}.$$

Dimostrazione. Indichiamo con W l'insieme descritto a destra del segno di uguaglianza nell'enunciato, e dimostriamo che soddisfa le proprietà (1)-(3) che caratterizzano $\text{Span}(S)$, cioè quelle della Proposizione 4.1.3. Cominciamo con la (2). Volendo provare che $S \subset W$, dobbiamo prendere un generico elemento s di S e provare che sta in W . Basta infatti prendere $n = 1$, $\lambda_1 = 1$ e $s_1 = s$, e si ha $s = \lambda_1 \cdot s_1 \in W$. Provata la (2), che comporta in particolare che $W \neq \emptyset$, per dimostrare la (1) basta vedere che se $\mu, \mu' \in \mathbb{R}$ e $w, w' \in W$, allora $\mu \cdot w + \mu' \cdot w' \in W$. Per ipotesi si ha:

$$w = \lambda_1 \cdot s_1 + \dots + \lambda_n \cdot s_n, \quad w' = \lambda'_1 \cdot s'_1 + \dots + \lambda'_{n'} \cdot s'_{n'}$$

dunque

$$\begin{aligned} \mu \cdot w + \mu' \cdot w' &= (\mu \cdot \lambda_1) \cdot s_1 + \dots + (\mu \cdot \lambda_n) \cdot s_n + (\mu' \cdot \lambda'_1) \cdot s'_1 + \dots + (\mu' \cdot \lambda'_{n'}) \cdot s'_{n'} . \end{aligned}$$

Da questa espressione è evidente che $\mu \cdot w + \mu' \cdot w'$ è una combinazione lineare di al più $n + n'$ elementi di S (sono meno di $n + n'$ se qualcuno degli s_1, \dots, s_n è uguale a qualcuno degli $s'_1, \dots, s'_{n'}$), dunque appartiene a W .

Resta da verificare la proprietà (3). Supponiamo dunque che Z sia un sottospazio vettoriale che contiene S , e proviamo che contiene W . Per fare ciò prendiamo un generico elemento $w = \lambda_1 \cdot s_1 + \dots + \lambda_n \cdot s_n$ di W e dimostriamo che sta in Z . Infatti Z contiene S , dunque contiene s_1, \dots, s_n . Inoltre Z , essendo un sottospazio, contiene anche $\lambda_1 \cdot s_1, \dots, \lambda_n \cdot s_n$, e la loro somma w . Il nostro argomento è completo. \square

Corollario 4.1.7.

$$\text{Span}(\{v_1, \dots, v_n\}) = \left\{ \lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n : \lambda_1, \dots, \lambda_n \in \mathbb{R} \right\}.$$

4.2 Indipendenza lineare e basi

Il Corollario 4.1.7 afferma che se $S = \{v_1, \dots, v_n\}$ è un insieme finito di vettori, il sottospazio generato da S consiste delle combinazioni lineari di v_1, \dots, v_n . Possono però accadere due fenomeni qualitativamente ben distinti: l'espressione di un elemento di S come combinazione lineare di v_1, \dots, v_n può essere unica oppure non esserlo, e vedremo a breve che la stessa situazione si dà per tutti i vettori di $\text{Span}(S)$. Concentrandoci sul vettore nullo siamo dunque condotti alla nozione seguente:

Definizione 4.2.1. I vettori v_1, \dots, v_n si dicono *linearmente indipendenti* se l'unica loro combinazione lineare che ha come risultato il vettore nullo è quella con coefficienti tutti nulli.

Per chiarire la definizione discutiamo due casi concreti:

Esempio 4.2.2. Consideriamo in \mathbb{R}^2 i vettori

$$v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Affermiamo che v_1, v_2 sono linearmente indipendenti mentre v_1, v_2, v_3 non lo sono. Per provare la prima affermazione dobbiamo considerare una combinazione lineare generica $\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2$, supporre che essa dia il vettore nullo, e verificare che allora λ_1 e λ_2 sono per forza nulli. Infatti per ipotesi si ha

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 = \lambda_1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2\lambda_1 - \lambda_2 \\ \lambda_1 + 3\lambda_2 \end{pmatrix}.$$

Da questo segue che $\lambda_2 = 2\lambda_1$ e $7\lambda_1 = 0$, dunque $\lambda_1 = \lambda_2 = 0$. Ciò prova l'indipendenza lineare di v_1, v_2 . Per verificare invece che v_1, v_2, v_3 sono linearmente *dependenti* (l'opposto che indipendenti) basta vedere che l'equazione

$\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \lambda_3 \cdot v_3 = 0$ ha qualche soluzione $(\lambda_1, \lambda_2, \lambda_3)$ diversa da $(0, 0, 0)$. È facile vedere che una tale soluzione è ad esempio $(2, -1, -1)$.

Tornando alla questione che ha motivato la definizione di indipendenza lineare, fissiamo v_1, \dots, v_n e denotiamo per semplicità il loro generato con $\text{Span}(v_1, \dots, v_n)$, omettendo il simbolo di insieme.

Proposizione 4.2.3. *I vettori v_1, \dots, v_n sono linearmente indipendenti se e soltanto se ogni vettore di $\text{Span}(v_1, \dots, v_n)$ ha un'espressione unica come combinazione lineare di v_1, \dots, v_n .*

Dimostrazione. Se ogni vettore ha un'espressione unica, in particolare 0 ha un'espressione unica, dunque v_1, \dots, v_n sono linearmente indipendenti. Viceversa, supponiamo che v_1, \dots, v_n siano linearmente indipendenti, e per assurdo che un elemento di $\text{Span}(v_1, \dots, v_n)$ abbia due espressioni distinte

$$a_1 \cdot v_1 + \dots + a_n \cdot v_n, \quad b_1 \cdot v_1 + \dots + b_n \cdot v_n.$$

Allora

$$(a_1 - b_1) \cdot v_1 + \dots + (a_n - b_n) \cdot v_n = 0$$

e non tutti i coefficienti della combinazione lineare sono nulli. Questo contraddice l'indipendenza lineare di v_1, \dots, v_n e conclude la dimostrazione. \square

Definizione 4.2.4. Chiameremo *base* dello spazio vettoriale V un insieme finito e ordinato (v_1, \dots, v_n) di vettori che generino V e che siano linearmente indipendenti.

Il Corollario 4.1.7 e la Proposizione 4.2.3 comportano la seguente caratterizzazione:

Proposizione 4.2.5. *Sia $\mathcal{B} = (v_1, \dots, v_n)$ un insieme finito e ordinato di vettori di V . Sono fatti equivalenti:*

- \mathcal{B} è una base di V ;
- ogni vettore $v \in V$ si scrive in uno e un solo modo come combinazione lineare degli elementi di \mathcal{B} .

Spieghiamo subito la ragione per la quale, nella definizione di base, insistiamo sulla richiesta che l'insieme sia "ordinato", cioè ad esempio che (v_1, v_2, \dots, v_n) sia da considerare distinto da (v_n, \dots, v_2, v_1) , a meno che $n = 1$. La proposizione precedente garantisce che se $\mathcal{B} = (v_1, \dots, v_n)$ è una base di V allora ogni vettore v di V si scrive in modo unico come combinazione lineare $x_1 \cdot v_1 + \dots + x_n \cdot v_n$.

Chiamiamo ora $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ il vettore delle *coordinate* di v rispetto alla

base \mathcal{B} , e lo indichiamo con $[v]_{\mathcal{B}}$. Naturalmente, se la base \mathcal{B} non avesse uno specificato ordine, non si potrebbe definire $[v]_{\mathcal{B}}$ se non a meno di ambiguità nell'ordine delle coordinate.

Proposizione 4.2.6. *Data una base \mathcal{B} di V l'applicazione $\phi : V \rightarrow \mathbb{R}^n$ data da $\phi(v) = [v]_{\mathcal{B}}$ è bigettiva e trasforma le operazioni di spazio vettoriale di V in quelle di \mathbb{R}^n .*

Dimostrazione. La bigettività segue subito dalla Proposizione 4.2.5: l'inversa è data da $\phi^{-1}(x) = x_1 \cdot v_1 + \dots + x_n \cdot v_n$. La seconda affermazione significa che:

- $\phi(0) = 0$;
- $\phi(v + v') = \phi(v) + \phi(v')$ per ogni $v, v' \in V$;
- $\phi(\lambda \cdot v) = \lambda \cdot \phi(v)$ per ogni $\lambda \in \mathbb{R}$ e $v \in V$.

Proviamo ad esempio il secondo fatto. Se $\phi(v) = x$ e $\phi(v') = x'$ allora

$$v = x_1 \cdot v_1 + \dots + x_n \cdot v_n, \quad v' = x'_1 \cdot v_1 + \dots + x'_n \cdot v_n,$$

dunque

$$v + v' = (x_1 + x'_1) \cdot v_1 + \dots + (x_n + x'_n) \cdot v_n,$$

pertanto $[v + v']_{\mathcal{B}} = x + x'$, ovvero $\phi(v + v') = x + x'$. □

Esempio 4.2.7. Si può verificare che i vettori $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ sopra considerati costituiscono una base $\mathcal{B} = (v_1, v_2)$ di \mathbb{R}^2 . Inoltre le coordinate del vettore v_3 risultano come segue:

$$\left[\begin{pmatrix} 5 \\ -1 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Per diversi degli spazi vettoriali che abbiamo finora considerato è definita una *base canonica*, ovvero una base che riteniamo la più naturale e quindi privilegiata. Cominciamo con \mathbb{R}^n , nel quale abbiamo la base canonica $\mathcal{E}_n =$

(e_1, \dots, e_n) definita come segue:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad e_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Talvolta, per sottolineare che il vettore e_i fa parte della base di \mathbb{R}^n , dunque ha n componenti, scriveremo $e_i^{(n)}$ invece che semplicemente e_i . Notiamo che le coordinate di un vettore x di \mathbb{R}^n rispetto alla base canonica \mathcal{E}_n sono semplicemente le componenti di x .

Ci rivolgiamo ora allo spazio $\mathcal{M}_{n \times m}(\mathbb{R})$ delle matrici con n righe e m colonne. Per introdurre su $\mathcal{M}_{n \times m}(\mathbb{R})$ una base canonica, chiamiamo E_{ij} la matrice che ha coefficiente 1 al posto (i, j) , cioè sulla riga i -esima e colonna j -esima, e 0 altrove. Definiamo quindi $\mathcal{E}_{n \times m}$ come l'insieme delle matrici E_{ij} per $i = 1, \dots, n$ e $j = 1, \dots, m$, ordinate in modo *lessicografico* (ovvero come in un dizionario) rispetto alla coppia (i, j) .

Consideriamo infine lo spazio dei polinomi $\mathbb{R}_{\leq d}[x]$ di grado al più d nella indeterminata x . In esso scegliamo la base canonica $1, x, x^2, \dots, x^d$. Concludiamo il paragrafo provando che $\mathbb{R}[x]$ non ammette invece alcuna base, anzi dimostrando il seguente più forte fatto:

Proposizione 4.2.8. *Lo spazio $\mathbb{R}[x]$ non ammette alcun insieme finito di generatori.*

Dimostrazione. Supponiamo per assurdo che $\mathbb{R}[x]$ sia generato da polinomi $p_1(x), \dots, p_k(x)$. Indichiamo con d_i il grado di $p_i(x)$, e definiamo d come il massimo di d_1, \dots, d_k . Affermiamo che x^{d+1} non appartiene allo spazio generato da $p_1(x), \dots, p_k(x)$, il che è assurdo perché $x^{d+1} \in \mathbb{R}[x]$. Infatti ogni polinomio che sta nello spazio generato da $p_1(x), \dots, p_k(x)$ è della forma $\lambda_1 \cdot p_1(x) + \dots + \lambda_k \cdot p_k(x)$, dunque ha grado al più d , mentre x^{d+1} ha grado $d+1$. \square

4.3 Dimensione

In questo paragrafo dimostriamo il seguente risultato:

Teorema 4.3.1. *Se V è uno spazio vettoriale che ammette basi, due qualsiasi basi di V hanno lo stesso numero di elementi.*

Ricordando che una base consiste di un insieme (ordinato) di generatori linearmente indipendenti, è molto facile dedurre questo teorema dalla seguente proposizione, che invece ha dimostrazione non immediata:

Proposizione 4.3.2. *Sia V uno spazio vettoriale e siano dati $n, m \geq 0$ e vettori $w_1, \dots, w_m, v_1, \dots, v_n$ in V tali che:*

- w_1, \dots, w_m sono linearmente indipendenti;
- $w_1, \dots, w_m \in \text{Span}(v_1, \dots, v_n)$.

Allora $m \leq n$.

Dimostrazione. Proviamo il risultato per induzione su m , introducendo la seguente affermazione $\mathcal{P}(m)$ relativa a un intero m generico: *dati w_1, \dots, w_m linearmente indipendenti in V , un intero $n \geq 0$ e vettori v_1, \dots, v_n tali che $w_1, \dots, w_m \in \text{Span}(v_1, \dots, v_n)$, si ha $m \leq n$.*

Per il passo base potremmo iniziare dimostrando $\mathcal{P}(m)$ per $m = 0$, nel qual caso la conclusione è ovvia perché $n \geq 0 = m$ per ipotesi, ma dimostriamo direttamente anche $\mathcal{P}(m)$ per $m = 1$ per meglio illustrare il ragionamento. In questo caso l'insieme di vettori contenente il solo w_1 è linearmente indipendente, il che significa precisamente che w_1 è non nullo. Ora, dovendo mostrare che $n \geq 1$, supponiamo per assurdo che $n = 0$. In questo caso l'insieme dei v_i è vuoto, dunque il suo generato è $\{0\}$. Tuttavia per ipotesi w_1 appartiene al generato dei v_i , e questo è assurdo poiché w_1 è non nullo.

Ci rivolgiamo ora al passo induttivo, supponendo nota $\mathcal{P}(m)$ per un generico $m \geq 1$. Dobbiamo dimostrare $\mathcal{P}(m+1)$, dunque consideriamo $m+1$ vettori linearmente indipendenti x_0, \dots, x_m che stanno nel generato di altri vettori y_1, \dots, y_k . Si noti che stiamo per comodità usando gli indici tra 0 e m per gli $m+1$ vettori x_0, \dots, x_m . Stiamo inoltre usando nuovi simboli x_i, y_j e k , per avere a disposizione v_i, w_j e n per quando useremo l'ipotesi induttiva. Dobbiamo provare che $m+1 \leq k$.

Poiché $x_0, \dots, x_m \in \text{Span}(y_1, \dots, y_k)$ esistono coefficienti $\lambda_{ij} \in \mathbb{R}$ tali che

$$x_i = \sum_{j=1}^k \lambda_{ij} \cdot y_j, \quad i = 0, \dots, m.$$

Ora osserviamo che x_0 non è il vettore nullo, altrimenti x_0, \dots, x_m non potrebbero essere linearmente indipendenti. Perciò qualcuno dei coefficienti λ_{0j} è non nullo. Poiché l'ordine dei vettori y_1, \dots, y_k non ha alcuna rilevanza ai

fini della proprietà da verificare, possiamo modificare questo ordine e supporre per comodità che sia λ_{0k} il coefficiente non nullo. Definiamo ora:

$$w_1 = x_1 - \frac{\lambda_{1k}}{\lambda_{0k}} \cdot x_0, \quad \dots, \quad w_m = x_m - \frac{\lambda_{mk}}{\lambda_{0k}} \cdot x_0.$$

Poniamo inoltre $n = k - 1$ e $v_1 = y_1, \dots, v_n = y_{k-1}$, e affermiamo che l'ipotesi induttiva si applica ai vettori w_1, \dots, w_m e v_1, \dots, v_n . Questo è sufficiente per concludere, dato che $\mathcal{P}(m)$ garantisce allora che $m \leq n$, ma $n = k - 1$ dunque $m + 1 \leq k$, come dovevasi dimostrare.

Resta pertanto da provare che w_1, \dots, w_m sono linearmente indipendenti e stanno nel generato di v_1, \dots, v_n . Cominciamo con la prima affermazione, considerando una combinazione lineare che dia il vettore nullo:

$$\begin{aligned} 0 &= a_1 \cdot w_1 + \dots + a_m \cdot w_m \\ &= a_1 \cdot \left(x_1 - \frac{\lambda_{1k}}{\lambda_{0k}} \cdot x_0 \right) + \dots + a_m \cdot \left(x_m - \frac{\lambda_{mk}}{\lambda_{0k}} \cdot x_0 \right) \\ &= a_0 \cdot x_0 + a_1 \cdot x_1 + \dots + a_m \cdot x_m \end{aligned}$$

dove a_0 è un coefficiente che dipende dagli altri a_i e dai λ_{ik} . Poiché per ipotesi x_0, \dots, x_m sono linearmente indipendenti, concludiamo che gli a_i , in particolare a_1, \dots, a_m come dovevamo dimostrare, sono nulli.

Proviamo ora che $w_i \in \text{Span}(v_1, \dots, v_n)$. Infatti

$$\begin{aligned} w_i &= x_i - \frac{\lambda_{ik}}{\lambda_{0k}} \cdot x_0 \\ &= \sum_{j=1}^k \lambda_{ij} \cdot y_j - \frac{\lambda_{ik}}{\lambda_{0k}} \cdot \sum_{j=1}^k \lambda_{0j} \cdot y_j \\ &= \sum_{j=1}^{k-1} \left(\lambda_{ij} - \frac{\lambda_{ik}}{\lambda_{0k}} \cdot \lambda_{0j} \right) \cdot y_j + \underbrace{\left(\lambda_{ik} - \frac{\lambda_{ik}}{\lambda_{0k}} \cdot \lambda_{0k} \right)}_0 \cdot y_k. \end{aligned}$$

Poiché $v_1 = y_1, \dots, v_n = y_{k-1}$ la dimostrazione è conclusa. \square

Grazie al teorema appena dimostrato possiamo ora dare la seguente:

Definizione 4.3.3. Se V è uno spazio vettoriale reale che ammette basi, chiamiamo *dimensione* di V il numero di elementi di una sua qualsiasi base. Per sottolineare che stiamo parlando di spazi vettoriali su \mathbb{R} , mentre più avanti

considereremo anche spazi su \mathbb{C} , chiamiamo più precisamente *dimensione reale* di V l'intero appena introdotto, e lo indichiamo con $\dim_{\mathbb{R}}(V)$. Gli spazi vettoriali che ammettono basi saranno nel seguito detti *di dimensione finita*.

Dalla Proposizione 4.3.2 segue facilmente il seguente:

Corollario 4.3.4. *Sia V uno spazio vettoriale di dimensione finita n e siano v_1, \dots, v_k vettori di V . Si ha:*

- *Se v_1, \dots, v_k sono linearmente indipendenti, allora $k \leq n$;*
- *Se v_1, \dots, v_k generano V , allora $k \geq n$.*

Concludiamo il paragrafo provando come annunciato che è la dimensione che consente di decidere quando due spazi distinti siano “uguali” o meno dal punto di vista delle operazioni di spazio vettoriale:

Proposizione 4.3.5. *Siano V e W spazi vettoriali reali di dimensione finita. Allora sono fatti equivalenti:*

1. *Esiste un'applicazione bigettiva $\phi : V \rightarrow W$ che rispetta le operazioni;*
2. $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(W)$.

Dimostrazione. Supponiamo che valga la proprietà (1) e consideriamo una base (v_1, \dots, v_n) di V . Affermiamo che $(\phi(v_1), \dots, \phi(v_n))$ è una base di W , il che comporta la (2). Per provarlo, verifichiamo che ogni vettore di w di W si esprime in uno e un solo modo come combinazione lineare degli elementi di $(\phi(v_1), \dots, \phi(v_n))$. Infatti $\phi^{-1}(w)$ ha un'unica espressione come

$$a_1 \cdot v_1 + \dots + a_n \cdot v_n,$$

da cui segue che $w = a_1 \cdot \phi(v_1) + \dots + a_n \cdot \phi(v_n)$, ed è facile vedere che l'espressione è ancora unica.

Supponiamo ora la (2) e proviamo la (1). La Proposizione 4.2.6 comporta che abbiamo applicazioni bigettive $\phi_V : V \rightarrow \mathbb{R}^n$ e $\phi_W : W \rightarrow \mathbb{R}^n$ che rispettano le operazioni. Basta allora porre $\phi = \phi_W^{-1} \circ \phi_V$. \square

4.4 Estrazione e completamento

Ricordiamo che una base di uno spazio V è un insieme finito e ordinato di vettori che hanno le due proprietà di generare e di essere linearmente indipendenti. In questo paragrafo spieghiamo come si può modificare un insieme che abbia una sola di queste proprietà a un altro che le abbia entrambe. Cominciamo con il semplice ma fondamentale:

Lemma 4.4.1. *Siano dati $v_0, v_1, \dots, v_n \in V$ con v_1, \dots, v_n linearmente indipendenti. Sono fatti equivalenti:*

1. v_0, v_1, \dots, v_n sono linearmente indipendenti;
2. $v_0 \notin \text{Span}(v_1, \dots, v_n)$.

Dimostrazione. Supponiamo che valga la (1) e proviamo la (2) per assurdo, ipotizzando che $v_0 = \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$. Allora $(-1) \cdot v_0 + \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n = 0$, il che fornisce una contraddizione.

Supponiamo ora invece la (2) e consideriamo una combinazione lineare $\alpha_0 \cdot v_0 + \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$ che dia il vettore nullo. Dobbiamo provare che allora per forza tutti gli α_i sono nulli. Se $\alpha_0 = 0$ naturalmente tutti gli altri α_i sono nulli, perché v_1, \dots, v_n sono linearmente indipendenti per ipotesi. Supponiamo invece che α_0 sia non nullo. Allora

$$v_0 = \left(-\frac{\alpha_1}{\alpha_0}\right) \cdot v_1 + \dots + \left(-\frac{\alpha_n}{\alpha_0}\right) \cdot v_n.$$

Ciò contraddice la (2) e conclude la dimostrazione. \square

Corollario 4.4.2. *Se V non ha dimensione finita allora in V esistono insiemi di vettori linearmente indipendenti con un numero arbitrariamente grande di elementi.*

Dimostrazione. Supponiamo per assurdo che il numero massimo di elementi di un insieme di vettori di V linearmente indipendenti sia finito, e chiamiamolo k . Se v_1, \dots, v_k sono linearmente indipendenti, la massimalità di k e il lemma precedente comportano che v_1, \dots, v_k generano, dunque V ha dimensione k . Assurdo. \square

Proviamo ora che un insieme di vettori linearmente indipendenti in uno spazio di dimensione finita si può sempre *completare a una base*, e che da un insieme di generatori è possibile *estrarre una base*. Lo facciamo nel caso più generale di un sottospazio:

Teorema 4.4.3. *Sia V uno spazio vettoriale di dimensione finita n e sia W un suo sottospazio vettoriale diverso da $\{0\}$. Allora:*

1. W ha dimensione finita m ; inoltre $m \leq n$ e se $m = n$ si ha $W = V$;
2. Dati w_1, \dots, w_k in W linearmente indipendenti, esistono w_{k+1}, \dots, w_m tali che (w_1, \dots, w_m) sia una base di W ;

3. Dati w_1, \dots, w_h generatori di W , esistono indici i_1, \dots, i_m con

$$1 \leq i_1 < i_2 < \dots < i_{m-1} < i_m \leq h$$

tali che $(w_{i_1}, \dots, w_{i_m})$ sia una base di W .

Dimostrazione. Cominciamo dal punto (1). Il fatto che W abbia dimensione finita segue dal Corollario 4.4.2. Infatti se W avesse dimensione infinita allora conterrebbe insiemi di vettori linearmente indipendenti con più di n elementi, il che contraddice il Corollario 4.3.4. La successiva affermazione segue ancora dal Corollario 4.3.4. Infatti se (w_1, \dots, w_m) è una base di W in particolare w_1, \dots, w_m sono linearmente indipendenti, dunque $m \leq n$. Inoltre se $m = n$ ma W è diverso dall'intero V , si può applicare il Lemma 4.4.1 deducendo l'assurdo che in V ci sono $n + 1$ vettori linearmente indipendenti.

Rivolgiamoci dunque al punto (2). Se w_1, \dots, w_k sono già un insieme di generatori di W non c'è nulla da dimostrare. Supponiamo dunque che esista $w_{k+1} \in W \setminus \text{Span}(w_1, \dots, w_k)$. Il lemma precedente garantisce allora che w_1, \dots, w_k, w_{k+1} sono linearmente indipendenti, e possiamo applicare di nuovo il medesimo ragionamento, trovando che o w_1, \dots, w_k, w_{k+1} generano o è possibile estenderli a un sistema $w_1, \dots, w_k, w_{k+1}, w_{k+2}$ di vettori linearmente indipendenti. L'iterazione di questo ragionamento ha in ogni caso termine, e dunque conduce a una base di W , dato che in W non si possono trovare più di m vettori linearmente indipendenti, grazie al Corollario 4.3.4.

Passiamo al punto (3). L'idea è in questo caso di esaminare uno dopo l'altro i vettori w_1, \dots, w_h decidendo per ciascuno se vada tenuto o debba essere scartato. Andranno scartati quei w_i che appartengono al generato dei vettori già tenuti, mentre andranno tenuti quelli che sono linearmente indipendenti dai precedenti. Più precisamente, procediamo come segue:

- Scartiamo tutti i primi vettori w_1, w_2, \dots finché non troviamo il primo vettore w_{i_1} non nullo; naturalmente se già w_1 è non nullo si pone $i_1 = 1$ e non si scarta nessun vettore;
- Dopo w_{i_1} , scartiamo tutti i vettori $w_{i_1+1}, w_{i_1+2}, \dots$ che appartengono al generato di w_{i_1} , finché non ne troviamo uno w_{i_2} che non appartiene al generato di w_{i_1} . Il Lemma 4.4.1 garantisce che w_{i_1}, w_{i_2} sono linearmente indipendenti. Naturalmente può succedere che già w_{i_1+1} non stia nel generato di w_{i_1} , oppure che scorriamo tutta la lista senza trovare vettori che non stiano nel generato di w_{i_1} . Nel primo caso non scartiamo alcun vettore e poniamo $i_2 = i_1 + 1$, nel secondo (w_{i_1}) è già la base cercata;

- Procediamo similmente, considerando il primo vettore w_{i_3} dopo w_{i_2} che non appartiene al generato di w_{i_1}, w_{i_2} , oppure fermandoci se non ne troviamo, e così via.

È chiaro dalla costruzione che i vettori w_{i_1}, \dots, w_{i_m} risultanti sono linearmente indipendenti. Inoltre è facile vedere che $w_j \in \text{Span}(w_{i_1}, \dots, w_{i_p})$ se w_{i_p} è l'ultimo vettore non scartato con $i_p \leq j$. Questo comporta che w_{i_1}, \dots, w_{i_m} generano tutti i w_j , dunque l'intero W , perciò sono una base di W . \square

Osservazione 4.4.4. Il procedimento di estrazione di una base da un insieme finito di generatori, come descritto nella dimostrazione del punto (3) del teorema precedente, ha una natura *algoritmica*, cioè può essere facilmente eseguito da un calcolatore.

Deduciamo ora dal Teorema 4.4.3 che, avendo già il numero giusto di vettori, è sufficiente richiedere “mezza definizione” di base per avere anche “l'altra mezza”:

Corollario 4.4.5. *Sia V uno spazio vettoriale di dimensione n . Per un insieme v_1, \dots, v_n di vettori sono fatti equivalenti:*

1. v_1, \dots, v_n sono linearmente indipendenti;
2. v_1, \dots, v_n generano V .

Dimostrazione. Se sono linearmente indipendenti ma non generano, possiamo completarli a una base. Se generano ma non sono linearmente indipendenti, possiamo estrarne una base. In entrambi i casi sarebbe contraddetto il Corollario 4.3.4. \square

4.5 Formula di Grassmann

Sia fissato uno spazio vettoriale V e siano W e Z suoi sottospazi vettoriali. Abbiamo già osservato che $W \cup Z$ non è un sottospazio vettoriale, a meno che non sia uguale a W oppure a Z . Chiamiamo allora *somma* di W e Z , e indichiamo con $W + Z$, il sottospazio $\text{Span}(W \cup Z)$, ovvero il più piccolo sottospazio vettoriale che contiene sia W sia Z . La notazione è giustificata dal seguente:

Lemma 4.5.1. $W + Z = \{w + z : w \in W, z \in Z\}$.

Dimostrazione. Basta osservare che l'insieme descritto a destra del segno di uguaglianza è un sottospazio vettoriale, che contiene $W \cup Z$, e che è contenuto in $\text{Span}(W \cup Z)$. \square

Supponiamo ora che V abbia dimensione finita. Allora i sottospazi W , Z , $W \cap Z$, $W + Z$ hanno tutti una ben definita dimensione. La seguente uguaglianza, detta formula di Grassmann, stabilisce un legame tra le dimensioni di questi quattro sottospazi:

Teorema 4.5.2. $\dim_{\mathbb{R}}(W) + \dim_{\mathbb{R}}(Z) = \dim_{\mathbb{R}}(W \cap Z) + \dim_{\mathbb{R}}(W + Z)$.

Dimostrazione. Indichiamo con k la dimensione di $W \cap Z$ e notiamo che W e Z contengono $W \cap Z$, dunque hanno dimensione almeno k . Poniamo dunque $\dim_{\mathbb{R}}(W) = k + p$ e $\dim_{\mathbb{R}}(Z) = k + q$, dove $p, q \geq 0$. Dobbiamo allora provare che $Z + W$ ha dimensione $k + p + q$.

Fissiamo ora una base (y_1, \dots, y_k) di $Z \cap W$. Notando che y_1, \dots, y_k sono vettori linearmente indipendenti e stanno in W , utilizziamo il punto (2) del Teorema 4.4.3 ottenendo una base $(y_1, \dots, y_k, w_1, \dots, w_p)$ di W . Similmente costruiamo una base $(y_1, \dots, y_k, z_1, \dots, z_q)$ di Z . Affermiamo ora che $(y_1, \dots, y_k, w_1, \dots, w_p, z_1, \dots, z_q)$ è una base di $W + Z$, il che comporta la conclusione desiderata.

Cominciamo col mostrare che $y_1, \dots, y_k, w_1, \dots, w_p, z_1, \dots, z_q$ sono generatori di $W + Z$. Consideriamo un generico elemento di $W + Z$, che grazie al lemma precedente si può scrivere come $w + z$ con $w \in W$ e $z \in Z$. Dobbiamo provare che $w + z$ si può esprimere come combinazione lineare di $y_1, \dots, y_k, w_1, \dots, w_p, z_1, \dots, z_q$.

Ora $y_1, \dots, y_k, w_1, \dots, w_p$ generano W , dunque

$$w = \alpha_1 \cdot y_1 + \dots + \alpha_k \cdot y_k + \beta_1 \cdot w_1 + \dots + \beta_p \cdot w_p$$

per costanti α_i e β_j opportune. Similmente

$$z = \alpha'_1 \cdot y_1 + \dots + \alpha'_k \cdot y_k + \gamma_1 \cdot z_1 + \dots + \gamma_q \cdot z_q.$$

Dunque

$$w + z = (\alpha_1 + \alpha'_1) \cdot y_1 + \dots + (\alpha_k + \alpha'_k) \cdot y_k + \beta_1 \cdot w_1 + \dots + \beta_p \cdot w_p + \gamma_1 \cdot z_1 + \dots + \gamma_q \cdot z_q$$

e abbiamo ottenuto quanto voluto.

Passiamo a dimostrare che $y_1, \dots, y_k, w_1, \dots, w_p, z_1, \dots, z_q$ sono linearmente indipendenti. Supponiamo dunque di avere una combinazione lineare con risultato nullo

$$(\heartsuit) \quad \alpha_1 \cdot y_1 + \dots + \alpha_k \cdot y_k + \beta_1 \cdot w_1 + \dots + \beta_p \cdot w_p + \gamma_1 \cdot z_1 + \dots + \gamma_q \cdot z_q = 0.$$

Dobbiamo provare che tutti gli α_i , β_j e γ_l sono nulli. Dall'identità precedente deduciamo che

$$\alpha_1 \cdot y_1 + \dots + \alpha_k \cdot y_k + \beta_1 \cdot w_1 + \dots + \beta_p \cdot w_p = -\gamma_1 \cdot z_1 - \dots - \gamma_q \cdot z_q.$$

Ora in questa identità il membro sinistro descrive un elemento di W , mentre il membro destro descrive un elemento di Z . L'identità comporta dunque che entrambi i membri appartengono a $W \cap Z$. In particolare il membro destro si può scrivere come combinazione lineare dei vettori y_1, \dots, y_k , che costituiscono una base di $W \cap Z$:

$$\begin{aligned} -\gamma_1 \cdot z_1 - \dots - \gamma_q \cdot z_q &= \alpha'_1 \cdot y_1 + \dots + \alpha'_k \cdot y_k \\ \Rightarrow \alpha'_1 \cdot y_1 + \dots + \alpha'_k \cdot y_k + \gamma_1 \cdot z_1 + \dots + \gamma_q \cdot z_q &= 0. \end{aligned}$$

Ne segue che tutti i coefficienti α'_i e γ_l sono nulli, poiché $y_1, \dots, y_k, z_1, \dots, z_q$ sono linearmente indipendenti (costituiscono una base di Z). La (\heartsuit) allora comporta che

$$\alpha_1 \cdot y_1 + \dots + \alpha_k \cdot y_k + \beta_1 \cdot w_1 + \dots + \beta_p \cdot w_p = 0$$

dunque anche gli α_i e β_j sono nulli, poiché $y_1, \dots, y_k, w_1, \dots, w_p$ sono linearmente indipendenti (costituiscono una base di W). La dimostrazione è completa. \square

Capitolo 5

Applicazioni lineari

Abbiamo già incontrato, nel capitolo precedente, applicazioni tra spazi vettoriali che rispettano le operazioni. Questo capitolo è dedicato per intero alle applicazioni aventi questa proprietà.

5.1 Definizione, esempi, nucleo, immagine

Siano fissati in questo paragrafo spazi vettoriali reali V e W . Una applicazione $f : V \rightarrow W$ è detta *lineare* se:

- $f(0) = 0$;
- $f(v_1 + v_2) = f(v_1) + f(v_2)$ per ogni $v_1, v_2 \in V$;
- $f(\lambda \cdot v) = \lambda \cdot f(v)$ per ogni $\lambda \in \mathbb{R}$ e $v \in V$.

Osservazione 5.1.1. In realtà la prima delle tre proprietà è conseguenza di ciascuna delle altre due, ma di fatto è conveniente, per verificare la linearità di una f assegnata, partire dalla verifica che $f(0) = 0$.

Osservazione 5.1.2. Una applicazione $f : V \rightarrow W$ è lineare se e soltanto se *rispetta le combinazioni lineari*, ovvero:

- $f(\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2) = \lambda_1 \cdot f(v_1) + \lambda_2 \cdot f(v_2)$ per ogni $\lambda_1, \lambda_2 \in \mathbb{R}$ e $v_1, v_2 \in V$.

Elenchiamo subito diversi esempi di applicazioni che sono (oppure non sono) lineari, specificandone la natura solo alla fine della lista:

1. $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = -7x$;

2. $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \begin{cases} x & \text{se } x \in \mathbb{Q}, \\ 0 & \text{altrimenti;} \end{cases}$
3. $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \operatorname{sgn}(x) \cdot \sqrt{|x \cdot y|}$, dove $\operatorname{sgn}(x)$ è il *segno* di x , dato da

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{se } x > 0, \\ 0 & \text{se } x = 0, \\ -1 & \text{se } x < 0; \end{cases}$$
4. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \begin{pmatrix} x^2 \\ y^2 \\ x \cdot y \end{pmatrix}$;
5. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \begin{pmatrix} x - y \\ y + 2x \\ x - 3y \end{pmatrix}$;
6. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \begin{pmatrix} |x| \\ |y| \end{pmatrix}$;
7. $f : \mathcal{F}(S, \mathbb{R}) \rightarrow \mathbb{R}$, $f(x) = x(s_0)$ dove $s_0 \in S$ è fissato;
8. $f : \mathcal{F}(S, \mathbb{R}) \rightarrow \mathbb{R}$, $f(x) = x(s_0) - x(s_1)$ dove $s_0, s_1 \in S$ sono fissati e distinti;
9. $f : \mathcal{F}(S, \mathbb{R}) \rightarrow \mathbb{R}$, $f(x) = x(s_0) + 1$ dove $s_0 \in S$ è fissato;
10. $f : \{x \in \mathcal{F}([0, 1], \mathbb{R}) : x \text{ continua}\} \rightarrow \mathbb{R}$, $f(x) = \int_0^1 x(t) dt$;
11. $f : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$, $f\left(\sum_{k=0}^d a_k \cdot x^k\right) = \sum_{k=1}^d k \cdot a_k \cdot x^{k-1}$, ovvero la *derivazione* di polinomi;
12. $\operatorname{tr} : \mathcal{M}_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$, $\operatorname{tr}(A) = \sum_{i=1}^n (A)_{ii}$ (tale applicazione è detta *traccia*, e associa a una matrice quadrata la somma degli elementi sulla diagonale *principale*);
13. $\mathcal{M}_{n \times m}(\mathbb{R}) \rightarrow \mathcal{M}_{m \times n}(\mathbb{R})$, $A \mapsto {}^t A$, dove $({}^t A)_{ij} = (A)_{ji}$. Tale applicazione è detta *trasposizione*, e la ${}^t A$ è detta *trasposta* di A .

Sono lineari le applicazioni descritte ai punti (1), (5), (7), (8), (10), (11), (12), (13). Non lo sono le altre (ma è consigliabile verificare per ciascuna quali siano gli assiomi violati).

Prima di procedere abbiamo un risultato molto semplice, da dimostrare per esercizio:

Proposizione 5.1.3. *Date applicazioni $f : V \rightarrow W$ e $g : W \rightarrow Z$ lineari, la loro composizione $g \circ f : V \rightarrow Z$ è lineare.*

Ad un'applicazione lineare $f : V \rightarrow W$ associamo ora due sottospazi vettoriali, l'uno di V e l'altro di W :

- Il *nucleo* di f , definito come $\text{Ker}(f) = \{v \in V : f(v) = 0\}$;
- L'*immagine* di f , denotata con $\text{Im}(f)$.

Un primo facile risultato consente di dare una chiara interpretazione, in termini delle proprietà di f , del fatto che uno di questi sottospazi sia banale:

Proposizione 5.1.4. *Sia $f : V \rightarrow W$ lineare. Allora:*

- f è *iniettiva* se e soltanto se $\text{Ker}(f) = \{0\}$;
- f è *surgettiva* se e soltanto se $\text{Im}(f) = W$;
- f è *l'applicazione identicamente nulla* se e soltanto se $\text{Ker}(f) = V$, oppure equivalentemente $\text{Im}(f) = \{0\}$.

Dimostrazione. Solo la prima affermazione non è evidente. Se f è iniettiva, essendo $f(0) = 0$, si ha $f^{-1}(0) = \{0\}$, dunque $\text{Ker}(f) = \{0\}$. Viceversa, siano $v, v' \in V$ tali che $f(v) = f(v')$. Allora $f(v - v') = 0$, dunque $v - v' \in \text{Ker}(f)$. Se quest'ultimo sottospazio è quello banale $\{0\}$, ne segue che $v = v'$, dunque f è iniettiva. \square

Corollario 5.1.5. *Sia $f : V \rightarrow W$ lineare con V e W di dimensione finita. Allora:*

- f è *iniettiva* se e soltanto se $\dim_{\mathbb{R}}(\text{Ker}(f)) = 0$;
- f è *surgettiva* se e soltanto se $\dim_{\mathbb{R}}(\text{Im}(f)) = \dim_{\mathbb{R}}(W)$.

Abbiamo ora un risultato, chiamato *formula della dimensione*, che lega tra loro la dimensione del nucleo, quella dell'immagine e quella dello spazio di partenza.

Teorema 5.1.6. *Sia $f : V \rightarrow W$ un'applicazione lineare con V di dimensione finita. Allora:*

$$\dim_{\mathbb{R}}(\text{Ker}(f)) + \dim_{\mathbb{R}}(\text{Im}(f)) = \dim_{\mathbb{R}}(V).$$

Dimostrazione. Poiché V ha dimensione finita e $\text{Ker}(f)$ ne è un sottospazio, anche quest'ultimo ha dimensione finita, che indichiamo con k . Possiamo dunque scegliere una base (v_1, \dots, v_k) di $\text{Ker}(f)$, e completarla a una base (v_1, \dots, v_n) di V . Asseriamo ora che $(f(v_{k+1}), \dots, f(v_n))$ è una base di $\text{Im}(f)$, il che comporta immediatamente la tesi.

Proviamo dunque che $f(v_{k+1}), \dots, f(v_n)$ generano $\text{Im}(f)$ e sono linearmente indipendenti.

Supponiamo che una combinazione lineare $\alpha_{k+1} \cdot f(v_{k+1}) + \dots + \alpha_n \cdot f(v_n)$ sia il vettore nullo. Dalla linearità di f deduciamo che

$$f(\alpha_{k+1} \cdot v_{k+1} + \dots + \alpha_n \cdot v_n) = 0 \quad ,$$

dunque

$$\alpha_{k+1} \cdot v_{k+1} + \dots + \alpha_n \cdot v_n \in \text{Ker}(f).$$

Avendo scelto (v_1, \dots, v_k) come base di $\text{Ker}(f)$, ne deduciamo che esistono coefficienti $\alpha_1, \dots, \alpha_k$ tali che

$$\begin{aligned} & \alpha_{k+1} \cdot v_{k+1} + \dots + \alpha_n \cdot v_n = \alpha_1 \cdot v_1 + \dots + \alpha_k \cdot v_k \\ \Rightarrow & (-\alpha_1) \cdot v_1 + \dots + (-\alpha_k) \cdot v_k + \alpha_{k+1} \cdot v_{k+1} + \dots + \alpha_n \cdot v_n = 0 \\ \Rightarrow & \alpha_1 = \dots = \alpha_k = \alpha_{k+1} = \dots = \alpha_n = 0 \quad . \end{aligned}$$

Ciò prova l'indipendenza lineare di $f(v_{k+1}), \dots, f(v_n)$. Per dimostrare che generano $\text{Im}(f)$, consideriamo un vettore $w \in \text{Im}(f)$ generico. Per definizione di immagine, esiste $v \in V$ tale che $f(v) = w$. Ora, (v_1, \dots, v_n) è una base di V , dunque

$$v = \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n.$$

Usiamo ora la linearità di f e il fatto che $v_1, \dots, v_k \in \text{Ker}(f)$, deducendo che

$$w = f(v) = \alpha_{k+1} \cdot f(v_{k+1}) + \dots + \alpha_n \cdot f(v_n)$$

e dunque la conclusione desiderata. □

Corollario 5.1.7. *Sia $f : V \rightarrow W$ lineare con V e W di dimensione finita.*

- *Se f è iniettiva allora $\dim_{\mathbb{R}}(V) \leq \dim_{\mathbb{R}}(W)$;*

- Se f è surgettiva allora $\dim_{\mathbb{R}}(V) \geq \dim_{\mathbb{R}}(W)$;
- Se f è bigettiva allora $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(W)$.

Inoltre, se V e W hanno la stessa dimensione, sono fatti equivalenti:

- f è iniettiva;
- f è surgettiva;
- f è bigettiva.

Mostriamo nel prossimo capitolo che la conoscenza della dimensione dell'immagine di un'assegnata applicazione lineare f (un numero intero che chiameremo *rango* di f) ha cruciale importanza in molti problemi (ad esempio quello della risoluzione dei sistemi lineari). Illustreremo anche un metodo efficace di calcolo del rango.

5.2 Somme dirette e proiezioni

Se V è uno spazio vettoriale e W, Z sono sottospazi di V tali che $W + Z = V$ e $W \cap Z = \{0\}$, diremo che V è decomposto nella *somma diretta* di W e Z e scriveremo $V = W \oplus Z$.

Lemma 5.2.1. *Se $V = W \oplus Z$ allora ogni vettore v di V si scrive in uno e un solo modo come $v = w + z$ con $w \in W$ e $z \in Z$.*

Dimostrazione. Per la caratterizzazione di $W + Z$ data nel Lemma 4.5.1 sappiamo che almeno un'espressione $v = w + z$ esiste. Supponiamo che ce ne siano due:

$$v = w_1 + z_1, \quad v = w_2 + z_2.$$

Ne deduciamo che $w_1 - w_2 = z_2 - z_1$. L'espressione sinistra di questo vettore comporta che esso appartiene a W , mentre quella destra comporta che appartiene a Z . Tuttavia per ipotesi si ha $W \cap Z = \{0\}$, dunque $w_1 - w_2 = z_2 - z_1 = 0$. Ne segue che $w_1 = w_2$ e $z_1 = z_2$, dunque le due espressioni di v erano in realtà uguali. \square

Nella situazione del lemma precedente, ponendo $p(v) = w$ e $q(v) = z$, si definiscono due applicazioni $p, q : V \rightarrow V$, dette *proiezioni* relative alla decomposizione in somma diretta $V = W \oplus Z$. Vale la seguente:

Proposizione 5.2.2. *1. p e q sono lineari;*

2. $\text{Im}(p) = \text{Ker}(q) = W$, $\text{Im}(q) = \text{Ker}(p) = Z$;
3. $p \circ q = q \circ p = 0$;
4. $p \circ p = p$, $q \circ q = q$;
5. $p + q = \text{Id}_V$.

Dimostrazione. Per il primo punto, verifichiamo ad esempio che $p(v_1 + v_2) = p(v_1) + p(v_2)$. Per definizione di p si ha $p(v_1) = w_1$ e $p(v_2) = w_2$ se $w_1, w_2 \in W$ ed esistono $z_1, z_2 \in Z$ tali che $v_1 = w_1 + z_1$ e $v_2 = w_2 + z_2$. Da queste relazioni segue che $v_1 + v_2 = (w_1 + w_2) + (z_1 + z_2)$. Inoltre $w_1 + w_2 \in W$ e $z_1 + z_2 \in Z$. Ancora dalla definizione di p segue allora che $p(v_1 + v_2) = w_1 + w_2$, ovvero la conclusione desiderata.

È chiaro dalla definizione che $\text{Im}(p) \subset W$. Inoltre se $w \in W$ si ha $w = w + 0$ e $0 \in Z$, dunque $p(w) = w$. Ne segue che $\text{Im}(p) = W$. Se $z \in Z$ si ha $z = 0 + z$ e $0 \in W$, dunque $p(z) = 0$, cioè $z \in \text{Ker}(p)$. Viceversa se $p(v) = 0$ allora $v = p(v) + z = 0 + z = z \in Z$. Dunque $\text{Ker}(p) = Z$. Per $\text{Im}(q)$ e $\text{Ker}(q)$ si procede analogamente. Il terzo punto segue immediatamente dal secondo.

Se $v \in V$ si ha $(p \circ p)(v) = p(p(v)) = p(v)$ poiché $p(v) \in W$ e abbiamo già osservato che $p(w) = w$ se $w \in W$. Ciò prova il quarto punto.

L'ultimo punto segue immediatamente dalla definizione. (Si noti che stiamo usando qui l'operazione di somma tra applicazioni a valori in V ; torneremo ampiamente su questo concetto nella Sezione 5.4). \square

Osservazione 5.2.3. La proiezione p da V sul fattore W della decomposizione $V = W \oplus Z$ dipende non dal solo W , ma anche da Z . Ad esempio se $V = \mathbb{R}^2$ e $W = \text{Span}(e_1)$, $Z_1 = \text{Span}(e_2)$, $Z_2 = \text{Span}(e_1 + e_2)$, abbiamo $V = W \oplus Z_1 = W \oplus Z_2$, e le due proiezioni su W relative alle diverse decomposizioni si esprimono come segue:

$$p_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \quad p_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ 0 \end{pmatrix}.$$

Nel Paragrafo 9.3 vedremo che in presenza di un'ulteriore struttura su V (un prodotto scalare) si può associare a un singolo sottospazio una ben definita proiezione.

Il seguente risultato, insieme alla Proposizione 5.2.2, caratterizza le proiezioni.

Proposizione 5.2.4. *Sia V uno spazio vettoriale di dimensione finita. Data $p : V \rightarrow V$ lineare tale che $p \circ p = p$, posto $W = \text{Im}(p)$ e $Z = \text{Ker}(p)$, si ha $V = W \oplus Z$ e la proiezione su W associata a questa decomposizione è p .*

Dimostrazione. Sia $x \in W \cap Z$. Poiché $x \in W$ si ha $x = p(v)$. Dato che $p \circ p = p$ abbiamo allora $p(x) = p(p(v)) = p(v)$, dunque $x = p(x)$. Ma $x \in Z$, dunque $p(x) = 0$, perciò $x = 0$. Resta provato che $W \cap Z = \{0\}$, e dalla formula della dimensione insieme alla formula di Grassmann segue che $V = W \oplus Z$.

Se $v = w + z$ è la decomposizione di un generico $v \in V$, dalle definizioni di W e Z e dalla relazione $p \circ p = p$ segue che $p(w) = w$ e $p(z) = 0$, dunque $p(v) = w$ e la dimostrazione è completa. \square

Descriviamo ora un esempio importante di decomposizione in somma diretta. Consideriamo lo spazio $\mathcal{M}_{n \times n}(\mathbb{R})$ delle matrici reali quadrate con n righe e n colonne. Diciamo che $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è *simmetrica* se i coefficienti di A con posti simmetrici rispetto alla diagonale principale sono uguali, ovvero se $(A)_{ij} = (A)_{ji}$ per ogni i e j , ovvero ancora, più sinteticamente, se $A = {}^tA$. Diciamo invece che A è *antisimmetrica* se $A = -{}^tA$. Indichiamo ora con \mathcal{S}_n l'insieme delle matrici $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ simmetriche, e con \mathcal{A}_n quello delle antisimmetriche. La linearità della funzione trasposizione comporta facilmente che \mathcal{S}_n e \mathcal{A}_n sono sottospazi di $\mathcal{M}_{n \times n}(\mathbb{R})$. Inoltre se $A \in \mathcal{S}_n \cap \mathcal{A}_n$ si ha $A = {}^tA$ e $A = -{}^tA$, dunque $A = -A$, da cui segue che $A = 0$. Abbiamo provato che $\mathcal{S}_n \cap \mathcal{A}_n = \{0\}$. Asseriamo ora che in realtà $\mathcal{S}_n + \mathcal{A}_n = \mathcal{M}_{n \times n}(\mathbb{R})$, dunque che si ha una somma diretta $\mathcal{M}_{n \times n}(\mathbb{R}) = \mathcal{S}_n \oplus \mathcal{A}_n$.

Proviamo che $\mathcal{S}_n + \mathcal{A}_n = \mathcal{M}_{n \times n}(\mathbb{R})$ con due metodi. Il primo si basa sulla formula di Grassmann, grazie alla quale, sapendo che $\mathcal{S}_n \cap \mathcal{A}_n = \{0\}$, basta vedere che $\dim_{\mathbb{R}}(\mathcal{S}_n) + \dim_{\mathbb{R}}(\mathcal{A}_n) = \dim_{\mathbb{R}}(\mathcal{M}_{n \times n}(\mathbb{R}))$. Sappiamo che quest'ultima dimensione vale n^2 , ed è facile verificare (usando la Proposizione 2.3.1) che

$$\dim_{\mathbb{R}}(\mathcal{S}_n) = \frac{n(n+1)}{2}, \quad \dim_{\mathbb{R}}(\mathcal{A}_n) = \frac{n(n-1)}{2},$$

da cui la conclusione.

Il secondo metodo che usiamo per provare che $\mathcal{M}_{n \times n}(\mathbb{R}) = \mathcal{S}_n \oplus \mathcal{A}_n$ consiste nell'esibire, per ogni matrice $A \in \mathcal{M}_{n \times n}(\mathbb{R})$, una sua espressione come $A = X + Y$ con X simmetrica e Y antisimmetrica. Dalla linearità della trasposizione e dal fatto che una matrice trasposta due volte torna uguale a se stessa segue subito che una tale espressione si ha con

$$X = \frac{1}{2}(A + {}^tA), \quad Y = \frac{1}{2}(A - {}^tA).$$

Questo argomento ci consente anche di esibire le proiezioni p, q relative alla decomposizione $\mathcal{M}_{n \times n}(\mathbb{R}) = \mathcal{S}_n \oplus \mathcal{A}_n$, che sono le seguenti:

$$p(A) = \frac{1}{2}(A + {}^tA), \quad q(A) = \frac{1}{2}(A - {}^tA).$$

Il lettore è invitato per esercizio a verificare direttamente da queste espressioni che $p \circ p = p$, $q \circ q = q$, $p \circ q = q \circ p = 0$.

5.3 Matrici e applicazioni associate

Fissiamo in questo paragrafo interi n ed m . Se $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e $x \in \mathbb{R}^n$, definiamo ora il prodotto $A \cdot x$ come

$$(A \cdot x)_i = \sum_{j=1}^n (A)_{ij} \cdot x_j, \quad i = 1, \dots, m.$$

Consideriamo dunque l'applicazione $\mathbb{R}^n \rightarrow \mathbb{R}^m$ che a un vettore $x \in \mathbb{R}^n$ associa il vettore $A \cdot x \in \mathbb{R}^m$. Affermiamo che questa applicazione è lineare. Si tratta di verificare le seguenti identità:

- $A \cdot (x + y) = A \cdot x + A \cdot y$ per ogni $x, y \in \mathbb{R}^n$;
- $A \cdot (\lambda \cdot x) = \lambda \cdot (A \cdot x)$ per ogni $\lambda \in \mathbb{R}$ e $x \in \mathbb{R}^n$.

La verifica di queste proprietà, una volta che le si sia scritte in coordinate, è pressoché immediata. La faremo tra poco, in un contesto più generale. L'applicazione $x \mapsto A \cdot x$ appena descritta è detta *applicazione associata alla matrice* A e viene indicata con f_A oppure, quando non si rischia di fare confusione, con la stessa lettera A , per ragioni che saranno spiegate ben presto.

Supponiamo ora dato un terzo intero k e matrici $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e $B \in \mathcal{M}_{n \times k}(\mathbb{R})$. Definiamo allora una nuova matrice $A \cdot B \in \mathcal{M}_{m \times k}(\mathbb{R})$ come segue:

$$(A \cdot B)_{ij} = \sum_{p=1}^n a_{ip} \cdot b_{pj}, \quad i = 1, \dots, m, \quad j = 1, \dots, k.$$

L'operazione \cdot ora introdotta è detta *prodotto righe per colonne*, poiché il coefficiente di posto (i, j) della matrice prodotto si ottiene “moltiplicando” tra loro la riga i -esima della matrice di sinistra e la colonna j -esima della matrice di destra. Il “prodotto” di una riga (a sinistra) per una colonna (a destra) si ottiene moltiplicando tra loro i coefficienti di posto corrispondente, e sommando

i risultati. Questa operazione ha naturalmente senso solo quando la “lunghezza” della riga coincide con la “altezza” della colonna. In particolare, come si desume comunque dalla definizione che abbiamo dato, ha senso eseguire il prodotto $A \cdot B$ solo quando il numero delle colonne di A è uguale a quello delle righe di B . Si osservi dunque che, anche quando il prodotto $A \cdot B$ ha senso, può non averlo $B \cdot A$. I prodotti hanno entrambi senso solo quando sia A che B sono matrici quadrate $n \times n$, ed è importante notare che anche in tal caso *non è in generale vero che $A \cdot B = B \cdot A$* (a meno che non si abbia $n = 1$). Siamo dunque in presenza di un’operazione binaria di prodotto su $\mathcal{M}_{n \times n}(\mathbb{R})$ che non gode della proprietà commutativa. Mostriamo ora che questa operazione gode invece di molte delle altre ragionevoli proprietà di un prodotto, cominciando dalle proprietà distributive:

Proposizione 5.3.1.

- Dati $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, $B_1, B_2 \in \mathcal{M}_{n \times k}(\mathbb{R})$ e $\lambda_1, \lambda_2 \in \mathbb{R}$, si ha:

$$A \cdot (\lambda_1 \cdot B_1 + \lambda_2 \cdot B_2) = \lambda_1 \cdot (A \cdot B_1) + \lambda_2 \cdot (A \cdot B_2) ;$$

- Dati $A_1, A_2 \in \mathcal{M}_{m \times n}(\mathbb{R})$, $B \in \mathcal{M}_{n \times k}(\mathbb{R})$ e $\lambda_1, \lambda_2 \in \mathbb{R}$, si ha:

$$(\lambda_1 \cdot A_1 + \lambda_2 \cdot A_2) \cdot B = \lambda_1 \cdot (A_1 \cdot B) + \lambda_2 \cdot (A_2 \cdot B).$$

Dimostrazione. Proviamo la prima proprietà, l’altra è analoga. L’uguaglianza da provare è tra due matrici $m \times k$, dunque per verificarla basta vedere che esse hanno lo stesso coefficiente nel posto (i, j) per ogni $i = 1, \dots, m$ e $j = 1, \dots, k$. Usando le definizioni delle varie operazioni coinvolte si trova infatti:

$$\begin{aligned} & \left(A \cdot (\lambda_1 \cdot B_1 + \lambda_2 \cdot B_2) \right)_{ij} \\ &= \sum_{p=1}^n (A)_{ip} \cdot \left(\lambda_1 \cdot B_1 + \lambda_2 \cdot B_2 \right)_{pj} \\ &= \sum_{p=1}^n (A)_{ip} \cdot (\lambda_1 \cdot (B_1)_{pj} + \lambda_2 \cdot (B_2)_{pj}) \\ &= \lambda_1 \cdot \sum_{p=1}^n (A)_{ip} \cdot (B_1)_{pj} + \lambda_2 \cdot \sum_{p=1}^n (A)_{ip} \cdot (B_2)_{pj} \\ &= \lambda_1 \cdot (A \cdot B_1)_{ij} + \lambda_2 \cdot (A \cdot B_2)_{ij} = \left(\lambda_1 \cdot (A \cdot B_1) + \lambda_2 \cdot (A \cdot B_2) \right)_{ij} . \end{aligned}$$

□

Il risultato appena provato si interpreta, nel nostro contesto degli spazi vettoriali, come *linearità del prodotto righe per colonne in entrambi gli argomenti*. Ciò significa che fissata $B \in \mathcal{M}_{n \times k}(\mathbb{R})$, l'applicazione

$$\begin{aligned} \mathcal{M}_{m \times n}(\mathbb{R}) &\rightarrow \mathcal{M}_{m \times k}(\mathbb{R}) \\ A &\mapsto A \cdot B \end{aligned}$$

è lineare. Analogamente, fissata una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, è lineare la

$$\begin{aligned} \mathcal{M}_{n \times k}(\mathbb{R}) &\rightarrow \mathcal{M}_{m \times k}(\mathbb{R}) \\ B &\mapsto A \cdot B \end{aligned}$$

Questo fatto si applica in particolare quando $k = 1$ e comporta la linearità dell'applicazione $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ associata a una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$.

Veniamo ora agli elementi neutri e all'associatività del prodotto righe per colonne. Per ogni n definiamo la *matrice identità* di ordine n come la matrice quadrata $I_n \in \mathcal{M}_{n \times n}(\mathbb{R})$ che ha coefficienti 1 sulla diagonale principale e 0 altrove. Specifichiamo meglio tale matrice introducendo la funzione δ di Kronecker, che ha come argomenti due indici i e j e restituisce il valore:

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

Abbiamo allora che $(I_n)_{ij}$ è dato esattamente da δ_{ij} .

Proposizione 5.3.2. • *Data $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ si ha*

$$A \cdot I_n = I_m \cdot A = A ;$$

• *Date $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, $B \in \mathcal{M}_{n \times k}(\mathbb{R})$ e $C \in \mathcal{M}_{k \times h}(\mathbb{R})$ si ha*

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

Dimostrazione. La prima affermazione è un semplice esercizio. La seconda asserisce l'uguaglianza tra due matrici $m \times h$. Basta allora verificare che in un generico posto (i, j) le matrici hanno il medesimo coefficiente:

$$\begin{aligned} \left(A \cdot (B \cdot C) \right)_{ij} &= \sum_{p=1}^n (A)_{ip} \cdot (B \cdot C)_{pj} = \sum_{p=1}^n (A)_{ip} \cdot \left(\sum_{q=1}^k (B)_{pq} \cdot (C)_{qj} \right) \\ &= \sum_{p=1}^n \sum_{q=1}^k (A)_{ip} \cdot (B)_{pq} \cdot (C)_{qj} = \sum_{q=1}^k \left(\sum_{p=1}^n (A)_{ip} \cdot (B)_{pq} \right) \cdot (C)_{qj} \\ &= \sum_{q=1}^k (A \cdot B)_{iq} \cdot (C)_{qj} = \left((A \cdot B) \cdot C \right)_{ij}. \end{aligned}$$

□

Si osservi che, tra le proprietà del prodotto righe per colonne di matrici analoghe a quelle del prodotto tra numeri, non discutiamo per ora l'esistenza dell'inverso. Lo faremo in parte nel Paragrafo 5.5 e più diffusamente nel capitolo successivo.

5.4 Spazio delle applicazioni lineari e matrice di un'applicazione

Torniamo in questo paragrafo alla situazione generale di due spazi vettoriali qualsiasi V e W . Abbiamo già osservato nel Capitolo 3 che l'insieme $\mathcal{F}(V, W)$ delle applicazioni da V in W può essere dotato di una naturale struttura di spazio vettoriale, utilizzando "punto per punto" le operazioni di W . Ricordiamo infatti che questa struttura non richiede a V di essere uno spazio vettoriale: un semplice insieme sarebbe già sufficiente. Quando però V è uno spazio vettoriale, tra le applicazioni da V in W possiamo selezionare quelle lineari. Indichiamo dunque con $\mathcal{L}_{\mathbb{R}}(V, W)$ l'insieme che consiste di tutte tali applicazioni:

$$\mathcal{L}_{\mathbb{R}}(V, W) = \{f : V \rightarrow W : f \text{ lineare}\}.$$

Proposizione 5.4.1. $\mathcal{L}_{\mathbb{R}}(V, W)$ è un sottospazio vettoriale di $\mathcal{F}(V, W)$.

Dimostrazione. Si tratta di vedere che una combinazione lineare di applicazioni lineari è lineare. Siano dunque $\lambda_1, \lambda_2 \in \mathbb{R}$ e $f_1, f_2 : V \rightarrow W$ lineari. Per vedere che $\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2$ è lineare dobbiamo valutarla su una combinazione lineare di vettori. Siano dunque $\mu_1, \mu_2 \in \mathbb{R}$ e $v_1, v_2 \in V$. Usando il fatto che le operazioni tra funzioni sono punto per punto quelle di W e la linearità di f_1 e f_2 abbiamo ora:

$$\begin{aligned} & (\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2)(\mu_1 \cdot v_1 + \mu_2 \cdot v_2) \\ &= \lambda_1 \cdot f_1(\mu_1 \cdot v_1 + \mu_2 \cdot v_2) + \lambda_2 \cdot f_2(\mu_1 \cdot v_1 + \mu_2 \cdot v_2) \\ &= \lambda_1 \cdot \mu_1 \cdot f_1(v_1) + \lambda_1 \cdot \mu_2 \cdot f_1(v_2) + \lambda_2 \cdot \mu_1 \cdot f_2(v_1) + \lambda_2 \cdot \mu_2 \cdot f_2(v_2) \\ &= \mu_1 \cdot (\lambda_1 \cdot f_1(v_1) + \lambda_2 \cdot f_2(v_1)) + \mu_2 \cdot (\lambda_1 \cdot f_1(v_2) + \lambda_2 \cdot f_2(v_2)) \\ &= \mu_1 \cdot (\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2)(v_1) + \mu_2 \cdot (\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2)(v_2). \end{aligned}$$

□

Uno degli scopi di questo paragrafo è quello di dimostrare il seguente:

Teorema 5.4.2. *Se $\dim_{\mathbb{R}}(V) = n$ e $\dim_{\mathbb{R}}(W) = m$ si ha*

$$\dim_{\mathbb{R}}(\mathcal{L}_{\mathbb{R}}(V, W)) = n \cdot m.$$

Proveremo questo fatto in ben tre modi distinti (se pur strettamente collegati tra loro), che servono a chiarire meglio la struttura di $\mathcal{L}_{\mathbb{R}}(V, W)$. Nel caso in cui V e W sono spazi di vettori numerici abbiamo intanto una caratterizzazione molto esplicita di $\mathcal{L}_{\mathbb{R}}(V, W)$, che enunciamo ricordando che ad una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ avevamo associato un'applicazione lineare $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$:

Proposizione 5.4.3. *L'applicazione*

$$\begin{aligned} \mathcal{M}_{m \times n}(\mathbb{R}) &\rightarrow \mathcal{L}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m) \\ A &\mapsto f_A \end{aligned}$$

è lineare e bigettiva.

Dimostrazione. Per la linearità, dobbiamo verificare che

$$f_{(\lambda_1 \cdot A_1 + \lambda_2 \cdot A_2)} = \lambda_1 \cdot f_{A_1} + \lambda_2 \cdot f_{A_2}.$$

Due applicazioni $\mathbb{R}^n \rightarrow \mathbb{R}^m$ coincidono se hanno lo stesso valore su ogni $x \in \mathbb{R}^n$, dunque dobbiamo vedere che

$$(\lambda_1 \cdot A_1 + \lambda_2 \cdot A_2) \cdot x = \lambda_1 \cdot (A_1 \cdot x) + \lambda_2 \cdot (A_2 \cdot x)$$

il che è stato visto nella Proposizione 5.3.1.

Proviamo l'iniettività dimostrando che il nucleo è il sottospazio banale $\{0\}$. Supponiamo dunque che f_A sia l'applicazione nulla, e proviamo che A è la matrice nulla. Per ipotesi abbiamo $A \cdot x = 0$ per ogni $x \in \mathbb{R}^n$. In particolare $A \cdot e_i = 0$ se e_i è il vettore i -esimo della base canonica di \mathbb{R}^n . Ma è facile vedere che $A \cdot e_i$ non è altro che la colonna i -esima di A , dunque deduciamo subito che $A = 0$.

Per la surgettività fissiamo $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineare e cerchiamo una matrice A tale che f_A sia f . Per l'argomento appena esposto, la A non può essere altro che la matrice la cui colonna i -esima è il vettore $f(e_i)$. Proviamo che tale A effettivamente va bene. Per come abbiamo definito A , indicando $(A)_{ji}$ con a_{ji} , si ha

$$f(e_i) = \sum_{j=1}^m a_{ji} \cdot e_j.$$

Se $x \in \mathbb{R}^n$ è qualsiasi, notando che $x = \sum_{i=1}^n x_i \cdot e_i$, abbiamo allora:

$$\begin{aligned} f(x) &= f\left(\sum_{i=1}^n x_i \cdot e_i\right) = \sum_{i=1}^n x_i \cdot f(e_i) = \sum_{i=1}^n x_i \cdot \sum_{j=1}^m a_{ji} \cdot e_j \\ &= \sum_{j=1}^m \left(\sum_{i=1}^n a_{ji} \cdot x_i\right) \cdot e_j = \sum_{j=1}^m (A \cdot x)_j \cdot e_j = A \cdot x . \end{aligned}$$

□

Corollario 5.4.4. $\dim_{\mathbb{R}}(\mathcal{L}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)) = n \cdot m$.

Dimostrazione. $\dim_{\mathbb{R}}(\mathcal{M}_{m \times n}(\mathbb{R})) = n \cdot m$ e le applicazioni lineari bigettive conservano la dimensione. □

Prima dimostrazione del Teorema 5.4.2. Fissiamo basi \mathcal{B} di V e \mathcal{C} di W , dunque abbiamo applicazioni lineari bigettive $\phi_{\mathcal{B}} : V \rightarrow \mathbb{R}^n$ e $\phi_{\mathcal{C}} : W \rightarrow \mathbb{R}^m$ date da $v \mapsto [v]_{\mathcal{B}}$ e $w \mapsto [w]_{\mathcal{C}}$ rispettivamente. Definiamo allora l'applicazione

$$\begin{aligned} \mathcal{L}_{\mathbb{R}}(V, W) &\rightarrow \mathcal{L}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m) \\ f &\mapsto \phi_{\mathcal{C}} \circ f \circ \phi_{\mathcal{B}}^{-1} . \end{aligned}$$

È facile vedere che tale mappa è lineare e bigettiva, onde la conclusione per il corollario precedente. □

Il prossimo semplice risultato giustifica, insieme alla Proposizione 5.4.3, la scelta, già fatta nel paragrafo precedente, di indicare con A stessa l'applicazione lineare $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ associata ad una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. Infatti al livello delle matrici abbiamo un'operazione di "moltiplicazione" (il prodotto righe per colonne), mentre al livello delle applicazioni lineari abbiamo la "moltiplicazione" data dalla composizione. Affinché l'identificazione di f_A con A sia ragionevole, occorre dunque che la "moltiplicazione" di due matrici A e B corrisponda precisamente alla "moltiplicazione" delle corrispondenti f_A ed f_B . (Questo tema era già stato anticipato nel Paragrafo 1.2). Il prossimo risultato asserisce esattamente che ciò in effetti avviene:

Proposizione 5.4.5. *Date $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e $B \in \mathcal{M}_{n \times k}(\mathbb{R})$ si ha*

$$f_{A \cdot B} = f_A \circ f_B .$$

Dimostrazione. L'uguaglianza da verificare è quella tra due applicazioni da \mathbb{R}^k in \mathbb{R}^m . Proviamo dunque esse hanno lo stesso valore su ogni $x \in \mathbb{R}^k$:

$$f_{A \cdot B}(x) = (A \cdot B) \cdot x = A \cdot (B \cdot x) = f_A(f_B(x)) = (f_A \circ f_B)(x).$$

□

Affrontiamo ora la questione di come si possa definire un'applicazione lineare tra due spazi vettoriali qualsiasi, ottenendo una seconda e più esplicita dimostrazione del Teorema 5.4.2.

Proposizione 5.4.6. *Se (v_1, \dots, v_n) è una base di V e sono assegnati vettori $w_1, \dots, w_n \in W$ qualsiasi, esiste una e una sola applicazione lineare $f : V \rightarrow W$ tale che $f(v_i) = w_i$ per $i = 1, \dots, n$.*

Dimostrazione. Se f esiste il suo valore su un vettore del tipo $\alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$ deve risultare $\alpha_1 \cdot w_1 + \dots + \alpha_n \cdot w_n$. Poiché ogni vettore di V si scrive come $\alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$, l'unicità di f è provata. Per dimostrare che f esiste, usiamo il fatto che ogni v in V si scrive *in modo unico* come $v = \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$, e poniamo $f(v) = \alpha_1 \cdot w_1 + \dots + \alpha_n \cdot w_n$. Si tratta ora di provare che tale f è lineare.

Siano $\lambda, \lambda' \in \mathbb{R}$ e $v, v' \in V$. Vogliamo provare che $f(\lambda \cdot v + \lambda' \cdot v') = \lambda \cdot f(v) + \lambda' \cdot f(v')$. Supponiamo dunque che $v = \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$ e $v' = \alpha'_1 \cdot v_1 + \dots + \alpha'_n \cdot v_n$. Allora

$$\lambda \cdot v + \lambda' \cdot v' = (\lambda \cdot \alpha_1 + \lambda' \cdot \alpha'_1) \cdot v_1 + \dots + (\lambda \cdot \alpha_n + \lambda' \cdot \alpha'_n) \cdot v_n$$

dunque per definizione

$$\begin{aligned} f(v) &= \alpha_1 \cdot w_1 + \dots + \alpha_n \cdot w_n \\ f(v') &= \alpha'_1 \cdot w_1 + \dots + \alpha'_n \cdot w_n \\ f(\lambda \cdot v + \lambda' \cdot v') &= (\lambda \cdot \alpha_1 + \lambda' \cdot \alpha'_1) \cdot w_1 + \dots + (\lambda \cdot \alpha_n + \lambda' \cdot \alpha'_n) \cdot w_n \end{aligned}$$

e la tesi segue subito. □

Seconda dimostrazione del Teorema 5.4.2. Siano (v_1, \dots, v_n) e (w_1, \dots, w_m) basi di V e W rispettivamente. Definiamo applicazioni lineari $f_{ij} : V \rightarrow W$ per $i = 1, \dots, m$ e $j = 1, \dots, n$ con la scelta:

$$f_{ij}(v_p) = \delta_{jp} \cdot w_i, \quad p = 1, \dots, n.$$

La proposizione precedente garantisce appunto che tali condizioni determinano f_{ij} univocamente. Affermiamo ora che l'insieme $\{f_{ij} : i = 1, \dots, m, j = 1, \dots, n\}$ opportunamente ordinato (ad esempio in modo lessicografico) costituisce una base di $\mathcal{L}_{\mathbb{R}}(V, W)$. Ciò naturalmente implica la conclusione.

Dobbiamo dunque provare che ogni applicazione $f : V \rightarrow W$ lineare si esprime in modo unico come combinazione lineare delle f_{ij} . Non è difficile vedere che, data f , si ha $f = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \cdot f_{ij}$, dove i coefficienti a_{ij} sono quelli univocamente determinati dalle condizioni:

$$f(v_p) = \sum_{i=1}^m a_{ip} \cdot w_i.$$

□

Secondo quanto visto nella dimostrazione appena conclusa, se $f : V \rightarrow W$ è lineare e sono date basi $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{C} = (w_1, \dots, w_m)$ di V e W rispettivamente, resta univocamente determinata una matrice $m \times n$. Questa idea avrà nel seguito una certa importanza, quindi ripetiamo la costruzione e introduciamo una notazione per la matrice.

Essendo \mathcal{C} una base di W , esistono e sono uniche costanti $a_{ij} \in \mathbb{R}$ per $i = 1, \dots, m$ e $j = 1, \dots, n$ tali che:

$$\begin{aligned} f(v_1) &= a_{11} \cdot w_1 + \dots + a_{m1} \cdot w_m \\ &\vdots \\ f(v_n) &= a_{1n} \cdot w_1 + \dots + a_{mn} \cdot w_m . \end{aligned}$$

Consideriamo dunque la matrice $A = (a_{ij}) \in \mathcal{M}_{m \times n}(\mathbb{R})$ che, essendo univocamente determinata da f e dalle basi \mathcal{B} e \mathcal{C} , possiamo denotare con $[f]_{\mathcal{B}}^{\mathcal{C}}$ e chiamare la *matrice associata all'applicazione f rispetto alle basi \mathcal{B} (in partenza) e \mathcal{C} (in arrivo)*. Possiamo anche descrivere $[f]_{\mathcal{B}}^{\mathcal{C}}$ efficacemente se pur solo a parole: *nella colonna i -esima di $[f]_{\mathcal{B}}^{\mathcal{C}}$ si trovano le coordinate rispetto a \mathcal{C} dell'immagine tramite f dell' i -esimo elemento di \mathcal{B} .*

Osservazione 5.4.7. Sia $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, siano $\mathcal{E}_n, \mathcal{E}_m$ le basi canoniche di \mathbb{R}^n e \mathbb{R}^m , e sia $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ l'applicazione lineare associata ad A . Allora $[f_A]_{\mathcal{E}_n}^{\mathcal{E}_m}$ è data da A stessa. Questo è un ulteriore motivo per indicare f_A con A stessa, quando non si rischi di fare confusione.

L'utilità e naturalità della matrice associata ad un'applicazione rispetto a basi assegnate risiede nel seguente teorema, di cui diamo prima due enunciati informali e poi uno formale (seguito dalla facile dimostrazione):

- l'azione di f sui vettori corrisponde esattamente all'azione della moltiplicazione a sinistra per $[f]_{\mathcal{B}}^{\mathcal{C}}$ sulle coordinate dei vettori (rispetto alle basi \mathcal{B} in partenza e \mathcal{C} in arrivo);
- identificando V a \mathbb{R}^n tramite la base \mathcal{B} e W a \mathbb{R}^m tramite \mathcal{C} , la f viene identificata all'applicazione associata alla matrice $[f]_{\mathcal{B}}^{\mathcal{C}}$.

Teorema 5.4.8. $[f(v)]_{\mathcal{C}} = [f]_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}}$ per ogni $v \in V$.

Dimostrazione. Sia $[v]_{\mathcal{B}} = x$, dunque $v = x_1 \cdot v_1 + \dots + x_n \cdot v_n$. Allora:

$$\begin{aligned} f(v) &= \sum_{j=1}^n x_j \cdot f(v_j) = \sum_{j=1}^n x_j \cdot \left(\sum_{i=1}^m a_{ij} \cdot w_i \right) \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \cdot x_j \right) \cdot w_i = \sum_{i=1}^m ([f]_{\mathcal{B}}^{\mathcal{C}} \cdot x)_i \cdot w_i \\ \Rightarrow [f(v)]_{\mathcal{C}} &= [f]_{\mathcal{B}}^{\mathcal{C}} \cdot x . \end{aligned}$$

□

Una terza dimostrazione del Teorema 5.4.2 segue ora immediatamente dalla seguente:

Proposizione 5.4.9. Se \mathcal{B} e \mathcal{C} sono basi di V e W rispettivamente, l'applicazione

$$\begin{aligned} \mathcal{L}_{\mathbb{R}}(V, W) &\rightarrow \mathcal{M}_{m \times n}(\mathbb{R}) \\ f &\mapsto [f]_{\mathcal{B}}^{\mathcal{C}} \end{aligned}$$

è lineare e bigettiva.

Dimostrazione. La linearità della mappa segue facilmente dalla definizione di matrice associata. Per verificare la bigettività basta esibire l'inversa. Grazie al teorema precedente, l'inversa è la mappa che associa ad una matrice A la mappa $\phi_{\mathcal{C}}^{-1} \circ f_A \circ \phi_{\mathcal{B}}$, dove $\phi_{\mathcal{B}} : V \rightarrow \mathbb{R}^n$ e $\phi_{\mathcal{C}} : W \rightarrow \mathbb{R}^m$ sono le bigezioni date da $v \mapsto [v]_{\mathcal{B}}$ e $w \mapsto [w]_{\mathcal{C}}$. □

5.5 Cambiamenti di base e inversione

Discuteremo in questo paragrafo come cambino le coordinate di un vettore e la matrice associata a un'applicazione lineare quando cambiano le basi rispetto a cui sono calcolate. Per fare ciò introdurremo una simbologia che consente di definire coordinate e matrice associata in modo sintetico e incisivo. Daremo inoltre la definizione di inversa di una matrice e di un'applicazione lineare, lasciando al capitolo successivo la discussione su come si possano determinare in pratica l'invertibilità e l'eventuale inversa di una matrice o applicazione assegnate.

Cominciamo ricordando che se è data una base $\mathcal{B} = (v_1, \dots, v_n)$ di uno spazio vettoriale V , il vettore $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ delle coordinate di v rispetto a

\mathcal{B} , che denotiamo con $[v]_{\mathcal{B}}$, è definito dall'identità $v = x_1 \cdot v_1 + \dots + x_n \cdot v_n$. Rammentiamo anche che abbiamo definito il prodotto di un vettore numerico riga con un vettore numerico colonna, che ha come risultato un numero. Se applichiamo la stessa definizione alla riga $\mathcal{B} = (v_1, \dots, v_n)$, i cui elementi sono

ora vettori e non più numeri, e alla colonna $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, troviamo esatta-

mente $x_1 \cdot v_1 + \dots + x_n \cdot v_n$ (la convenzione per il prodotto tra uno scalare e un vettore è che lo scalare vada sempre scritto prima). Possiamo dunque esprimere in maniera alternativa la definizione delle coordinate: *il vettore $[v]_{\mathcal{B}}$ delle coordinate di v rispetto a \mathcal{B} è quel vettore numerico x tale che $v = \mathcal{B} \cdot x$* . Oppure, in forma più concisa:

$$\boxed{v = \mathcal{B} \cdot [v]_{\mathcal{B}}.}$$

In modo analogo possiamo ora ridefinire $[f]_{\mathcal{B}}^{\mathcal{C}}$ se $f : V \rightarrow W$ è lineare e $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{C} = (w_1, \dots, w_m)$ sono basi di V , W rispettivamente. Ricordiamo che si ha $[f]_{\mathcal{B}}^{\mathcal{C}} = (a_{ij})$ se $f(v_j) = \sum a_{ij} \cdot w_i$, ovvero:

$$\begin{aligned} (\diamond) \quad & (f(v_1), \dots, f(v_n)) \\ & = (a_{11} \cdot w_1 + \dots + a_{m1} \cdot w_m, \dots, a_{1n} \cdot w_1 + \dots + a_{mn} \cdot w_m) . \end{aligned}$$

Per ipotesi la f è lineare, cioè si distribuisce sulle combinazioni lineari dei vettori, e allora è ragionevole interpretare l'applicazione di f a un vettore v come un prodotto $f \cdot v$ (così come accade quando V e W sono spazi di vettori

numerici ed f è associata a una matrice). Possiamo allora scrivere il primo membro della (\diamond) come $(f \cdot v_1, \dots, f \cdot v_n)$, ovvero $f \cdot (v_1, \dots, v_n)$, cioè $f \cdot \mathcal{B}$ (si noti che abbiamo qui ancora esteso il significato dell'operazione \cdot , sempre avendo in mente il caso in cui f è associata a una matrice). Passando al secondo membro della (\diamond) , possiamo ora interpretarlo come il prodotto tra la riga $\mathcal{C} = (w_1, \dots, w_m)$ e la matrice $[f]_{\mathcal{B}}^{\mathcal{C}} = (a_{ij})$ che ha m righe e n colonne. Abbiamo dunque la seguente definizione sintetica di $[f]_{\mathcal{B}}^{\mathcal{C}}$:

$$f \cdot \mathcal{B} = \mathcal{C} \cdot [f]_{\mathcal{B}}^{\mathcal{C}}.$$

Illustriamo l'utilità della definizione sintetica di matrice associata a un'applicazione dimostrando in modo molto semplice un risultato che richiederebbe invece qualche sforzo in più se si usasse la definizione originale.

Proposizione 5.5.1. *Se $f : V \rightarrow W$ e $g : W \rightarrow Z$ sono applicazioni lineari e sono date basi $\mathcal{B}, \mathcal{C}, \mathcal{D}$ di V, W, Z rispettivamente, si ha:*

$$[g \circ f]_{\mathcal{B}}^{\mathcal{D}} = [g]_{\mathcal{C}}^{\mathcal{D}} \cdot [f]_{\mathcal{B}}^{\mathcal{C}}.$$

Dimostrazione. $(g \circ f) \cdot \mathcal{B} = g \cdot (f \cdot \mathcal{B}) = g \cdot \mathcal{C} \cdot [f]_{\mathcal{B}}^{\mathcal{C}} = \mathcal{D} \cdot [g]_{\mathcal{C}}^{\mathcal{D}} \cdot [f]_{\mathcal{B}}^{\mathcal{C}}$. □

Prima di affrontare la questione di come cambino le coordinate e le matrici cambiando basi, dobbiamo ricordare una definizione e analizzarla nel contesto degli spazi vettoriali. Una applicazione $f : A \rightarrow B$ tra insiemi si dice *invertibile* se esiste una $g : B \rightarrow A$ tale che $f \circ g$ sia l'identità di B e $g \circ f$ sia l'identità di A . Tale g se esiste è unica, è denotata con f^{-1} ed è detta *inversa* di f . Notiamo che f è invertibile se e soltanto se è bigettiva. Per le applicazioni lineari, abbiamo la seguente:

Osservazione 5.5.2.

- Se $f : V \rightarrow W$ è lineare e invertibile, la f^{-1} è lineare;
- Se $f : V \rightarrow W$ è lineare e invertibile e V ha dimensione finita, allora W ha dimensione finita e uguale a quella di V ;
- Se V e W hanno dimensioni finite e uguali tra loro, allora esiste $f : V \rightarrow W$ invertibile;
- Se V e W hanno dimensioni finite e uguali tra loro ed è data $f : V \rightarrow W$ lineare, allora sono fatti equivalenti:
 - f è invertibile;

- f è iniettiva;
- f è surgettiva;
- Se V e W hanno dimensioni finite e uguali tra loro e sono date $f : V \rightarrow W$ e $g : W \rightarrow V$ lineari, allora sono fatti equivalenti:
 - f è invertibile e g è la sua inversa;
 - $g \circ f$ è l'identità di V ;
 - $f \circ g$ è l'identità di W .

Passando agli spazi di vettori numerici, notiamo intanto che l'identità di \mathbb{R}^n è l'applicazione lineare associata alla matrice I_n già introdotta in precedenza. Diciamo ora che una matrice $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è invertibile se lo è l'applicazione lineare a essa associata, ovvero se esiste una matrice A^{-1} tale che $A \cdot A^{-1} = I_n$ oppure equivalentemente $A^{-1} \cdot A = I_n$. Nel prossimo capitolo vedremo un criterio semplice ed efficace per stabilire se una data matrice ammetta inversa, e per trovarla quando esiste.

Rivolgiamoci dunque alla questione di come siano legate tra loro le coordinate $[v]_{\mathcal{B}}$ e $[v]_{\mathcal{B}'}$ di uno stesso vettore v rispetto a due basi diverse $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{B}' = (v'_1, \dots, v'_n)$ di uno spazio vettoriale V . Poiché \mathcal{B} è base, possiamo scrivere gli elementi di \mathcal{B}' come combinazioni lineari degli elementi di \mathcal{B} , ovvero esistono costanti a_{ij} tali che:

$$v'_j = \sum_{i=1}^n a_{ij} \cdot v_i \quad j = 1, \dots, n.$$

Introducendo la matrice $A = (a_{ij})$ possiamo ora scrivere questa identità sinteticamente come $\mathcal{B}' = \mathcal{B} \cdot A$. Si noti che A non è altro che la matrice $[\text{Id}_V]_{\mathcal{B}'}^{\mathcal{B}}$ che rappresenta l'identità di V rispetto alle basi \mathcal{B}' in partenza e \mathcal{B} in arrivo.

Lemma 5.5.3. *Se $\mathcal{B}' = \mathcal{B} \cdot A$ allora esiste A^{-1} e si ha $\mathcal{B} = \mathcal{B}' \cdot A^{-1}$.*

Dimostrazione. Sia \tilde{A} la matrice per la quale si ha $\mathcal{B} = \mathcal{B}' \cdot \tilde{A}$. Allora:

$$A \cdot \tilde{A} = [\text{Id}_V]_{\mathcal{B}'}^{\mathcal{B}} \cdot [\text{Id}_V]_{\mathcal{B}}^{\mathcal{B}'} = [\text{Id}_V]_{\mathcal{B}}^{\mathcal{B}} = I_n.$$

□

Possiamo ora ricavare la regola di cambiamento delle coordinate:

$$\begin{aligned} v &= \mathcal{B} \cdot [v]_{\mathcal{B}} = \mathcal{B}' \cdot A^{-1} \cdot [v]_{\mathcal{B}} \\ \Rightarrow [v]_{\mathcal{B}'} &= A^{-1} \cdot [v]_{\mathcal{B}}. \end{aligned}$$

Riassumendo:

$$\boxed{\mathcal{B}' = \mathcal{B} \cdot A \quad \Rightarrow \quad [v]_{\mathcal{B}'} = A^{-1} \cdot [v]_{\mathcal{B}}.}$$

Altrettanto facile è ora discutere come cambia la matrice di un'applicazione. Supponiamo di avere $f : V \rightarrow W$ lineare e basi $\mathcal{B}, \mathcal{B}'$ di V e $\mathcal{C}, \mathcal{C}'$ di W . Allora esistono matrici invertibili A e M tali che $\mathcal{B}' = \mathcal{B} \cdot A$ e $\mathcal{C}' = \mathcal{C} \cdot M$. Usando la definizione sintetica di matrice associata abbiamo:

$$\begin{aligned} f \cdot \mathcal{B}' &= f \cdot \mathcal{B} \cdot A = \mathcal{C} \cdot [f]_{\mathcal{B}}^{\mathcal{C}} \cdot A = \mathcal{C}' \cdot M^{-1} \cdot [f]_{\mathcal{B}}^{\mathcal{C}} \cdot A \\ &\Rightarrow [f]_{\mathcal{B}'}^{\mathcal{C}'} = M^{-1} \cdot [f]_{\mathcal{B}}^{\mathcal{C}} \cdot A. \end{aligned}$$

Dunque:

$$\boxed{\mathcal{B}' = \mathcal{B} \cdot A, \mathcal{C}' = \mathcal{C} \cdot M \quad \Rightarrow \quad [f]_{\mathcal{B}'}^{\mathcal{C}'} = M^{-1} \cdot [f]_{\mathcal{B}}^{\mathcal{C}} \cdot A.}$$

Capitolo 6

Sistemi lineari, rango, determinante

Un *sistema lineare* di m equazioni nelle n incognite x_1, \dots, x_n è uno del tipo seguente

$$\begin{cases} a_{11} \cdot x_1 + \dots + a_{1n} \cdot x_n & = & b_1 \\ & \vdots & \\ a_{m1} \cdot x_1 + \dots + a_{mn} \cdot x_n & = & b_m, \end{cases}$$

dove le a_{ij} e b_i sono costanti reali assegnate. Introducendo la matrice $A =$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m \times n}(\mathbb{R}) \text{ e il vettore } b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m, \text{ ed interpre-$$

tando le incognite x_1, \dots, x_n come un vettore $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, possiamo riscrivere

il sistema nella forma

$$A \cdot x = b.$$

Ricordando che la matrice A rappresenta anche un'applicazione da \mathbb{R}^n a \mathbb{R}^m , la teoria già discussa comporta i seguenti semplici fatti:

- Il sistema $A \cdot x = b$ ha soluzione se e soltanto se $b \in \text{Im}(A)$;
- Se $m = n$ ed esiste la matrice A^{-1} inversa di A , allora la soluzione esiste ed è unica, ed è data da $x = A^{-1} \cdot b$.

Emergono quindi in modo naturale le questioni seguenti:

- Come determinare se un dato vettore appartiene all'immagine dell'applicazione associata a una matrice assegnata?
- Come determinare se una matrice assegnata è invertibile e, se lo è, come trovarne l'inversa?

Queste domande avranno nelle prossime pagine risposte complete e di natura 'algoritmica' (cioè facilmente traducibili in programmi di un calcolatore).

6.1 Rango e soluzioni di un sistema

Se $f : V \rightarrow W$ è un'applicazione lineare tra spazi vettoriali ed $\text{Im}(f)$ ha dimensione finita, chiamiamo tale dimensione *rango* di f e la indichiamo con $\text{rank}(f)$. Notiamo che se almeno uno tra V e W ha dimensione finita, automaticamente anche $\text{Im}(f)$ ha dimensione finita, dunque si può definire $\text{rank}(f)$.

Se A è una matrice, chiamiamo *rango* di A e indichiamo con $\text{rank}(A)$ il rango della applicazione lineare associata. Notiamo ora che i vettori che costituiscono le colonne di A sono certamente un sistema di generatori per $\text{Im}(A)$. Dal procedimento di estrazione di una base da un insieme di generatori deduciamo allora la seguente semplice caratterizzazione del rango di una matrice:

Proposizione 6.1.1. *Il rango di una matrice A è il numero massimo di colonne di A che costituiscono un sistema di vettori linearmente indipendenti.*

È ora semplice provare il seguente risultato, detto Teorema di Rouché-Capelli:

Teorema 6.1.2. *Dati $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$, il sistema $A \cdot x = b$ ha soluzioni $x \in \mathbb{R}^n$ se e soltanto se la matrice A e la matrice $(A \ b)$, ottenuta da A aggiungendo la colonna b , hanno lo stesso rango.*

Dimostrazione. Ognuno dei fatti seguenti è facilmente equivalente al successivo:

- Esiste una soluzione;
- b appartiene all'immagine di A ;
- b è combinazione lineare delle colonne di A ;
- Aggiungendo b alle colonne di A la dimensione del generato rimane la stessa;

- A e $(A \ b)$ hanno lo stesso rango.

□

Introduciamo ora un ulteriore elemento di terminologia. Un sistema lineare si dice *omogeneo* se è del tipo $A \cdot x = 0$, ovvero se il vettore b dei *termini noti* è nullo. Naturalmente un sistema omogeneo ammette sempre la soluzione $x = 0$, ovvero $x_1 = \dots = x_n = 0$. Più esattamente, usando la formula della dimensione, si verifica che *l'insieme delle soluzioni di un sistema lineare omogeneo $A \cdot x = 0$ in n incognite è lo spazio vettoriale $\text{Ker}(A)$ di dimensione $n - \text{rank}(A)$.*

Il prossimo risultato si può esprimere nel modo seguente: *se un sistema lineare ammette soluzioni, la sua generica soluzione è la somma di una qualsiasi fissata soluzione particolare e della generica soluzione del sistema lineare omogeneo avente la stessa matrice.*

Proposizione 6.1.3. *Dati $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$ e $x_0 \in \mathbb{R}^n$ tale che $A \cdot x_0 = b$, si ha*

$$\{x \in \mathbb{R}^n : A \cdot x = b\} = \{x_0 + u : A \cdot u = 0\}.$$

Dimostrazione. Dato $x \in \mathbb{R}^n$ si ha:

$$\begin{aligned} A \cdot x = b &\Leftrightarrow A \cdot x = A \cdot x_0 \\ &\Leftrightarrow A \cdot (x - x_0) = 0 \\ &\Leftrightarrow x - x_0 = u \quad \text{con } A \cdot u = 0 \\ &\Leftrightarrow x = x_0 + u \quad \text{con } A \cdot u = 0. \end{aligned}$$

□

Un sistema lineare di m equazioni in n incognite si dice:

- *quadrato* se $m = n$;
- *sovradeterminato* se $m > n$;
- *sottodeterminato* se $m < n$.

Dall'ultimo risultato e la discussione che lo precede sui sistemi omogenei deduciamo ora che:

- un sistema lineare non omogeneo quadrato oppure sovradeterminato può avere una sola soluzione, infinite soluzioni, oppure nessuna soluzione;

- un sistema lineare non omogeneo sottodeterminato può avere infinite soluzioni oppure nessuna soluzione.

Per l'effettiva risoluzione di un sistema lineare sono disponibili semplici algoritmi, che non descriviamo qui esplicitamente. Essi sono basati sul fatto che da un sistema se ne possono ricavare altri *equivalenti* (cioè aventi le stesse soluzioni) con operazioni del tipo seguente:

- riordino delle equazioni;
- sostituzione dell'equazione i -esima con una combinazione lineare delle equazioni nella quale la i -esima abbia coefficiente non nullo.

6.2 Determinante

Il teorema di Rouché-Capelli traduce il problema della risolubilità di un sistema lineare in un calcolo di rango di matrici. In questo paragrafo affrontiamo il problema del calcolo del rango in un caso particolare, mostrando come si possa stabilire se una matrice quadrata $n \times n$ abbia rango n . Nello prossimo paragrafo vedremo come trattare il caso generale.

Notiamo intanto che una matrice A quadrata $n \times n$ ha rango n se e soltanto se è invertibile, e cominciamo a esaminare il caso 2×2 . Sia dunque data $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. La A è invertibile se e soltanto se per ogni ξ, η il sistema $\begin{cases} a \cdot x + b \cdot y = \xi \\ c \cdot x + d \cdot y = \eta \end{cases}$ ammette soluzione unica x, y . Se $a \cdot d = 0$ e $b \cdot c = 0$ è immediato vedere che la soluzione o non esiste o non è unica. Altrimenti il sistema è equivalente al seguente:

$$\begin{cases} d \cdot (a \cdot x + b \cdot y) - b \cdot (c \cdot x + d \cdot y) = d \cdot \xi - b \cdot \eta \\ -c \cdot (a \cdot x + b \cdot y) + a \cdot (c \cdot x + d \cdot y) = -c \cdot \xi + a \cdot \eta. \end{cases}$$

(Abbiamo sostituito la prima equazione con d volte se stessa più $-b$ volte la seconda, e la seconda con a volte se stessa più $-c$ volte la prima; stiamo supponendo che $a \cdot d \neq 0$ oppure $b \cdot c \neq 0$, dunque il sistema è equivalente al precedente).

Abbiamo allora il sistema:

$$\begin{cases} (a \cdot d - b \cdot c) \cdot x = d \cdot \xi - b \cdot \eta \\ (a \cdot d - b \cdot c) \cdot y = -c \cdot \xi + a \cdot \eta. \end{cases}$$

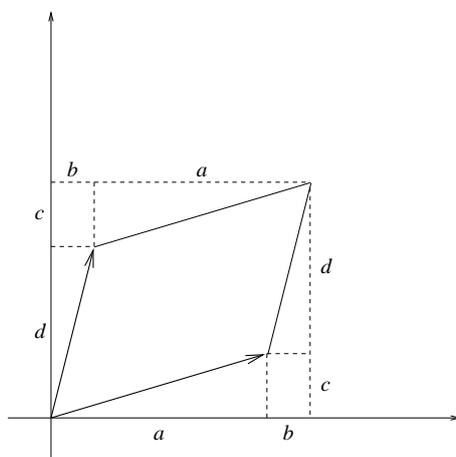


Figura 6.1: Area di un parallelogramma

Scopriamo dunque che $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è invertibile se e soltanto se $a \cdot d - b \cdot c$ è non nullo (infatti tale condizione implica che $a \cdot d \neq 0$ oppure $b \cdot c \neq 0$). Anzi, dalla forma della soluzione del sistema deduciamo che se $a \cdot d - b \cdot c$ è non nullo si ha

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{a \cdot d - b \cdot c} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

La quantità $a \cdot d - b \cdot c$ trovata è detta *determinante* della matrice A , e indicata con $\det(A)$ oppure $\det_2(A)$ quando si voglia sottolineare che A è una matrice 2×2 .

Possiamo ora dare un'interpretazione geometrica del determinante, che ci suggerirà come generalizzare la sua definizione ad una dimensione qualsiasi.

Ricordiamo infatti che la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è invertibile se e soltanto se le

sue colonne $\begin{pmatrix} a \\ c \end{pmatrix}$ e $\begin{pmatrix} b \\ d \end{pmatrix}$ sono linearmente indipendenti. Due vettori sono linearmente dipendenti se e soltanto se sono multipli l'uno dell'altro, dunque una quantità geometrica che misura se i due vettori siano indipendenti è l'area del parallelogramma che ha tali vettori come lati. Usando la Fig. 6.1 possiamo ora calcolare l'area di tale parallelogramma come

$$(a + b) \cdot (c + d) - 2 \cdot b \cdot c - 2 \cdot \frac{1}{2} \cdot a \cdot c - 2 \cdot \frac{1}{2} \cdot b \cdot d = a \cdot d - b \cdot c.$$

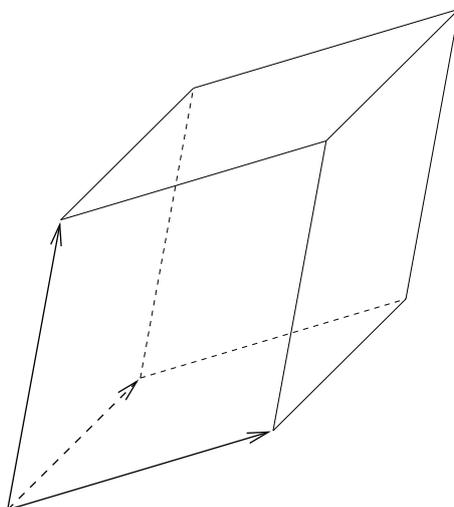


Figura 6.2: Un solido determinato da tre vettori

Troviamo dunque precisamente il determinante, già ottenuto algebricamente.

Osserviamo anche che il determinante, potendo assumere valore negativo, non è esattamente l'area, ma una sorta di “area con segno”, che possiamo interpretare come segue. Dati vettori $u, v \in \mathbb{R}^2$ il numero $\det(u \ v)$ ha valore assoluto uguale all'area del parallelogrammo di lati u e v ; il suo segno è positivo se l'angolo minore di π da u a v è diretto in senso antiorario, negativo altrimenti.

Passando ora alla dimensione tre, abbiamo una naturale definizione della nozione di determinante di una matrice con colonne u, v, w , ovvero il volume del solido determinato da u, v, w come in Fig. 6.2. Un calcolo analogo a quello fatto nel caso bidimensionale prova che tale volume è dato da

$$\begin{aligned} & u_1 \cdot v_2 \cdot w_3 + u_2 \cdot v_3 \cdot w_1 + u_3 \cdot v_1 \cdot w_2 \\ - & u_1 \cdot v_3 \cdot w_2 - u_2 \cdot v_1 \cdot w_3 - u_3 \cdot v_2 \cdot w_1. \end{aligned}$$

Questa espressione sarà allora chiamata determinante della matrice $(u \ v \ w)$ ed indicata con $\det(u \ v \ w)$ oppure $\det_3(u \ v \ w)$. La proprietà fondamentale, che segue subito dall'interpretazione geometrica, è che $\det(u \ v \ w)$ è non nullo se e solo se u, v, w costituiscono una base di \mathbb{R}^3 .

In realtà anche in questo caso stiamo calcolando un “volume con segno”. Il segno sarà positivo se è possibile disporre il dito pollice, l'indice ed il medio della mano destra rispettivamente nelle direzioni di u, v e w .

verificare grafia Sarrus C'è una regola mnemonica per il calcolo del determinante di una matrice 3×3 , detta *regola di Sarrus*, che funziona come segue. Data $A \in \mathcal{M}_{3 \times 3}(\mathbb{R})$ si ricopiano le prime due colonne di A alla sua destra:

$$A = (a_{ij})_{i,j=1,2,3} \longrightarrow \begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array} .$$

Quindi si prendono, con segno positivo, i prodotti dei termini sulle tre diagonali “nord-ovest/sud-est” e, con segno negativo, i prodotti dei termini sulle diagonali “nord-est/sud-ovest”, ottenendo:

$$\begin{aligned} \det(A) &= a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} \\ &- (a_{13} \cdot a_{22} \cdot a_{31} + a_{11} \cdot a_{23} \cdot a_{32} + a_{12} \cdot a_{21} \cdot a_{33}) . \end{aligned}$$

Procedendo in modo analogo in ogni dimensione si può definire una funzione

$$\det_n : \mathcal{M}_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$$

con la proprietà che $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è invertibile se e soltanto se $\det_n(A) \neq 0$. Per evitare di dover fare riferimento al calcolo del “volume n -dimensionale” di un oggetto in \mathbb{R}^n si possono però utilizzare altri metodi per costruire la funzione determinante, che conducono tra l'altro a esplicite regole di calcolo. Il primo metodo è di tipo *assiomatico*, cioè basato sull'individuazione di alcune proprietà fondamentali che bastano a caratterizzare la funzione determinante. Si può stabilire il seguente:

Teorema 6.2.1. *Per ogni $n \geq 1$ esiste una e una sola funzione*

$$\det_n : \mathcal{M}_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$$

tale che:

1. *fissati arbitrariamente $v_2, \dots, v_n \in \mathbb{R}^n$, la funzione*

$$\mathbb{R}^n \ni v_1 \mapsto \det_n(v_1 \ v_2 \ \dots \ v_n)$$

è lineare;

2. *scambiando tra loro due colonne di una matrice $n \times n$, il determinante cambia di segno;*
3. $\det_n(I_n) = 1$.

La dimostrazione di questo risultato è lunga e difficile, e non la faremo. È però un utile esercizio la verifica che per $n = 2, 3$ le proprietà elencate nell'enunciato valgono per la funzione \det_n costruita esplicitamente sopra.

Il secondo metodo di costruzione della funzione \det_n si basa sull'esame della formula algebrica di tale funzione per $n = 2, 3$ e sull'estensione della formula per ogni n . Si osserva infatti per $n = 2, 3$ che \det_n è la somma di $n!$ termini, ognuno dei quali è, a meno del segno, il prodotto di n coefficienti della matrice A , che stanno su righe e colonne diverse. Per descrivere con precisione quando il segno sia positivo e quando sia negativo dobbiamo fare una breve digressione sulle *permutazioni*.

Una permutazione dell'insieme $\{1, \dots, n\}$ è una funzione

$$\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$$

bigettiva. Ogni permutazione σ ha una segnatura $\text{sgn}(\sigma) \in \{-1, +1\}$ definita come

$$\text{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Questo numero è ± 1 perché si può scrivere come una frazione in cui sia denominatore che numeratore sono prodotto di $\frac{1}{2}n(n-1)$ interi, e i fattori al denominatore e al numeratore sono gli stessi salvo che per il segno (sempre positivo al denominatore, variabile al numeratore); dunque $\text{sgn}(\sigma)$ è il prodotto di $\frac{1}{2}n(n-1)$ fattori ciascuno uguale ± 1 .

La segnatura gode della seguente proprietà fondamentale, che non dimostriamo:

$$\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau).$$

Questa formula serve anche per calcolare più facilmente la segnatura di una σ assegnata. Ricordiamo intanto che si chiama *trasposizione* una permutazione che consista nel solo scambio tra loro di una coppia di elementi di $\{1, \dots, n\}$. È ora un fatto, facile, che ogni permutazione è composizione di trasposizioni, dunque avremo che $\text{sgn}(\sigma)$ è $+1$ se σ è composizione di un numero pari di trasposizioni, -1 altrimenti. Nella pratica quindi per calcolare $\text{sgn}(\sigma)$ si parte dalla stringa $(\sigma(1), \dots, \sigma(n))$ e si eseguono delle trasposizioni fino a ricondurla a $(1, \dots, n)$, quindi si contano le trasposizioni eseguite.

È importante osservare che ogni permutazione si può esprimere in più modi come prodotto di trasposizioni, ma la parità del numero di trasposizioni che compaiono è sempre la stessa. Facciamo un esempio con una permutazione su

cinque elementi:

$$\begin{array}{ccccccccc} (4, 5, 1, 2, 3) & \rightarrow & (1, 5, 4, 2, 3) & \rightarrow & (1, 3, 4, 2, 5) & \rightarrow & (1, 2, 4, 3, 5) & \rightarrow & (1, 2, 3, 4, 5) \\ & & \downarrow & & & & & & \uparrow \\ (4, 1, 5, 2, 3) & \rightarrow & (1, 4, 5, 2, 3) & \rightarrow & (1, 4, 2, 5, 3) & \rightarrow & (1, 2, 4, 5, 3) & \rightarrow & (1, 2, 4, 3, 5). \end{array}$$

La permutazione in questione ha dunque segnatura $+1$.

Tornando al determinante, si può verificare che per $n = 2, 3$ ciascuno degli $n!$ termini che compaiono nell'espressione di $\det_n(A)$ è della forma

$$\pm a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)}$$

e che il segno \pm è esattamente uguale a $\text{sgn}(\sigma)$. Lo stesso fatto vale in generale:

Teorema 6.2.2. Data $A = (a_{ij})_{i,j=1,\dots,n} \in \mathcal{M}_{n \times n}(\mathbb{R})$ si ha

$$\det_n(A) = \sum_{\sigma} \text{sgn}(\sigma) \cdot \prod_{i=1}^n a_{i,\sigma(i)}$$

dove la somma è estesa a tutte le $n!$ permutazioni di $\{1, \dots, n\}$.

Di nuovo non dimostriamo questo risultato. Enunciamo però una sua conseguenza non difficile:

Proposizione 6.2.3. $\det_n({}^t A) = \det_n(A)$ per ogni $A \in \mathcal{M}_{n \times n}(\mathbb{R})$.

Una seconda importante conseguenza è la validità delle *formule di sviluppo di Laplace*:

Teorema 6.2.4. Data $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ sia $A^{ij} \in \mathcal{M}_{(n-1) \times (n-1)}(\mathbb{R})$ la matrice ottenuta da A cancellando la riga i -esima e la colonna j -esima. Allora per ogni $k \in \{1, \dots, n\}$ si ha:

$$\begin{aligned} \det_n(A) &= \sum_{j=1}^n (-1)^{k+j} \cdot a_{kj} \cdot \det_{n-1}(A^{kj}) \\ &= \sum_{i=1}^n (-1)^{i+k} \cdot a_{ik} \cdot \det_{n-1}(A^{ik}) . \end{aligned}$$

Le formule precedenti sono dette rispettivamente di *sviluppo lungo la riga k -esima* e di *sviluppo lungo la colonna k -esima*. Osserviamo che su di esse si potrebbe basare una definizione di tipo *induttivo* del determinante \det_n , dopo averlo definito per $n = 1$ come $\det_1(a) = a$.

Illustriamo ora altre proprietà fondamentali del determinante, riconducendole (quando non sia troppo complicato) alle proprietà caratterizzanti (Teorema 6.2.1) oppure alla formula esplicita (Teorema 6.2.2). D'ora in poi scriveremo sempre $\det(A)$ in luogo di $\det_n(A)$.

Proposizione 6.2.5. *Se le colonne di $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ sono linearmente dipendenti allora $\det(A) = 0$.*

Dimostrazione. Una delle colonne è combinazione lineare delle altre. A meno di scambiare tra loro le colonne, il che cambia eventualmente il segno del determinante ma non il fatto che esso sia o meno nullo, possiamo supporre che sia la prima colonna v_1 ad essere una combinazione lineare $\alpha_2 \cdot v_2 + \dots + \alpha_n \cdot v_n$ delle altre v_2, \dots, v_n . Ora per la linearità nel primo argomento abbiamo

$$\begin{aligned} \det(v_1 \ v_2 \ \dots \ v_n) &= \alpha_2 \cdot \det(v_2 \ v_2 \ v_3 \ \dots \ v_n) \\ &+ \alpha_3 \cdot \det(v_3 \ v_2 \ v_3 \ \dots \ v_n) + \dots \\ &+ \alpha_n \cdot \det(v_n \ v_2 \ v_3 \ \dots \ v_n) . \end{aligned}$$

Siccome scambiando tra loro due colonne il determinante cambia segno, il determinante di una matrice con due colonne uguali è nullo. Tutte le matrici che compaiono a destra nell'espressione precedente hanno questa caratteristica, onde la conclusione. \square

Vale anche il viceversa della proposizione precedente, ma la dimostrazione è più complicata e la omettiamo:

Teorema 6.2.6. *Se le colonne di $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ sono linearmente indipendenti allora $\det(A) \neq 0$.*

Ne deduciamo allora che una matrice quadrata è invertibile se e soltanto se ha determinante non nullo. Ammettendo questo risultato, dimostriamo invece il prossimo fondamentale *teorema di Binet*:

Teorema 6.2.7. *Date $A, B \in \mathcal{M}_{n \times n}(\mathbb{R})$ si ha $\det(A \cdot B) = \det(A) \cdot \det(B)$.*

Dimostrazione. Se $\det(B) = 0$, cioè se B non è invertibile, certamente non è invertibile $A \cdot B$ (perché come applicazione lineare non è iniettiva), dunque $\det(A \cdot B) = 0$ e il teorema vale. Supponiamo invece che $\det(B) \neq 0$ e consideriamo l'applicazione

$$D_B : \mathcal{M}_{n \times n}(\mathbb{R}) \ni A \mapsto \frac{\det(A \cdot B)}{\det(B)} \in \mathbb{R}.$$

È immediato verificare che la D_B soddisfa le proprietà caratterizzanti della funzione \det_n elencate nel Teorema 6.2.1, dunque $D_B = \det_n$ e la conclusione segue subito. \square

Dal fatto che il determinante non cambia per trasposizione otteniamo invece la seguente:

Proposizione 6.2.8. *Sia A una matrice quadrata. Allora le righe di A sono linearmente indipendenti se e soltanto se lo sono le colonne.*

Lo stesso fatto insieme alle proprietà assiomatiche comporta anche subito il seguente risultato sul quale conviene spesso basare il calcolo del determinante:

Proposizione 6.2.9. • *Se in una matrice quadrata si scambiano tra loro due colonne (oppure due righe) il determinante cambia segno.*

- *Se in una matrice quadrata si sostituisce una colonna (o riga) con se stessa più un multiplo di un'altra colonna (o riga), il valore del determinante non cambia.*
- *Se in una matrice quadrata si sostituisce una colonna (o riga) con λ volte se stessa, il nuovo valore del determinante risulta λ volte il vecchio valore.*

6.3 Matrice inversa e rango

Il determinante ci consente non solo di stabilire se una matrice sia invertibile, ma anche di calcolarne l'inversa. Ricordiamo che si indica con A^{ij} la matrice ottenuta da A cancellando la riga i -esima e la colonna j -esima.

Proposizione 6.3.1. *Se A è quadrata e $\det(A) \neq 0$ allora*

$$(A^{-1})_{ij} = \frac{(-1)^{i+j}}{\det(A)} \cdot \det(A^{ji}).$$

Dimostrazione. Sia b_{ij} il coefficiente descritto nell'enunciato, e sia $B = (b_{ij})$. Dobbiamo verificare che $B = A^{-1}$ ovvero che $\sum_{j=1}^n b_{ij} \cdot a_{jk} = \delta_{ik}$, dove la δ è la funzione di Kronecker. Se $i = k$ abbiamo infatti:

$$\sum_{j=1}^n b_{kj} \cdot a_{jk} = \frac{1}{\det(A)} \cdot \sum_{j=1}^n (-1)^{j+k} \cdot a_{jk} \cdot \det(A^{jk})$$

che risulta 1 per la formula di sviluppo del determinante lungo la colonna k -esima. Se invece $i \neq k$ abbiamo

$$\sum_{j=1}^n b_{ij} \cdot a_{jk} = \frac{1}{\det(A)} \cdot \sum_{j=1}^n (-1)^{j+i} \cdot a_{jk} \cdot \det(A^{ji}).$$

Questa formula descrive lo sviluppo lungo la colonna i -esima del determinante di una matrice ottenuta da A sostituendo la colonna i -esima con la k -esima. Poiché tale matrice ha due colonne uguali il suo determinante è nullo, e la dimostrazione è completa. \square

Una conseguenza della formula dell'inversa è il seguente *teorema di Cramer*:

Teorema 6.3.2. *Se A è una matrice quadrata e $\det(A) \neq 0$ allora la soluzione del sistema $A \cdot x = b$ è data da*

$$x_i = \frac{\det(A_i)}{\det(A)}$$

dove A_i è ottenuta da A sostituendo la colonna i -esima con b .

Dimostrazione. Si tratta di un calcolo:

$$\begin{aligned} x_i &= (A^{-1} \cdot b)_i = \sum_{j=1}^n (A^{-1})_{ij} \cdot b_j = \sum_{j=1}^n \frac{(-1)^{i+j}}{\det(A)} \cdot \det(A^{ji}) \cdot b_j \\ &= \frac{1}{\det(A)} \cdot \sum_{j=1}^n (-1)^{i+j} \cdot b_j \cdot \det(A^{ji}). \end{aligned}$$

Nell'ultima formula si riconosce lo sviluppo di $\det(A_i)$ lungo la colonna i -esima (quella nella quale si trova il vettore b) e la tesi segue subito. \square

Veniamo ora al calcolo del rango. Data una matrice A (quadrata o rettangolare) chiamiamo *sottomatrice* di A una matrice ottenuta da A cancellando alcune righe e alcune colonne. Diciamo che una sottomatrice A' è una *orlata* di una sottomatrice A'' se A'' si ottiene da A' cancellando un'ulteriore riga e un'ulteriore colonna. Osserviamo che se $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, una sua sottomatrice $k \times h$ ha $(m - k) \cdot (n - h)$ orlate.

Il seguente risultato, detto *teorema degli orlati*, consente di ricondurre la determinazione del rango di una matrice a un calcolo di determinanti:

Teorema 6.3.3. *Sia $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e sia $r \geq 1$ un intero tale che:*

- esiste una sottomatrice A' di A con r righe ed r colonne avente determinante non nullo;
- tutte le orlate di A' hanno determinante nullo.

Allora $\text{rank}(A) = r$.

Dimostrazione. Ricordiamo che il rango è la dimensione dello spazio generato dalle colonne, dunque non cambia permutando le colonne. Inoltre una permutazione delle righe trasforma A nella matrice della composizione $f \circ A$ dove $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ è una funzione lineare bigettiva (permutazione delle coordinate). Poiché A ed $f \circ A$ hanno lo stesso rango, ne deduciamo che anche le permutazioni di righe non modificano il rango.

Possiamo perciò supporre che la sottomatrice $r \times r$ invertibile sia quella delle prime r righe ed r colonne. Scriviamo allora la matrice A nella forma

$$A = \left(\begin{array}{ccc|ccc} v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ \hline y_1^{r+1} & \cdots & y_r^{r+1} & y_{r+1}^{r+1} & \cdots & y_n^{r+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ y_1^n & \cdots & y_r^n & y_{r+1}^n & \cdots & y_n^n \end{array} \right).$$

Per ipotesi v_1, \dots, v_r costituiscono una base di \mathbb{R}^r . Ne segue subito che le prime r colonne di A sono linearmente indipendenti, dunque $\text{rank}(A) \geq r$. Per stabilire la disuguaglianza inversa dobbiamo mostrare che per $i > r$ la colonna i -esima di A è linearmente dipendente dalle prime r . Fissiamo tale i . Consideriamo ora un qualsiasi $j > r$ e ricordiamo che per ipotesi la sottomatrice di A ottenuta orlando la $A' = (v_1 \dots v_r)$ con la colonna i -esima e la riga j -esima ha determinante nullo. Poiché le prime r colonne di questa matrice sono linearmente indipendenti, è l'ultima colonna che deve essere una combinazione lineare delle precedenti. Esistono dunque costanti $\alpha_1^j, \dots, \alpha_r^j$ tali che:

$$\begin{cases} v_i &= \alpha_1^j \cdot v_1 + \dots + \alpha_r^j \cdot v_r \\ y_i^j &= \alpha_1^j \cdot y_1^j + \dots + \alpha_r^j \cdot y_r^j. \end{cases}$$

Abbiamo sottolineato che le costanti dipendono a priori da j , perché i è stato fissato mentre j è uno qualsiasi degli interi maggiori di r . Dalle formule segue tuttavia che le costanti α_k^j in realtà *non dipendono da j* , perché l'espressione di v_i come combinazione lineare di v_1, \dots, v_r è unica.

Dal fatto che gli scalari α_k^j sono indipendenti da j segue subito che la colonna i -esima di A è combinazione lineare con coefficienti $\alpha_1, \dots, \alpha_r$ delle prime r colonne. Ciò conclude il nostro argomento. \square

Spieghiamo ora come il teorema degli orlati si utilizzi in pratica per trovare il rango di una matrice A :

- Se tutti i coefficienti di A sono nulli, il rango di A è nullo. Altrimenti si fissa una sottomatrice A_1 di A che contenga un solo coefficiente non nullo e si procede;
- Si esaminano tutte le orlate di A_1 , cioè le sottomatrici 2×2 che contengono il coefficiente fissato; se tutte queste matrici hanno determinante nullo, il rango di A è 1; altrimenti, alla prima che si trova con determinante non nullo, la si fissa chiamandola A_2 , e si procede;
- Si esaminano le orlate di A_2 e si procede allo stesso modo.

Osservazione 6.3.4. *Nel corso dell'esecuzione del procedimento appena descritto, al passo p -esimo si devono considerare solo le sottomatrici $(p+1) \times (p+1)$ che sono orlate di una fissata $p \times p$, non tutte le $(p+1) \times (p+1)$. Il vantaggio numerico di questo fatto è consistente. Ad esempio per una matrice in $\mathcal{M}_{6 \times 5}(\mathbb{R})$ le orlate di una 2×2 sono $4 \cdot 3 = 12$, mentre tutte le 3×3 sono $\binom{6}{3} \cdot \binom{5}{3} = 20 \cdot 10 = 200$.*

Una conseguenza notevole del teorema degli orlati e del fatto che il determinante non cambia per trasposizione è il seguente:

Corollario 6.3.5. *In ogni matrice, il numero massimo di righe linearmente indipendenti è uguale al numero massimo di colonne linearmente indipendenti.*

Capitolo 7

Sottospazi affini

Abbiamo visto nel capitolo precedente che l'insieme delle soluzioni di un sistema lineare è ottenuto da un sottospazio vettoriale di \mathbb{R}^n per traslazione. Questo capitolo è dedicato a uno studio sistematico degli insiemi di questo tipo.

7.1 Equazioni cartesiane e parametriche

Sia V uno spazio vettoriale su \mathbb{R} . Un sottoinsieme E di V si dice *sottospazio affine* se esistono un vettore v_0 in V ed un sottospazio vettoriale W di V tali che

$$E = \{v_0 + w : w \in W\} ,$$

che scriveremo anche come $E = v_0 + W$. Si osservi che per definizione un sottospazio affine è non vuoto.

Proposizione 7.1.1. *Se $E = v_0 + W$ allora W è univocamente determinato da E , mentre v_0 può essere sostituito da un qualsiasi punto di E .*

Dimostrazione. Supponiamo che $v_0 + W = v'_0 + W'$. Allora $v'_0 = v_0 + w_0$ per qualche $w_0 \in W$. Inoltre dato $w' \in W'$ qualsiasi si ha $v'_0 + w' = v_0 + w$ per qualche $w \in W$, dunque $w' = v_0 + w - (v_0 + w_0) = w - w_0 \in W$. Ciò prova che $W' \subset W$, e per simmetria si ha $W' = W$. Inoltre:

$$v'_0 + W = v_0 + W \Leftrightarrow v'_0 - v_0 \in W \Leftrightarrow v'_0 \in v_0 + W = E.$$

□

Se E è un sottospazio affine di V ed $E = v_0 + W$, diciamo che W è il sottospazio vettoriale *associato* a E . Definiamo inoltre la *dimensione* $\dim_{\mathbb{R}}(E)$ di E come quella di W .

La proposizione precedente comporta subito la seguente:

Osservazione 7.1.2. Se E è un sottospazio affine di V sono fatti equivalenti:

- V contiene 0;
- E è un sottospazio vettoriale di V .

Come già annunciato, vale anche la:

Osservazione 7.1.3. Se non è vuoto, l'insieme delle soluzioni di un sistema lineare è un sottospazio affine di \mathbb{R}^n .

Diremo ora che un sottospazio affine E di \mathbb{R}^n è presentato in *forma cartesiana* se è descritto come insieme delle soluzioni di un sistema, cioè come

$$E = \{x \in \mathbb{R}^n : A \cdot x = b\}$$

per qualche $m \geq 1$, $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$. In tal caso, diremo anche che E ha *equazioni cartesiane* $A \cdot x = b$ (si tratta di m equazioni in n incognite). Vedremo tra poco che ogni sottospazio affine di \mathbb{R}^n può essere descritto tramite equazioni cartesiane, ma può accadere che alcune equazioni siano sovrabbondanti. Notiamo intanto che si ha

$$\dim_{\mathbb{R}}(E) = \dim_{\mathbb{R}}(\text{Ker}(A)) = n - \text{rank}(A)$$

(si ricordi che E è non vuoto per ipotesi).

Diremo invece che E è presentato in *forma parametrica* (ovvero tramite *equazioni parametriche*) se è descritto come

$$E = \{x_0 + B \cdot t : t \in \mathbb{R}^p\}$$

per qualche $x_0 \in \mathbb{R}^n$, $p \geq 1$ e $B \in \mathcal{M}_{n \times p}(\mathbb{R})$. La ragione di questa terminologia è che, se v_1, \dots, v_p sono le colonne di B , l'uguaglianza $E = \{x_0 + B \cdot t : t \in \mathbb{R}^p\}$ si riscrive come

$$E = \{x_0 + t_1 \cdot v_1 + \dots + t_p \cdot v_p : t_1, \dots, t_p \in \mathbb{R}\} \quad ,$$

dunque al fatto che un punto di E è descritto dai *parametri* t_1, \dots, t_p . Vedremo che ogni sottospazio affine di \mathbb{R}^n può essere descritto in forma parametrica, ma può accadere che alcuni parametri siano sovrabbondanti. Notiamo che $\dim_{\mathbb{R}}(E) = \text{rank}(B)$.

Proposizione 7.1.4. *Sia $E \subset \mathbb{R}^n$ un sottospazio affine di dimensione d . Allora:*

- *Esistono $A \in \mathcal{M}_{(n-d) \times n}(\mathbb{R})$ e $b \in \mathbb{R}^{n-d}$ tali che*

$$E = \{x \in \mathbb{R}^n : A \cdot x = b\} \quad ,$$

ovvero E ammette una presentazione tramite $n - d$ equazioni cartesiane;

- *Esistono $x_0 \in \mathbb{R}^n$ e $B \in \mathcal{M}_{n \times d}(\mathbb{R})$ tali che*

$$E = \{x_0 + B \cdot t : t \in \mathbb{R}^d\} \quad ,$$

ovvero E ammette una presentazione parametrica con d parametri.

Dimostrazione. Sia $E = x_0 + W$ e scegliamo una base w_1, \dots, w_d di W . Per la presentazione parametrica basta allora porre $B = (w_1 \ \dots \ w_d)$. Per quella cartesiana dobbiamo esibire una matrice $A \in \mathcal{M}_{n \times (n-d)}(\mathbb{R})$ tale che $\text{Ker}(A) = \text{Span}(w_1, \dots, w_d)$, perché trovata tale A basterà scegliere b come $A \cdot x_0$. Definiamo la matrice completando w_1, \dots, w_d ad una base w_1, \dots, w_n di \mathbb{R}^n e ponendo $A = [f]_{\mathcal{E}_n}^{\mathcal{E}_{n-d}}$ dove $f : \mathbb{R}^n \rightarrow \mathbb{R}^{n-d}$ è definita dalle relazioni

$$f(w_i) = \begin{cases} 0 & \text{per } i = 1, \dots, d \\ e_{i-d}^{(n-d)} & \text{per } i = d+1, \dots, n. \end{cases}$$

□

Questa proposizione dimostra che un sottospazio affine definito da equazioni cartesiane $A \cdot x = b$ con $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e $\text{rank}(A) < m$ può essere definito anche da meno di m equazioni, vale a dire esattamente da $\text{rank}(A) = n - \dim_{\mathbb{R}}(E)$ equazioni. È inoltre facile verificare che si possono semplicemente scartare $m - \text{rank}(A)$ delle equazioni $A \cdot x = b$, scegliendo un insieme di $\text{rank}(A)$ righe di A linearmente indipendenti. Si osservi tuttavia che bisogna dapprima verificare che il sottospazio affine sia non vuoto, cioè che il sistema di equazioni abbia soluzioni.

Analogamente, se $E = \{x_0 + B \cdot t : t \in \mathbb{R}^p\}$ con $B \in \mathcal{M}_{n \times p}(\mathbb{R})$ e $\text{rank}(B) < p$, si può sostituire B con una matrice avente $\text{rank}(B)$ colonne. Basterà in questo caso scegliere $\text{rank}(B)$ colonne di B che siano linearmente indipendenti.

7.2 Rette nel piano, rette e piani nello spazio

Con il linguaggio introdotto nel paragrafo precedente, risolvere un sistema lineare significa passare da una presentazione cartesiana a una parametrica di un sottospazio affine di \mathbb{R}^n (dopo avere verificato che il sistema ammette soluzioni). In questo paragrafo ci concentriamo su questo passaggio e il suo opposto (da equazioni parametriche a equazioni cartesiane) nel caso di sottospazi di \mathbb{R}^2 e \mathbb{R}^3 , provando che è possibile eseguirli in modo semplice e sistematico. Per \mathbb{R}^3 ci limiteremo ai sottospazi vettoriali, cioè ai sistemi omogenei.

Cominciamo con \mathbb{R}^2 , i cui unici sottospazi non banali sono le rette, ed osserviamo che la presentazione cartesiana

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : a \cdot x + b \cdot y = c \right\}$$

definisce una retta se e soltanto se $(a, b) \neq (0, 0)$, ovvero se $a^2 + b^2 > 0$. Analogamente la presentazione parametrica

$$\left\{ \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} : t \in \mathbb{R} \right\}$$

definisce una retta se e soltanto se $\alpha^2 + \beta^2 > 0$. Inoltre:

$$\begin{aligned} & \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : a \cdot x + b \cdot y = c \right\} \\ &= \left\{ \frac{c}{a^2 + b^2} \begin{pmatrix} a \\ b \end{pmatrix} + t \cdot \begin{pmatrix} -b \\ a \end{pmatrix} : t \in \mathbb{R} \right\}, \\ & \left\{ \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} : t \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \beta \cdot x - \alpha \cdot y = \beta \cdot x_0 - \alpha \cdot y_0 \right\}. \end{aligned}$$

Passando a \mathbb{R}^3 , consideriamo dapprima le rette. Una presentazione cartesiana di una retta passante per l'origine è del tipo

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z = 0 \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z = 0 \end{cases} \right\}$$

con $\text{rank} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2$, mentre una presentazione parametrica è del tipo

$$\left\{ t \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} : t \in \mathbb{R} \right\}$$

con $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \neq 0$. Il passaggio da equazione parametrica a equazione cartesiana è facile:

$$\left\{ t \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} : t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{cases} \beta \cdot x - \alpha \cdot y = 0 \\ \gamma \cdot x - \alpha \cdot z = 0 \\ \gamma \cdot y - \beta \cdot z = 0 \end{cases} \right\}.$$

Si osservi che le equazioni sono tre, invece che due, ma sono sempre dipendenti tra loro, anzi è sempre possibile scartare una delle tre ottenendone due indipendenti, infatti

$$\text{rank} \begin{pmatrix} \beta & -\alpha & 0 \\ \gamma & 0 & -\alpha \\ 0 & \gamma & -\beta \end{pmatrix} = 2.$$

Abbiamo scritto tre equazioni perché, al variare di α, β, γ , varia l'equazione che, dipendendo dalle altre due, si può scartare.

Il passaggio da equazione cartesiana a equazione parametrica si effettua invece come segue:

$$\begin{aligned} & \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z = 0 \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z = 0 \end{cases} \right\} \\ &= \left\{ t \cdot \begin{pmatrix} b_1 \cdot c_2 - b_2 \cdot c_1 \\ -a_1 \cdot c_2 + a_2 \cdot c_1 \\ a_1 \cdot b_2 - a_2 \cdot b_1 \end{pmatrix} : t \in \mathbb{R} \right\}. \end{aligned}$$

Spieghiamo come si giunga a questa regola. Stiamo cercando una soluzione non nulla del sistema

$$\begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z = 0 \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z = 0. \end{cases}$$

Osserviamo allora che si ha

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \det \begin{pmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 0,$$

dunque, sviluppando i due determinanti lungo la prima riga:

$$\begin{aligned} a_1 \cdot \det \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix} + b_1 \cdot \left(-\det \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \right) + c_1 \cdot \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} &= 0 \\ a_2 \cdot \det \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix} + b_2 \cdot \left(-\det \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \right) + c_2 \cdot \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} &= 0 \end{aligned}$$

ed abbiamo trovato la soluzione richiesta, infatti la condizione

$$\text{rank} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2$$

equivale esattamente al fatto che almeno uno dei determinanti delle tre sottomatrici 2×2 sia non nullo.

Ci rivolgiamo dunque ai piani per l'origine, rappresentati in forma cartesiana come

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : a \cdot x + b \cdot y + c \cdot z = 0 \right\}$$

con $(a, b, c) \neq (0, 0, 0)$, e in forma parametrica come

$$\left\{ t_1 \cdot \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} + t_2 \cdot \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} : t_1, t_2 \in \mathbb{R} \right\}$$

con $\text{rank} \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \\ \gamma_1 & \gamma_2 \end{pmatrix} = 2$. Il passaggio da presentazione cartesiana a parametrica si effettua come segue:

$$\begin{aligned} &\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : a \cdot x + b \cdot y + c \cdot z = 0 \right\} \\ &= \left\{ t_1 \cdot \begin{pmatrix} b \\ -a \\ 0 \end{pmatrix} + t_2 \cdot \begin{pmatrix} c \\ 0 \\ -a \end{pmatrix} + t_3 \cdot \begin{pmatrix} 0 \\ c \\ -b \end{pmatrix} : t_1, t_2, t_3 \in \mathbb{R} \right\} \end{aligned}$$

dove uno dei tre vettori si può sempre scartare perché linearmente dipendente dagli altri due. Concludiamo illustrando il passaggio inverso, che si spiega come

sopra:

$$\left\{ t_1 \cdot \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} + t_2 \cdot \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} : t_1, t_2 \in \mathbb{R} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{array}{l} (\beta_1 \cdot \gamma_2 - \beta_2 \cdot \gamma_1) \cdot x \\ -(\alpha_1 \cdot \gamma_2 - \alpha_2 \cdot \gamma_1) \cdot y \\ +(\alpha_1 \cdot \beta_2 - \alpha_2 \cdot \beta_1) \cdot z = 0 \end{array} \right\}.$$

7.3 Somma di sottospazi

Dimostriamo in questo paragrafo un analogo per sottospazi affini della formula di Grassmann. Notiamo intanto che se E, F sono sottospazi affini di V allora $E \cap F$, se è non vuoto, è a sua volta un sottospazio affine. Per analogia con il caso dei sottospazi vettoriali definiamo ora la *somma* $E + F$ come il più piccolo sottospazio affine di V che contiene $E \cup F$, ovvero come l'intersezione di tutti i sottospazi affini di V che contengono $E \cup F$. Vogliamo porre in relazione tra loro le dimensioni di $E, F, E \cap F, E + F$. C'è un primo facile caso:

Proposizione 7.3.1. *Se $E \cap F \neq \emptyset$ si ha*

$$\dim_{\mathbb{R}}(E) + \dim_{\mathbb{R}}(F) = \dim_{\mathbb{R}}(E \cap F) + \dim_{\mathbb{R}}(E + F).$$

Dimostrazione. Scelto $v_0 \in E \cap F$ si ha $E = v_0 + W$ e $F = v_0 + Z$ per W, Z opportuni. È allora immediato verificare che $E \cap F = v_0 + (W \cap Z)$ e $E + F = v_0 + (W + Z)$ onde la conclusione dalla formula di Grassmann. \square

Più complicato è invece il caso seguente:

Proposizione 7.3.2. *Siano $E = v_0 + W$ e $F = u_0 + Z$ sottospazi affini di V , con $E \cap F = \emptyset$. Allora*

$$\dim_{\mathbb{R}}(E + F) = 1 + \dim_{\mathbb{R}}(W + Z).$$

Dimostrazione. Posto $x_0 = u_0 - v_0$, la conclusione segue subito dalle seguenti asserzioni:

- $x_0 \notin W + Z$;
- $E + F = v_0 + \text{Span}(\{x_0\} \cup (W + Z))$.

Stabiliamo la prima per assurdo, supponendo che x_0 si possa scrivere come $w + z$ con $w \in W$ e $z \in Z$. Allora $v_0 - w = u_0 + z$, e questo punto appartiene all'intersezione di E ed F : contraddizione.

Per stabilire la seconda asserzione poniamo $X = \text{Span}(\{x_0\} \cup (W + Z))$ e $L = v_0 + X$ e verifichiamo che L , il quale ovviamente è un sottospazio affine di V , contiene $E \cup F$ ed è il più piccolo con tale proprietà, cioè è contenuto in qualsiasi altro sottospazio affine $T = v_0 + Y$ che contiene $E \cup F$. Il primo fatto è semplice:

$$\left. \begin{array}{l} X \supset W \Rightarrow L \supset v_0 + W = E \\ X \ni u_0 - v_0, X \supset Z \Rightarrow L \supset v_0 + (u_0 - v_0) + Z = F \end{array} \right\} \Rightarrow L \supset E \cup F.$$

Verifichiamo allora il secondo:

$$\left. \begin{array}{l} T \supset E \Rightarrow Y \supset W \\ T \supset F \Rightarrow Y \supset Z \\ T \ni u_0 \Rightarrow Y \ni u_0 - v_0 \end{array} \right\} \Rightarrow Y \supset X \Rightarrow T \supset L.$$

La dimostrazione è completa. \square

Prima di illustrare con un esempio il significato della proposizione precedente, diamo ancora una definizione generale: due sottospazi affini $E = v_0 + W$ e $F = u_0 + Z$ di V si dicono *paralleli* se $W \subset Z$ oppure $Z \subset W$. Si osservi che, se E ed F hanno la stessa dimensione, questa condizione significa che $W = Z$, ovvero che E ed F sono traslati del medesimo sottospazio vettoriale, il che naturalmente corrisponde alla nozione elementare di parallelismo.

Rivolgiamoci dunque al caso particolare di due rette nello spazio tridimensionale. Se si intersecano, le rette possono avere in comune un punto oppure coincidere, e la dimensione della loro somma è rispettivamente 2 e 1. Se invece non si intersecano possono essere parallele oppure *sghembe*, e la dimensione della loro somma è rispettivamente 2 e 3. La situazione delle rette sghembe è mostrata nella Fig. 7.1, dove viene illustrata la dimostrazione della Proposizione 7.3.2: in questo caso $W = \text{Span}(w)$, $Z = \text{Span}(z)$, ed i vettori $w, z, x = x_0$ costituiscono una base di \mathbb{R}^3 .

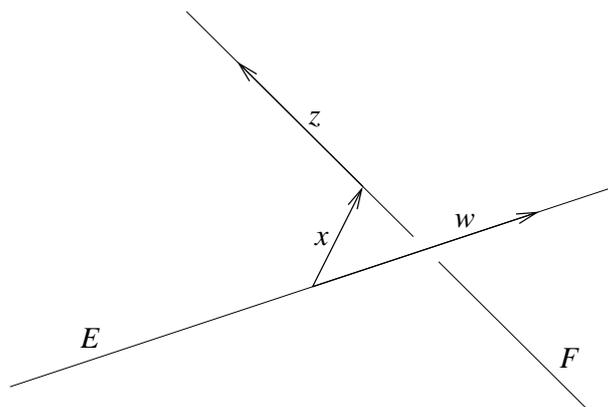


Figura 7.1: Due rette sghembe nello spazio

Capitolo 8

Numeri complessi

Nel Capitolo 2 abbiamo motivato i passaggi da \mathbb{N} a \mathbb{Z} e poi da \mathbb{Z} a \mathbb{Q} con l'esigenza di risolvere le equazioni $x + n = 0$ e $x \cdot n = 1$, altrimenti prive di radici. Invece nel passaggio da \mathbb{Q} a \mathbb{R} abbiamo aggiunto non solo le soluzioni di equazioni non aventi radici in \mathbb{Q} (come ad esempio $x^2 = 2$) ma anche molti altri oggetti, corrispondenti ai punti della retta.¹ Ora anche su \mathbb{R} ci sono equazioni che non hanno soluzioni, la più semplice delle quali è

$$x^2 + 1 = 0.$$

Si “inventa” allora una sua soluzione, detta unità immaginaria e denotata usualmente con i (dai matematici, e con j dagli ingegneri). Estendendo le naturali operazioni algebriche che sono date su \mathbb{R} si giunge allora a definire un nuovo insieme di numeri, denotato con \mathbb{C} , i cui elementi sono detti *numeri complessi*. Questo capitolo è dedicato alle principali proprietà di \mathbb{C} .

8.1 Il campo complesso

La “soluzione” i dell'equazione $x^2 + 1 = 0$ è un oggetto che non appartiene a \mathbb{R} , ma vogliamo considerarlo come un numero e dunque eseguire operazioni algebriche che lo coinvolgano, come ad esempio:

- sommargli un numero reale α , ottenendo un oggetto $\alpha + i$;

¹Si può infatti mostrare che i punti di \mathbb{R} che sono radici di polinomi a coefficienti razionali sono “infinitamente meno numerosi” di quelli che non lo sono. I due esempi più celebri di numeri che non risolvono equazioni polinomiali razionali sono il rapporto π tra la lunghezza di una circonferenza e quella del suo diametro, e il numero di Nepero e (base del logaritmo naturale).

- moltiplicarlo per un numero reale β , ottenendo un oggetto $i \cdot \beta$;
- fare insieme le due cose precedenti, ottenendo un oggetto $\alpha + i \cdot \beta$;
- sommare tra loro due oggetti del tipo ottenuto, il che è ragionevole fare nel modo seguente:

$$(\alpha_1 + i \cdot \beta_1) + (\alpha_2 + i \cdot \beta_2) = (\alpha_1 + \alpha_2) + i \cdot (\beta_1 + \beta_2) ;$$

- moltiplicare tra loro due oggetti del tipo ottenuto; poiché si ha $i^2 = -1$ il modo ragionevole per eseguire una tale moltiplicazione è il seguente:

$$(\alpha_1 + i \cdot \beta_1) \cdot (\alpha_2 + i \cdot \beta_2) = (\alpha_1 \cdot \alpha_2 - \beta_1 \cdot \beta_2) + i \cdot (\alpha_1 \cdot \beta_2 + \alpha_2 \cdot \beta_1).$$

Poiché somma e prodotto tra loro di oggetti del tipo $\alpha + i \cdot \beta$ con $\alpha, \beta \in \mathbb{R}$ danno oggetti dello stesso tipo, abbiamo ora un insieme su cui sono definite due operazioni binarie interne, cioè l'insieme

$$\mathbb{C} = \{\alpha + i \cdot \beta : \alpha, \beta \in \mathbb{R}\}.$$

Teorema 8.1.1. Ponendo $0 = 0 + i \cdot 0$ e $1 = 1 + i \cdot 0$ l'insieme \mathbb{C} con le operazioni sopra introdotte risulta un campo.

Dimostrazione. Dobbiamo verificare le proprietà (1)-(9) della definizione di campo, discusse nei Sottoparagrafi 2.1.2 e 2.1.3. Si tratta di verifiche molto semplici, tranne che per la distributività (9) e l'esistenza dell'inverso moltiplicativo (6). Proviamo dapprima la (9):

$$\begin{aligned} & (\alpha_1 + i \cdot \beta_1) \cdot ((\alpha_2 + i \cdot \beta_2) + (\alpha_3 + i \cdot \beta_3)) \\ = & (\alpha_1 + i \cdot \beta_1) \cdot ((\alpha_2 + \alpha_3) + i \cdot (\beta_2 + \beta_3)) \\ = & (\alpha_1 \cdot (\alpha_2 + \alpha_3) - \beta_1 \cdot (\beta_2 + \beta_3)) + i \cdot (\alpha_1 \cdot (\beta_2 + \beta_3) + \beta_1 \cdot (\alpha_2 + \alpha_3)) \\ = & ((\alpha_1 \cdot \alpha_2 - \beta_1 \cdot \beta_2) + (\alpha_1 \cdot \alpha_3 - \beta_1 \cdot \beta_3)) \\ & + i \cdot ((\alpha_1 \cdot \beta_2 + \beta_1 \cdot \alpha_2) + i \cdot (\alpha_1 \cdot \beta_3 + \beta_1 \cdot \alpha_3)) \\ = & (\alpha_1 + i \cdot \beta_1) \cdot (\alpha_2 + i \cdot \beta_2) + (\alpha_1 + i \cdot \beta_1) \cdot (\alpha_3 + i \cdot \beta_3) . \end{aligned}$$

Passando alla (6), supponiamo che $\alpha + i \cdot \beta$ sia diverso da $0 = 0 + i \cdot 0$, ovvero che uno almeno tra α e β sia non nullo. Dobbiamo risolvere nelle incognite $x, y \in \mathbb{R}$ l'equazione $(\alpha + i \cdot \beta) \cdot (x + i \cdot y) = 1 + i \cdot 0$, ovvero il sistema

$$\begin{cases} \alpha \cdot x - \beta \cdot y = 1 \\ \beta \cdot x + \alpha \cdot y = 0. \end{cases}$$

La matrice del sistema ha determinante $\alpha^2 + \beta^2 > 0$, dunque la soluzione esiste ed è data da $x = \frac{\alpha}{\alpha^2 + \beta^2}$ e $y = -\frac{\beta}{\alpha^2 + \beta^2}$. \square

Sottolineiamo la regola trovata nella dimostrazione precedente per il calcolo dell'inverso di un numero complesso:

$$\alpha + i \cdot \beta \neq 0 \Rightarrow (\alpha + i \cdot \beta)^{-1} = \frac{\alpha - i \cdot \beta}{\alpha^2 + \beta^2}$$

D'ora in poi useremo di frequente una singola lettera, spesso z , per indicare un numero complesso, cioè un elemento di \mathbb{C} . Se $z = \alpha + i \cdot \beta$ diremo inoltre che α è la *parte reale* di z , e la indicheremo con $\Re(z)$, mentre β sarà chiamata *parte immaginaria* e denotata con $\Im(z)$. Identificheremo inoltre un numero complesso della forma $x + i \cdot 0$ con il numero reale x , convenendo pertanto che \mathbb{R} sia un sottoinsieme di \mathbb{C} . Questa scelta è ragionevole anche perché le operazioni algebriche introdotte su \mathbb{C} coincidono con quelle definite su \mathbb{R} quando applicate a numeri complessi con parte immaginaria nulla. Infine, chiamiamo *immaginario puro* un numero complesso con parte reale nulla.

Poiché è facile sbagliarsi, sottolineiamo che la parte immaginaria di un numero complesso è un numero *reale*, non immaginario puro, e che un numero reale è *anche* un numero complesso.

8.2 Il piano complesso

Avendo definito \mathbb{C} come l'insieme degli oggetti della forma $x + i \cdot y$ con $x, y \in \mathbb{R}$, abbiamo una naturale corrispondenza tra \mathbb{C} e \mathbb{R}^2 , ottenuta associando al numero $x + i \cdot y$ il punto (o vettore) di coordinate x e y nel piano. Possiamo dunque anche rappresentare i numeri complessi graficamente nel piano, come suggerito in Fig. 8.1. Osserviamo subito che la somma tra numeri complessi, dato che viene eseguita separatamente per parte reale e immaginaria, non è altro che la consueta somma tra vettori nel piano, eseguita con la regola del parallelogrammo. Anche l'operazione di prodotto ha un'interpretazione geometrica, sulla quale torneremo.

Definiamo ora il coniugato \bar{z} di un numero complesso z come l'immagine di z rispetto alla riflessione nella retta reale. Se $z = x + i \cdot y$ si ha allora $\bar{z} = x - i \cdot y$. Sono inoltre di facile verifica le seguenti proprietà:

- $\Re(z) = \frac{1}{2}(z + \bar{z})$;
- $\Im(z) = \frac{1}{2i}(z - \bar{z})$;
- $\overline{\bar{z}} = z$;
- $\overline{z + w} = \bar{z} + \bar{w}$;

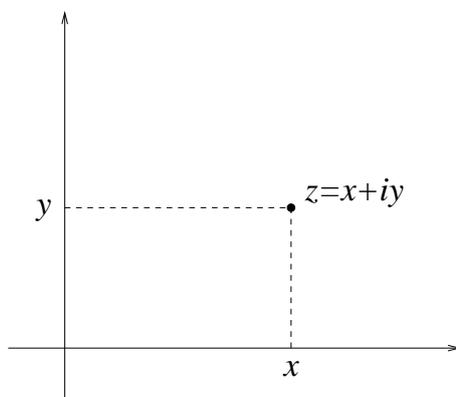


Figura 8.1: Il piano complesso

- $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$.

Usiamo ora più a fondo la corrispondenza tra \mathbb{C} e il piano, nonché la geometria elementare del piano, introducendo la misura della distanza tra un numero $z \in \mathbb{C}$ e l'origine del piano. Se $z = x + i \cdot y$, questa distanza si calcola con il teorema di Pitagora, e risulta uguale a $\sqrt{x^2 + y^2}$. Chiamiamo *modulo* di z questa quantità, e la denotiamo con $|z|$, poiché naturalmente su un numero reale uguaglia il valore assoluto. Si verifica immediatamente che

$$|\bar{z}| = |z|, \quad |z|^2 = z \cdot \bar{z}$$

il che conduce a una nuova regola di calcolo dell'inverso di un numero complesso, che consente di non introdurre esplicitamente la sua parte reale e la sua parte immaginaria:

$$z \neq 0 \Rightarrow z^{-1} = \frac{\bar{z}}{|z|^2}.$$

Usando questa formula si verifica anche facilmente che se $z \neq 0$ allora $\overline{z^{-1}} = \bar{z}^{-1}$ e $|\frac{1}{z}| = \frac{1}{|z|}$. Inoltre $|z \cdot w| = |z| \cdot |w|$.

Abbiamo per il modulo anche le seguenti elementari disuguaglianze:

- $|z| \geq 0$ per ogni z , e $|z| = 0$ se e soltanto se $z = 0$;
- $|z| \geq |\Re(z)|$;
- $|z| \geq |\Im(z)|$;

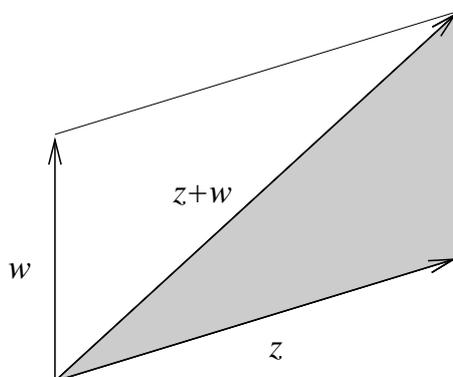


Figura 8.2: Disuguaglianza triangolare

- $|z| \leq |\Re(z)| + |\Im(z)|$.

Proviamo ora una disuguaglianza più forte, che avrà una ben chiara interpretazione geometrica:

Proposizione 8.2.1. $|z + w| \leq |z| + |w|$.

Dimostrazione.

$$\begin{aligned} |z + w|^2 &= (z + w) \cdot \overline{(z + w)} = (z + w) \cdot (\bar{z} + \bar{w}) \\ &= |z|^2 + |w|^2 + (z \cdot \bar{w} + \bar{z} \cdot w) = |z|^2 + |w|^2 + 2 \cdot \Re(z \cdot \bar{w}) \\ &\leq |z|^2 + |w|^2 + 2 \cdot |z \cdot \bar{w}| = |z|^2 + |w|^2 + 2 \cdot |z| \cdot |w| = (|z| + |w|)^2. \end{aligned}$$

□

L'interpretazione geometrica della disuguaglianza si ottiene considerando $z, w, z + w$ come i lati di un triangolo, secondo quanto suggerito in Fig. 8.2. La disuguaglianza è allora la classica *triangolare*: la lunghezza di un lato è non maggiore della somma delle lunghezze degli altri due.

8.3 Forma polare ed esponenziale complessa

Sfruttando ancora la corrispondenza tra \mathbb{C} e \mathbb{R}^2 , introduciamo ora un nuovo modo di rappresentare i numeri complessi, che ci consentirà tra l'altro di dare una chiara interpretazione geometrica del prodotto. Se $z \in \mathbb{C}$ è non nullo, si può esprimere il corrispondente punto $(\Re(z), \Im(z))$ di \mathbb{R}^2 tramite *coordinate*

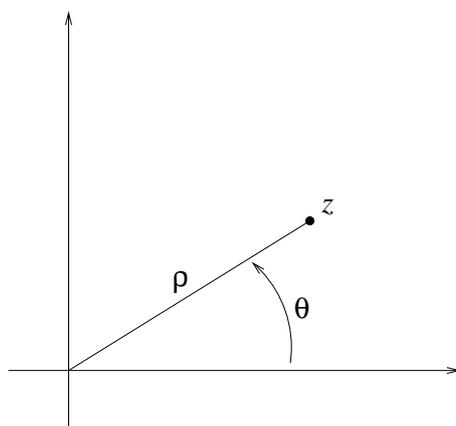


Figura 8.3: Forma polare di un numero complesso

polari (ρ, ϑ) come suggerito in Fig. 8.3, con $\rho > 0$ e $\vartheta \in \mathbb{R}$. Possiamo dunque scrivere

$$(\clubsuit) \quad z = \rho \cdot (\cos(\vartheta) + i \cdot \sin(\vartheta)),$$

che chiameremo *forma polare* del numero complesso z . Osserviamo che il numero ρ , che risulta univocamente determinato dalla (\clubsuit) , non è altro che il modulo $|z|$ di z . Chiamiamo invece ϑ l'*argomento* $\arg(z)$ di z , e notiamo che $\arg(z)$ è invece definito in modo ambiguo, cioè a meno di un addendo del tipo $2k\pi$ con $k \in \mathbb{Z}$. L'ambiguità può essere rimossa convenendo che $\arg(z)$ debba appartenere all'intervallo $[0, 2\pi)$ oppure a $[-\pi, \pi)$. Tuttavia questa scelta ha l'inconveniente che numeri complessi molto vicini tra loro risultano avere argomenti molto diversi, quindi è preferibile per lo più mantenere l'ambiguità e considerare $\arg(z)$ non come un solo numero, ma come un insieme del tipo $\{\vartheta_0 + 2k\pi : k \in \mathbb{Z}\}$. Osserviamo ancora che la scrittura (\clubsuit) si può estendere anche per z nullo. Abbiamo in questo caso $\rho = 0$ e ϑ completamente arbitrario.

Calcoliamo ora il prodotto di due numeri complessi espressi in forma polare:

$$\begin{aligned} & \left(\rho \cdot (\cos(\vartheta) + i \cdot \sin(\vartheta)) \right) \cdot \left(\lambda \cdot (\cos(\phi) + i \cdot \sin(\phi)) \right) \\ = & (\rho \cdot \lambda) \cdot \left((\cos(\vartheta) \cdot \cos(\phi) - \sin(\vartheta) \cdot \sin(\phi)) \right. \\ & \left. + i(\sin(\vartheta) \cdot \cos(\phi) + \cos(\vartheta) \cdot \sin(\phi)) \right) \\ = & (\rho \cdot \lambda) \cdot (\cos(\vartheta + \phi) + i \cdot \sin(\vartheta + \phi)) . \end{aligned}$$

Scopriamo pertanto che, moltiplicando un numero complesso per un altro numero della forma $\rho \cdot (\cos(\vartheta) + i \cdot \sin(\vartheta))$, modulo e argomento variano come segue:

- il modulo viene moltiplicato per un fattore ρ ;
- l'argomento viene aumentato di un addendo ϑ .

Ne deduciamo allora la seguente:

Proposizione 8.3.1. *La moltiplicazione per il numero $\rho \cdot (\cos(\vartheta) + i \cdot \sin(\vartheta))$ è l'applicazione $\mathbb{C} \rightarrow \mathbb{C}$ ottenuta componendo l'omotetia di ragione ρ e centro 0 con la rotazione di angolo ϑ intorno a 0.*

La formula trovata per il prodotto comporta inoltre che, posto $f(\vartheta) = \cos(\vartheta) + i \cdot \sin(\vartheta)$, si ha $f(\vartheta) \cdot f(\phi) = f(\vartheta + \phi)$. Poiché questa è una delle proprietà fondamentali soddisfatte dalle funzioni esponenziali, ciò suggerisce che $f(\vartheta)$ si possa scrivere come a^ϑ per un opportuno a (non reale, naturalmente). Di fatto si sceglie $a = e^i$, ovvero si conviene che

$$e^{i\vartheta} = \cos(\vartheta) + i \cdot \sin(\vartheta).$$

Questa scelta si può anche giustificare attraverso gli sviluppi in serie di Taylor delle funzioni esponenziale e trigonometriche. Ricordiamo infatti che per t reale risulta

$$e^t = \sum_{n=0}^{\infty} \frac{1}{n!} t^n.$$

Se ora sostituiamo t con $i \cdot \vartheta$ e ricordiamo che $i^2 = -1$, abbiamo

$$\begin{aligned} e^{i\vartheta} &= \sum_{n=0}^{\infty} \frac{i^{2n}}{(2n)!} \cdot \vartheta^{2n} + \sum_{n=0}^{\infty} \frac{i^{2n+1}}{(2n+1)!} \cdot \vartheta^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \cdot \vartheta^{2n} + i \cdot \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \cdot \vartheta^{2n+1} \end{aligned}$$

e nell'ultima riga riconosciamo gli sviluppi in serie di Taylor di $\cos(\vartheta)$ e $\sin(\vartheta)$.

8.4 Equazioni polinomiali

Il passaggio da \mathbb{R} a \mathbb{C} è stato dettato dall'incapacità di risolvere su \mathbb{R} l'equazione polinomiale $x^2 + 1 = 0$. Ora anche su \mathbb{C} possiamo considerare equazioni polinomiali, ovvero definire $\mathbb{C}[z]$ come l'insieme delle espressioni del tipo

$a_0 + a_1 \cdot z + \dots + a_d \cdot z^d$ dove $a_i \in \mathbb{C}$ (con le ovvie convenzioni, usate anche nel caso reale e discusse nel Paragrafo 2.2), e considerare equazioni del tipo $p(z) = 0$ con $p(z) \in \mathbb{C}[z]$. In questo paragrafo studiamo queste equazioni, cominciando con la $z^n = 1$ per $n \in \mathbb{N}$, le cui soluzioni sono dette *radici n -esime dell'unità*.

Risolviamo l'equazione $z^n = 1$ cercando le soluzioni z nella forma polare $z = \rho \cdot e^{i\vartheta}$. L'equazione risulta allora $\rho^n \cdot e^{i n \vartheta} = 1$. Ne deduciamo che ρ^n è il modulo del numero complesso 1, cioè $\rho^n = 1$, dunque $\rho = 1$ (dato che $\rho = |z| \geq 0$). Inoltre $n\vartheta$ è un argomento del numero complesso 1, cioè $n\vartheta$ è della forma $2k\pi$ con $k \in \mathbb{Z}$. Pertanto le soluzioni della $z^n = 1$ sono tutti e soli i numeri complessi del tipo $e^{i \cdot \frac{2k\pi}{n}}$ con $k \in \mathbb{Z}$. In realtà questi numeri non sono tutti distinti tra loro, perché due valori di k che differiscano per un multiplo di n definiscono la stessa soluzione. Dunque l'equazione $z^n = 1$ ha esattamente n soluzioni distinte, date da

$$z = e^{i \cdot \frac{2k\pi}{n}}, \quad k = 0, 1, \dots, n-1.$$

È anche facile ora interpretare geometricamente l'insieme delle soluzioni trovate: si tratta dei vertici del poligono regolare con n lati, inscritto nella circonferenza di centro 0 e raggio 1 e avente il numero $1 \in \mathbb{C}$ come uno dei vertici.

Se ora modifichiamo leggermente l'equazione, considerandone una del tipo $z^n = w$ con $w \neq 0$, lo stesso argomento prova che le sue soluzioni sono i numeri

$$\sqrt[n]{|w|} \cdot e^{i \cdot \frac{\phi + 2k\pi}{n}}, \quad k = 0, 1, \dots, n-1$$

dove $\sqrt[n]{|w|}$ rappresenta l'unica radice n -esima reale positiva del numero reale positivo $|w|$, e ϕ è un argomento di w . Di nuovo abbiamo n soluzioni distinte, che rappresentano i vertici di un poligono regolare con n vertici centrato nell'origine.

Questa discussione mostra che le più facili equazioni polinomiali di grado n , cioè quelle in cui compare soltanto una costante e la potenza n -esima dell'incognita, hanno n soluzioni in \mathbb{C} . Vale in realtà il seguente risultato, detto *teorema fondamentale dell'algebra*, la cui dimostrazione è fuori dagli scopi di questo corso:

Teorema 8.4.1. *Ogni polinomio complesso di grado positivo ammette radici complesse.*

Ammettendo questo fatto, ci limitiamo qui a mostrare che, in un senso opportuno, le radici di un polinomio di grado n sono sempre nel numero di

n . Cominciamo con l'osservare che la *divisione* tra polinomi (comunemente eseguita con la regola di Ruffini) si estende senza variazioni dal caso reale al caso complesso. Vale cioè la seguente:

Proposizione 8.4.2. *Dati $p(z), f(z) \in \mathbb{C}[z]$ con $f(z) \neq 0$ esistono e sono unici $q(z), r(z) \in \mathbb{C}[z]$ tali che $\deg(r(z)) < \deg(f(z))$ e*

$$p(z) = q(z) \cdot f(z) + r(z).$$

Osserviamo che nel caso in cui $f(z)$ sia una costante non nulla f , la condizione $\deg(r(z)) < \deg(f) = 0$ significa che il *resto* $r(z)$ è il polinomio nullo (come ovvio). Abbiamo ora il facile:

Lemma 8.4.3. *Dati $p(z) \in \mathbb{C}[z]$ e $z_0 \in \mathbb{C}$, si ha $p(z_0) = 0$ se e soltanto se $p(z)$ è multiplo di $(z - z_0)$.*

Dimostrazione. Dividendo $p(z)$ per $(z - z_0)$ troviamo un resto di grado minore di $\deg(z - z_0) = 1$, dunque una costante r . Perciò $p(z) = q(z) \cdot (z - z_0) + r$. Valutando questa uguaglianza per $z = z_0$ si trova che $r = p(z_0)$, dunque immediatamente la conclusione. \square

Se $z_0 \in \mathbb{C}$ è una radice di $p(z) \in \mathbb{C}[z]$, dividendo $p(z)$ per la più alta potenza di $(z - z_0)$ per cui il resto risulti nullo, possiamo ora scrivere $p(z)$ come $(z - z_0)^k \cdot g(z)$, con $k \geq 1$ e $g(z) \in \mathbb{C}[z]$ tale che $g(z_0) \neq 0$. L'intero k è detto *molteplicità* della radice z_0 . Il teorema fondamentale dell'algebra sopra enunciato comporta ora facilmente il seguente più preciso:

Teorema 8.4.4. *Un polinomio complesso di grado $n \geq 1$ ha esattamente n radici complesse se contate con la loro molteplicità.*

8.5 Spazi vettoriali complessi

Uno *spazio vettoriale complesso* è un insieme V dotato di un elemento privilegiato 0 e di due operazioni

$$\begin{array}{ll} V \times V & \rightarrow V \\ (v, w) & \mapsto v + w \end{array} \qquad \begin{array}{ll} \mathbb{C} \times V & \rightarrow V \\ (\lambda, v) & \mapsto \lambda \cdot v \end{array}$$

che soddisfano esattamente gli stessi 8 assiomi degli spazi vettoriali reali, discussi nel Paragrafo 3.1. I più facili esempi di spazio vettoriale complesso sono del tutto analoghi a quelli illustrati nel caso reale: lo spazio \mathbb{C}^n dei vettori colonna numerici a n coordinate, lo spazio $\mathbb{C}[z]$ dei polinomi a coefficienti

complessi in un'indeterminata z , lo spazio $\mathcal{M}_{n \times m}(\mathbb{C})$ delle matrici $n \times m$ a coefficienti complessi, lo spazio $\mathcal{F}(S, \mathbb{C})$ delle funzioni a valori complessi definite su un insieme S .

L'intera teoria finora sviluppata nel caso reale, dipendendo solo dalle proprietà di campo di \mathbb{R} , si estende senza alcuna variazione al caso di \mathbb{C} . In particolare, è possibile definire per ogni spazio vettoriale complesso V che ammetta insiemi finiti di generatori una *dimensione complessa* indicata con $\dim_{\mathbb{C}}(V)$. Discutiamo ora la ragione per la quale sia importante, parlando di dimensione, specificare se essa sia intesa in senso reale oppure complesso (cioè se si stia considerando $\dim_{\mathbb{R}}$ oppure $\dim_{\mathbb{C}}$).

Cominciamo con l'osservare che \mathbb{C} è il più facile esempio di spazio vettoriale complesso, con base costituita dal numero 1 e dunque avente dimensione complessa 1. Tuttavia \mathbb{C} , essendo identificato a \mathbb{R}^2 , è anche uno spazio vettoriale reale, e la sua dimensione reale è 2 (con base data dai numeri 1 e i). Più in generale, se V è uno spazio vettoriale complesso, osservando che $\mathbb{R} \subset \mathbb{C}$, si può dare a V anche una struttura di spazio vettoriale reale: basta considerare la medesima operazione di somma tra vettori e limitare l'operazione di prodotto tra scalari e vettori a quegli scalari che siano numeri reali, cioè complessi con parte immaginaria nulla. Vale ora la seguente:

Proposizione 8.5.1. *Se V è uno spazio vettoriale complesso di dimensione finita si ha $\dim_{\mathbb{R}}(V) = 2 \cdot \dim_{\mathbb{C}}(V)$.*

Dimostrazione. Scelta una base v_1, \dots, v_n di V su \mathbb{C} , si verifica subito che $v_1, i \cdot v_1, \dots, v_n, i \cdot v_n$ è una base di V su \mathbb{R} . \square

Sempre utilizzando il fatto che uno spazio vettoriale complesso è anche uno spazio vettoriale reale possiamo ora dire \mathbb{C} -lineare un'applicazione che sia lineare rispetto alla struttura complessa, e \mathbb{R} -lineare una che lo sia rispetto alla struttura reale. Abbiamo la seguente immediata:

Proposizione 8.5.2. *Siano V e W spazi vettoriali su \mathbb{C} e sia $f : V \rightarrow W$ un'applicazione. Allora:*

- *La moltiplicazione per lo scalare i definisce un'applicazione \mathbb{C} -lineare $J_V : V \rightarrow V$, e similmente una $J_W : W \rightarrow W$.*
- *La f è \mathbb{C} -lineare se e soltanto se è \mathbb{R} -lineare e $f \circ J_V = J_W \circ f$.*

Concludiamo osservando che una matrice reale $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ è anche una matrice complessa, dunque è possibile considerarne il rango sia in senso reale sia in senso complesso, e similmente per la dimensione del nucleo. Dal teorema degli orlati segue ora facilmente la seguente:

Proposizione 8.5.3. *Se $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ si ha*

$$\text{rank}_{\mathbb{R}}(A) = \text{rank}_{\mathbb{C}}(A), \quad \dim_{\mathbb{R}}(\text{Ker}(A)) = \dim_{\mathbb{C}}(\text{Ker}(A)).$$

Capitolo 9

Prodotti scalari e hermitiani

Ci proponiamo in questo capitolo di estendere agli spazi vettoriali alcune delle nozioni consuete della geometria piana elementare. Cominciamo considerando le nozioni di *distanza* (tra punti) e di *angolo* (in particolare, di *ortogonalità*, tra vettori o rette) e notando che esse sono naturalmente legate fra loro. Infatti le lunghezze dei lati di un triangolo determinano i suoi angoli. Inoltre, ad esempio, il segmento che realizza la minima distanza di un punto p da una retta r è quello contenuto nella retta ortogonale a r e passante per p . Illustriamo ora il legame tra le due nozioni ricavando le formule esplicite, e prendendo spunto da queste formule per una definizione che sarà il fulcro del capitolo.

Notiamo intanto che per il teorema di Pitagora la distanza $d(0, x)$ tra 0 e un punto x di \mathbb{R}^2 è data da $\sqrt{x_1^2 + x_2^2}$. Ci chiediamo quanto misuri l'ampiezza ϑ dell'angolo formato nell'origine da due rette generate rispettivamente da vettori x e y non nulli. A questo fine ricordiamo che per il teorema di Carnot si ha

$$\cos \vartheta = \frac{d(0, x)^2 + d(0, y)^2 - d(x, y)^2}{2d(0, x) \cdot d(0, y)}.$$

Per applicare questa formula dobbiamo ancora ricordare che la distanza $d(x, y)$ tra x e y è la lunghezza del segmento che li congiunge, ovvero del vettore $x - y$. Abbiamo dunque

$$\begin{aligned}d(0, x) &= \sqrt{x_1^2 + x_2^2}, \\d(0, y) &= \sqrt{y_1^2 + y_2^2}, \\d(x, y) &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.\end{aligned}$$

Sostituendo nella formula per $\cos \vartheta$ otteniamo ora

$$\begin{aligned} \cos \vartheta &= \frac{x_1^2 + x_2^2 + y_1^2 + y_2^2 - (x_1^2 - 2x_1y_1 + y_1^2 + x_2^2 - 2x_2y_2 + y_2^2)}{2\sqrt{x_1^2 + x_2^2} \cdot \sqrt{y_1^2 + y_2^2}} \\ &= \frac{x_1y_1 + x_2y_2}{\sqrt{x_1^2 + x_2^2} \cdot \sqrt{y_1^2 + y_2^2}}. \end{aligned}$$

Introduciamo allora la funzione $\langle \cdot | \cdot \rangle_{\mathbb{R}^2} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ data da

$$\langle x|y \rangle_{\mathbb{R}^2} = x_1 \cdot y_1 + x_2 \cdot y_2 \quad ,$$

che chiameremo *prodotto scalare*. Abbiamo che:

- La distanza tra x e 0 è data da

$$d(0, x) = \sqrt{\langle x|x \rangle_{\mathbb{R}^2}};$$

- L'ampiezza ϑ dell'angolo formato nell'origine da due rette generate da vettori x e y non nulli soddisfa l'uguaglianza

$$\cos \vartheta = \frac{\langle x|y \rangle_{\mathbb{R}^2}}{\sqrt{\langle x|x \rangle_{\mathbb{R}^2}} \cdot \sqrt{\langle y|y \rangle_{\mathbb{R}^2}}}.$$

Osserviamo che dalla seconda proprietà segue in particolare che le rette $\text{Span}(x)$ e $\text{Span}(y)$ sono *perpendicolari* (o *ortogonali*) tra loro, relazione che esprimiamo in formula come $\text{Span}(x) \perp \text{Span}(y)$, se e solo se $\langle x|y \rangle_{\mathbb{R}^2} = 0$. Dalla prima proprietà segue invece che

$$d(x, y) = \sqrt{\langle x - y|x - y \rangle_{\mathbb{R}^2}}.$$

Definendo ora la *norma* $\|x\|_{\mathbb{R}^2}$ di un vettore $x \in \mathbb{R}^2$ come il numero non negativo $\sqrt{\langle x|x \rangle_{\mathbb{R}^2}}$, possiamo ora riassumere quanto scoperto con le relazioni

$$\begin{aligned} d(x, y) &= \|x - y\|_{\mathbb{R}^2} \\ \cos(x, y) &= \frac{\langle x|y \rangle_{\mathbb{R}^2}}{\|x\|_{\mathbb{R}^2} \cdot \|y\|_{\mathbb{R}^2}} \end{aligned}$$

dove come al solito ϑ è l'angolo tra $\text{Span}(x)$ e $\text{Span}(y)$ se x e y sono non nulli.

Le considerazioni svolte finora nel caso del piano si trasportano senza variazioni al caso di \mathbb{R}^n , ponendo

$$\begin{aligned} \langle x|y \rangle_{\mathbb{R}^n} &= x_1 \cdot y_1 + \dots + x_n \cdot y_n = {}^t x \cdot y, \\ \|x\|_{\mathbb{R}^n} &= \sqrt{\langle x|x \rangle_{\mathbb{R}^n}} \in \mathbb{R}_+ . \end{aligned}$$

Con il primo paragrafo del capitolo estendiamo la nozione di prodotto scalare al caso di uno spazio vettoriale qualsiasi. Nel seguito stabiliamo molte utili proprietà del prodotto scalare, e lo estendiamo anche al caso complesso.

9.1 Prodotti scalari

Sia V uno spazio vettoriale *reale*. Una applicazione $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{R}$ si dice:

- *bilineare* se

$$\langle \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 | v_3 \rangle = \lambda_1 \cdot \langle v_1 | v_3 \rangle + \lambda_2 \cdot \langle v_2 | v_3 \rangle$$

$$\langle v_3 | \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 \rangle = \lambda_1 \cdot \langle v_3 | v_1 \rangle + \lambda_2 \cdot \langle v_3 | v_2 \rangle$$

per ogni $v_1, v_2, v_3 \in V$ e $\lambda_1, \lambda_2 \in \mathbb{R}$;

- *simmetrica* se $\langle v_1 | v_2 \rangle = \langle v_2 | v_1 \rangle$ per ogni $v_1, v_2 \in V$;
- *definita positiva* se $\langle v | v \rangle > 0$ per $v \neq 0$;
- un *prodotto scalare* su V se è bilineare, simmetrica e definita positiva.

Dalle proprietà del prodotto righe per colonne è facile dedurre che l'applicazione $\langle \cdot | \cdot \rangle_{\mathbb{R}^n}$ sopra definita è un prodotto scalare su \mathbb{R}^n , che sarà chiamato *prodotto scalare canonico* di \mathbb{R}^n .

Ricordando che $\langle x | y \rangle_{\mathbb{R}^n}$ è definito da ${}^t x \cdot y$, possiamo ora estendere la definizione e introdurre un prodotto scalare su $\mathcal{M}_{m \times n}(\mathbb{R})$. Poniamo cioè

$$\langle A | B \rangle = \text{tr}({}^t A \cdot B).$$

La verifica della bilinearità di questa applicazione è molto semplice: basta usare la linearità del prodotto righe per colonne in ciascuno dei suoi argomenti, la linearità della trasposizione e la linearità della traccia. Che l'applicazione sia definita positiva è semplice:

$$\langle A | A \rangle = \sum_{i=1}^n ({}^t A \cdot A)_{ii} = \sum_{i=1}^n \sum_{j=1}^m ({}^t A)_{ij} \cdot (A)_{ji} = \sum_{j=1}^m \sum_{i=1}^n ((A)_{ji})^2$$

e quest'ultima somma ha risultato positivo se $A \neq 0$. La simmetria segue invece dal fatto (ovvio) che una matrice quadrata e la sua trasposta hanno la stessa traccia, e dal seguente più generale:

Proposizione 9.1.1. $M \in \mathcal{M}_{p \times q}(\mathbb{R}), N \in \mathcal{M}_{q \times r}(\mathbb{R}) \Rightarrow {}^t(M \cdot N) = {}^tN \cdot {}^tM$.

Dimostrazione.

$$\begin{aligned} ({}^t(M \cdot N))_{ij} &= (M \cdot N)_{ji} = \sum_{s=1}^q (M)_{js} \cdot (N)_{si} \\ &= \sum_{s=1}^q ({}^tN)_{is} \cdot ({}^tM)_{sj} = ({}^tN \cdot {}^tM)_{ij} . \end{aligned}$$

□

Tornando a \mathbb{R}^n , caratterizziamo ora le applicazioni bilineari simmetriche. Cominciamo col definire, fissata $A \in \mathcal{M}_{n \times n}(\mathbb{R})$, l'applicazione $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ data da $\langle x|y \rangle_A = {}^tx \cdot A \cdot y$. È immediato verificare che tale applicazione è bilineare. Inoltre:

Proposizione 9.1.2. *Data $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ bilineare*

- *esiste ed è unica $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ tale che $f = \langle \cdot | \cdot \rangle_A$;*
- *f è simmetrica se e solo se A è una matrice simmetrica.*

Dimostrazione. Cominciamo con l'unicità di A . Se $f = \langle \cdot | \cdot \rangle_A$ e $A = (a_{ij})$, allora $a_{ij} = f(e_i, e_j)$, dunque in effetti la A è univocamente determinata da f . Per l'esistenza si tratta ora di vedere che, posto $a_{ij} = f(e_i, e_j)$ e definita A come la matrice (a_{ij}) , si ha $f(x, y) = \langle x|y \rangle_A$ per ogni $x, y \in \mathbb{R}^n$. Infatti:

$$\begin{aligned} f(x, y) &= f\left(\sum_{i=1}^n x_i \cdot e_i, \sum_{j=1}^n y_j \cdot e_j\right) \\ &= \sum_{i=1}^n x_i \cdot \left(\sum_{j=1}^n a_{ij} \cdot y_j\right) = \sum_{i=1}^n x_i \cdot (A \cdot y)_i = {}^tx \cdot A \cdot y = \langle x|y \rangle_A . \end{aligned}$$

Passando alla seconda asserzione notiamo che, essendo ${}^ty \cdot A \cdot x$ è un numero, si ha

$${}^ty \cdot A \cdot x = {}^t({}^ty \cdot A \cdot x) = {}^tx \cdot {}^tA \cdot y ,$$

dunque f è simmetrica se e soltanto se ${}^tx \cdot (A - {}^tA) \cdot y = 0$ per ogni $x, y \in \mathbb{R}^n$. Per $x = e_i$ e $y = e_j$ questa condizione significa che $a_{ij} = a_{ji}$, onde facilmente la conclusione. □

Rimandiamo al Capitolo 10 la discussione completa di quali condizioni una matrice A simmetrica debba soddisfare per dare luogo a un'applicazione $\langle \cdot | \cdot \rangle_A$ che sia definita positiva. Ci limitiamo qui a osservare quanto segue:

- Se A è diagonale con coefficienti $\lambda_1, \dots, \lambda_n$ sulla diagonale, allora $\langle \cdot | \cdot \rangle_A$ è definito positivo se e soltanto se ciascuno dei λ_i è positivo;
- Se $A = {}^t M \cdot M$ con $M \in \mathcal{M}_{n \times n}(\mathbb{R})$ invertibile, allora

$$\langle x | x \rangle_A = {}^t x \cdot {}^t M \cdot M \cdot x = \langle M \cdot x | M \cdot x \rangle_{\mathbb{R}^n}$$

e questo numero è positivo poiché $M \cdot x \neq 0$. Ne segue che in questo caso $\langle \cdot | \cdot \rangle_A$ è definita positiva.

Un esempio di prodotto scalare molto importante per le applicazioni è quello sullo spazio delle funzioni continue da $[0, 1]$ a \mathbb{R} , definito come

$$\langle f | g \rangle = \int_0^1 f(t)g(t) dt.$$

9.2 Norma e ortogonalità

Fissiamo in questo paragrafo un prodotto scalare $\langle \cdot | \cdot \rangle$ su uno spazio vettoriale reale V . Poiché si ha $\langle v | v \rangle > 0$ per $v \neq 0$ e naturalmente $\langle 0 | 0 \rangle = 0$, possiamo definire la *norma* di v , indicandola con $\|v\|$, come la radice quadrata non negativa di $\langle v | v \rangle$. Osserviamo intanto che dalla sola norma è possibile risalire a tutti i valori del prodotto scalare. Infatti:

$$\langle v | w \rangle = \frac{1}{2} (\|v + w\|^2 - \|v\|^2 - \|w\|^2).$$

Abbiamo inoltre le seguenti elementari proprietà della norma:

- $\|v\| > 0$ se $v \neq 0$ e $\|0\| = 0$;
- $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$.

Meno elementare è invece la seguente *disuguaglianza di Cauchy-Schwarz*:

Proposizione 9.2.1. *Per ogni $v, w \in V$ si ha*

$$|\langle v | w \rangle| \leq \|v\| \cdot \|w\|$$

e l'uguaglianza vale se e solo se v e w sono proporzionali tra loro.

Dimostrazione. Nel caso in cui v e w siano proporzionali tra loro, dunque in particolare nel caso in cui uno tra v e w sia nullo, l'uguaglianza è ovvia. Supponiamoli allora non proporzionali e al variare di t in \mathbb{R} consideriamo il vettore $t \cdot v + w$. Esso è sempre non nullo, dunque $\|t \cdot v + w\|^2 > 0$ per ogni t . Eseguendo il calcolo si trova che $\|v\|^2 \cdot t^2 + 2\langle v|w \rangle \cdot t + \|w\|^2 > 0$ per ogni t . Ora, un polinomio reale di grado 2 non si annulla mai se e soltanto se il suo discriminante è negativo. Troviamo dunque la disuguaglianza

$$\langle v|w \rangle^2 - \|v\|^2 \cdot \|w\|^2 < 0$$

e la conclusione segue subito. \square

Dal risultato precedente deduciamo ora la *disuguaglianza triangolare*, già incontrata e interpretata geometricamente nel Paragrafo 8.2:

Teorema 9.2.2. *Per ogni $v, w \in V$ si ha*

$$\|v + w\| \leq \|v\| + \|w\|$$

e l'uguaglianza vale solo se uno tra v e w si ottiene dall'altro tramite moltiplicazione per uno scalare non negativo.

Dimostrazione. Se v e w sono proporzionali tra loro la conclusione è molto semplice. Se invece non lo sono si ha

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + 2\langle v|w \rangle + \|w\|^2 \\ &< \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

\square

Le proprietà della norma comportano che il numero $\|v\|$ può essere sensatamente interpretato come la *distanza* di v dall'origine 0. Più in generale, ponendo $d(v, w) = \|v - w\|$, si verificano le seguenti ragionevoli proprietà per una funzione $V \times V \rightarrow \mathbb{R}$ che rappresenti la distanza tra due punti:

- $d(v, w) > 0$ se $v \neq w$, e $d(v, v) = 0$;
- $d(v, w) = d(w, v)$;
- $d(v, w) \leq d(v, z) + d(z, w)$.

L'unica proprietà non evidente è l'ultima, che verifichiamo come segue:

$$\begin{aligned} d(v, w) &= \|v - w\| = \|(v - z) + (z - w)\| \\ &\leq \|v - z\| + \|z - w\| = d(v, z) + d(z, w). \end{aligned}$$

Introduciamo ora alcuni nuovi termini. Un vettore u di V sarà detto *unitario* se $\|u\| = 1$. Due vettori v, w saranno detti *ortogonali* se $\langle v|w \rangle = 0$. Più in generale, un sistema v_1, \dots, v_k di vettori sarà detto *ortogonale* se $\langle v_i|v_j \rangle = 0$ per $i \neq j$, e sarà detto *ortonormale* se inoltre tutti i v_i sono unitari, ovvero se $\langle v_i|v_j \rangle = \delta_{ij}$ per ogni i, j .

Proposizione 9.2.3. *Un sistema ortogonale di vettori non nulli è linearmente indipendente.*

Dimostrazione. Siano v_1, \dots, v_k i vettori, e supponiamo che $\alpha_1 \cdot v_1 + \dots + \alpha_k \cdot v_k$ sia nullo. Allora per ogni j si ha

$$0 = \langle 0|v_j \rangle = \langle \alpha_1 \cdot v_1 + \dots + \alpha_k \cdot v_k | v_j \rangle = \alpha_j \cdot \|v_j\|^2.$$

Essendo $v_j \neq 0$ si ha $\|v_j\|^2 > 0$, onde $\alpha_j = 0$. □

Da questo risultato deduciamo che, se $\dim_{\mathbb{R}}(V) = n$, un sistema ortogonale di n vettori non nulli è una base. Il prossimo fondamentale risultato illustra come trovare le coordinate di un vettore rispetto a una base siffatta:

Teorema 9.2.4. *Se (v_1, \dots, v_n) è una base ortogonale di V e $v \in V$ allora*

$$v = \sum_{i=1}^n \frac{\langle v|v_i \rangle}{\|v_i\|^2} \cdot v_i.$$

Dimostrazione. Cominciamo notando che se $y \in V$ e $\langle y|v_j \rangle = 0$ per ogni j allora $y = 0$. Infatti $y = y_1 \cdot v_1 + \dots + y_n \cdot v_n$ per coordinate y_1, \dots, y_n opportune, e $\langle y|v_j \rangle = y_j \cdot \|v_j\|^2$ con $\|v_j\|^2 > 0$.

Posto allora $y = v - \sum_{i=1}^n \frac{\langle v|v_i \rangle}{\|v_i\|^2} \cdot v_i$ basta vedere che $\langle y|v_j \rangle = 0$ per ogni j , e questo è ovvio. □

L'importanza (anche ai fini delle applicazioni) del risultato appena dimostrato è dovuta alla seguente osservazione (che dall'enunciato segue immediatamente): *la coordinata j -esima di un vettore rispetto a una base ortogonale dipende solo dal j -esimo vettore della base, non dagli altri.* Notiamo che ciò è

falso in generale: cambiando anche un solo vettore di una base non ortogonale, le coordinate possono cambiare tutte.

Ci poniamo ora il problema di come ottenere basi ortogonali. Prima di tutto osserviamo che una base ortogonale (v_1, \dots, v_n) può essere facilmente sostituita da una base ortonormale, *normalizzando* ciascun vettore v_i , cioè sostituendolo con $\frac{v_i}{\|v_i\|}$. Inoltre se (v_1, \dots, v_n) è una base ortonormale, le coordinate rispetto a essa si trovano in modo ancora più facile: la j -esima coordinata di $v \in V$ è semplicemente $\langle v | v_j \rangle$. Il prossimo risultato, che descrive il cosiddetto *procedimento di ortonormalizzazione di Gram-Schmidt*, prova che ogni base può essere trasformata in una base ortonormale:

Teorema 9.2.5. *Se (x_1, \dots, x_n) è una base di V , il seguente procedimento ricorsivo produce una base ortonormale (v_1, \dots, v_n) :*

- Si pone $v_1 = \frac{x_1}{\|x_1\|}$;
- Se $1 \leq k < n$, supposti costruiti v_1, \dots, v_k , si pone

$$y_{k+1} = x_{k+1} - \sum_{i=1}^k \langle x_{k+1} | v_i \rangle \cdot v_i$$

$$\text{e } v_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|}.$$

Inoltre per ogni k si ha $\text{Span}(v_1, \dots, v_k) = \text{Span}(x_1, \dots, x_k)$.

Dimostrazione. Uno dei punti chiave dell'enunciato è espresso solo in modo tacito, ma conviene esplicitarlo. Dire che il "procedimento ricorsivo produce una base ortonormale" significa in particolare che esso può essere eseguito fino al passo n -esimo, cioè che i vettori y_2, \dots, y_n successivamente costruiti sono non nulli, dunque è lecito normalizzarli. Per avere la dimostrazione del teorema basterà ora provare per induzione su $k = 1, \dots, n$ l'asserto seguente:

- i vettori v_1, \dots, v_k possono essere costruiti, sono ortonormali, e generano $\text{Span}(x_1, \dots, x_k)$.

L'asserto è ovvio per $k = 1$. Supponiamolo vero per qualche $k < n$ e dimostriamolo per $k + 1$. Poiché

$$\sum_{i=1}^k \langle x_{k+1} | v_i \rangle \cdot v_i \in \text{Span}(v_1, \dots, v_k) = \text{Span}(x_1, \dots, x_k)$$

e $x_{k+1} \notin \text{Span}(x_1, \dots, x_k)$, si ha $y_{k+1} \neq 0$, dunque v_{k+1} può essere definito. Inoltre per $j \leq k$ si ha:

$$\begin{aligned} \langle y_{k+1} | v_j \rangle &= \langle x_{k+1} | v_j \rangle - \sum_{i=1}^k \langle x_{k+1} | v_i \rangle \cdot \langle v_i | v_j \rangle \\ &= \langle x_{k+1} | v_j \rangle - \sum_{i=1}^k \delta_{ij} \cdot \langle x_{k+1} | v_i \rangle = 0 . \end{aligned}$$

Ne segue immediatamente che v_1, \dots, v_k, v_{k+1} è un sistema ortonormale. Resta da vedere che $\text{Span}(v_1, \dots, v_k, v_{k+1}) = \text{Span}(x_1, \dots, x_k, x_{k+1})$. Sapendo che $\text{Span}(v_1, \dots, v_k) = \text{Span}(x_1, \dots, x_k)$, basta allora vedere che

$$v_{k+1} \in \text{Span}(v_1, \dots, v_k, x_{k+1}), \quad v_{k+1} \notin \text{Span}(v_1, \dots, v_k).$$

Entrambi questi fatti sono evidenti, ed il teorema è dimostrato. \square

Analizzando il procedimento di ortonormalizzazione ora descritto, è facile provare che la matrice di cambiamento di base dalla (x_1, \dots, x_k) alla sua ortonormalizzata (v_1, \dots, v_k) è triangolare superiore (cioè ha coefficienti nulli sotto la diagonale principale) con coefficienti positivi sulla diagonale principale. Abbiamo anche il seguente:

Corollario 9.2.6. *Un sistema ortonormale di vettori può essere completato a una base ortonormale.*

Dimostrazione. Prima si completa a una base qualsiasi, poi si ortonormalizza. \square

9.3 Decomposizioni ortogonali e proiezioni

Fissiamo in questo paragrafo uno spazio vettoriale V su \mathbb{R} con prodotto scalare $\langle \cdot | \cdot \rangle$. Se $S \subset V$ è un sottoinsieme qualsiasi, poniamo

$$S^\perp = \{v \in V : \langle v | s \rangle = 0 \ \forall s \in S\}.$$

È facile vedere che S^\perp è sempre un sottospazio vettoriale, che chiameremo *sottospazio ortogonale* a S . Se $v \in V$, indicheremo $\{v\}^\perp$ semplicemente con v^\perp .

Proposizione 9.3.1. *Se V ha dimensione finita e W è un sottospazio di V si ha $V = W \oplus W^\perp$. Inoltre $\dim_{\mathbb{R}}(W^\perp) = \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(W)$. Infine $(W^\perp)^\perp = W$.*

Dimostrazione. È chiaro che $W \cap W^\perp = \{0\}$. Per vedere che $V = W + W^\perp$, scegliamo una base x_1, \dots, x_n di V con x_1, \dots, x_k base di W , e ortonormalizziamola ottenendo v_1, \dots, v_n . Allora $v_1, \dots, v_k \in W$ e $v_{k+1}, \dots, v_n \in W^\perp$, dunque la conclusione. La seconda affermazione è ora ovvia e la terza segue osservando che $(W^\perp)^\perp$ è un sottospazio che contiene W e ha la medesima dimensione di W . \square

La proiezione su W (si veda il Paragrafo 5.2) associata alla decomposizione in somma diretta $V = W \oplus W^\perp$ è detta *proiezione ortogonale* di V su W e indicata con p_W . Il prossimo risultato comporta in particolare che p_W dipende in effetti solo da W (mentre sappiamo che la proiezione su W associata ad una decomposizione $V = W \oplus Z$ qualsiasi dipende anche da Z).

9.3.1 Rudimenti di calcolo differenziale in più variabili

Tra poco, a proposito delle proiezioni ortogonali, avremo bisogno di alcune nozioni di natura analitica, di cui ci sarà necessità altre volte (ma sporadicamente) nel resto di questo testo. Per questa ragione le riassumiamo brevemente in questo sottoparagrafo. Lo facciamo in modo informale e senza presentare dimostrazioni.

Definizione 9.3.2. Un sottoinsieme $\Omega \subset \mathbb{R}^n$ è *aperto* se per ogni $x \in \Omega$ esiste $r > 0$ tale che $\{y \in \mathbb{R}^n : \|y - x\| < r\} \subset \Omega$. Dati $\Omega \subset \mathbb{R}^n$ aperto e $f : \Omega \rightarrow \mathbb{R}$ chiamiamo *derivata parziale j -esima* di f in un punto $\bar{x} \in \Omega$ la derivata in \bar{x} di f (se esiste) ottenuta considerandola come funzione della sola x_j , con le altre variabili viste come parametri, ovvero

$$\frac{\partial f}{\partial x_j}(\bar{x}) = \lim_{h \rightarrow 0} \frac{f(\bar{x} + h e_j) - f(\bar{x})}{h}.$$

Esempio 9.3.3. Se $f(x, y) = \ln(1 + x^2 \cdot e^y) \cdot \cos(e^x \cdot y^3)$ allora

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{2x \cdot e^y}{1 + x^2 \cdot e^y} \cdot \cos(e^x \cdot y^3) + \\ &\quad - \ln(1 + x^2 \cdot e^y) \cdot e^x \cdot y^3 \cdot \sin(e^x \cdot y^3) \\ \frac{\partial f}{\partial y}(x, y) &= \frac{x^2 \cdot e^y}{1 + x^2 \cdot e^y} \cdot \cos(e^x \cdot y^3) + \\ &\quad - \ln(1 + x^2 \cdot e^y) \cdot e^x \cdot 3y^2 \cdot \sin(e^x \cdot y^3). \end{aligned}$$

Definizione 9.3.4. Chiamiamo *gradiente* di f in \bar{x} il vettore di \mathbb{R}^n che ha come componenti le derivate parziali di f in \bar{x} :

$$\text{grad}_{\bar{x}}(f) = \left(\frac{\partial f}{\partial x_j}(\bar{x}) \right)_{j=1, \dots, n} \in \mathbb{R}^n.$$

Per il calcolo delle derivate parziali (e dunque del gradiente) è spesso utile la seguente versione in più variabili della regola di derivazione di funzione composta, che per semplicità enunciamo senza specificare i domini delle funzioni coinvolte:

Proposizione 9.3.5. *Date funzioni reali f_1, \dots, f_k di n variabili e una funzione reale g di k variabili si ha*

$$\frac{\partial g(f_1, \dots, f_k)}{\partial x_i}(x) = \sum_{j=1}^k \frac{\partial g}{\partial y_j}(f_1(x), \dots, f_k(x)) \cdot \frac{\partial f_j}{\partial x_i}(x).$$

Passiamo ora alle derivate seconde:

Definizione 9.3.6. Chiamiamo *derivata parziale seconda* (i, j) -esima di f in \bar{x} la derivata parziale i -esima della derivata parziale j -esima

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) = \left(\frac{\partial f}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) \right) (\bar{x}),$$

con $\frac{\partial^2 f}{\partial x_i \partial x_i}$ indicata con $\frac{\partial^2 f}{\partial x_i^2}$. Infine chiamiamo *matrice hessiana* di f in \bar{x} la matrice $n \times n$ che ha come coefficienti le derivate parziali seconde di f in \bar{x} :

$$H_{\bar{x}}(f) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) \right)_{i, j=1, \dots, n} \in \mathcal{M}_{n \times n}(\mathbb{R}).$$

Il seguente risultato ha fondamentale importanza, e la dimostrazione non è immediata:

Proposizione 9.3.7. *Se $f : \Omega \rightarrow \mathbb{R}$ ha tutte le derivate parziali seconde continue su Ω allora calcolando le derivate parziali seconde miste non importa l'ordine con cui le due derivate vengono eseguite, ovvero*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

per ogni $i, j = 1, \dots, n$ e ogni $x \in \Omega$. Equivalentemente, la matrice $H_x(f)$ è simmetrica per ogni $x \in \Omega$.

Enunciamo ora la versione in più variabili della formula di approssimazione di Taylor del secondo ordine, che è precisamente tutto quello di cui avremo bisogno in questo testo:

Proposizione 9.3.8. *Dati $\Omega \subset \mathbb{R}^n$ aperto, $\bar{x} \in \Omega$ e $f : \Omega \rightarrow \mathbb{R}$ avente tutte le derivate parziali seconde continue, si ha*

$$\begin{aligned} f(\bar{x} + v) &= f(\bar{x}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\bar{x}) \cdot v_i + \frac{1}{2} \cdot \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) \cdot v_i \cdot v_j + o(\|v\|^2) \\ &= f(\bar{x}) + {}^t \text{grad}_{\bar{x}}(f) \cdot v + \frac{1}{2} \cdot {}^t v \cdot H_{\bar{x}}(f) \cdot v + o(\|v\|^2) \\ &= f(\bar{x}) + \langle \text{grad}_{\bar{x}}(f) | v \rangle_{\mathbb{R}^n} + \frac{1}{2} \langle v | v \rangle_{H_{\bar{x}}(f)} + o(\|v\|^2). \end{aligned}$$

Osserviamo che in questa catena di uguaglianze il terzo e il quarto membro rappresentano soltanto diverse grafie per esprimere il secondo membro: la formula effettivamente rilevante è l'uguaglianza tra il primo e il secondo membro. Pur senza fornire una dimostrazione completa della formula, la motiviamo ponendo $u = \frac{v}{\|v\|}$ e $t = \|v\|$. Facendo finta che u sia fisso e che a variare sia soltanto t , possiamo porre $g(t) = f(\bar{x} + tu)$ e osservare che per definizione $f(\bar{x} + v) = g(t)$. Dunque applicando la formula di approssimazione di Taylor nel caso di funzioni di una variabile abbiamo

$$g(t) = g(0) + g'(0) \cdot t + \frac{1}{2} g''(0) \cdot t^2 + o(t^2).$$

Abbiamo ora ovviamente $g(0) = f(\bar{x})$; inoltre

$$\begin{aligned} g'(t) &= \lim_{s \rightarrow 0} \frac{g(t+s) - g(t)}{s} \\ &= \lim_{s \rightarrow 0} \frac{f(\bar{x}_1 + (t+s)u_1, \dots, \bar{x}_n + (t+s)u_n) - f(\bar{x} + tu)}{t} \\ &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\bar{x} + tu) \cdot u_j \end{aligned}$$

(nell'ultimo passaggio abbiamo impiegato la Proposizione 9.3.5); ne segue che

$$g'(0) \cdot t = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\bar{x}) \cdot tu_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\bar{x}) \cdot v_j.$$

Con lo stesso metodo usato per calcolare $g'(t)$ otteniamo anche

$$g''(0) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) u_i u_j$$

da cui

$$\frac{1}{2} g''(0) \cdot t^2 = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) (t u_i)(t u_j) = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) v_i v_j.$$

A questo punto resta solo da notare che $t = \|v\|$, dunque $o(t^2) = o(\|v\|^2)$.

Prima di poter trarre conseguenze importanti dall'intera formula di approssimazione di Taylor del secondo ordine dovremo sviluppare notevolmente la teoria (lo faremo nel Capitolo 10, in particolare con la Proposizione 10.2.6); per ora ci limitiamo alla seguente immediata conseguenza dell'approssimazione al prim'ordine:

Proposizione 9.3.9. *Se \bar{x} è un punto di massimo o di minimo locale per f allora $\text{grad}_{\bar{x}}(f) = 0$.*

9.3.2 Proiezioni ortogonali

Grazie alle nozioni di analisi sviluppate nel sottoparagrafo precedente possiamo ora fornire un'importante caratterizzazione geometrica delle proiezioni ortogonali. Torniamo come in precedenza a considerare fissato uno spazio vettoriale V su \mathbb{R} con prodotto scalare $\langle \cdot | \cdot \rangle$, e supponiamo inoltre che V abbia dimensione finita.

Proposizione 9.3.10. *Se W un sottospazio di V e v_1, \dots, v_k è una base ortogonale di W allora*

$$p_W(v) = \sum_{i=1}^k \frac{\langle v | v_i \rangle}{\|v_i\|^2} \cdot v_i.$$

Inoltre $p_W(v)$ è il punto di W avente minima distanza da v e vale la seguente disuguaglianza di Bessel:

$$\sum_{i=1}^k \frac{|\langle v | v_i \rangle|^2}{\|v_i\|^2} \leq \|v\|^2.$$

Dimostrazione. Completando v_1, \dots, v_k a una base ortogonale v_1, \dots, v_n di V si ha

$$v = \sum_{i=1}^k \frac{\langle v|v_i \rangle}{\|v_i\|^2} \cdot v_i + \sum_{i=k+1}^n \frac{\langle v|v_i \rangle}{\|v_i\|^2} \cdot v_i.$$

Il primo di questi vettori appartiene a W e il secondo a W^\perp , onde la conclusione per definizione di proiezione associata alla decomposizione.

Per provare la seconda affermazione dobbiamo vedere che al variare di $t \in \mathbb{R}^k$ il minimo della funzione

$$f(t) = \|v - (t_1 \cdot v_1 + \dots + t_k \cdot v_k)\|$$

è assunto precisamente per $t_i = \frac{\langle v|v_i \rangle}{\|v_i\|^2}$. Poiché la f assume valori non negativi e l'elevamento al quadrato è una funzione monotona crescente su $[0, +\infty)$, basta verificare la stessa affermazione per la funzione g data da $g(t) = f(t)^2$. È chiaro che $g(t) \rightarrow +\infty$ per $\|t\| \rightarrow +\infty$, dunque grazie alla Proposizione 9.3.9 basta verificare che il solo punto nel quale le derivate parziali di g si annullano è quello con $t_i = \frac{\langle v|v_i \rangle}{\|v_i\|^2}$. Ora

$$g(t) = \|v\|^2 - \sum_{i=1}^k 2 \langle v|v_i \rangle \cdot t_i + \sum_{i=1}^k \|v_i\|^2 \cdot t_i^2$$

e calcolando $\frac{\partial}{\partial t_j}$ si trova $-2 \langle v|v_j \rangle + 2\|v_j\|^2 \cdot t_j$, onde la conclusione.

Per l'ultima affermazione, notiamo che

$$v = p_W(v) + p_{W^\perp}(v), \quad \langle p_W(v)|p_{W^\perp}(v) \rangle = 0 \quad ,$$

onde $\|v\|^2 = \|p_W(v)\|^2 + \|p_{W^\perp}(v)\|^2 \geq \|p_W(v)\|^2$ e la tesi segue dall'espressione trovata per $p_W(v)$. \square

9.3.3 Rette e piani nello spazio

Alla luce di quanto appena visto possiamo ora meglio interpretare alcuni dei fatti visti nel Paragrafo 7.2. Consideriamo in \mathbb{R}^2 e \mathbb{R}^3 i loro prodotti scalari canonici. L'espressione in \mathbb{R}^2 di una retta in forma cartesiana

$$r = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : a \cdot x + b \cdot y = 0 \right\}$$

equivale allora alla scrittura $r = \begin{pmatrix} a \\ b \end{pmatrix}^\perp$. Dunque il passaggio a forma parametrica equivale alla ricerca di un vettore non nullo ortogonale a $\begin{pmatrix} a \\ b \end{pmatrix}$, che si verifica subito essere $\begin{pmatrix} -b \\ a \end{pmatrix}$.

Passando a \mathbb{R}^3 , notiamo che per la Proposizione 9.3.1 possiamo associare a ogni retta il piano ad essa ortogonale, e viceversa. Abbiamo inoltre le seguenti interpretazioni delle presentazioni cartesiane:

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : a \cdot x + b \cdot y + c \cdot z = 0 \right\} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}^\perp,$$

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z = 0 \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z = 0 \end{cases} \right\} = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}^\perp \cap \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}^\perp.$$

Poiché

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}^\perp = \left(\text{Span} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right)^\perp,$$

ne deduciamo che una presentazione cartesiana di un piano corrisponde a una presentazione parametrica della retta ad esso ortogonale. Analogamente

$$\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}^\perp \cap \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}^\perp = \left(\text{Span} \left(\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} \right) \right)^\perp,$$

dunque una presentazione cartesiana di una retta corrisponde ad una presentazione parametrica del piano ad essa ortogonale. Ciò spiega la ragione della forte analogia tra il passaggio da forma parametrica a cartesiana per rette e quello da forma cartesiana a parametrica per piani. In entrambi i casi infatti

si tratta di trovare, dato $v = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ non nullo, un insieme di generatori per v^\perp , e un tale insieme è dato da $\begin{pmatrix} \beta \\ -\alpha \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma \\ 0 \\ -\alpha \end{pmatrix}, \begin{pmatrix} 0 \\ \gamma \\ -\beta \end{pmatrix}$.

Similmente si spiega l'analogia tra il passaggio da forma cartesiana a parametrica per rette e quello da forma parametrica a cartesiana per piani, infatti in entrambi i casi, dati $v_1 = \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix}$ linearmente indipendenti, si tratta di trovare una base per $v_1^\perp \cap v_2^\perp$, e questa è data dal vettore $\begin{pmatrix} \beta_1 \cdot \gamma_2 - \beta_2 \cdot \gamma_1 \\ -\alpha_1 \cdot \gamma_2 + \alpha_2 \cdot \gamma_1 \\ \alpha_1 \cdot \beta_2 - \alpha_2 \cdot \beta_1 \end{pmatrix}$, che denoteremo ora con $v_1 \wedge v_2$ e chiameremo *prodotto vettoriale* di v_1 e v_2 . Possiamo anche dare una migliore interpretazione geometrica di questo vettore. Si verifica infatti che esso, oltre a essere un generatore della retta ortogonale al piano generato da v_1 e v_2 , ha norma uguale all'area del parallelogrammo di lati v_1 e v_2 , e verso tale che $\det(v_1 \ v_2 \ v_1 \wedge v_2) > 0$, ovvero tale che è possibile disporre le dita pollice, indice e medio della mano destra secondo le direzioni di v_1 , v_2 e $v_1 \wedge v_2$ rispettivamente.

9.4 Matrici simmetriche e ortogonali

Dato come sopra V con un prodotto scalare $\langle \cdot | \cdot \rangle$, tra le applicazioni lineari $f : V \rightarrow V$ risultano di particolare interesse quelle tali che $\langle v | f(w) \rangle = \langle f(v) | w \rangle$ per ogni v, w , dette *autoaggiunte*, e quelle tali che $\langle f(v) | f(w) \rangle = \langle v | w \rangle$ per ogni v, w , dette *ortogonali*. Poiché il prodotto scalare determina la norma e viceversa, notiamo che f è ortogonale se e soltanto se $\|f(v)\| = \|v\|$ per ogni v , ovvero se f preserva la distanza tra i vettori, dunque se f è una *isometria*.

In relazione alla nozione generale di autoaggiunzione ci limitiamo al seguente risultato, da confrontare con le Proposizioni 5.2.2 e 5.2.4:

Proposizione 9.4.1. *Sia V di dimensione finita. Una applicazione lineare $p : V \rightarrow V$ è la proiezione ortogonale su un sottospazio se e soltanto se è autoaggiunta e $p \circ p = p$.*

Dimostrazione. Sia p la proiezione ortogonale su W . Sappiamo già che $p \circ p = p$. Dati $v_1, v_2 \in V$ sia $p(v_i) = w_i$, dunque $v_i = w_i + z_i$ con $w_i \in W$ e $z_i \in W^\perp$. Allora

$$\langle p(v_1) | v_2 \rangle = \langle w_1 + z_1 | w_2 \rangle = \langle w_1 | w_2 \rangle = \langle w_1 | w_2 + z_2 \rangle = \langle v_1 | p(v_2) \rangle$$

dunque p è autoaggiunta.

Supponiamo viceversa che p sia autoaggiunta e che $p \circ p = p$. Sappiamo che p è allora la proiezione su W associata alla decomposizione $W \oplus Z$ con

$W = \text{Im}(p)$ e $Z = \text{Ker}(p)$. Dobbiamo vedere che $Z = W^\perp$. Se $w \in W$ e $z \in Z$, dalla autoaggiunzione segue che

$$\langle w|z \rangle = \langle p(w)|z \rangle = \langle w|p(z) \rangle = \langle w|0 \rangle = 0.$$

Ciò prova che $Z \subset W^\perp$, e dalla formula di Grassmann segue l'uguaglianza desiderata. \square

Analizziamo ora le nozioni di autoaggiunzione e ortogonalità nel caso del prodotto scalare standard di \mathbb{R}^n . Il seguente è immediato:

Lemma 9.4.2. *Se $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ allora $\langle x|A \cdot y \rangle = \langle {}^tA \cdot x|y \rangle$.*

Ne segue allora che A è autoaggiunta se e soltanto se è simmetrica, ed è ortogonale se e soltanto se tA è l'inversa di A . Scrivendo A come insieme di colonne $(v_1 \dots v_n)$ e notando che $({}^tA \cdot A)_{ij} = {}^t v_i \cdot v_j = \langle v_i|v_j \rangle_{\mathbb{R}^n}$, ne deduciamo che A è ortogonale se e solo se le sue colonne sono una base ortonormale.

Esempi di matrici ortogonali 2×2 (dunque di isometrie di \mathbb{R}^2 rispetto alla distanza indotta dal prodotto scalare standard) sono le rotazioni

$$\begin{pmatrix} \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

e le riflessioni rispetto a rette, che in una base ortonormale opportuna si scrivono sempre come

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

In dimensione 3 sono ortogonali (cioè sono isometrie rispetto alla distanza indotta dal prodotto scalare standard di \mathbb{R}^3) ad esempio le rotazioni intorno a rette, che in una base ortonormale opportuna si scrivono come

$$\begin{pmatrix} \cos(\vartheta) & -\sin(\vartheta) & 0 \\ \sin(\vartheta) & \cos(\vartheta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

e le riflessioni rispetto a piani, che si scrivono come

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

9.5 Il caso complesso

Il modulo, che abbiamo introdotto su \mathbb{C} per misurare la lunghezza dei vettori, si esprime come $\|z\|^2 = z \cdot \bar{z}$. Ciò suggerisce che in \mathbb{C}^n l'analogo del prodotto scalare reale sia dato da

$$\langle v|w \rangle_{\mathbb{C}^n} = v_1 \cdot \bar{w}_1 + \dots + v_n \cdot \bar{w}_n = {}^t v \cdot \bar{w} = {}^t \bar{w} \cdot v,$$

che scriveremo anche come $w^* \cdot v$ definendo $A^* = {}^t \bar{A}$ per ogni $A \in \mathcal{M}_{n \times m}(\mathbb{C})$ e chiamando A^* l'*aggiunta* di A (ovviamente non importa quale operazione sia eseguita per prima tra la trasposizione e il coniugio). Notiamo subito che identificando \mathbb{C}^n a \mathbb{R}^{2n} si ha $\langle x|y \rangle_{\mathbb{R}^{2n}} = \Re(\langle x|y \rangle_{\mathbb{C}^n})$.

Analizzando le proprietà di $\langle \cdot | \cdot \rangle_{\mathbb{C}^n}$ si perviene alla seguente definizione. Se V è uno spazio vettoriale complesso, si chiama *prodotto scalare hermitiano* un'applicazione $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{C}$ che sia:

- *sesquilineare*, cioè lineare a sinistra e anti-lineare a destra, ovvero tale che

$$\langle \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 | v_3 \rangle = \lambda_1 \cdot \langle v_1 | v_3 \rangle + \lambda_2 \cdot \langle v_2 | v_3 \rangle$$

$$\langle v_3 | \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 \rangle = \bar{\lambda}_1 \cdot \langle v_3 | v_1 \rangle + \bar{\lambda}_2 \cdot \langle v_3 | v_2 \rangle$$

per ogni $v_1, v_2, v_3 \in V$ e $\lambda_1, \lambda_2 \in \mathbb{C}$;

- *hermitiana*, ovvero tale che

$$\langle v_2 | v_1 \rangle = \overline{\langle v_1 | v_2 \rangle}$$

per ogni $v_1, v_2 \in V$;

- *definita positiva*, ovvero tale che $\langle v|v \rangle > 0$ se $v \neq 0$ (si noti che la condizione precedente garantisce che $\langle v|v \rangle \in \mathbb{R}$).

Oltre a $\langle \cdot | \cdot \rangle_{\mathbb{C}^n}$, che naturalmente chiameremo *prodotto scalare hermitiano canonico* di \mathbb{C}^n , sono esempi di prodotti scalari $\langle A|B \rangle = \text{tr}(B^* \cdot A) = \text{tr}({}^t A \cdot \bar{B})$ su $\mathcal{M}_{n \times m}(\mathbb{C})$, e

$$\langle f|g \rangle = \int_0^1 f(t) \cdot \overline{g(t)} dt$$

sullo spazio delle funzioni da $[0, 1]$ in \mathbb{C} continue (cioè aventi parte reale ed immaginaria continua).

Sia ora fissato V spazio vettoriale complesso con un prodotto scalare hermitiano $\langle \cdot | \cdot \rangle$. Notiamo intanto che $\Re(\langle \cdot | \cdot \rangle)$ è un prodotto scalare su V pensato

come spazio reale. Inoltre, tutte le definizioni e i risultati descritti nel caso reale si estendono con semplici modifiche al caso complesso. In particolare si definisce la norma $\|v\|$ di un vettore come $\sqrt{\langle v|v \rangle}$ e si ha $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$. Valgono poi le disuguaglianze di Cauchy-Schwarz e triangolare:

$$\begin{aligned} |\langle z|w \rangle| &\leq \|z\| \cdot \|w\|, \\ \|z + w\| &\leq \|z\| + \|w\|. \end{aligned}$$

Le nozioni di vettore unitario, vettori ortogonali e sistema ortonormale si trasportano direttamente al caso complesso, e vale la seguente:

Proposizione 9.5.1. *Se V ha dimensione finita, $W \subset V$ è un sottospazio vettoriale e v_1, \dots, v_k è una base ortogonale di W , allora $V = W \oplus W^\perp$ e la proiezione p_W su W relativa a tale decomposizione (detta proiezione ortogonale su W) è data da*

$$p_W(v) = \sum_{j=1}^k \frac{\langle v|v_j \rangle}{\|v_j\|^2} \cdot v_j.$$

Inoltre vale la disuguaglianza di Bessel:

$$\sum_{j=1}^k \frac{|\langle v|v_j \rangle|^2}{\|v_j\|^2} \leq \|v\|^2.$$

A proposito dell'enunciato precedente notiamo che, a differenza del caso reale, non si può scrivere $\langle v_j|v \rangle$ in luogo di $\langle v|v_j \rangle$ nella formula di p_W . Per ricordare che bisogna usare $\langle v|v_j \rangle$ basta notare che si sta esprimendo la proiezione p_W , che è *lineare*, non anti-lineare, in v , e la nostra convenzione è che $\langle \cdot | \cdot \rangle$ sia lineare a sinistra.

Anche il procedimento di ortonormalizzazione di Gram-Schmidt di una base x_1, \dots, x_n è invariato: si suppongono ricorsivamente costruiti v_1, \dots, v_k e si pone

$$y_{k+1} = x_{k+1} - \sum_{i=1}^k \langle x_{k+1}|v_i \rangle \cdot v_i, \quad v_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|}.$$

Si estende inoltre alla lettera la nozione di applicazione *autoaggiunta* da V in se stesso, e vale la medesima caratterizzazione delle proiezioni ortogonali provata nel caso reale: *una $p : V \rightarrow V$ lineare è una proiezione ortogonale se e soltanto se è autoaggiunta e $p \circ p = p$.*

Passiamo ora al caso dello spazio \mathbb{C}^n . Se $A \in \mathcal{M}_{n \times n}(\mathbb{C})$ e $z, w \in \mathbb{C}^n$, poniamo $\langle z|w \rangle_A = w^* \cdot A \cdot z$. Come nel caso reale si verifica che le applicazioni

$\mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ lineari a sinistra e antilineari a destra sono precisamente quelle della forma $\langle \cdot | \cdot \rangle_A$. Inoltre una tale applicazione è hermitiana se e soltanto se $A^* = A$, nel qual caso diremo che A stessa è una *matrice hermitiana*. Infine si ha

$$\langle z | A \cdot w \rangle_{\mathbb{C}^n} = \langle A^* \cdot z | w \rangle_{\mathbb{C}^n}$$

da cui segue facilmente che:

- $\langle z | A \cdot w \rangle_{\mathbb{C}^n} = \langle A \cdot z | w \rangle_{\mathbb{C}^n}$ per ogni $z, w \in \mathbb{C}^n$ se e soltanto se $A^* = A$, dunque l'applicazione associata a una matrice è autoaggiunta se e soltanto se la matrice è hermitiana;
- $\langle A \cdot z | A \cdot w \rangle_{\mathbb{C}^n} = \langle z | w \rangle_{\mathbb{C}^n}$ per ogni $z, w \in \mathbb{C}^n$ se e soltanto se A^* è l'inversa di A , il che equivale al fatto che le colonne di A siano una base ortonormale di \mathbb{C}^n : diremo in tal caso che A è *unitaria*, osservando che è allora un'isometria rispetto alla distanza indotta dal prodotto scalare standard di \mathbb{C}^n .

Capitolo 10

Autovalori e diagonalizzazione

In questo capitolo ci proponiamo di analizzare l'azione di un'assegnata applicazione lineare $f : V \rightarrow V$ cercando una base \mathcal{B} rispetto alla quale la matrice $[f]_{\mathcal{B}}^{\mathcal{B}}$ sia semplice ed espressiva. Per motivare l'interesse nei confronti di questo problema iniziamo da una situazione pratica.

Supponiamo che un corpo materiale rigido C sia soggetto a varie forze e che il baricentro sia in quiete, cioè che la risultante di tutte le forze applicate nel baricentro sia nulla. Scegliamo un sistema di riferimento cartesiano ortogonale con l'origine nel baricentro del corpo C . Chiamando $f(x)$ la risultante di tutte le forze applicate a un punto $x \in C$ otteniamo una funzione $f : C \rightarrow \mathbb{R}^3$ tale che $f(0) = 0$. Possiamo allora operare una *approssimazione lineare* di f , supponendo che essa abbia la forma $f(x) = F \cdot x$ per qualche matrice $F \in \mathcal{M}_{3 \times 3}(\mathbb{R})$. A questo punto decomponiamo F come $F = S + A$, con $S = \frac{1}{2}(F + {}^tF)$ simmetrica e $A = \frac{1}{2}(F - {}^tF)$ antisimmetrica; si tratta esattamente delle proiezioni di F rispetto alla decomposizione in somma diretta $\mathcal{M}_{n \times n}(\mathbb{R}) = \mathcal{S}_n \oplus \mathcal{A}_n$. È ora possibile dimostrare (non lo faremo del tutto, ma proveremo un risultato in questo senso nella Proposizione 14.3.8) che A è responsabile di un movimento rigido di C (più esattamente, di una rotazione intorno a un asse), mentre S , che si chiama *tensori degli sforzi* è responsabile della deformazione di C . Supponiamo adesso che S sia la matrice seguente:

$$S = \frac{1}{8} \begin{pmatrix} 1 & -3\sqrt{3} & 6 \\ -3\sqrt{3} & -5 & -2\sqrt{3} \\ 6 & -2\sqrt{3} & 12 \end{pmatrix}.$$

Esaminandola appare decisamente poco chiaro quale sia qualitativamente l'azione su C degli sforzi associati a S ; ad esempio non è neppure evidente se ci

siano direzioni nelle quali lo sforzo sia prevalentemente di compressione oppure di trazione. Se però prendiamo la base $\mathcal{B} = (v_1, v_2, v_3)$ di \mathbb{R}^3 data da

$$v_1 = \frac{1}{4} \begin{pmatrix} \sqrt{3} \\ -1 \\ 2\sqrt{3} \end{pmatrix}, \quad v_2 = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{3} \\ 0 \end{pmatrix}, \quad v_3 = \frac{1}{4} \begin{pmatrix} -3 \\ \sqrt{3} \\ 2 \end{pmatrix}$$

possiamo verificare con un calcolo diretto che

$$Av_1 = 2v_1, \quad Av_2 = -v_2, \quad Av_3 = 0.$$

Questo significa che nella direzione v_1 lo sforzo è puramente di trazione, nella direzione v_2 è puramente di compressione, mentre nella direzione v_3 è nullo. Osserviamo inoltre che la base \mathcal{B} è ortonormale, dunque le direzioni dei vettori v_1, v_2, v_3 sono ortogonali tra loro. Ne segue anche che se scegliamo v_1, v_2, v_3 come versori degli assi di un nuovo riferimento cartesiano ortogonale monometrico in \mathbb{R}^3 , rispetto alle nuove coordinate il tensore degli sforzi è rappresentato dalla matrice

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

che è decisamente più semplice della S originale.

Esempio 10.0.2. Ecco un'altra matrice simmetrica la cui azione si comprende molto meglio dopo un opportuno cambiamento di base. Prendiamo la matrice

$$S = \begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{pmatrix}$$

e consideriamo l'applicazione lineare $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ associata a S . Sia ora $\mathcal{B} = (v_1, v_2, v_3)$ la base data da

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}.$$

Si verifica allora che $Av_1 = 6v_1$, $Av_2 = 6v_2$, $Av_3 = 0$, dunque

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Osservando anche che la base \mathcal{B} è ortogonale, ne deduciamo che l'azione di f consiste nella proiezione ortogonale sul piano generato da v_1 e v_2 seguita da un'omotetia di ragione 6 su tale piano. Si tratta indubbiamente di una descrizione di f più efficace di quella fornita dalla sola matrice $A = [f]_{\mathcal{E}_3}^{\mathcal{E}_3}$.

10.1 Matrici coniugate e diagonalizzabilità

Ci riferiamo per un attimo a una situazione più generale di quella descritta, ovvero ad un'applicazione lineare $f : V \rightarrow W$ tra spazi vettoriali di dimensione finita, entrambi su \mathbb{K} , dove \mathbb{K} è \mathbb{R} oppure \mathbb{C} . In tal caso è facile vedere che esistono basi \mathcal{B} e \mathcal{C} tali che

$$[f]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

dove $k = \text{rank}(f)$ e ciascuno 0 indica un'opportuna matrice con coefficienti tutti nulli. Dunque l'unica caratteristica di f che è possibile ricavare analizzando le matrici $[f]_{\mathcal{B}}^{\mathcal{C}}$ è il suo rango. Questo suggerisce di ridursi al caso in cui $W = V$ e di analizzare soltanto le matrici del tipo $[f]_{\mathcal{B}}^{\mathcal{B}}$, come anticipato sopra.

Consideriamo dunque nel resto del paragrafo uno spazio vettoriale V di dimensione n su \mathbb{K} e $f : V \rightarrow V$ lineare. Per trovare la matrice $[f]_{\mathcal{B}}^{\mathcal{B}}$ più espressiva ne fissiamo una a caso e ci chiediamo come si ricavino da essa le altre. Fissiamo cioè una base \mathcal{B}_0 e consideriamo l'associata matrice $A_0 = [f]_{\mathcal{B}_0}^{\mathcal{B}_0}$. Se prendiamo ora un'altra base \mathcal{B} sappiamo che esiste una matrice M invertibile tale che $\mathcal{B} = \mathcal{B}_0 \cdot M$, e che la matrice $A = [f]_{\mathcal{B}}^{\mathcal{B}}$ è data da $M^{-1} \cdot A_0 \cdot M$. Notiamo ora che se M' è una qualsiasi matrice $n \times n$ invertibile allora $\mathcal{B}_0 \cdot M'$ è sempre una base di V . Definendo allora una matrice A coniugata ad A_0 se esiste M invertibile tale che $A = M^{-1} \cdot A_0 \cdot M$, abbiamo provato la seguente:

Proposizione 10.1.1. *Se A_0 è una matrice di f , le altre matrici di f sono tutte e sole le coniugate di A_0 .*

(Alcuni usano il termine *simile* in luogo del termine *coniugata* che useremo sempre noi). Poiché le matrici di cui è più facile comprendere l'azione sono quelle diagonali, diciamo che $f : V \rightarrow V$ è *diagonalizzabile* se esiste una base \mathcal{B} tale che $[f]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale. Dunque una matrice è diagonalizzabile se è coniugata a una diagonale. Notiamo che f è diagonalizzabile se e soltanto se esistono una base (v_1, \dots, v_n) di V e costanti $\lambda_1, \dots, \lambda_n$ tali che $f(v_j) = \lambda_j \cdot v_j$ per ogni j . Ciò suggerisce di partire, nell'analisi di una qualsiasi f , dall'equazione $f(v) = \lambda \cdot v$, che naturalmente è banalmente soddisfatta per ogni λ se $v = 0$. Escludendo

tale caso siamo condotti alla seguente definizione: uno scalare $\lambda \in \mathbb{K}$ si dice *autovalore* per f se esiste $v \in V$ non nullo tale che $f(v) = \lambda \cdot v$; ogni v siffatto è detto *autovettore* di f relativo a λ . Notiamo ora che, se λ è un autovalore di f , unendo 0 all'insieme di tutti gli autovettori relativi a λ otteniamo l'insieme

$$\begin{aligned} \{v \in V : f(v) = \lambda \cdot v\} &= \{v \in V : (\lambda \cdot \text{Id}_V - f)(v) = 0\} \\ &= \text{Ker}(\lambda \cdot \text{Id}_V - f) \end{aligned}$$

che è dunque un sottospazio vettoriale di V ed è chiamato *autospatio* relativo all'autovalore λ . Vale inoltre la seguente:

Osservazione 10.1.2. Uno scalare $\lambda \in \mathbb{K}$ è autovalore di f se e solo se

$$\dim_{\mathbb{K}}(\text{Ker}(\lambda \cdot \text{Id}_V - f)) > 0.$$

Nel caso in cui f sia l'applicazione $\mathbb{K}^n \rightarrow \mathbb{K}^n$ associata ad una matrice $A \in \mathcal{M}_{n \times n}(\mathbb{K})$, deduciamo dall'osservazione precedente un semplice criterio per trovare gli autovalori di f , infatti λ lo è se e solo se $\det(\lambda \cdot I_n - A) = 0$. Proviamo nella prossima proposizione che questo criterio si estende in generale. Ricordiamo dapprima che un polinomio in $\mathbb{K}[t]$ si dice *monico* se il coefficiente della sua potenza più alta è l'unità.

Proposizione 10.1.3. Se \mathcal{B} è una base di V si ha che

$$p_f(t) = \det(t \cdot I_n - [f]_{\mathcal{B}}^{\mathcal{B}})$$

è indipendente da \mathcal{B} ed è un polinomio monico di grado n in t . Inoltre $\lambda \in \mathbb{K}$ è un autovalore di f se e soltanto se $p_f(\lambda) = 0$.

Dimostrazione. Notiamo intanto che $t \cdot I_n - [f]_{\mathcal{B}}^{\mathcal{B}} = [t \cdot \text{Id}_V - f]_{\mathcal{B}}^{\mathcal{B}}$, da cui segue subito l'ultima affermazione. Inoltre, se \mathcal{B}' è un'altra base di V e $\mathcal{B}' = \mathcal{B} \cdot M$, si ha

$$\begin{aligned} t \cdot I_n - [f]_{\mathcal{B}'}^{\mathcal{B}'} &= [t \cdot \text{Id}_V - f]_{\mathcal{B}'}^{\mathcal{B}'} = M^{-1} \cdot [t \cdot \text{Id}_V - f]_{\mathcal{B}}^{\mathcal{B}} \cdot M \\ &= M^{-1} \cdot (t \cdot I_n - [f]_{\mathcal{B}}^{\mathcal{B}}) \cdot M \end{aligned}$$

e la prima affermazione discende allora dal Teorema di Binet. La seconda affermazione è ovvia. \square

Il polinomio $p_f(t)$ descritto nella proposizione precedente è detto *polinomio caratteristico* di f . L'argomento sviluppato nella dimostrazione implica

in particolare che si può definire il *determinante* $\det(f)$ di f come $\det([f]_{\mathcal{B}}^{\mathcal{B}})$ dove \mathcal{B} è una base arbitraria di V . Inoltre il termine noto $p_f(0)$ del polinomio caratteristico risulta uguale a $(-1)^n \cdot \det(f)$. Anche il coefficiente del termine di grado $n - 1$ in $p_f(t)$ ammette una chiara interpretazione, che richiede il seguente semplice:

Lemma 10.1.4. *Se A e B sono matrici quadrate si ha $\text{tr}(A \cdot B) = \text{tr}(B \cdot A)$.*

Usando questo fatto è facile vedere che si può definire la *traccia* di f come la traccia di una sua qualsiasi matrice, e che il coefficiente di t^{n-1} in $p_f(t)$ è l'opposto della traccia di f .

Tornando alla diagonalizzabilità abbiamo ora la seguente:

Proposizione 10.1.5. *Se $p_f(t)$ ha n radici distinte che appartengono a \mathbb{K} allora f è diagonalizzabile.*

Dimostrazione. Siano $\lambda_1, \dots, \lambda_n$ le radici di $p_f(t)$. L'ipotesi garantisce che esistono v_1, \dots, v_n non nulli tali che $f(v_j) = \lambda_j \cdot v_j$. Basta allora vedere che essi costituiscono una base. Lo facciamo verificando induttivamente che v_1, \dots, v_k sono linearmente indipendenti, il che è ovvio se $k = 1$. Sia dunque $k < n$, supponiamo dimostrato che v_1, \dots, v_k sono linearmente indipendenti, e proviamo che $v_{k+1} \notin \text{Span}(v_1, \dots, v_k)$. Lo facciamo ragionando per assurdo:

$$\begin{aligned} v_{k+1} &= \alpha_1 \cdot v_1 + \dots + \alpha_k \cdot v_k \\ \Rightarrow f(v_{k+1}) &= \alpha_1 \cdot f(v_1) + \dots + \alpha_k \cdot f(v_k) \\ \Rightarrow \lambda_{k+1} \cdot v_{k+1} &= \alpha_1 \cdot \lambda_1 \cdot v_1 + \dots + \alpha_k \cdot \lambda_k \cdot v_k \\ \Rightarrow \lambda_{k+1} \cdot (\alpha_1 \cdot v_1 + \dots + \alpha_k \cdot v_k) &= \alpha_1 \cdot \lambda_1 \cdot v_1 + \dots + \alpha_k \cdot \lambda_k \cdot v_k \\ \Rightarrow \alpha_1 \cdot (\lambda_1 - \lambda_{k+1}) \cdot v_1 + \dots + \alpha_k \cdot (\lambda_k - \lambda_{k+1}) \cdot v_k &= 0 \\ \Rightarrow \alpha_j \cdot (\lambda_j - \lambda_{k+1}) &= 0 \quad \forall j \\ \Rightarrow \alpha_j &= 0 \quad \forall j . \end{aligned}$$

Contraddizione. □

A proposito dell'enunciato appena dimostrato ricordiamo che un polinomio di grado n ha sempre n radici (con molteplicità) su \mathbb{C} , mentre può averne meno di n su \mathbb{R} . Dunque nel caso $\mathbb{K} = \mathbb{C}$ l'ipotesi è soltanto che le radici del polinomio caratteristico siano distinte. Invece nel caso $\mathbb{K} = \mathbb{R}$ l'ipotesi è che siano reali e distinte.

Nel caso in cui il polinomio caratteristico non abbia radici distinte esiste un più elaborato criterio per la diagonalizzabilità, basato sulle nozioni seguenti:

se λ è un autovalore di f , chiamiamo *molteplicità algebrica* la molteplicità di λ come radice di $p_f(t)$, e *molteplicità geometrica* la dimensione su \mathbb{K} del relativo autospazio, cioè la dimensione del nucleo di $\lambda \cdot \text{Id}_V - f$. Indichiamo le due molteplicità con $\text{m.a.}(\lambda)$ e $\text{m.g.}(\lambda)$ rispettivamente. Abbiamo intanto:

Lemma 10.1.6. *Se λ è autovalore di f si ha $1 \leq \text{m.g.}(\lambda) \leq \text{m.a.}(\lambda)$.*

Dimostrazione. Per definizione di autovalore si ha $\text{m.g.}(\lambda) \geq 1$. Sia ora $k = \text{m.g.}(\lambda)$ e scegliamo una base v_1, \dots, v_k di $\text{Ker}(\lambda \cdot \text{Id}_V - f)$, completandola poi a una base $\mathcal{B} = (v_1, \dots, v_n)$ di V . Si ha allora:

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \left(\begin{array}{cc|c} \lambda & 0 & \\ & \ddots & A' \\ 0 & & \lambda \\ \hline & 0 & A'' \end{array} \right)$$

per matrici $A' \in \mathcal{M}_{k \times (n-k)}(\mathbb{K})$ e $A'' \in \mathcal{M}_{(n-k) \times (n-k)}(\mathbb{K})$ opportune. Ne segue che $p_f(t) = (t - \lambda)^k \cdot p_{A''}(t)$ dunque che $\text{m.a.}(\lambda) \geq k$, onde la conclusione. \square

Con un argomento simile a quello usato per la Proposizione 10.1.5 si può dimostrare ora il seguente:

Teorema 10.1.7. *Siano V uno spazio vettoriale su \mathbb{K} e $f : V \rightarrow V$ un'applicazione lineare.*

- Se $\mathbb{K} = \mathbb{C}$ allora f è diagonalizzabile se e solo se per ogni autovalore λ di f si ha $\text{m.g.}(\lambda) = \text{m.a.}(\lambda)$;
- Se $\mathbb{K} = \mathbb{R}$ allora f è diagonalizzabile se e solo se ha tutti gli autovalori reali e per ogni autovalore λ si ha $\text{m.g.}(\lambda) = \text{m.a.}(\lambda)$.

Osserviamo che, essendo $\mathbb{R} \subset \mathbb{C}$, a una matrice $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ sono associate sia un'applicazione $\mathbb{R}^n \rightarrow \mathbb{R}^n$ sia una $\mathbb{C}^n \rightarrow \mathbb{C}^n$. Quindi quando ci si chiede se A sia diagonalizzabile è necessario specificare se lo si chieda in senso reale o in senso complesso. Dal teorema precedente e dal fatto che per una matrice reale la dimensione del nucleo reale è uguale a quella del nucleo complesso (si veda la Proposizione 8.5.3) segue che A è *diagonalizzabile su \mathbb{R}* , cioè esiste M reale invertibile tale che $M^{-1} \cdot A \cdot M$ sia diagonale, se e solo se A è *diagonalizzabile su \mathbb{C} e ha autovalori reali*.

Concludiamo il paragrafo interpretando in termini di autovalori il determinante e la traccia di una matrice $A \in \mathcal{M}_{n \times n}(\mathbb{C})$. Abbiamo infatti già visto che il determinante è (a meno del segno) il termine noto di $p_A(t)$, mentre la traccia

è l'opposto del coefficiente di t^{n-1} . Se A ha autovalori $\lambda_1, \dots, \lambda_n$ contati con la loro molteplicità algebrica abbiamo ora $p_A(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$. Il termine noto di questo polinomio è $(-1)^n \lambda_1 \cdot \dots \cdot \lambda_n$, mentre il coefficiente di t^{n-1} è $-(\lambda_1 + \dots + \lambda_n)$. Ne deduciamo la seguente:

Proposizione 10.1.8. *Per una qualsiasi matrice quadrata il determinante è il prodotto degli autovalori, e la traccia ne è la somma. In entrambi i casi gli autovalori devono essere considerati con la loro molteplicità algebrica.*

Notiamo che questa proposizione vale anche per matrici reali aventi autovalori non tutti reali.

10.2 Il teorema spettrale

Molte delle matrici che emergono dai problemi della fisica e dell'ingegneria appartengono alle classi speciali già introdotte delle simmetriche, delle antisimmetriche, delle ortogonali, delle hermitiane, oppure delle unitarie. In questo paragrafo mostriamo che tutte le matrici di questi tipi hanno una "forma canonica" particolarmente semplice. Cominciamo col risultato più importante, detto *teorema spettrale*:

Teorema 10.2.1. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. Allora A è simmetrica se e soltanto se A si diagonalizza tramite una matrice ortogonale.*

Dimostrazione. La seconda condizione dell'enunciato significa che esiste una matrice $M \in \mathcal{M}_{n \times n}(\mathbb{R})$ ortogonale, cioè tale che ${}^tM = M^{-1}$, per la quale $M^{-1} \cdot A \cdot M$ è diagonale; la condizione equivale anche al fatto che A ammetta una base *ortonormale* di autovettori. Nell'enunciato l'implicazione interessante è naturalmente

$$(\heartsuit) \quad A \text{ simmetrica} \Rightarrow A \text{ si diagonalizza tramite una } M \text{ ortogonale,}$$

tuttavia conviene iniziare dimostrando l'implicazione inversa. Supponiamo allora che $M^{-1} \cdot A \cdot M = D$ con D diagonale e M ortogonale. Usando il fatto che $M^{-1} = {}^tM$ si ricava allora $A = M \cdot D \cdot {}^tM$, e trasponendo questa uguaglianza si ottiene ${}^tA = M \cdot {}^tD \cdot {}^tM$, da cui segue subito che ${}^tA = A$ poiché naturalmente ${}^tD = D$.

Passando all'implicazione (\heartsuit) , supponiamo allora che A sia simmetrica e cominciamo a dimostrare che A ammette almeno un autovalore reale. Poiché su \mathbb{C} l'equazione $p_A(t) = 0$ ammette n soluzioni con molteplicità, basterà vedere

che ogni autovalore $\lambda \in \mathbb{C}$ in realtà è reale. Fissato tale λ , notiamo che esiste $z \in \mathbb{C}^n$ non nullo tale che $A \cdot z = \lambda \cdot z$. Dunque:

$$\begin{aligned} \lambda \cdot \|z\|_{\mathbb{C}^n}^2 &= \langle \lambda \cdot z | z \rangle_{\mathbb{C}^n} = \langle A \cdot z | z \rangle_{\mathbb{C}^n} = \langle z | A^* \cdot z \rangle_{\mathbb{C}^n} \\ &= \langle z | A \cdot z \rangle_{\mathbb{C}^n} = \langle z | \lambda \cdot z \rangle_{\mathbb{C}^n} = \bar{\lambda} \cdot \|z\|_{\mathbb{C}^n}^2 . \end{aligned}$$

Poiché $z \neq 0$, ne segue che $\bar{\lambda} = \lambda$, dunque $\lambda \in \mathbb{R}$.

Avendo provato che A ammette autovalori reali, ne fissiamo uno che denotiamo con λ_1 , e scegliamo un relativo autovettore reale v_1 che sia unitario. Possiamo allora completare v_1 a una base che sia ortonormale. Esiste dunque una matrice M_1 ortogonale tale che

$$M_1^{-1} \cdot A \cdot M_1 = \begin{pmatrix} \lambda_1 & x_1 \\ 0 & A_1 \end{pmatrix}$$

dove x_1 è un vettore riga e $A_1 \in \mathcal{M}_{(n-1) \times (n-1)}(\mathbb{R})$. Essendo $M_1^{-1} = {}^t M_1$, è facile vedere che $M_1^{-1} \cdot A \cdot M_1$ è ancora simmetrica (l'argomento è esattamente quello con il quale abbiamo dimostrato l'implicazione inversa di (♥)), dunque $x_1 = 0$ e A_1 è simmetrica. Ci siamo pertanto ricondotti al caso di una matrice simmetrica $(n-1) \times (n-1)$, e un semplice ragionamento per induzione consente di concludere la dimostrazione. \square

(L'origine del nome *spettrale* che abbiamo attribuito al teorema precedente è (in termini estremamente semplicistici) questa: il consueto spettro prismatico che decompone la luce bianca nei varî colori sta trovando autovalori e autovettori di una certa applicazione, detta *funzione d'onda* che è in un opportuno spazio lineare e simmetrica.)

Il primo passo della dimostrazione precedente, consistente nella verifica che una matrice simmetrica ha almeno un autovalore reale, può essere compiuto anche senza passare ai numeri complessi, usando invece metodi analitici:

Proposizione 10.2.2. *Se $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è simmetrica allora la funzione*

$$f : \mathbb{R}^n \setminus \{0\} \ni x \mapsto \frac{\langle A \cdot x | x \rangle_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}^2} \in \mathbb{R}$$

ammette massimo. Tale massimo è un autovalore di A , e il punto di massimo di f è un autovettore a esso relativo.

Dimostrazione. Poiché $f(k \cdot x) = f(x)$ se $k > 0$, i valori presi dalla f sono quelli presi sulla sfera unitaria. Tale sfera è un chiuso limitato sul quale la f è

continua, dunque il massimo esiste. Sia \bar{x} un punto di massimo ed m il valore di tale massimo. Notando che

$$f(x) = \frac{\sum_{i,j=1}^n (A)_{ij} \cdot x_i \cdot x_j}{\sum_{i=1}^n x_i^2}$$

ed imponendo che per $x = \bar{x}$ la derivata parziale $\frac{\partial f}{\partial x_k}$ si annulli (vedi anche il successivo Sottoparagrafo 10.2.1) otteniamo

$$\left(2 \sum_{j=1}^n (A)_{kj} \cdot \bar{x}_j \right) \cdot \left(\sum_{i=1}^n \bar{x}_i^2 \right) - \left(\sum_{i,j=1}^n (A)_{ij} \cdot \bar{x}_i \cdot \bar{x}_j \right) \cdot 2\bar{x}_k = 0$$

ovvero $\sum_{j=1}^n (A)_{kj} \cdot \bar{x}_j = m \cdot \bar{x}_k$, da cui segue la tesi. \square

Il teorema spettrale ci consente ora di rispondere alla domanda, rimasta in sospenso nel capitolo precedente, di quali matrici simmetriche definiscano prodotti scalari:

Corollario 10.2.3. *Se $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è simmetrica, allora $\langle \cdot | \cdot \rangle_A$ è definito positivo se e solo se A ha autovalori positivi.*

Dimostrazione. Esiste M ortogonale tale che ${}^t M \cdot A \cdot M$ è una matrice diagonale con coefficienti $\lambda_1, \dots, \lambda_n$. Poiché M è invertibile, si ha $\langle x | x \rangle_A > 0$ per ogni $x \in \mathbb{R}^n$ se e solo se $\langle M \cdot x | M \cdot x \rangle_A$ per ogni $x \in \mathbb{R}^n$. Ora

$$\langle M \cdot x | M \cdot x \rangle_A = \lambda_1 \cdot x_1^2 + \dots + \lambda_n \cdot x_n^2$$

da cui segue subito la tesi. \square

Un criterio utile nella pratica è fornito dal seguente:

Teorema 10.2.4. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ simmetrica, sia A_k la sottomatrice delle prime k righe e k colonne di A , e sia $d_k = \det(A_k)$. Allora $\langle \cdot | \cdot \rangle_A$ è definito positivo se e soltanto se $\det(A_k) > 0$ per $k = 1, \dots, n$.*

Dimostrazione. Se $\langle \cdot | \cdot \rangle_A$ è definito positivo ogni A_k definisce un prodotto scalare su \mathbb{R}^k , dunque ha autovalori positivi, e la conclusione segue poiché d_k è il prodotto degli autovalori di A_k .

Proviamo il viceversa dimostrando per induzione su $k = 1, \dots, n$ che A_k ha autovalori tutti positivi, il che fornisce la conclusione desiderata per $k = n$. Per $k = 1$ l'affermazione è ovvia. Sia $k \geq 1$ e supponiamo la conclusione vera per A_k . Poiché $d_{k+1} = \det(A_{k+1}) > 0$, la A_{k+1} ha (contando la molteplicità) un numero pari di autovalori negativi. Supponiamo per assurdo che ce ne siano almeno due. Poiché A_{k+1} è diagonalizzabile, definendo W come il sottospazio di \mathbb{R}^{k+1} che è somma di tutti gli autospazi di A_{k+1} relativi agli autovalori negativi, si ha $\dim_{\mathbb{R}}(W) \geq 2$ e $\langle w|w \rangle_{A_{k+1}} < 0$ per ogni $w \in W$ con $w \neq 0$. Osserviamo ora che $\mathbb{R}^k \times \{0\}$ e W sono sottospazi di \mathbb{R}^{k+1} di dimensioni rispettivamente k e almeno 2, da cui segue che la loro intersezione ha dimensione almeno $k+2 - (k+1) = 1$, dunque contiene un vettore $w = \begin{pmatrix} u \\ 0 \end{pmatrix}$ non nullo. Abbiamo allora $\langle w|w \rangle_{A_{k+1}} = \langle u|u \rangle_{A_k}$, e per ipotesi induttiva $\langle u|u \rangle_{A_k} > 0$, mentre $\langle w|w \rangle_{A_{k+1}} < 0$. Abbiamo trovato un assurdo e la dimostrazione è completa. \square

Il seguente risultato generalizza il precedente e si dimostra in modo simile (ma un po' più laborioso):

Teorema 10.2.5. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ simmetrica, sia A_k la sottomatrice delle prime k righe e k colonne di A , e sia $d_k = \det(A_k)$. Poniamo $d_0 = 1$ e supponiamo che d_1, \dots, d_{k-1} siano non nulli. Allora i segni degli autovalori di A (contati con la loro molteplicità) sono i segni di*

$$\frac{d_1}{d_0}, \frac{d_2}{d_1}, \dots, \frac{d_n}{d_{n-1}}.$$

10.2.1 Teorema spettrale e matrice hessiana

Oltre ad avere una dimostrazione di natura almeno in parte analitica, il teorema spettrale ha anche un'importante applicazione all'analisi in più variabili. Sia Ω un aperto di \mathbb{R}^n e sia $f : \Omega \rightarrow \mathbb{R}$ una funzione con tutte le derivate parziali continue fino all'ordine 2 (vedi il Sottoparagrafo 9.3.1). Ricordiamo che abbiamo definito gradiente e matrice hessiana di f in un punto x come

$$\begin{aligned} \text{grad}_x(f) &= \left(\frac{\partial f}{\partial x_j}(x) \right)_{j=1, \dots, n} \in \mathbb{R}^n, \\ H_x(f) &= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i, j=1, \dots, n} \in \mathcal{M}_{n \times n}(\mathbb{R}) \end{aligned}$$

e che la Proposizione 9.3.7 garantisce che $H_x(f)$ è simmetrica. Abbiamo allora la seguente estensione della Proposizione 9.3.9:

Proposizione 10.2.6. *Nelle ipotesi precedenti, sia $\bar{x} \in \Omega$.*

- *Se \bar{x} è un punto di minimo locale per f allora $\text{grad}_{\bar{x}}(f) = 0$ e $H_{\bar{x}}(f)$ ha autovalori non negativi.*
- *Se $\text{grad}_{\bar{x}}(f) = 0$ e $H_{\bar{x}}(f)$ ha autovalori positivi allora \bar{x} è un punto di minimo locale per f .*

Dimostrazione. La formula di approssimazione di Taylor del secondo ordine garantisce che

$$f(\bar{x} + v) = f(\bar{x}) + \langle \text{grad}_{\bar{x}}(f) | v \rangle_{\mathbb{R}^n} + \frac{1}{2} \langle v | v \rangle_{H_{\bar{x}}(f)} + o(\|v\|^2)$$

dalla quale abbiamo già dedotto che se \bar{x} è un punto di minimo locale allora $\text{grad}_{\bar{x}}(f) = 0$. Inoltre, grazie al teorema spettrale, in qualsiasi punto \bar{x} tale che $\text{grad}_{\bar{x}}(f) = 0$ si ha, a meno di cambiamenti ortogonali delle coordinate, che

$$f(\bar{x} + h) = f(\bar{x}) + \frac{1}{2} \cdot (\lambda_1 \cdot h_1^2 + \dots + \lambda_n \cdot h_n^2) + o(\|h\|^2)$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di $H_{\bar{x}}(f)$. La conclusione è allora immediata. \square

10.2.2 Versione complessa del teorema spettrale

Tornando alla diagonalizzabilità delle matrici, passiamo ora al caso complesso definendo una $A \in \mathcal{M}_{n \times n}(\mathbb{C})$ normale se $A^* \cdot A = A \cdot A^*$. Abbiamo la seguente estensione del teorema spettrale:

Teorema 10.2.7. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{C})$. Allora A è normale se e soltanto se si diagonalizza tramite una matrice unitaria.*

Dimostrazione. Si procede esattamente come per il teorema spettrale reale. Bisogna provare che

$$A \text{ è normale} \Leftrightarrow \text{esiste } U \text{ con } U^* = U^{-1} \text{ e } U^{-1} \cdot A \cdot U \text{ diagonale.}$$

L'implicazione \Leftarrow è poco interessante e facile: se $U^{-1} \cdot A \cdot U = D$ con D diagonale e $U^* = U^{-1}$ allora $A = U \cdot D \cdot U^*$ e si ottiene

$$\begin{aligned} A^* \cdot A &= U \cdot D^* \cdot U^* \cdot U \cdot D \cdot U^* = U \cdot \bar{D} \cdot D \cdot U^* \\ A \cdot A^* &= U \cdot D \cdot U^* \cdot U \cdot D^* \cdot U^* = U \cdot D \cdot \bar{D} \cdot U^* \end{aligned}$$

e la conclusione che A è normale segue subito, poiché $\overline{D} \cdot D = D \cdot \overline{D}$.

Per provare l'implicazione \Rightarrow scegliamo invece una base ortonormale della quale il primo vettore sia un autovalore di A rispetto a qualche autovalore λ_1 . Ciò conduce a una matrice U_1 unitaria tale che

$$U_1^{-1} \cdot A \cdot U_1 = \begin{pmatrix} \lambda_1 & z_1 \\ 0 & A_1 \end{pmatrix}.$$

Usando il fatto che $U_1^{-1} = U_1^*$ e il fatto che A è normale si verifica subito che anche $U_1^{-1} \cdot A \cdot U_1$ è normale. Imponendo questa condizione alla matrice $\begin{pmatrix} \lambda_1 & z_1 \\ 0 & A_1 \end{pmatrix}$ si trova che $z_1 = 0$ e A_1 è normale, dunque si conclude facilmente per induzione. \square

Questo teorema si applica in particolare alle matrici hermitiane e conduce al seguente risultato, che si dimostra essenzialmente come nel caso reale:

Teorema 10.2.8. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{C})$ hermitiana. Allora gli autovalori di A sono reali. Inoltre sono fatti equivalenti:*

- $\langle \cdot | \cdot \rangle_A$ è definito positivo;
- A ha autovalori tutti positivi;
- Se A_k è la sottomatrice delle prime k righe e k colonne di A , si ha $\det(A_k) > 0$ per $k = 1, \dots, n$.

10.2.3 Forma canonica di matrici antisimmetriche e ortogonali

Il Teorema 10.2.7 si applica anche alle matrici antisimmetriche reali, alle antihermitiane (quelle opposte alla propria aggiunta), alle ortogonali reali, e alle unitarie. È intanto facile il seguente:

Lemma 10.2.9. *Se $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è antisimmetrica oppure $A \in \mathcal{M}_{n \times n}(\mathbb{C})$ è antihermitiana allora gli autovalori di A sono immaginari puri. Se $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ è ortogonale oppure $A \in \mathcal{M}_{n \times n}(\mathbb{C})$ è unitaria allora gli autovalori di A sono di modulo unitario.*

Dimostrazione. Per A antihermitiana (in particolare, se 'e antisimmetrica reale), se $\lambda \in \mathbb{C}$ è un autovalore e $z \in \mathbb{C}^n$ è un autovettore relativo, si ha

$$\begin{aligned} \lambda \cdot \|z\|_{\mathbb{C}^n}^2 &= \langle \lambda \cdot z | z \rangle_{\mathbb{C}^n} = \langle A \cdot z | z \rangle_{\mathbb{C}^n} = \langle z | A^* \cdot z \rangle_{\mathbb{C}^n} \\ &= \langle z | -A \cdot z \rangle_{\mathbb{C}^n} = -\langle z | \lambda \cdot z \rangle_{\mathbb{C}^n} = -\bar{\lambda} \cdot \|z\|_{\mathbb{C}^n}^2 \end{aligned}$$

dunque $\bar{\lambda} = -\lambda$, da cui la conclusione. Se invece A è unitaria (in particolare, se è ortogonale reale) si ha

$$|\lambda|^2 \cdot \|z\|_{\mathbb{C}^n}^2 = \langle \lambda \cdot z | \lambda \cdot z \rangle_{\mathbb{C}^n} = \langle A \cdot z | A \cdot z \rangle_{\mathbb{C}^n} = \langle z | A^* \cdot A \cdot z \rangle_{\mathbb{C}^n} = \|z\|_{\mathbb{C}^n}^2$$

dunque $|\lambda| = 1$. □

Per le antihermitiane e le unitarie il lemma precedente e il Teorema 10.2.7 comportano subito la:

Proposizione 10.2.10. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{C})$.*

- *Se A è antihermitiana allora esiste $M \in \mathcal{M}_{n \times n}(\mathbb{C})$ unitaria tale che $M^{-1} \cdot A \cdot M$ è una matrice diagonale con coefficienti immaginari puri sulla diagonale.*
- *Se A è unitaria allora esiste $M \in \mathcal{M}_{n \times n}(\mathbb{C})$ unitaria tale che $M^{-1} \cdot A \cdot M$ è una matrice diagonale con coefficienti di modulo unitario sulla diagonale.*

Con qualche sforzo in più si trova invece per le antisimmetriche e le ortogonali il seguente:

Teorema 10.2.11. *Sia $A \in \mathcal{M}_{n \times n}(\mathbb{R})$.*

- *Se A è antisimmetrica allora esiste $M \in \mathcal{M}_{n \times n}(\mathbb{R})$ ortogonale tale che $M^{-1} \cdot A \cdot M$ è una matrice diagonale a blocchi con blocchi del tipo (0) oppure $\begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix}$ con $\lambda \neq 0$.*
- *Se A è ortogonale allora esiste $M \in \mathcal{M}_{n \times n}(\mathbb{R})$ ortogonale tale che $M^{-1} \cdot A \cdot M$ è una matrice diagonale a blocchi con blocchi del tipo (1) oppure (-1) oppure $\begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}$ con $\vartheta \in \mathbb{R}$.*

Non forniremo tutti i dettagli della dimostrazione di questo risultato, ma diamo almeno l'idea nel caso di A ortogonale (il caso di A antisimmetrica è analogo e anche più semplice). Ad autovalori reali ± 1 corrispondono autovettori reali, da cui i blocchi ± 1 sulla diagonale. Supponiamo di avere invece un autovalore $e^{i\vartheta} = \cos \vartheta + i \sin \vartheta$ non reale. Sappiamo che allora anche $e^{-i\vartheta} = \cos \vartheta - i \sin \vartheta$ è un autovalore di A , e prendiamo un autovettore $z \in \mathbb{C}^n$ a esso relativo. Scrivendo $z = x + iy$ con $x, y \in \mathbb{R}^n$ abbiamo allora

$$\begin{aligned} A \cdot (x + iy) &= (\cos \vartheta - i \sin \vartheta) \cdot (x + iy) \\ \Rightarrow A \cdot x + i(A \cdot y) &= \cos \vartheta \cdot x + \sin \vartheta \cdot y + i(-\sin \vartheta \cdot x + \cos \vartheta \cdot y) \\ \Rightarrow \begin{cases} A \cdot x = \cos \vartheta \cdot x + \sin \vartheta \cdot y \\ A \cdot y = -\sin \vartheta \cdot x + \cos \vartheta \cdot y. \end{cases} \end{aligned}$$

Dunque, ammesso che x e y siano linearmente indipendenti, sul sottospazio da essi generato la A agisce appunto come il blocco quadrato $\begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}$ previsto dall'enunciato. Possiamo ora vedere che x e y sono ortogonali e hanno la stessa norma, da cui segue che in effetti sono linearmente indipendenti e anzi che a meno di normalizzazione costituiscono una base ortonormale del sottospazio da essi generato (in accordo con il fatto che, nell'enunciato, è una matrice M ortogonale a coniugare A a una matrice diagonale a blocchi). Infatti è immediato vedere che $x - iy$ è un autovettore di A relativo all'autovalore $e^{i\vartheta}$, che è distinto da $e^{-i\vartheta}$, dunque per il Teorema 10.2.7 si ha $\langle x + iy | x - iy \rangle_{\mathbb{C}^n} = 0$, da cui

$$({}^t x + i {}^t y) \cdot (x + iy) = 0 \Rightarrow {}^t x \cdot x - {}^t y \cdot y + 2i {}^t x \cdot y = 0 \Rightarrow \begin{cases} \|x\| = \|y\| \\ \langle x | y \rangle_{\mathbb{R}^n} = 0. \end{cases}$$

Se gli autovalori di A sono tutti distinti l'argomento appena presentato dimostra completamente il Teorema 10.2.11 (per A ortogonale), mentre nel caso di autovalori con molteplicità maggiore di 1 serve qualche ulteriore considerazione.

Capitolo 11

Coniche metriche e affini

In questo capitolo incontreremo la definizione di alcuni luoghi geometrici nel piano, detti *coniche non degeneri*, descrivendone i prototipi dal punto di vista *metrico* (cioè a meno di movimenti rigidi del piano) e scoprendo che sono definiti da equazioni polinomiali di secondo grado. Vedremo poi che con una piccola restrizione qualsiasi sottoinsieme del piano definito da un'equazione polinomiale di secondo grado è una conica non degenera, e mostreremo come classificarlo a meno di *trasformazioni affini*. A quest'ultimo fine useremo un argomento del tutto elementare, rimandando una comprensione più concettuale della classificazione al Capitolo 13

11.1 Modelli metrici delle coniche

In questo paragrafo trattiamo brevemente le nozioni di isometria e di trasformazione affine del piano e introduciamo i modelli metrici dei luoghi di cui ci occuperemo nell'intero capitolo. Una trattazione più sistematica e dettagliata delle isometrie è contenuta nel Capitolo 16.

11.1.1 Piano euclideo, isometrie e trasformazioni affini

Nel piano euclideo la distanza tra due punti di coordinate (x_1, y_1) e (x_2, y_2) si calcola usando il teorema di Pitagora e vale

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Le *isometrie* del piano sono le bigezioni $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ che preservano la distanza. Geometricamente è chiaro che sono isometrie le trasformazioni dei tipi seguenti:

- traslazioni;
- rotazioni intorno a punti;
- riflessioni rispetto a rette;
- composizioni dei tipi precedenti.

In realtà vedremo nel Capitolo 16 che questo elenco contiene tutti i tipi di isometria del piano. Dal punto di vista analitico è possibile dimostrare (vedi ancora il Capitolo 16) che ogni isometria di \mathbb{R}^2 , vista come cambiamento di coordinate, è una trasformazione del tipo

$$\begin{pmatrix} X \\ Y \end{pmatrix} = M \cdot \begin{pmatrix} x \\ y \end{pmatrix} + v_0$$

con M matrice ortogonale e $v_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ vettore di traslazione. Dando per buono questo fatto possiamo ora facilmente provare la:

Proposizione 11.1.1. *Un'isometria di \mathbb{R}^2 è la composizione di una rotazione o di una riflessione con una traslazione.*

Dimostrazione. È immediato verificare che ogni matrice ortogonale 2×2 è di uno dei seguenti tipi:

$$R_\vartheta = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}, \quad S_\vartheta = \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{pmatrix}.$$

Sappiamo già che R_ϑ è la rotazione di angolo ϑ intorno a 0; essendo $\text{tr}(S_\vartheta) = 0$ e $\det(S_\vartheta) = -1$ il Teorema 10.2.11 comporta invece che S_ϑ è coniugata tramite una matrice ortogonale a $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, dunque è la riflessione rispetto a una retta per 0; in realtà con le formule trigonometriche di duplicazione è facile vedere che

$$S_\vartheta \cdot \begin{pmatrix} \cos \frac{\vartheta}{2} \\ \sin \frac{\vartheta}{2} \end{pmatrix} = \begin{pmatrix} \cos \frac{\vartheta}{2} \\ \sin \frac{\vartheta}{2} \end{pmatrix}, \quad S_\vartheta \cdot \begin{pmatrix} \sin \frac{\vartheta}{2} \\ -\cos \frac{\vartheta}{2} \end{pmatrix} = \begin{pmatrix} -\sin \frac{\vartheta}{2} \\ \cos \frac{\vartheta}{2} \end{pmatrix}$$

dunque S_ϑ rappresenta la riflessione rispetto alla retta per 0 di inclinazione $\frac{\vartheta}{2}$. \square

Come già anticipato, oltre alle isometrie impiegheremo un tipo più generale di cambiamento di coordinate, che chiameremo *affine*, dato dalla composizione di una bigezione lineare con una traslazione, dunque anch'esso descritto dalla formula

$$\begin{pmatrix} X \\ Y \end{pmatrix} = M \cdot \begin{pmatrix} x \\ y \end{pmatrix} + v_0$$

nella quale però si richiede soltanto che $\det(M) \neq 0$.

11.1.2 Equazioni metriche di circonferenza, ellisse, parabola e iperbole

Le coniche non degeneri sono particolari luoghi geometrici nel piano, definiti attraverso la nozione di distanza. Le descriveremo *a meno di isometrie*, cioè scriveremo le loro equazioni rispetto a opportuni sistemi di coordinate ottenuti con una trasformazione isometrica a partire dalle coordinate originali.

Circonferenza La circonferenza è il luogo dei punti aventi una fissata distanza $r > 0$ da un punto fissato, detto *centro*. Applicando un'isometria (in questo caso, una traslazione) possiamo supporre che il centro sia l'origine, quindi l'equazione della circonferenza è

$$x^2 + y^2 = r^2.$$

Ellisse L'ellisse è l'insieme dei punti che hanno fissata somma delle distanze da due punti fissati, detti *fuochi*. Applicando la traslazione che porta nell'origine il punto medio del segmento che congiunge i fuochi, seguita da una rotazione intorno all'origine, possiamo supporre che i fuochi siano $(-h, 0)$ e $(h, 0)$ per qualche $h > 0$. Se $2k > 0$ è il valore costante sull'ellisse della somma delle distanze dai fuochi, vediamo che $k > h$, in quanto la minima possibile somma delle distanze da $(-h, 0)$ e $(h, 0)$ per un punto del piano si ha sul segmento $[-h, h] \times \{0\}$ e vale $2h$. L'equazione dell'ellisse è allora

$$\sqrt{(x+h)^2 + y^2} + \sqrt{(x-h)^2 + y^2} = 2k$$

dalla quale con semplici manipolazioni algebriche si ottiene l'equazione equivalente

$$\frac{x^2}{k^2} + \frac{y^2}{k^2 - h^2} = 1$$

che possiamo riscrivere per opportuni $a > b > 0$ come

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

e che consideriamo l'equazione canonica metrica di un'ellisse. I numeri a, b sono detti *semiassi* perché gli assi coordinati sono assi di simmetria per l'ellisse e i segmenti interni all'ellisse su di essi hanno lunghezze $2a$ e $2b$.

Iperbole L'iperbole è l'insieme dei punti che hanno fissata differenza (in valore assoluto) delle distanze da due punti fissati, detti *fuochi*. Come per l'ellisse possiamo supporre a meno di isometrie che i fuochi siano $(-h, 0)$ e $(h, 0)$ per qualche $h > 0$. Se $2k > 0$ è il valore costante sull'iperbole della differenza delle distanze dai fuochi, vediamo che $k < h$, in quanto la massima possibile differenza delle distanze da $(-h, 0)$ e $(h, 0)$ per un punto del piano si ha sull'asse delle ascisse fuori dal segmento $(-h, h) \times \{0\}$ e vale $2h$. L'equazione dell'iperbole è allora

$$\left| \sqrt{(x+h)^2 + y^2} - \sqrt{(x-h)^2 + y^2} \right| = 2k.$$

Sviluppando i calcoli si ottiene esattamente la medesima equazione trovata per l'ellisse (infatti i segni per cui differiscono le equazioni iniziali svaniscono con i due elevamenti al quadrato), che conviene però scrivere come

$$\frac{x^2}{k^2} - \frac{y^2}{h^2 - k^2} = 1$$

e che possiamo riscrivere per opportuni $a, b > 0$ come

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

che consideriamo l'equazione canonica metrica di un'iperbole. I numeri a, b sono ancora detti *semiassi*.

Parabola La parabola è l'insieme dei punti equidistanti da una retta fissata e da un punto fissato fuori dalla retta. Traslando l'origine degli assi nel punto medio del segmento di minima distanza tra il punto e la retta ed eseguendo una rotazione, vediamo che a meno di isometria possiamo supporre che la retta abbia equazione $y = -h$ e il punto sia $(0, h)$ per qualche $h > 0$. L'equazione della parabola è dunque

$$|y+h| = \sqrt{x^2 + (y-h)^2}$$

da cui elevando al quadrato si trova $4hy = x^2$ ovvero

$$y = ax^2$$

per qualche $a > 0$, che consideriamo l'equazione canonica metrica della parabola.

Sezioni coniche L'equazione $x^2 + y^2 = z^2$ definisce in \mathbb{R}^3 un cono. Infatti $x^2 + y^2$ rappresenta il quadrato della distanza del punto (x, y, z) dall'asse z , dunque il luogo di equazione $x^2 + y^2 = z^2$ si ottiene ruotando intorno all'asse z l'unione delle rette $x \pm z = 0$ nel piano xz (o anche una sola di esse), da cui risulta appunto un cono. Si può verificare che le coniche metriche sopra descritte sono tutte e sole le curve che si ottengono intersecando il cono $x^2 + y^2 = z^2$ con un piano che non contenga l'origine. Ad esempio se intersechiamo con $z = 1$ otteniamo la circonferenza unitaria $x^2 + y^2 = 1$ (una particolare ellisse). Se invece intersechiamo con $y = 1$ troviamo l'iperbole $z^2 - x^2 = 1$. E infine se intersechiamo con il piano $x + z = \sqrt{2}$ sul quale poniamo coordinate $Y = \frac{1}{\sqrt{2}}(z - x)$ e $X = y$ otteniamo la parabola $Y = X^2$; le coordinate X, Y sul piano $x + z = \sqrt{2}$ danno un sistema di riferimento ortogonale perché la formula

$$\begin{cases} X = y \\ Y = \frac{1}{\sqrt{2}}(z - x) \\ Z = \frac{1}{\sqrt{2}}(z + x) \end{cases}$$

definisce un'isometria di \mathbb{R}^3 .

11.2 Matrice di una conica e classificazione affine delle coniche non degeneri

Consideriamo ora le coniche sopra descritte solo a meno di trasformazioni affini. Abbiamo allora le equazioni canoniche $x^2 + y^2 = 1$ (ellisse), $x^2 - y^2 = 1$ (iperbole) e $y = x^2$ (parabola). Convieni anche osservare che con un cambio di coordinate l'equazione dell'iperbole prende la forma $xy = 1$, che può anch'essa essere impiegata come equazione canonica. Notiamo inoltre che l'insieme vuoto \emptyset può essere definito (ad esempio) dall'equazione $x^2 + y^2 + 1 = 0$. Osservando che tutte le equazioni appena descritte sono polinomiali di secondo grado, ci poniamo ora il problema di classificare dal punto di vista affine il luogo di zeri di un qualsiasi polinomio di secondo grado, ovvero un luogo di equazione

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0.$$

La ragione della notazione scelta per i coefficienti è che l'equazione può essere riscritta come

$${}^t \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$$

L'utilità di questa scrittura emergerà pienamente nell'ambito dello studio delle coniche tramite la geometria proiettiva (vedi il Capitolo 13). Ci limitiamo qui invece a una verifica "a mano" del fatto che si può classificare il luogo definito dall'equazione precedente impiegando tecniche di algebra lineare. Una spiegazione più concettuale della ragione per cui queste tecniche funzionano è

rimandata al futuro. Poniamo dunque $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$ e proponiamoci

di classificare dal punto di vista affine il luogo definito dall'equazione di secondo grado cui è associata la matrice A . A questo fine per $i = 1, 2, 3$ definiamo d_i come il determinante della sottomatrice quadrata di A formata dalle prime i righe e dalle prime i colonne. Dunque d_1 è il coefficiente di x^2 , poi d_2 è il determinante della parte di A relativa alla porzione puramente quadratica dell'equazione, e infine d_3 non è altro che il determinante di A .

Teorema 11.2.1. *Sia \mathcal{L} il luogo in \mathbb{R}^2 definito da un'equazione polinomiale di secondo grado a cui è associata una matrice simmetrica $A \in \mathcal{M}_{3 \times 3}(\mathbb{R})$. Se $\det(A) \neq 0$ allora \mathcal{L} è l'insieme vuoto oppure esiste una trasformazione affine che manda \mathcal{L} in uno dei luoghi di equazione $x^2 + y^2 = 1$ (ellisse), oppure $x^2 - y^2 = 1$ (iperbole), oppure $y = x^2$ (parabola). Più esattamente, il seguente elenco esaurisce tutte le possibilità e fornisce una classificazione affine di \mathcal{L} :*

- (1) *Se $d_2 > 0$ e $d_1 \cdot d_3 > 0$ allora \mathcal{L} è vuoto;*
- (2) *Se $d_2 > 0$ e $d_1 \cdot d_3 < 0$ allora \mathcal{L} è un'ellisse;*
- (3) *Se $d_2 < 0$ allora \mathcal{L} è un'iperbole;*
- (4) *Se $d_2 = 0$ allora \mathcal{L} è una parabola.*

Sottolineiamo che nei quattro casi descritti alla fine dell'enunciato precedente si dà sempre come ipotesi che d_3 sia non nullo, nel qual caso diremo che \mathcal{L} è una *conica non degenera*. Per dimostrare il Teorema 11.2.1 iniziamo con la seguente:

Proposizione 11.2.2. *I casi (1)-(4) del Teorema 11.2.1 coprono tutte le possibilità con $d_3 \neq 0$. Inoltre le matrici di due equazioni rientrano nel medesimo caso se le due equazioni sono ottenute l'una dall'altra tramite moltiplicazione per una costante non nulla oppure tramite una trasformazione affine.*

Dimostrazione. Per provare la prima affermazione basta vedere che se $d_2 > 0$, ovvero $a_{11} \cdot a_{22} - a_{12}^2 > 0$, allora $d_1 = a_{11}$ non può annullarsi, il che è ovvio.

Per la seconda affermazione, vediamo facilmente che il caso in cui una matrice rientra non cambia se moltiplichiamo l'equazione per una costante $\lambda \neq 0$ (per $\lambda < 0$ nei casi (1) e (2) bisogna notare che d_1 e d_3 cambiano segno entrambi). Esaminiamo ora l'effetto sulla matrice A di un'equazione di una trasformazione affine del tipo $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto M \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Per iniziare possiamo riscrivere questa trasformazione come

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} M & v_0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \text{con} \quad v_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Ponendo $N = \begin{pmatrix} M & v_0 \\ 0 & 1 \end{pmatrix}$ vediamo ora che N è invertibile, dato che lo è M ;

inoltre poiché l'equazione corrispondente ad A è data da $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ vediamo facilmente che la matrice della nuova equazione dopo avere applicato la trasformazione affine è $A' = {}^t N \cdot A \cdot N$. Inoltre se $A = \begin{pmatrix} Q & \ell \\ {}^t \ell & k \end{pmatrix}$ si ha $A' = \begin{pmatrix} {}^t M \cdot Q \cdot M & \ell' \\ {}^t \ell' & k' \end{pmatrix}$. Abbiamo allora:

- $d'_2 = \det({}^t M) \cdot d_2 \cdot \det(M) = (\det(M))^2 \cdot d_2$ è concorde con d_2 ;
- $d'_3 = \det({}^t N) \cdot d_3 \cdot \det(N) = (\det(N))^2 \cdot d_3$ è concorde con d_3 .

Per concludere resta solo da verificare che se $d_2 > 0$ allora $d'_1 = a'_{11}$ è concorde con $d_1 = a_{11}$ (il quale è concorde con a_{22} , poiché $d_2 = a_{11}a_{22} - a_{12}^2 > 0$). Cominciamo a notare che $a'_{11} = a_{11}m_{11}^2 + 2a_{12}m_{11}m_{21} + a_{22}m_{21}^2$ e osserviamo che m_{11} e m_{21} non possono annullarsi entrambi; dividendo per il quadrato di quello non nullo vediamo che per qualche $t \in \mathbb{R}$ il coefficiente a'_{11} è concorde con $a_{11}t^2 + 2a_{12}t + a_{22}$ oppure con $a_{22}t^2 + 2a_{12}t + a_{11}$, e dunque è concorde con a_{11} e con a_{22} , dato che questi due polinomi in t hanno discriminante $-4d_2 < 0$. \square

Dimostrazione del Teorema 11.2.1. L'argomento dipende in modo essenziale dalla Proposizione 11.2.2. Infatti in ciascuno dei casi (1)-(4) modifichiamo l'equazione o moltiplicando per un numero non nullo o con una trasformazione affine e *sfruttiamo* il fatto che la nuova matrice continua a ricadere nel medesimo caso (1)-(4) per provare che la nuova equazione è del tipo desiderato

oppure che possiamo operare un'ulteriore trasformazione per giungere al tipo desiderato.

Iniziamo dal caso (4), ovvero $d_2 = 0$, e notiamo che non si può avere $a_{11} = a_{22} = 0$, altrimenti avremmo anche $a_{12} = 0$ e le prime due colonne di A sarebbero linearmente dipendenti, contro l'ipotesi $\det(A) \neq 0$. A meno di scambiare x e y possiamo allora supporre $a_{11} \neq 0$, e anzi $a_{11} = 1$ dividendo l'equazione per a_{11} . Dunque $a_{22} = a_{12}^2$, e l'equazione diviene

$$x^2 + 2a_{12}xy + a_{12}^2y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$$

ovvero

$$(x + a_{12}y)^2 = -2a_{13}x - 2a_{23}y - a_{33}.$$

Questa equazione diviene quella della parabola ponendo

$$\begin{cases} \xi = x + a_{12}y \\ \eta = -2a_{13}x - 2a_{23}y - a_{33}, \end{cases}$$

ma per verificare che questo cambiamento di coordinate corrisponda a una trasformazione affine lecita dobbiamo provare che $\det \begin{pmatrix} 1 & a_{12} \\ -2a_{13} & -2a_{23} \end{pmatrix}$ è non nullo, cioè che $a_{12}a_{13} - a_{23} \neq 0$, e ciò segue dall'invertibilità della matrice A originale, così come di quella ottenuta dopo il primo passaggio, poiché

$$\begin{aligned} \det \begin{pmatrix} 1 & a_{12} & a_{13} \\ a_{12} & a_{12}^2 & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} &= a_{13}(a_{12}a_{23} - a_{12}^2a_{13}) - a_{23}(a_{23} - a_{12}a_{13}) \\ &= -(a_{12}a_{13} - a_{23})^2. \end{aligned}$$

Passiamo al caso (3), ovvero $d_2 < 0$. Distinguiamo dapprima il sottocaso $a_{11} = a_{22} = 0$; possiamo allora supporre $a_{12} = 1$ ottenendo l'equazione

$$2xy + 2a_{13}x + 2a_{23}y + a_{33} = 0$$

ovvero

$$(2x + 2a_{23})(y + a_{13}) = 2a_{23}a_{13} - a_{33},$$

la quale certamente definisce un'iperbole a meno di trasformazioni affini purché il secondo membro sia non nullo, il che è vero poiché

$$0 \neq \det(A) = \det \begin{pmatrix} 0 & 1 & a_{13} \\ 1 & 0 & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = 2a_{23}a_{13} - a_{33}.$$

Liquidato il sottocaso $a_{11} = a_{22} = 0$, possiamo come sopra ricondurci ad $a_{11} = 1$. L'equazione è allora

$$x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0,$$

ovvero

$$(x + a_{12}y)^2 - (a_{12}^2 - a_{22})y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$$

dove $a_{12}^2 - a_{22} = -d_2 > 0$. Sostituendo $x + a_{12}y$ con x e $\sqrt{-d_2}y$ con y , il che certamente corrisponde a una trasformazione affine, ci riconduciamo a

$$x^2 - y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$$

(con nuovi coefficienti a_{ij} , ovviamente), dunque a

$$(x + a_{13})^2 - (y - a_{23})^2 = a_{13}^2 - a_{23}^2 - a_{33}$$

la quale definisce un'iperbole perché

$$0 \neq d_3 = \det \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & -1 & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = a_{13}^2 - a_{23}^2 - a_{33}.$$

Restano infine i casi (1) e (2), con $d_2 > 0$. Supponiamo allora $a_{11} = 1$, ottenendo l'equazione

$$(x + a_{12}y)^2 + (a_{22} - a_{12}^2)y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0,$$

e ancora

$$\begin{aligned} x^2 + y^2 + 2a_{13}x + 2a_{23}y + a_{33} &= 0, \\ (x + a_{13})^2 + (y + a_{23})^2 &= a_{33} - a_{13}^2 - a_{23}^2. \end{aligned}$$

Abbiamo ora

$$d_3 = \det \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 1 & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = a_{33} - a_{13}^2 - a_{23}^2$$

dunque nel caso (1), con $d_3 > 0$ (infatti $d_1 = 1$) l'equazione definisce l'insieme vuoto, mentre nel caso (2) definisce un'ellisse. \square

Per concludere il capitolo menzioniamo i tipi affini delle coniche degeneri, cioè dei luoghi definiti da equazioni polinomiali di secondo grado con associata matrice avente determinante nullo, escludendo però il caso in cui l'equazione si riduca a una lineare:

- due rette parallele, di equazione canonica $x^2 = 1$;
- due rette incidenti, di equazione canonica $xy = 0$;
- una singola retta (doppia), di equazione canonica $x^2 = 0$;
- un singolo punto, di equazione canonica $x^2 + y^2 = 0$.

La classificazione di una conica degenera si ottiene sempre facilmente cercando un cambiamento affine di coordinate che porti l'equazione in una delle quattro canoniche appena scritte.

Capitolo 12

Spazi proiettivi

In questo capitolo vedremo come lo spazio ordinario n -dimensionale \mathbb{R}^n si possa estendere a un oggetto più grande, lo *spazio proiettivo* n -dimensionale. Esso consente una comprensione molto profonda di fatti già incontrati (come la classificazione affine delle coniche non degeneri) nonché di nuovi aspetti della geometria di \mathbb{R}^n . L'idea fondamentale dell'estensione di \mathbb{R}^n allo spazio proiettivo è quella di aggiungere a \mathbb{R}^n i suoi "punti all'infinito". La formalizzazione di questa idea necessita tuttavia del richiamo di alcuni preliminari e ha alcune sottigliezze tecniche.

12.1 Relazioni di equivalenza e insiemi quoziente

Se X è un insieme, una *relazione* R su X è un sottoinsieme del prodotto cartesiano $X \times X$. Per sottolineare il fatto che il significato del termine corrisponde a quello del linguaggio naturale, invece che $(x, y) \in R$ scriveremo sempre xRy .

Esempio 12.1.1. Se $X = \{1, 2, 3, 4\}$ una relazione R su X è il sottoinsieme di $X \times X$ dato da $R = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$. Tuttavia osserviamo facilmente che $(x, x') \in R$ se e soltanto se $x < x'$, dunque possiamo dare alla relazione R il significato "è minore di", e facendo questo è ovviamente più naturale scrivere $1R3$ invece che $(1, 3) \in R$.

Per procedere ci servirà la:

Definizione 12.1.2. Una relazione R su un insieme X si dice:

- *riflessiva* se xRx per ogni $x \in X$;
- *simmetrica* se xRy implica yRx ;

- *transitiva* se xRy e yRz implicano xRz
- *di equivalenza* se è riflessiva, simmetrica e transitiva.

Anche se non ci servirà, per completezza ricordiamo la:

Definizione 12.1.3. Una relazione R su un insieme X si dice:

- *antisimmetrica* se xRy e yRx implicano $y = x$;
- *d'ordine* se è riflessiva, antisimmetrica e transitiva.

Esempio 12.1.4. Se X è l'insieme dei sottoinsiemi finiti di \mathbb{N} , la relazione "essere contenuto in", ovvero \subset , è una relazione d'ordine. Osserviamo che non tutti gli elementi di X sono confrontabili tra loro rispetto alla relazione \subset : ad esempio $\{1, 3\} \subset \{1, 2, 7\}$ e $\{1, 2, 7\} \subset \{1, 3\}$ sono entrambe false.

Tornando invece al filone principale diamo la:

Definizione 12.1.5. Fissata una relazione di equivalenza \sim su un insieme X , chiameremo *classe di equivalenza* di un elemento x di X l'insieme

$$[x] = \{y \in X : y \sim x\}.$$

Diremo poi *insieme quoziente* di X rispetto alla relazione \sim l'insieme X/\sim di tutte le classi di equivalenza, e *proiezione nel quoziente* la funzione (ovviamente surgettiva) da X in X/\sim data da $x \mapsto [x]$. Infine chiameremo *insieme di rappresentanti* un sottoinsieme di X che contenga un elemento in ciascuna classe di equivalenza di \sim .

È molto semplice vedere che le classi di equivalenza costituiscono una *partizione* di X , cioè che la loro unione è tutto X e che sono due a due disgiunte o coincidenti. Possiamo interpretare allora il quoziente X/\sim come l'oggetto ottenuto da X identificando tra loro tutti gli elementi di X appartenenti a ciascuna classe di equivalenza. Per definizione la restrizione della proiezione nel quoziente a un insieme Y di rappresentanti per \sim su X è una bigezione tra Y e il quoziente X/\sim . In alcuni casi tuttavia un insieme di rappresentanti non è adeguato a dare una rappresentazione significativa di un quoziente, e altre descrizioni possono essere preferibili.

Esempio 12.1.6. Sull'insieme $X = \mathbb{R}^{n+1} \setminus \{0\}$ consideriamo la relazione seguente: $x \sim y$ se esiste $\lambda > 0$ tale che $y = \lambda \cdot x$. È facile vedere che \sim è di equivalenza e che le classi di equivalenza sono le semirette aperte uscenti

dall'origine. La sfera n -dimensionale $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ è un insieme di rappresentanti per \sim , e in effetti offre una buona descrizione dell'insieme quoziente X/\sim : si può pensare che identificare tra loro tutti i punti di una semiretta uscente dall'origine corrisponda a collassare la semiretta sul suo punto a distanza 1 dall'origine.

Esempio 12.1.7. Sull'insieme X degli iscritti all'Università di Pisa consideriamo la relazione \sim data da $x \sim y$ se x e y compiono gli anni lo stesso giorno. Anch'essa è di equivalenza, e potremmo scegliere come insieme di rappresentanti un gruppo di 366 studenti nati in giorni diversi, ma la scelta sarebbe poco significativa. Una descrizione più sensata del quoziente X/\sim è invece "l'insieme dei 366 giorni dell'anno".

12.2 Definizione di spazio proiettivo

Incontriamo ora la nozione chiave del capitolo:

Definizione 12.2.1. Se \mathbb{K} è \mathbb{R} o \mathbb{C} e V è uno spazio vettoriale di dimensione almeno 1 su \mathbb{K} , consideriamo su $V \setminus \{0\}$ la relazione di equivalenza \sim così definita: $x \sim y$ se esiste $\lambda \in \mathbb{K}$ tale che $y = \lambda \cdot x$ (e notiamo che tale λ è automaticamente non nullo). Chiamiamo allora *spazio proiettivo* associato a V l'insieme quoziente $(V \setminus \{0\})/\sim$, che indichiamo con $\mathbb{P}(V)$. Chiamiamo in particolare *spazio proiettivo n -dimensionale su \mathbb{K}* l'insieme $\mathbb{P}^n(\mathbb{K}) = \mathbb{P}(\mathbb{K}^{n+1})$.

Le classi di equivalenza di \sim su $V \setminus \{0\}$ sono le rette per l'origine private dell'origine, ma c'è un'ovvia corrispondenza biunivoca tra l'insieme delle rette per l'origine in V e l'insieme delle rette per l'origine in V private dell'origine, dunque possiamo dire che $\mathbb{P}(V)$ è l'insieme delle rette per l'origine in V , cioè l'insieme ottenuto da V identificando tra loro tutti i punti di ciascuna retta. In particolare:

- $\mathbb{P}^0(\mathbb{R})$ consiste dell'unica retta in \mathbb{R} , dunque è un punto;
- $\mathbb{P}^1(\mathbb{R})$, che diremo *retta proiettiva reale*, è l'insieme delle rette in \mathbb{R}^2 ;
- $\mathbb{P}^2(\mathbb{R})$, che diremo *piano proiettivo reale*, è l'insieme delle rette in \mathbb{R}^3 ;
- $\mathbb{P}^3(\mathbb{R})$, che diremo *spazio proiettivo reale*, è l'insieme delle rette in \mathbb{R}^4 .

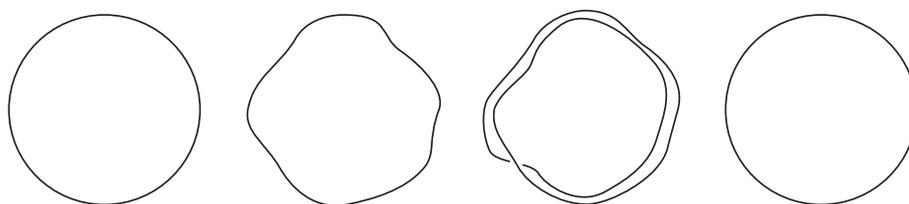


Figura 12.1: Identificando tra loro i punti antipodali su una circonferenza si ottiene una nuova circonferenza

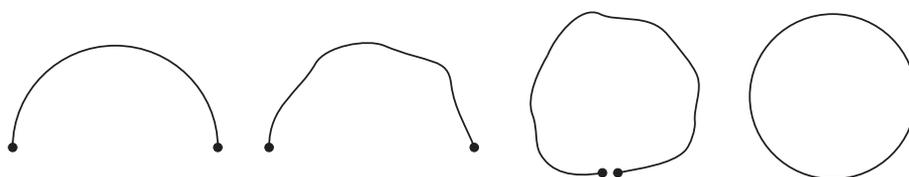


Figura 12.2: Identificando tra loro gli estremi di una semicirconferenza si ottiene una circonferenza

12.2.1 La retta proiettiva reale

Cerchiamo ora di dare una descrizione geometrica degli spazi proiettivi reali di dimensione più bassa, cominciando da $\mathbb{P}^1(\mathbb{R})$, che è l'insieme delle rette per l'origine in \mathbb{R}^2 . Se in \mathbb{R}^2 prendiamo la circonferenza $S^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$, vediamo che essa interseca ciascuna retta in due punti, tra loro antipodali, di S^1 ; dunque $\mathbb{P}^1(\mathbb{R})$ si ottiene da S^1 identificando tra loro le coppie di punti tra loro antipodali. Se pensiamo S^1 non come a una circonferenza rigida ma come a un elastico di gomma, è chiaro che possiamo deformarla fino a far coincidere ogni coppia di punti che erano originariamente antipodali, e il risultato è di nuovo una circonferenza, come illustrato in Fig. 12.1.

Un altro modo per giungere alla stessa descrizione di $\mathbb{P}^1(\mathbb{R})$ è il seguente. Invece che tutta la circonferenza S^1 prendiamo in considerazione la semicirconferenza $\alpha = \{x \in S^1 : x_2 \geq 0\}$ e osserviamo che essa interseca ogni retta per l'origine in \mathbb{R}^2 in un solo punto, tranne la retta orizzontale, che α interseca nei suoi due estremi $(-1, 0)$ e $(1, 0)$. Dunque possiamo pensare che $\mathbb{P}^1(\mathbb{R})$ si ottenga da α identificando tra loro i due estremi. Se di nuovo si considera α come un arco flessibile, è evidente che questa identificazione conduce a una circonferenza, come illustrato in Fig. 12.2.

Osservazione 12.2.2. Togliendo alla semicirconferenza α uno dei suoi estre-

mi, ad esempio $(-1, 0)$, si ottiene un insieme β che si può identificare a un intervallo semiaperto e che interseca ogni retta per l'origine in \mathbb{R}^2 in un solo punto, dunque β è in corrispondenza biunivoca con $\mathbb{P}^1(\mathbb{R})$. Tuttavia questa corrispondenza non è geometricamente significativa, perché trascura il fatto che per $\varepsilon > 0$ molto piccolo il punto $(-\sqrt{1-\varepsilon^2}, \varepsilon)$ rappresenta in $\mathbb{P}^1(\mathbb{R})$ un punto molto vicino a quello corrispondente a $(1, 0)$.

12.2.2 Il piano proiettivo reale

Procedendo come per $\mathbb{P}^1(\mathbb{R})$ vediamo facilmente che $\mathbb{P}^2(\mathbb{R})$ corrisponde naturalmente all'insieme ottenuto dalla sfera bidimensionale $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ identificando tra loro a coppie i punti antipodali. Raffigurare questo insieme però non è semplice, quindi procediamo subito con il passaggio successivo, prendendo la semisfera $E = \{x \in S^2 : x_3 \geq 0\}$. Abbiamo ora che ogni retta per l'origine in \mathbb{R}^3 incontra E in un solo punto, tranne le rette orizzontali (sul piano $x_3 = 0$), che incontrano E in due punti antipodali dell'equatore Γ che delimita E . Quindi $\mathbb{P}^2(\mathbb{R})$ si ottiene da E identificando tra loro i punti antipodali su Γ . Possiamo semplificare la figura (rendendola effettivamente bidimensionale) osservando che la proiezione ortogonale $(x_1, x_2, x_3) \mapsto (x_1, x_2, 0)$ sul piano $x_3 = 0$ fornisce una bigezione tra E e il disco bidimensionale $D = \{(x_1, x_2, 0) : x_1^2 + x_2^2 \leq 1\}$, e questa bigezione non ha alcun effetto su Γ . Dunque $\mathbb{P}^2(\mathbb{R})$ è l'oggetto che si ottiene dal disco D identificando tra loro i punti antipodali sul suo bordo Γ . Raffiguriamo allora $\mathbb{P}^2(\mathbb{R})$ come in Fig. 12.3, ovvero come un disco la cui circonferenza di bordo è suddivisa in due semicirconferenze, con una freccia su ciascuna delle due che ci ricorda che esse vanno punto per punto identificate tra loro. Non è possibile semplificare ulteriormente questa descrizione, anzi si può dimostrare (ma non è affatto facile) che $\mathbb{P}^2(\mathbb{R})$ non si può realizzare come sottoinsieme di \mathbb{R}^3 . Dalla descrizione segue comunque, come illustrato in Fig. 12.4, che un oggetto che si sposta dentro $\mathbb{P}^2(\mathbb{R})$ può tornare al punto di inizio del suo spostamento trovandosi a essere la copia speculare di quello che era quando è partito. Inoltre è possibile verificare (Fig. 12.5) che togliendo a $\mathbb{P}^2(\mathbb{R})$ un disco (senza bordo) si ottiene il nastro di Möbius, una superficie che si può realizzare in \mathbb{R}^3 e ha la fondamentale proprietà di essere *monolatera*: se si inizia a dipingerla di un colore da qualche parte, inevitabilmente si finisce per dipingerla tutta, compreso il punto che (apparentemente) sta dal lato opposto di quello da cui si è iniziato.

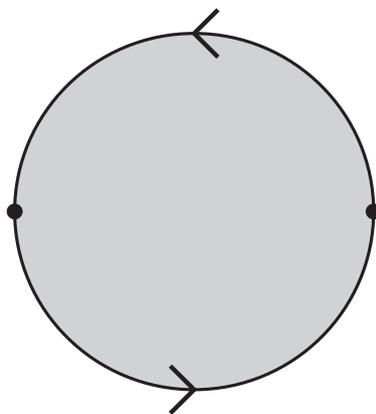


Figura 12.3: Il piano proiettivo reale

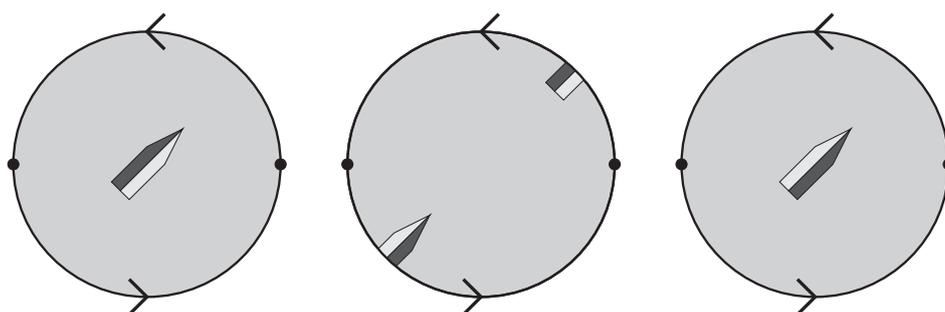


Figura 12.4: Un viaggio nel piano proiettivo reale

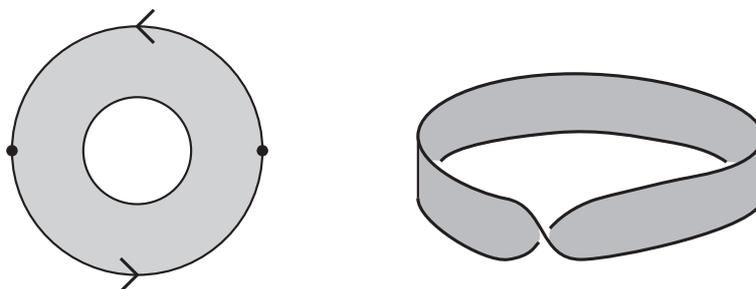


Figura 12.5: Togliendo un disco al piano proiettivo reale si ottiene un nastro di Möbius

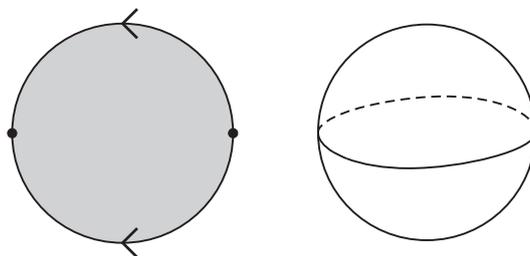


Figura 12.6: Costruzione di una sfera a partire da un disco

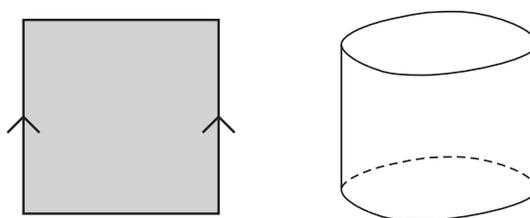


Figura 12.7: Costruzione di un cilindro a partire da un quadrato

12.2.3 Superfici ottenute incollando lati di una figura piana

Abbiamo visualizzato $\mathbb{P}^2(\mathbb{R})$ come l'oggetto ottenuto dal disco identificando i punti antipodali sul bordo. La costruzione di superfici a partire da figure piane tramite identificazione di porzioni di bordo è un procedimento molto generale, e consente di ottenere varie superfici note. Ad esempio se si suddivide il bordo di un disco in due semicirconferenze che vengono poi identificate in verso discorde, come in Fig. 12.6, si ottiene una sfera. Invece identificando due lati opposti di un quadrato si ottiene un cilindro (Fig. 12.7) oppure un nastro di Möbius (Fig. 12.8) a seconda del verso di identificazione.

12.2.4 Spazio proiettivo come completamento all'infinito dello spazio affine

Nelle precedenti descrizioni di $\mathbb{P}^1(\mathbb{R})$ e $\mathbb{P}^2(\mathbb{R})$ avevamo un insieme di rappresentanti sovrabbondante (cioè tale da intersecare tutte le classi di equivalenza, ma alcune in più di un punto), sul quale pertanto abbiamo poi dovuto compiere ulteriori identificazioni. Adottiamo ora un punto di vista opposto, ovvero descriviamo un insieme di rappresentanti deficitario (cioè tale da intersecare

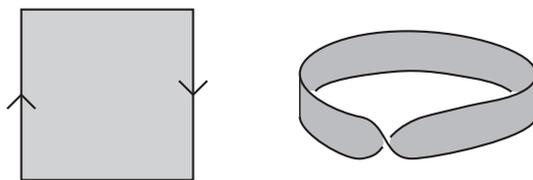


Figura 12.8: Costruzione di un nastro di Möbius a partire da un quadrato

ciascuna classe di equivalenza in al più un punto, ma qualcuna in zero punti); la restrizione a un tale insieme della proiezione nel quoziente è allora iniettiva, dunque otteniamo un sottoinsieme di $\mathbb{P}^n(\mathbb{R})$, e ci chiediamo cosa manchi a ottenere tutto $\mathbb{P}^n(\mathbb{R})$.

Proposizione 12.2.3. *Identificando lo spazio \mathbb{R}^n al sottospazio affine $\mathbb{R}^n \times \{1\}$ di \mathbb{R}^{n+1} , si ottiene un'identificazione di \mathbb{R}^n con un sottoinsieme di $\mathbb{P}^n(\mathbb{R})$. Il complementare di \mathbb{R}^n in $\mathbb{P}^n(\mathbb{R})$ è in modo naturale identificato a $\mathbb{P}^{n-1}(\mathbb{R})$.*

Dimostrazione. Ogni retta per l'origine interseca $\mathbb{R}^n \times \{1\}$ in al più un punto, il che prova la prima affermazione. Il complementare di \mathbb{R}^n in $\mathbb{P}^n(\mathbb{R})$ è l'insieme delle rette che non intersecano $\mathbb{R}^n \times \{1\}$, ovvero quelle che giacciono su $\mathbb{R}^n \times \{0\}$, da cui la seconda affermazione. \square

I punti di $\mathbb{P}^{n-1}(\mathbb{R})$ che si aggiungono a \mathbb{R}^n per ottenere $\mathbb{P}^n(\mathbb{R})$ hanno una chiara interpretazione geometrica. Fissiamo infatti un tale punto di $\mathbb{P}^{n-1}(\mathbb{R})$, ovvero una retta ℓ per l'origine in $\mathbb{R}^n \times \{0\}$, e consideriamo una retta ℓ' in \mathbb{R}^n , ovvero contenuta in $\mathbb{R}^n \times \{1\}$, e parallela a ℓ , ma non necessariamente passante per l'origine di \mathbb{R}^n . Se prendiamo un punto P su ℓ' e lo facciamo tendere all'infinito (non importa da quale delle due parti), vediamo che il punto di $\mathbb{P}^n(\mathbb{R})$ corrispondente a P , cioè la retta per l'origine e per P , si avvicina sempre più a ℓ . Possiamo dunque considerare il punto ℓ di $\mathbb{P}^{n-1}(\mathbb{R})$ come il comune *punto all'infinito* di tutte le rette ℓ' contenute in \mathbb{R}^n e parallele a ℓ . Riassumendo enunciamo la:

Proposizione 12.2.4. *$\mathbb{P}^n(\mathbb{R})$ si ottiene aggiungendo a \mathbb{R}^n i punti di $\mathbb{P}^{n-1}(\mathbb{R})$, ciascuno dei quali rappresenta il comune punto all'infinito di un fascio di rette parallele in \mathbb{R}^n .*

Osservazione 12.2.5. Poiché nella descrizione precedente intervengono le rette in \mathbb{R}^n non necessariamente passanti per l'origine, conviene pensare allo spazio \mathbb{R}^n che viene completato a $\mathbb{P}^n(\mathbb{R})$ come *spazio affine* piuttosto che come spazio vettoriale. Questa idea verrà ulteriormente motivata nel Paragrafo 12.4.

Nel caso $n = 2$ la descrizione di $\mathbb{P}^n(\mathbb{R})$ appena fornita comporta immediatamente la:

Proposizione 12.2.6. *Due rette disgiunte in \mathbb{R}^2 hanno il medesimo punto all'infinito.*

Ecco un nuovo punto di vista sull'interpretazione di $\mathbb{P}^n(\mathbb{R})$ come completamento di \mathbb{R}^n con l'aggiunta di un $\mathbb{P}^{n-1}(\mathbb{R})$ costituito dai punti all'infinito di \mathbb{R}^n . Se a \mathbb{R}^n applichiamo la trasformazione $x \mapsto \frac{x}{1+\|x\|}$ otteniamo una bigezione con il disco n -dimensionale aperto $\{x \in \mathbb{R}^n : \|x\| < 1\}$. Questo oggetto si completa in modo naturale al disco chiuso $\{x \in \mathbb{R}^n : \|x\| \leq 1\}$ con l'aggiunta della sfera $(n-1)$ -dimensionale $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$; ma una retta in \mathbb{R}^n diviene nel disco aperto un arco (o un diametro, se passa per l'origine) i cui estremi sono punti antipodali di S^{n-1} . I due versi nei quali tende all'infinito una retta rappresentano in realtà la stessa direzione, dunque l'oggetto con cui completare all'infinito \mathbb{R}^n è ciò che si ottiene da S^{n-1} identificando tra loro i punti antipodali, dunque proprio $\mathbb{P}^{n-1}(\mathbb{R})$.

12.2.5 La retta proiettiva complessa

Anche se il nostro interesse sarà rivolto in modo esclusivo agli spazi proiettivi reali, forniamo qui una descrizione geometrica della *retta proiettiva complessa* $\mathbb{P}^1(\mathbb{C})$. Analogamente a quanto visto nel caso reale, essa può essere interpretata come il completamento della retta complessa \mathbb{C} con il suo unico punto all'infinito ∞ . (Dal punto di vista reale \mathbb{C} è un piano, e di direzioni per andare all'infinito ce ne sono infinite, ma dal punto di vista complesso è una retta, dunque di direzione ce n'è una sola.) Geometricamente $\mathbb{P}^1(\mathbb{C})$ corrisponde in modo naturale alla sfera S^2 ; una bigezione tra S^2 e $\mathbb{C} \cup \{\infty\}$ si ottiene con la *proiezione stereografica*, illustrata in Fig. 12.9: si identifica \mathbb{C} al piano tangente alla sfera S^2 nel polo sud, e si proietta un punto P di S^2 diverso dal polo nord N sul punto $z \in \mathbb{C}$ ottenuto intersecando con \mathbb{C} la retta passante per N e per P . In questo modo $S^2 \setminus \{N\}$ è in bigezione con \mathbb{C} , mentre N corrisponde a ∞ , e questa corrispondenza è ragionevole perché per $|z|$ che diventa molto grande, cioè per z che si avvicina a ∞ , si ha che P si avvicina a N .

12.3 Coordinate omogenee e sottospazi proiettivi

Per definizione, un punto di $\mathbb{P}^n(\mathbb{R})$ è la classe di equivalenza di un vettore $x = (x_1, x_2, \dots, x_{n+1}) \in \mathbb{R}^{n+1}$ rispetto alla relazione $x \sim y$ se $y = \lambda x$ con $\lambda \in \mathbb{R}$. Sappiamo che la classe di equivalenza di x si indica in generale con $[x]$, ma in

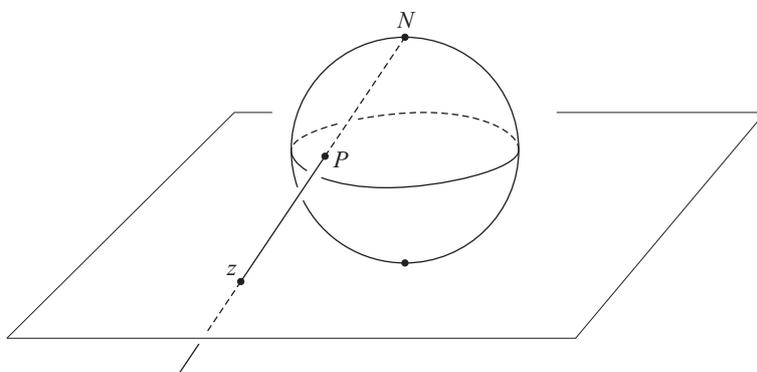


Figura 12.9: La proiezione stereografica

questo caso specializziamo la notazione indicandola con $[x_1 : x_2 : \dots : x_{n+1}]$, dove i segni “:” che separano le coordinate ci ricordano che di x non contano i valori numerici delle singole coordinate, ma soltanto i rapporti $x_i : x_j$, infatti per ogni $\lambda \neq 0$ il punto di $\mathbb{P}^n(\mathbb{R})$ dato da $[\lambda x_0 : \lambda x_1 : \dots : \lambda x_{n+1}]$ coincide con $[x_1 : x_2 : \dots : x_{n+1}]$. Quando un elemento di $\mathbb{P}^n(\mathbb{R})$ è presentato come $[x_1 : x_2 : \dots : x_{n+1}]$, diremo che è scritto in *coordinate omogenee*. Il prossimo risultato motiva questa definizione:

Proposizione 12.3.1. *Se $p(x) \in \mathbb{R}[x_1, x_2, \dots, x_{n+1}]$ è un polinomio omogeneo di grado d , ovvero una somma di monomi tutti di grado d nelle indeterminate x_1, x_2, \dots, x_{n+1} , allora l’equazione $p(x) = 0$ è ben definita in $\mathbb{P}^n(\mathbb{R})$, ovvero il fatto che sia soddisfatta o meno non dipende dalle coordinate omogenee usate per descrivere un punto di $\mathbb{P}^n(\mathbb{R})$.*

Dimostrazione. Dobbiamo verificare che se $[x] = [y]$ allora vale $p(x) = 0$ se e soltanto se vale $p(y) = 0$. Infatti l’ipotesi $[x] = [y]$ significa che esiste $\lambda \neq 0$ tale che $y = \lambda x$, e l’ipotesi che $p(x)$ sia omogeneo di grado d comporta che $p(y) = \lambda^d \cdot p(x)$, da cui la conclusione. \square

Definizione 12.3.2. Chiamiamo *sottospazio proiettivo* di $\mathbb{P}^n(\mathbb{R})$ un sottoinsieme definito da equazioni omogenee di grado 1 nelle coordinate omogenee.

Dalla definizione segue immediatamente che $T \subset \mathbb{P}^n(\mathbb{R})$ è un sottospazio proiettivo se e soltanto se esiste un sottospazio vettoriale $W \subset \mathbb{R}^{n+1}$ tale che T consiste delle classi di equivalenza dei punti di $W \setminus \{0\}$, ovvero dei punti di $\mathbb{P}^n(\mathbb{R})$ corrispondenti a rette contenute in W . Se W ha dimensione $k+1$ diremo

allora che T ha *dimensione* k . Questa osservazione ci consente la seguente rivisitazione della Proposizione 12.2.6:

Proposizione 12.3.3. *Due rette proiettive r_1 e r_2 distinte in $\mathbb{P}^2(\mathbb{R})$ si intersecano sempre in un punto. Se né r_1 né r_2 sono la retta all'infinito dello spazio affine \mathbb{R}^2 , dunque se r_j è il completamento all'infinito di una retta affine $\ell_j \subset \mathbb{R}^2$, allora il punto di intersezione di r_1 e r_2 è:*

- *nel caso di ℓ_1 e ℓ_2 incidenti, il loro punto di intersezione;*
- *nel caso di ℓ_1 e ℓ_2 parallele, il loro comune punto all'infinito.*

Dimostrazione. Supponiamo che r_j si ottenga da un sottospazio W_j di \mathbb{R}^3 . Allora W_1 e W_2 sono piani distinti in \mathbb{R}^3 , pertanto hanno in comune una retta, e questa retta è il punto di $\mathbb{P}^2(\mathbb{R})$ in cui si intersecano r_1 ed r_2 . Il resto della proposizione è evidente. \square

12.4 Cambiamenti di coordinate affini e proiettivi

Nello spazio vettoriale \mathbb{R}^n i cambiamenti di coordinate che è naturale considerare sono quelli lineari, ovvero quelli del tipo $x' = B \cdot x$ con $B \in \mathcal{M}_{n \times n}(\mathbb{R})$ invertibile. Analogamente in $\mathbb{P}^n(\mathbb{R})$ sono quelli che in coordinate omogenee si esprimono come $[x'] = [M \cdot x]$ con $M \in \mathcal{M}_{(n+1) \times (n+1)}(\mathbb{R})$ invertibile. Notiamo che se $\lambda \neq 0$ la matrice $\lambda \cdot M$ definisce su $\mathbb{P}^n(\mathbb{R})$ lo stesso cambiamento di coordinate definito da M .

Ricordiamo ora che nel Sottoparagrafo 12.2.4 abbiamo interpretato $\mathbb{P}^n(\mathbb{R})$ come il completamento all'infinito di \mathbb{R}^n , identificando \mathbb{R}^n al sottospazio affine $\{x \in \mathbb{R}^{n+1} : x_{n+1} = 1\}$ di \mathbb{R}^{n+1} e osservando che la restrizione a tale sottospazio della proiezione da $\mathbb{R}^{n+1} \setminus \{0\}$ nel quoziente $\mathbb{P}^n(\mathbb{R})$ è iniettiva. Possiamo allora chiederci quali cambiamenti di coordinate proiettivi su $\mathbb{P}^n(\mathbb{R})$ mandino \mathbb{R}^n in sé stesso, ovvero “lascino il finito al finito”. Se il cambiamento di coordinate avviene tramite una matrice $M = \begin{pmatrix} B & v \\ {}^t w & c \end{pmatrix}$, la condizione è che

$$\left[M \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} \right] = \begin{bmatrix} Bx + v \\ {}^t w \cdot x + c \end{bmatrix}$$

abbia ultima coordinata non nulla per ogni $x \in \mathbb{R}^n$ (non nulla ma non necessariamente uguale 1, dato che si tratta di coordinate omogenee), ed è chiaro che ciò accade se e soltanto se $w = 0$ e $c \neq 0$ (altrimenti basta prendere

$x = -\frac{c}{\|w\|^2} \cdot w$). Possiamo allora sostituire M con $\frac{1}{c} \cdot M$, e vediamo che il cambiamento di coordinate prende la forma

$$\begin{bmatrix} x' \\ 1 \end{bmatrix} = \begin{bmatrix} Bx + v \\ 1 \end{bmatrix}, \quad \text{ovvero} \quad x' = Bx + v$$

con $B \in \mathcal{M}_{n \times n}(\mathbb{R})$ invertibile e $v \in \mathbb{R}^n$. Chiameremo *affine* un tale cambiamento di coordinate su \mathbb{R}^n , dato che manda sottospazi affini in sottospazi affini. (Non è semplice, ma si può provare anche il viceversa: una bigezione di \mathbb{R}^n con sé stesso che manda sottospazi affini in sottospazi affini è necessariamente del tipo $x \mapsto B \cdot x + v$ con $B \in \mathcal{M}_{n \times n}(\mathbb{R})$ invertibile e $v \in \mathbb{R}^n$.) Riassumendo, abbiamo provato la:

Proposizione 12.4.1. *I cambiamenti di coordinate di $\mathbb{P}^n(\mathbb{R})$ che mandano la parte al finito in sé stessa sono i cambiamenti di coordinate affini di \mathbb{R}^n .*

Capitolo 13

Coniche e quadriche affini e proiettive

In questo capitolo riprendiamo lo studio delle coniche, iniziato nel Capitolo 11, alla luce della teoria degli spazi proiettivi vista nel Capitolo 12. Presenteremo una seconda dimostrazione più concettuale del teorema di classificazione, quindi estenderemo la teoria ai sottoinsiemi dello spazio tridimensionale definiti da equazioni polinomiali di secondo grado. Anche di questi ultimi daremo una descrizione sia dal punto di vista affine sia da quello proiettivo, dimostrando in particolare un teorema di classificazione affine.

13.1 Completamento proiettivo e punti all'infinito

In questo paragrafo presentiamo i primi semplici risultati generali sui sottoinsiemi di \mathbb{R}^n e di $\mathbb{P}^n(\mathbb{R})$ definiti da equazioni polinomiali di secondo grado, quindi li applichiamo al caso delle coniche.

Proposizione 13.1.1.

- Un'equazione di secondo grado omogenea in $x \in \mathbb{R}^{n+1}$ si scrive in modo unico come ${}^t x \cdot A \cdot x = 0$ con $A \in \mathcal{M}_{(n+1) \times (n+1)}(\mathbb{R})$ simmetrica;
- Un'equazione di secondo grado in $x \in \mathbb{R}^n$ si scrive in modo unico come ${}^t \begin{pmatrix} x \\ 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} = 0$ con $A \in \mathcal{M}_{(n+1) \times (n+1)}(\mathbb{R})$ simmetrica.

Dimostrazione. Per la prima affermazione, notiamo che l'equazione è del tipo

$$\sum_{1 \leq i < j \leq n+1} \beta_{ij} x_i x_j = 0,$$

e basta porre $a_{ii} = \beta_{ii}$ e $a_{ij} = a_{ji} = \frac{1}{2}\beta_{ij}$ per $i < j$. Per la seconda affermazione partiamo da un'equazione polinomiale di secondo grado in $x \in \mathbb{R}^n$ e moltiplichiamo per x_{n+1} i monomi di grado 1, e per x_{n+1}^2 il termine noto. Abbiamo adesso un'equazione omogenea in $x \in \mathbb{R}^{n+1}$ da cui quella originale si riottiene ponendo $x_{n+1} = 1$, e la conclusione segue allora dalla prima affermazione. \square

In entrambi i casi della proposizione precedente diremo che A è la matrice associata all'equazione.

Osservazione 13.1.2. Per qualsiasi matrice $B \in \mathcal{M}_{(n+1) \times (n+1)}(\mathbb{R})$ l'equazione ${}^t \begin{pmatrix} x \\ 1 \end{pmatrix} \cdot B \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} = 0$ è polinomiale di secondo grado in $x \in \mathbb{R}^n$, ma matrici B diverse possono dare luogo alla stessa equazione. La matrice associata a un'equazione è unica soltanto insistendo che essa sia simmetrica. La matrice associata all'equazione ${}^t \begin{pmatrix} x \\ 1 \end{pmatrix} \cdot B \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} = 0$ è $\frac{1}{2}(B + {}^tB)$.

Se nel secondo punto della proposizione precedente scriviamo $A = \begin{pmatrix} Q & \ell \\ {}^t\ell & c \end{pmatrix}$ abbiamo che $Q \in \mathcal{M}_{n \times n}(\mathbb{R})$ rappresenta la parte quadratica dell'equazione, mentre $2\ell \in \mathcal{M}_{1 \times n}(\mathbb{R})$ rappresenta la parte lineare, infine $c \in \mathbb{R}$ è il termine noto. Diamo ora la:

Definizione 13.1.3. Sia $\mathcal{L} \subset \mathbb{R}^n$ il luogo di equazione ${}^t \begin{pmatrix} x \\ 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} = 0$.

Chiamiamo allora *completamento proiettivo* di \mathcal{L} il luogo $\overline{\mathcal{L}} \subset \mathbb{P}^n(\mathbb{R})$ definito dall'equazione omogenea ${}^t x \cdot A \cdot x = 0$, ovvero $\overline{\mathcal{L}} = \{[x] \in \mathbb{P}^n(\mathbb{R}) : {}^t x \cdot A \cdot x = 0\}$. Definiamo poi l'insieme dei *punti all'infinito* di \mathcal{L} come $\mathcal{L}_\infty = \overline{\mathcal{L}} \setminus \mathcal{L}$, visto come sottoinsieme dell'insieme $\mathbb{P}^{n-1}(\mathbb{R})$ dei punti all'infinito di \mathbb{R}^n ; abbiamo allora $\mathcal{L}_\infty = \{[x] \in \mathbb{P}^{n-1}(\mathbb{R}) : {}^t x \cdot Q \cdot x = 0\}$.

Per una conica non degenera e non vuota la costruzione precedente ha una chiarissima interpretazione geometrica. Impiegando le equazioni canoniche abbiamo infatti i casi seguenti:

Conica \mathcal{L}	Ellisse	Iperbole	Parabola
Equazione di \mathcal{L}	$x^2 + y^2 = 1$	$x^2 - y^2 = 1$	$y = x^2$
Equazione di $\overline{\mathcal{L}}$	$x^2 + y^2 = z^2$	$x^2 - y^2 = z^2$	$yz = x^2$
Equazione di \mathcal{L}_∞	$x^2 + y^2 = 0$	$x^2 - y^2 = 0$	$x^2 = 0$
Punti all' ∞	Nessuno	$[1 : 1]$ e $[1 : -1]$	$[0 : 1]$

Abbiamo allora che:

- Un'ellisse non ha punti all'infinito;
- Un'iperbole ha due punti all'infinito, dati dalle direzioni dei suoi asintoti;
- Una parabola ha un punto all'infinito, dato dalla direzione del suo asse.

13.2 Classificazione delle coniche proiettive e affini

Supponiamo di avere in $\mathbb{P}^n(\mathbb{R})$ un luogo \mathcal{P} definito da un'equazione polinomiale omogenea di secondo grado a cui è associata una matrice simmetrica A . Operando un cambiamento di coordinate proiettivo $[x'] = [M \cdot x]$ con $M \in \mathcal{M}_{(n+1) \times (n+1)}(\mathbb{R})$ invertibile, vediamo allora che la matrice dell'equazione nelle nuove coordinate è ${}^tM \cdot A \cdot M$. Ne deduciamo il:

Teorema 13.2.1. *Consideriamo il luogo $\mathcal{P} = \{[x] \in \mathbb{P}^n(\mathbb{R}) : {}^t x \cdot A \cdot x = 0\}$ con $A \in \mathcal{M}_{(n+1) \times (n+1)}(\mathbb{R})$ simmetrica e $\det(A) \neq 0$. Allora a meno di cambiamenti proiettivi delle coordinate si ha*

$$\mathcal{P} = \{[x] \in \mathbb{P}^n(\mathbb{R}) : x_1^2 + \dots + x_p^2 = x_{p+1}^2 + \dots + x_{n+1}^2\}$$

con $p \geq n + 1 - p$.

Dimostrazione. Grazie al teorema spettrale possiamo impiegare dapprima una M ortogonale per sostituire A con una ${}^tM \cdot A \cdot M$ diagonale con autovalori $\lambda_1, \dots, \lambda_{n+1}$. Poniamo ora $x'_j = \sqrt{|\lambda_j|} \cdot x_j$ e ci riconduciamo ad una A diagonale con ± 1 sulla diagonale principale. Basta infine riordinare le variabili ed eventualmente cambiare di segno all'equazione. \square

Questo risultato comporta in particolare che ci sono a meno di cambiamenti di coordinate proiettivi soltanto due coniche proiettive non degeneri in $\mathbb{P}^2(\mathbb{R})$: l'insieme vuoto $x^2 + y^2 + z^2 = 0$ e la conica $x^2 + y^2 = z^2$. In particolare l'ellisse, l'iperbole e la parabola hanno la stessa conica proiettiva non vuota come completamento proiettivo. In altre parole, l'ellisse, l'iperbole e la parabola si ottengono dalla medesima conica proiettiva non degenera ma con scelte diverse della retta all'infinito. Più esattamente, se partiamo dalla conica proiettiva $x^2 + y^2 = z^2$ abbiamo:

- Scegliendo $z = 0$ come retta all'infinito otteniamo la conica affine di equazione $x^2 + y^2 = 1$, l'ellisse;

- Scegliendo come retta all'infinito quella che passa per $[0 : 1 : 1]$ e per $[0 : -1 : 1]$ otteniamo un'iperbole; ad esempio vediamo facilmente che il cambiamento di coordinate proiettivo $[x : y : z] \mapsto [z : y : x]$ fissa $[1 : 0 : 1]$ e trasforma $[0 : 1 : 1]$ in $[1 : 1 : 0]$ e $[0 : -1 : 1]$ in $[1 : -1 : 0]$, inoltre manda $x^2 + y^2 = z^2$ in $z^2 + y^2 = x^2$, ovvero $x^2 - y^2 = z^2$, e a questo punto prendendo la parte affine $z = 1$ otteniamo l'iperbole canonica $x^2 - y^2 = 1$;
- Scegliendo come retta all'infinito quella che passa per i punti $[0 : 1 : 1]$ e $[1 : 0 : 0]$ otteniamo una parabola; ad esempio il cambiamento di coordinate proiettivo $[x : y : z] \mapsto [x : y + z : z - y]$ manda $[0 : 1 : 1]$ in $[0 : 1 : 0]$, fissa $[1 : 0 : 0]$ e manda $[0 : -1 : 1]$ in $[0 : 0 : 1]$, inoltre manda $x^2 + y^2 = z^2$ in $x^2 = yz$, e a questo punto prendendo la parte affine $z = 1$ otteniamo la parabola canonica $y = x^2$.

Come già annunciato, ripetiamo ora la classificazione delle coniche affini (Teorema 11.2.1) in modo più semplice ed elegante. Sfrutteremo il seguente fatto, che ci servirà anche più avanti:

Lemma 13.2.2. *Date $A, B \in \mathcal{M}_{n \times n}(\mathbb{R})$ con A simmetrica e B invertibile, i segni degli autovalori di ${}^tB \cdot A \cdot B$, contati con la loro molteplicità, coincidono con i segni degli autovalori di A .*

Dimostrazione. Non è difficile vedere che il numero di autovalori positivi di A coincide con la massima dimensione possibile di un sottospazio W di \mathbb{R}^n tale che $\langle w|w \rangle_A > 0$ per ogni $w \in W \setminus \{0\}$. Poiché $\langle w|w \rangle_{{}^tB \cdot A \cdot B} = \langle Bw|Bw \rangle_A$ concludiamo subito che ${}^tB \cdot A \cdot B$ ha tanti autovalori positivi quanti ne ha A . Analogamente per quelli negativi, e allora anche per quelli nulli. \square

Teorema 13.2.3. *Sia \mathcal{L} la conica in \mathbb{R}^2 alla cui equazione è associata una matrice simmetrica $A \in \mathcal{M}_{3 \times 3}(\mathbb{R})$ con $\det(A) \neq 0$. Sia d_j il determinante della sottomatrice di A formata dalle prime j righe e j colonne. Allora A rientra in uno delle seguenti casi:*

- | | |
|---------------------------------------|---------------------------------------|
| (1) $d_2 > 0$ e $d_1 \cdot d_3 > 0$; | (2) $d_2 > 0$ e $d_1 \cdot d_3 < 0$; |
| (3) $d_2 < 0$; | (4) $d_2 = 0$. |

Inoltre modificando l'equazione di \mathcal{L} tramite moltiplicazione per una costante non nulla oppure tramite una trasformazione affine la nuova matrice A rientra nello stesso caso (1)-(4). Infine a seconda del caso in cui A rientra la conica ℓ ha il seguente tipo affine:

- (1) insieme vuoto; (2) ellisse; (3) iperbole; (4) parabola.

Dimostrazione. Sia $A = \begin{pmatrix} Q & \ell \\ {}^t\ell & c \end{pmatrix}$. Grazie al Teorema 10.2.5 possiamo intrinsecamente interpretare i casi (1)-(4) come segue:

- (1) Gli autovalori di A sono concordi tra loro;
- (2) Gli autovalori di Q sono concordi tra loro ma quelli di A non lo sono;
- (3) Gli autovalori di Q sono discordi tra loro
- (4) Q ha almeno un autovalore nullo.

Osservando che un cambiamento affine di coordinate associato a $M = \begin{pmatrix} B & v \\ 0 & 1 \end{pmatrix}$ trasforma A in

$${}^tM \cdot A \cdot M = \begin{pmatrix} {}^tB \cdot Q \cdot B & {}^tB(Qv + \ell) \\ ({}^tvQ + {}^t\ell)B & \langle v|v \rangle_Q + 2\langle v|\ell \rangle + c \end{pmatrix}$$

il Lemma 13.2.2 comporta subito che il caso in cui la matrice rientra non cambia tramite moltiplicazione dell'equazione per una costante non nulla oppure tramite una trasformazione affine. Osserviamo prima di procedere che nel caso (4) la matrice Q ha almeno un autovalore non nullo, altrimenti A non potrebbe avere rango 3.

Usando un cambiamento di coordinate con $v = 0$ ed eventualmente cambiando segno all'equazione possiamo ora ricondurci, come nella dimostrazione del Teorema 13.2.1, ai casi seguenti:

$$\begin{aligned} (1) \quad Q &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; & (2) \quad Q &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \\ (3) \quad Q &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ e } d_3 > 0; & (4) \quad Q &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Nei casi (1)-(3) possiamo operare un cambiamento di coordinate con $B = I_2$ e $v = -Q^{-1}\ell$, riconducendoci al caso $\ell = 0$, e allora eventualmente dopo un'omotetia troviamo le equazioni canoniche

$$(1) \quad x^2 + y^2 + 1 = 0; \quad (2) \quad x^2 + y^2 = 1; \quad (3) \quad x^2 - y^2 = 1.$$

Nel caso (4) sempre con $B = I_2$ e v opportuno ci riconduciamo alla matrice

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & k \\ 0 & k & c \end{pmatrix} \text{ e poi cambiamo coordinate con } M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2k} \\ 0 & 0 & \frac{c}{4k^2} \end{pmatrix} \text{ e}$$

troviamo l'equazione canonica $y = x^2$. □

13.3 Modelli affini delle quadriche

Introduciamo in questo paragrafo i luoghi nello spazio tridimensionale analoghi alle coniche nel piano, fornendo quelli che nel prossimo paragrafo si riveleranno essere i modelli affini.

Definizione 13.3.1. Chiamiamo *quadrica non degenera* un sottoinsieme dello spazio definito a un'equazione polinomiale di secondo grado a cui è associata una matrice $A \in \mathcal{M}_{4 \times 4}(\mathbb{R})$ simmetrica con determinante non nullo.

13.3.1 Insieme vuoto, ellissoide, paraboloidi ellittico

I primi e più semplici esempi di quadrica non degenera sono:

- L'insieme vuoto, di equazione $x^2 + y^2 + z^2 + 1 = 0$;
- L'ellissoide, di equazione $x^2 + y^2 + z^2 = 1$;
- Il paraboloidi ellittico, di equazione $z = x^2 + y^2$.

Naturalmente l'equazione $x^2 + y^2 + z^2 = 1$ non definisce altro che la sfera unitaria S^2 di \mathbb{R}^3 , ma dal punto di vista affine l'equazione equivale a qualsiasi altra del tipo $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, che per analogia con il caso dell'ellisse è appunto detto *ellissoide* di *semiassi* a , b e c . A proposito del paraboloidi ellittico $z = x^2 + y^2$ osserviamo che la quantità $x^2 + y^2$ si interpreta intrinsecamente come il quadrato della distanza di un punto $(x; y; z)$ dal terzo asse coordinato Oz . Ne segue che il paraboloidi ellittico ha una simmetria rotazionale intorno a tale asse, anzi più esattamente è ottenuto ruotando intorno all'asse Oz la parabola $z = x^2$ contenuta nel piano Oxz .

13.3.2 Iperboloidi a una e a due falde

Altre due quadriche di un certo interesse sono:

- L'iperboloidi ellittico o a due falde, di equazione $x^2 + y^2 = z^2 - 1$;
- L'iperboloidi iperbolico o a una falda, di equazione $x^2 + y^2 = z^2 + 1$.

Poiché di nuovo le coordinate x e y appaiono solo nell'espressione $x^2 + y^2$ al primo membro, osserviamo che queste due quadriche hanno una simmetria rotazionale intorno al terzo asse coordinato. Più esattamente:

- L'iperboloide ellittico $x^2 + y^2 = z^2 - 1$ si ottiene ruotando l'iperbole $z^2 - x^2 = 1$ nel piano Oxz intorno all'asse Oz , che passa per i vertici dell'iperbole;
- L'iperboloide iperbolico $x^2 + y^2 = z^2 + 1$ si ottiene ruotando l'iperbole $x^2 - z^2 = 1$ nel piano Oxz intorno all'asse Oz , che è l'asse del segmento che ha come estremi i vertici dell'iperbole.

Questa costruzione degli iperboloidi giustifica la terminologia: quello ellittico consiste di due porzioni distinte (le due *falde*) ognuna generata dalla rotazione di uno dei rami dell'iperbole; quello iperbolico consiste di una sola porzione (la *falda*) generata dalla rotazione di uno qualsiasi dei rami.

L'iperboloide iperbolico $x^2 + y^2 = z^2 + 1$ ha una proprietà geometrica molto notevole che vale la pena di sottolineare. Cominciamo osservando che la sua intersezione con il piano $y = 1$ ha equazione $x^2 = z^2$, dunque consiste delle rette $z = \pm x$. A partire da ciò non è difficile riconoscere i fatti seguenti:

- Per ogni punto dell'iperboloide iperbolico passa una retta interamente contenuta nell'iperboloide (anzi, due tali rette);
- L'iperboloide è ottenuto ruotando intorno all'asse Oz la retta $z = x$ (oppure la retta $z = -x$) contenuta nel piano $y = 1$.

La prima proprietà appena descritta si esprime dicendo che l'iperboloide iperbolico è una superficie *rigata*. Geometricamente è piuttosto notevole il fatto l'unione di un reticolo di rette (anzi, di due tali reticoli) possa dar luogo a una superficie indubbiamente curva come quella dell'iperboloide a una falda. In generale se r e s sono rette sghembe, la superficie generata dalla rotazione di r intorno a s è un iperboloide iperbolico (cioè equivale alla superficie di equazione $x^2 + y^2 = z^2 + 1$ a meno di un cambiamento affine di coordinate).

13.3.3 Paraboloide iperbolico

L'ultima quadrica che ci servirà da modello affine è:

- Il paraboloide iperbolico o a sella, di equazione $z = x^2 - y^2$.

Anche questa superficie ha caratteristiche geometriche decisamente notevoli:

- Prendendo le sezioni del paraboloide iperbolico con piani della forma $z = k$ con $k \in \mathbb{R}$ si osserva che esso è unione di una famiglia di iperboli equilateri, insieme a una coppia di rette perpendicolari tra loro (iperbole equilatera degenera);

- Prendendone le sezioni con piani del tipo $y = k$ con $k \in \mathbb{R}$, si osserva che esso è unione di una famiglia di parabole tutte con la stessa distanza focale, asse verticale e concavità rivolta verso l'alto, con i vertici delle parabole nella famiglia che descrivono una parabola avente la stessa distanza focale, asse verticale e concavità rivolta verso il basso;
- Prendendone le sezioni con i piani $x + y = k$ oppure $x - y = k$ con $k \in \mathbb{R}$ si osserva che per ogni punto del paraboloido iperbolico passano due rette in esso interamente contenute; in particolare il paraboloido iperbolico è una superficie rigata.

13.3.4 Punti all'infinito

Vediamo ora come descrivere l'insieme dei punti all'infinito dei modelli di quadrica non degeneri introdotti in precedenza. Come al solito si tratta di aggiungere una variabile rendendo omogenea l'equazione, quindi di porre questa variabile uguale a 0. Ecco allora le possibilità, dove indichiamo con \mathcal{C} l'unica conica proiettiva non degenera (vedi il Teorema 13.2.1) e tralasciamo la quadrica vuota:

Quadrica \mathcal{L}	Equazione di $\overline{\mathcal{L}}$	Equazione di \mathcal{L}_∞	\mathcal{L}_∞
Ellissoide $x^2 + y^2 + z^2 = 1$	$x^2 + y^2 + z^2 = u^2$	$x^2 + y^2 + z^2 = 0$	\emptyset
Iperboloide a 2 falde $x^2 + y^2 = z^2 - 1$	$x^2 + y^2 = z^2 - u^2$	$x^2 + y^2 = z^2$	\mathcal{C}
Iperboloide a 1 falda $x^2 + y^2 = z^2 + 1$	$x^2 + y^2 = z^2 + u^2$	$x^2 + y^2 = z^2$	\mathcal{C}
Paraboloido ellittico $z = x^2 + y^2$	$zu = x^2 + y^2$	$x^2 + y^2 = 0$	Un punto
Paraboloido iperbolico $z = x^2 - y^2$	$zu = x^2 - y^2$	$x^2 - y^2 = 0$	Due rette

Dalla tabella precedente estraiamo un'altra interessante informazione, che giustifica anche la terminologia introdotta. Iniziamo col ricordare che grazie al Teorema 13.2.1 esistono in $\mathbb{P}^3(\mathbb{R})$ solo due quadriche proiettive non vuote e

distinte, di equazioni

$$x^2 + y^2 + z^2 = u^2, \quad x^2 + y^2 = z^2 + u^2$$

che chiameremo rispettivamente *ellissoide proiettivo* e *iperboloide proiettivo*. Abbiamo allora la:

Proposizione 13.3.2. *Il completamento proiettivo di una quadrica è:*

- *L'ellissoide proiettivo per l'ellissoide, l'iperboloide ellittico e il paraboloide ellittico;*
- *L'iperboloide proiettivo per l'iperboloide iperbolico e il paraboloide iperbolico.*

Dimostrazione. Per l'ellissoide e per gli iperboloidi la conclusione è evidente. Per i paraboloidi si opera il cambio di variabile $z = w - v$ e $u = w + v$, ottenendo rispettivamente le equazioni

$$x^2 + y^2 + v^2 = w^2, \quad x^2 + v^2 = y^2 + w^2.$$

□

13.4 Classificazione affine delle quadriche non degeneri

In questo paragrafo mostriamo che ciascuna quadrica non degenera può essere ricondotta tramite un cambiamento di coordinate affine a uno dei 6 modelli discussi in precedenza. Operiamo la classificazione in due tappe: prima in modo intrinseco ma poco operativo, successivamente in modo più pratico:

Teorema 13.4.1. *Sia $A = \begin{pmatrix} Q & \ell \\ {}_t\ell & c \end{pmatrix} \in \mathcal{M}_{4 \times 4}(\mathbb{R})$ la matrice simmetrica dell'equazione che definisce una quadrica non degenera \mathcal{L} . Allora i segni degli autovalori di Q e quelli di A non cambiano operando un cambiamento di coordinate affine o moltiplicando l'equazione per un numero positivo, mentre cambiano tutti contemporaneamente moltiplicando l'equazione per un numero negativo. Inoltre a meno di moltiplicare l'equazione per un numero negativo le possibilità per questi segni sono le seguenti, nel qual caso esiste una trasformazione affine di \mathbb{R}^3 che trasforma \mathcal{L} nella quadrica indicata:*

Segni autovalori Q	Segni autovalori A	Tipo affine \mathcal{L}
+, +, +	+, +, +, +	Insieme vuoto
	+, +, +, -	Ellissoide
+, +, -	+, +, -, +	Iperboloide ellittico
	+, +, -, -	Iperboloide iperbolico
+, +, 0	+, +, +, -	Paraboloide ellittico
+, -, 0	+, -, +, -	Paraboloide iperbolico

Dimostrazione. La prima affermazione segue banalmente dal Lemma 13.2.2. Per vedere che i segni degli autovalori di Q , a meno di cambiamenti affini di coordinate e moltiplicazione dell'equazione per un numero non nullo, sono solo quelli in tabella, basta mostrare che Q non può avere due autovalori nulli, il che segue dal fatto che in tal caso essendo diagonalizzabile avrebbe rango al più 1, dunque le prime tre colonne di A sarebbero linearmente dipendenti. Proviamo ora che, dati i segni degli autovalori di Q , quelli di A possono essere solo come indicato. Nel caso +, +, + per Q , se A avesse almeno due autovalori negativi allora esisterebbe W di dimensione 2 con $\langle w|w \rangle_A < 0$ per ogni $w \in W \setminus \{0\}$, mentre $\langle u|u \rangle_Q > 0$ per ogni $u \in \mathbb{R}^3 \setminus \{0\}$, dunque $\langle y|y \rangle_A > 0$ per ogni $y \in (\mathbb{R}^3 \times \{0\}) \setminus \{0\}$, e questo è assurdo perché W e $\mathbb{R}^3 \times \{0\}$ si intersecano almeno in una retta. Nel caso +, +, - per Q è chiaro che A ha almeno un autovalore negativo, ma non può averne 3 per un argomento analogo a quello appena presentato. Nel caso +, +, 0 per Q è di nuovo chiaro che A ha almeno un autovalore negativo, e non può averne più di uno perché $\langle y|y \rangle_A \geq 0$ per ogni $y \in \mathbb{R}^3 \times \{0\}$. Infine nel caso +, -, 0 per Q si ha $\langle y|y \rangle_A \geq 0$ per ogni $y \in \mathbb{R} \times \{0\} \times \mathbb{R} \times \{0\}$, mentre $\langle y|y \rangle_A \leq 0$ per ogni $y \in \{0\} \times \mathbb{R} \times \mathbb{R} \times \{0\}$, da cui segue che A non può avere né tre autovalori positivi né tre negativi.

Resta da vedere che a seconda di quale dei 6 casi si dà per i segni degli autovalori di Q e quelli di A , è possibile trasformare \mathcal{L} tramite cambiamenti di coordinate affini a uno dei 6 modelli. Innanzitutto cambiamo coordinate con $M = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}$ con B ortogonale, in modo da diagonalizzare Q . Poi permutiamo le coordinate, eseguiamo un'omotetia su ciascuna ed eventualmente cambiamo segno all'equazione riconducendoci al caso in cui Q , a seconda del

segno degli autovalori, ha una delle seguenti forme:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Nei primi due casi cambiando coordinate con $B = I_3$ e $v = -Q^{-1}\ell$ ci si riconduce al caso $\ell = 0$, e la conclusione è immediata. Negli ultimi due casi si ottiene invece $a_{14} = a_{24} = 0$, dunque $a_{34} \neq 0$, e si ricavano le equazioni canoniche dei paraboloidi con l'ultimo cambiamento di coordinate di matrice

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2a_{34}} & -\frac{a_{44}}{2a_{34}} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

□

Vediamo ora come il Teorema 13.4.1 si applichi in pratica per classificare una quadrica alla cui equazione è associata una matrice A con $\det(A) \neq 0$. La situazione più semplice discende direttamente dal Teorema 10.2.5:

Proposizione 13.4.2. *Se d_j è il determinante della sottomatrice di A formata dalle prime j righe e j colonne, e $d_j \neq 0$ per $j = 1, 2, 3, 4$, allora i segni degli autovalori di Q sono i segni di $d_1, \frac{d_2}{d_1}, \frac{d_3}{d_2}$, mentre i segni degli autovalori di A sono i segni di $d_1, \frac{d_2}{d_1}, \frac{d_3}{d_2}, \frac{d_4}{d_3}$.*

Il prossimo risultato spiega invece come procedere quando la proposizione precedente non si applica:

Proposizione 13.4.3. *Se $d_4 \neq 0$ allora si danno due casi:*

- *a meno di permutazione delle coordinate si ha $d_2 \neq 0$ e in tal caso i segni di d_1, d_2, d_3, d_4 consentono sempre la classificazione;*
- *per qualsiasi permutazione delle coordinate si ha sempre $d_2 = 0$ e in tal caso la quadrica è un iperboloide, iperbolico (a una falda) se $d_4 > 0$ ed ellittico (a due falde) se $d_4 < 0$.*

Dimostrazione. Supponiamo che si dia il primo caso, ovvero che $d_2 \neq 0$ dopo una permutazione delle coordinate, e illustriamo come concludere la classificazione. Se $d_2 < 0$ otteniamo che Q ha i primi due autovalori discordi; se $d_3 = 0$ concludiamo subito che abbiamo un paraboide iperbolico, altrimenti l'ultimo

autovalore di Q ha il segno di $\frac{d_3}{d_2}$, per A i primi tre autovalori hanno i segni di quelli di Q e l'ultimo ha il segno di $\frac{d_4}{d_3}$. Se $d_2 > 0$ necessariamente si ha $d_1 \neq 0$, dunque i primi due autovalori di Q hanno il segno di d_1 ; se $d_3 = 0$ concludiamo subito che abbiamo un paraboloide ellittico, altrimenti concludiamo esattamente come nel caso $d_2 < 0$.

Supponiamo invece che dopo qualsiasi permutazione delle coordinate si abbia $d_2 = 0$. Se uno dei coefficienti sulla diagonale principale di Q è nullo allora sono interamente nulle la riga e la colonna che lo contengono, e la restante sottomatrice 2×2 ha determinante nullo, da cui segue che Q ha rango al più 1, il che è impossibile poiché $\det(A) \neq 0$. I coefficienti sulla diagonale principale di A sono dunque non nulli, e anzi concordi, dunque possiamo supporre che siano tutti positivi. Ne deduciamo che Q deve avere la forma

$$\begin{pmatrix} a & \varepsilon_{12} \cdot \sqrt{ab} & \varepsilon_{13} \cdot \sqrt{ac} \\ \varepsilon_{12} \cdot \sqrt{ab} & b & \varepsilon_{23} \cdot \sqrt{bc} \\ \varepsilon_{13} \cdot \sqrt{ac} & \varepsilon_{23} \cdot \sqrt{bc} & c \end{pmatrix}$$

dove ciascuno dei ε_{ij} vale $+1$ o -1 . Se $\varepsilon_{12} \cdot \varepsilon_{13} \cdot \varepsilon_{23} = +1$ allora è facile vedere che Q ha rango 1, di nuovo contraddicendo l'ipotesi $\det(A) \neq 0$; ad esempio se $\varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23} = +1$ si ha

$$\begin{pmatrix} a & \sqrt{ab} & \sqrt{ac} \\ \sqrt{ab} & b & \sqrt{bc} \\ \sqrt{ac} & \sqrt{bc} & c \end{pmatrix} = \begin{pmatrix} \sqrt{a} \\ \sqrt{b} \\ \sqrt{c} \end{pmatrix} \begin{pmatrix} \sqrt{a} \\ \sqrt{b} \\ \sqrt{c} \end{pmatrix} \begin{pmatrix} \sqrt{a} \\ \sqrt{b} \\ \sqrt{c} \end{pmatrix};$$

se due tra $\varepsilon_{12}, \varepsilon_{13}, \varepsilon_{23}$ sono -1 e uno è $+1$ la conclusione è analoga. Supponiamo al contrario che $\varepsilon_{12} \cdot \varepsilon_{13} \cdot \varepsilon_{23} = -1$; allora si verifica con qualche calcolo che in ogni caso $\det(Q) = -4a^2b^2c^2 < 0$; poiché $d_1 > 0$ e $d_3 < 0$ concludiamo che Q ha due autovalori positivi e uno negativo, e la classificazione dell'enunciato segue subito. \square

Capitolo 14

Curve nel piano e nello spazio

In questo capitolo tradurremo matematicamente l'idea intuitiva di curva come la traiettoria unidimensionale nel piano o nello spazio percorsa da un punto che si muove in modo “regolare”. Quindi indagheremo le principali proprietà geometriche delle curve.

14.1 Curve parametrizzate e regolari

Una “curva” (come ad esempio il tracciato di un sentiero su una cartina topografica) è un sottoinsieme del piano, ma per matematizzare questa nozione è necessario impiegare il concetto di funzione. Lo facciamo in dimensione qualsiasi perché la teoria generale non presenta differenze significative rispetto al caso bidimensionale.

Definizione 14.1.1. Chiameremo *curva parametrizzata* una funzione continua $\alpha : [a, b] \rightarrow \mathbb{R}^n$, dove $[a, b]$ è un intervallo di \mathbb{R} . Diremo inoltre che α è *regolare* se le componenti di α hanno tutte derivata continua e per ogni t in $[a, b]$ si ha che $\alpha'(t)$ è non nullo. Diremo poi che α è *regolare a tratti* se è possibile suddividere $[a, b]$ in un numero finito di intervalli su ciascuno dei quali la restrizione di α è regolare. Infine diremo che α è:

- *semplice* se è iniettiva su $[a, b]$;
- *chiusa* se $\alpha(a) = \alpha(b)$;
- *semplice e chiusa* se $\alpha(a) = \alpha(b)$ e α è iniettiva su $[a, b]$.

Osservazione 14.1.2. La regolarità corrisponde alla condizione geometrica che per ogni punto dell'immagine di α esista una retta tangente. A questo fine *non* basta richiedere che esista $\alpha'(t)$, come mostra la curva

$$\alpha(t) = \begin{cases} (0, -t^3) & \text{per } t \in [-1, 0] \\ (t^3, 0) & \text{per } t \in [0, 1] \end{cases}$$

che è derivabile due volte, con $\alpha'(0) = (0, 0)$, e ha come immagine l'unione dei segmenti unitari sui due semiassi coordinati positivi. Un altro esempio è la curva $\alpha(t) = (t^2, t^3)$ per $t \in [-1, 1]$, che è derivabile con $\alpha'(0) = (0, 0)$, e ha come immagine una cuspidine, unione dei grafici su $[0, 1]$ delle funzioni $y = \pm \sqrt[3]{x^2}$.

Con il prossimo risultato formalizziamo il significato di retta tangente già menzionato nell'osservazione precedente:

Proposizione 14.1.3. *Sia $\alpha : [a, b] \rightarrow \mathbb{R}^n$ una curva regolare e sia $t_0 \in [a, b]$. Per t che tende a t_0 l'unica retta di \mathbb{R}^n passante per $\alpha(t_0)$ e per $\alpha(t)$ tende alla retta*

$$\alpha(t_0) + \text{Span}(\alpha'(t_0)),$$

detta retta tangente ad α in $\alpha(t_0)$. Inoltre tale retta è l'unica avente contatto almeno duplice con α in $\alpha(t_0)$.

Dimostrazione. La retta di \mathbb{R}^n passante per $\alpha(t_0)$ e per $\alpha(t)$ è

$$\alpha(t_0) + \text{Span}(\alpha(t) - \alpha(t_0)),$$

dunque la sua giacitura è generata da $\frac{1}{t}(\alpha(t) - \alpha(t_0))$, che tende appunto ad $\alpha'(t_0)$. Passiamo alla seconda parte dell'enunciato limitandoci per semplicità al caso $n = 2$. Se prendiamo una qualsiasi retta r di equazione $ax + by + c = 0$, e chiamiamo X e Y le componenti di α , la distanza da r di un punto $\alpha(t)$ su α vale

$$d(t) = \frac{|aX(t) + bY(t) + c|}{\sqrt{a^2 + b^2}}$$

e la richiesta che r abbia contatto almeno duplice con α in $\alpha(t_0)$ significa che la funzione d è infinitesima di ordine almeno 1 in t_0 , ovvero che $d(t_0) = d'(t_0) = 0$. La condizione $d(t_0) = 0$ significa che r contiene $\alpha(t_0)$, e per ogni a e b esiste un unico c per il quale ciò accade. Resta dunque solo da provare che la condizione $d'(t_0) = 0$ corrisponde al fatto che $ax + by = 0$ sia l'equazione della giacitura

$\text{Span}(\alpha'(t_0))$ della retta già trovata. Possiamo ora supporre che $a^2 + b^2 = 1$ e notare che

$$d(t) = \sqrt{(aX(t) + bY(t) + c)^2}$$

$$\Rightarrow d'(t) = \frac{aX(t) + bY(t) + c}{|aX(t) + bY(t) + c|} \cdot (aX'(t) + bY'(t)).$$

Nell'ultima espressione la frazione vale sempre ± 1 dove è definita, dunque $d'(t_0)$ può valere 0 solo se $aX'(t_0) + bY'(t_0) = 0$, che è precisamente quanto volevamo verificare. \square

14.1.1 Cambiamento di parametro

Per formalizzare l'idea di curva siamo stati costretti a definirla come funzione piuttosto che come sottoinsieme del piano o dello spazio. Ma geometricamente una curva è un sottoinsieme del piano o dello spazio; ecco come recuperare questa idea:

Definizione 14.1.4. Siano $\alpha : [a, b] \rightarrow \mathbb{R}^n$ e $\beta : [c, d] \rightarrow \mathbb{R}^n$ curve regolari. Diciamo che β è ottenuta da α tramite un *cambiamento di parametro* se esiste una bigezione derivabile $\tau : [c, d] \rightarrow [a, b]$ con inversa derivabile tale che $\beta(s) = \alpha(\tau(s))$ per ogni $s \in [c, d]$. La τ stessa sarà allora detta *funzione di cambiamento di parametro*.

Osservazione 14.1.5. Supponiamo che β sia ottenuta da α tramite un cambiamento di parametro τ . Allora:

- α e β hanno la medesima immagine;
- α è ottenuta da β tramite il cambiamento di parametro

$$\tau^{-1} : [a, b] \rightarrow [c, d];$$

- Si dà uno dei casi seguenti:
 - (i) τ è crescente, $\tau'(s) > 0$ per ogni $s \in [c, d]$, $\tau(c) = a$ e $\tau(d) = b$;
 - (ii) τ è decrescente, $\tau'(s) < 0$ per ogni $s \in [c, d]$, $\tau(c) = b$ e $\tau(d) = a$.

L'osservazione precedente giustifica l'idea di considerare due curve ottenute l'una dall'altra tramite un cambiamento di parametrizzazione come *due parametrizzazioni della stessa curva*. Nella situazione (i) diremo inoltre che il cambiamento di parametro *conserva l'orientazione della curva*, nella situazione (ii)

che *inverte l'orientazione*. Per enfatizzare l'idea di una curva come sottoinsieme del piano o dello spazio, chiameremo *supporto* di una curva parametrizzata la sua immagine, e considereremo tale supporto a meno di cambiamenti di parametrizzazione. Diremo inoltre che il supporto è *orientato* se è visto a meno di cambiamenti di parametrizzazione che conservano l'orientazione.

Notiamo che se $\alpha : [a, b] \rightarrow \mathbb{R}^n$ rappresenta una curva orientata, allora $\beta : [a, b] \rightarrow \mathbb{R}^n$ data da $\beta(s) = \alpha(a + b - s)$ rappresenta la stessa curva con l'orientazione opposta.

14.1.2 Lunghezza

La lunghezza di un segmento rettilineo in \mathbb{R}^n è la distanza tra i suoi estremi, ovvero la norma della loro differenza. Per una curva regolare qualsiasi invece la formalizzazione della nozione di lunghezza richiede qualche sforzo. Cominciamo a notare che se suddividiamo l'intervallo $[a, b]$ tramite punti $a = t_0 < t_1 < t_2 < \dots < t_{N-1} < t_N = b$, e supponiamo che la suddivisione sia molto fitta, possiamo bene approssimare α con la spezzata di vertici $\alpha(t_0), \alpha(t_1), \alpha(t_2), \dots, \alpha(t_{N-1}), \alpha(t_N)$, la quale ha lunghezza

$$\sum_{j=1}^N \|\alpha(t_j) - \alpha(t_{j-1})\|. \quad (14.1)$$

Possiamo quindi notare che $\alpha(t_j)$ è approssimato da

$$\alpha(t_{j-1}) + (t_j - t_{j-1}) \cdot \alpha'(t_{j-1}),$$

dunque la lunghezza (14.1) è approssimata da

$$\sum_{j=1}^N (t_j - t_{j-1}) \cdot \|\alpha'(t_{j-1})\|. \quad (14.2)$$

Per l'ipotesi di regolarità di α abbiamo ora che $\|\alpha'\|$ è continua, dunque integrabile; inoltre la somma (14.2), al tendere di N a $+\infty$ e di $\max_{1 \leq j \leq N} (t_j - t_{j-1})$ a 0, tende all'integrale di $\|\alpha'\|$ su $[a, b]$. Ciò giustifica la:

Definizione 14.1.6. Se $\alpha : [a, b] \rightarrow \mathbb{R}^n$ è una curva regolare, chiamiamo *lunghezza* di α la quantità

$$L(\alpha) = \int_a^b \|\alpha'(t)\| dt.$$

Notiamo che la definizione di lunghezza si estende in modo evidente al caso di curve regolari a tratti, sommando gli integrali sui tratti regolari.

Esempio 14.1.7. Se parametrizziamo la circonferenza di centro O e raggio $r > 0$ come $\alpha(\vartheta) = (r \cos \vartheta, r \sin \vartheta)$ con $\vartheta \in [0, 2\pi]$, abbiamo $\|\alpha'(\vartheta)\| = r$ per ogni ϑ , dunque ritroviamo il fatto noto che $L(\alpha) = 2\pi r$. Ma possiamo riottenere lo stesso risultato notando che la circonferenza è unione delle curve $\beta_{\pm}(x) = \left(x, \sqrt{r^2 - x^2}\right)$ con $x \in [-r, r]$. Abbiamo infatti

$$\|\beta'_{\pm}(x)\| = \sqrt{1 + \left(\frac{1}{2} \cdot \frac{-2x}{\sqrt{r^2 - x^2}}\right)^2} = \frac{r}{\sqrt{r^2 - x^2}} = \frac{1}{\sqrt{1 - \frac{x^2}{r^2}}}$$

e con la sostituzione $u = \frac{x}{r}$ troviamo

$$L(\beta_+) + L(\beta_-) = 2 \int_{-1}^1 \frac{1}{\sqrt{1 - u^2}} r \, du = 2r \cdot \arcsin u \Big|_{-1}^1 = 2r \left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right) \right) = 2\pi r.$$

Il prossimo risultato traduce il fatto del tutto intuitivo che la lunghezza è una proprietà del supporto di una curva e non di una sua parametrizzazione:

Proposizione 14.1.8. Se β è ottenuta da α tramite un cambiamento di parametro allora $L(\beta) = L(\alpha)$.

Dimostrazione. Sia $\beta(s) = \alpha(\tau(s))$ con $\tau : [c, d] \rightarrow [a, b]$. Allora

$$\beta'(s) = \alpha'(\tau(s)) \cdot \tau'(s) \quad \Rightarrow \quad \|\beta'(s)\| = \|\alpha'(\tau(s))\| \cdot |\tau'(s)|$$

e il risultato segue immediatamente dalle formule di integrazione per sostituzione:

$$L(\beta) = \int_c^d \|\beta'(s)\| \, ds = \int_c^d \|\alpha'(\tau(s))\| \cdot |\tau'(s)| \, ds = \int_a^b \|\alpha'(t)\| \, dt = L(\alpha).$$

□

14.1.3 Integrale su una curva di una funzione

Esaminando la formula $L(\alpha) = \int_a^b \|\alpha'(t)\| \, dt$ che fornisce la lunghezza di una curva α , possiamo ora interpretare $\|\alpha'(t)\| \, dt$ come l'*elemento di lunghezza* di α

nel punto $\alpha(t)$. Supponendo che il supporto di α giaccia nel dominio Ω di una funzione $f : \Omega \rightarrow \mathbb{R}$, possiamo poi estendere tale interpretazione considerando $f(\alpha(t)) \cdot \|\alpha'(t)\| dt$ come il “costo elementare” dell’elemento di lunghezza di α in $\alpha(t)$ contato con il “peso” $f(\alpha(t))$. Integrando questo “costo elementare” otteniamo allora il “costo totale” di α nel quale in ciascun punto x di Ω un elemento di lunghezza viene contato con il peso $f(x)$:

Definizione 14.1.9. Data $f : \Omega \rightarrow \mathbb{R}$ continua e $\alpha : [a, b] \rightarrow \Omega$ regolare, chiamiamo *integrale di f su α* la quantità

$$\int_{\alpha} f = \int_a^b f(\alpha(t)) \cdot \|\alpha'(t)\| dt.$$

Il prossimo risultato si dimostra con un argomento del tutto analogo a quello che abbiamo impiegato per la Proposizione 14.1.8:

Proposizione 14.1.10. Se β è ottenuta da α tramite un cambiamento di parametro allora $\int_{\alpha} f = \int_{\beta} f$.

14.1.4 Parametro d’arco

Interpretando una curva come un tracciato unidimensionale nel piano o nello spazio, vediamo che la norma della derivata di una sua parametrizzazione rappresenta la velocità scalare con cui la curva viene percorsa tramite quella parametrizzazione. È del tutto intuitivo che è sempre possibile percorrere qualsiasi tracciato a velocità scalare costante unitaria; come vedremo nei prossimi paragrafi, si tratta di una situazione particolarmente importante dal punto di vista geometrico, a cui conviene dare un nome e di cui conviene dimostrare con cura l’esistenza:

Definizione 14.1.11. Diciamo che una curva parametrizzata $\alpha : [a, b] \rightarrow \mathbb{R}^n$ regolare è in *parametro d’arco* se $\|\alpha'(t)\| = 1$ per ogni $t \in [a, b]$.

Teorema 14.1.12. Data $\alpha : [a, b] \rightarrow \mathbb{R}^n$ di lunghezza L esiste un cambiamento di parametro $\tau : [0, L] \rightarrow [a, b]$ tale che $\beta = \alpha \circ \tau$ è in *parametro d’arco*. Inoltre tale cambiamento di parametro conserva l’orientazione.

Dimostrazione. Ponendo $\sigma(t) = \int_a^t \|\alpha'(u)\| du$ otteniamo una bigezione derivabile e crescente $\sigma : [a, b] \rightarrow [0, L]$ con $\sigma'(t) = \|\alpha'(t)\|$. Poniamo allora $\tau = \sigma^{-1}$

e $\beta = \alpha \circ \tau$. Poiché $\tau'(s) = \frac{1}{\sigma'(\tau(s))} = \frac{1}{\|\alpha'(\tau(s))\|}$ e $\beta'(s) = \alpha'(\tau(s)) \cdot \tau'(s)$ concludiamo immediatamente che $\|\beta'(s)\| = 1$ per ogni $s \in [0, L]$. \square

Osservazione 14.1.13. La costruzione del cambiamento di parametro operata nella dimostrazione precedente ha una chiara interpretazione. Ponendo infatti $\sigma(t) = \int_a^t \|\alpha'(u)\| du$ si sta scegliendo di usare come parametro $s = \sigma(t)$ per parametrizzare la curva proprio la lunghezza della curva tra il punto $\alpha(a)$ e il punto $\alpha(t)$. Ed è chiaro che se come parametro (cioè come “tempo”) si usa la lunghezza, allora la velocità scalare (lunghezza fratto tempo) risulterà unitaria.

14.1.5 Teoremi di Jordan e Dini

Questo sottoparagrafo è dedicato a tre importanti teoremi che non dimostriamo, o almeno non completamente. Il primo risultato, il teorema di Jordan, è del tutto intuitivo ma in realtà decisamente difficile da verificare formalmente. Iniziamo ricordando due nozioni, la prima delle quali abbiamo già impiegato nei Sottoparagrafi 9.3.1 e 10.2.1.

Definizione 14.1.14. Un sottoinsieme $A \subset \mathbb{R}^n$ è *aperto* se per ogni $x \in A$ esiste $r > 0$ tale che $\{y \in \mathbb{R}^n : \|y - x\| < r\} \subset A$. Inoltre A è *connesso* se per ogni $x_0, x_1 \in A$ esiste una curva $\gamma : [0, 1] \rightarrow A$ tale che $\gamma(0) = x_0$ e $\gamma_1 = x_1$.

Enunciamo dunque il Teorema di Jordan:

Teorema 14.1.15. *Se $\alpha : [a, b] \rightarrow \mathbb{R}^2$ è una curva semplice e chiusa e J è il suo supporto allora $\mathbb{R}^2 \setminus J$ consiste di due aperti connessi disgiunti, uno limitato e uno illimitato.*

Passando al secondo argomento del sottoparagrafo, analizziamo il rapporto tra la nozione di curva e quella di grafico di una funzione. Osserviamo intanto che certamente sono curve il grafico di una funzione $y = f(x)$ e quello di una funzione $x = f(y)$ con $f : [a, b] \rightarrow \mathbb{R}$ derivabile con derivata continua, dato che possono essere parametrizzati rispettivamente come $\alpha(t) = (t, f(t))$ e come $\alpha(t) = (f(t), t)$ con $t \in [a, b]$. È certamente falso che ogni curva sia un grafico (ad esempio, una circonferenza non lo è), ma il seguente risultato (che non dimostriamo formalmente) consente di affermare che lo è *localmente*:

Proposizione 14.1.16. *Sia $\alpha : [a, b] \rightarrow \mathbb{R}^2$ una curva semplice e regolare di componenti $\alpha = (X, Y)$, e sia $t_0 \in (a, b)$. Se $X'(t_0) \neq 0$ esistono $\varepsilon, \delta > 0$ tali che l'immagine di α intersecata con il rettangolo*

$$[X(t_0) - \varepsilon, X(t_0) + \varepsilon] \times [Y(t_0) - \delta, Y(t_0) + \delta]$$

è il grafico di una funzione $y = f(x)$ con $f : [X(t_0) - \varepsilon, X(t_0) + \varepsilon] \rightarrow \mathbb{R}$. Se invece $Y'(t_0) \neq 0$ esistono $\varepsilon, \delta > 0$ tali che l'immagine di α intersecata con il rettangolo

$$[X(t_0) - \delta, X(t_0) + \delta] \times [Y(t_0) - \varepsilon, Y(t_0) + \varepsilon]$$

è il grafico di una funzione $x = f(y)$ con $f : [Y(t_0) - \varepsilon, Y(t_0) + \varepsilon] \rightarrow \mathbb{R}$.

L'idea geometrica di questo risultato è molto naturale: ad esempio nel caso $X'(t_0) \neq 0$ si ha che la funzione X è crescente oppure decrescente vicino a t_0 , quindi vicino a t_0 è invertibile, e allora basta prendere $f = Y \circ X^{-1}$ opportunamente ristretta. I dettagli però sono abbastanza laboriosi.

Veniamo dunque al terzo argomento del sottoparagrafo, ovvero alle relazioni tra la nozione di curva e quella di luogo di zeri (o insieme di livello) di una funzione. Cominciamo notando che sono curve (ad esempio perché sono localmente grafici) tutte le coniche non degeneri (ellissi, iperboli, parabole), e che tali coniche sono inoltre definite da equazioni, ovvero sono luoghi di zeri di funzioni $F : \mathbb{R}^2 \rightarrow \mathbb{R}$. Si pone dunque in modo naturale il problema di stabilire quali luoghi di zeri di funzioni sono curve. Certamente, anche per funzioni F molto regolari, non lo sono tutti:

Esempio 14.1.17. Se $F(x, y) = x^2 + y^2$ allora il luogo di equazione $F(x, y) = 0$ è il solo punto $(0, 0)$. Se $F(x, y) = xy$ allora il luogo di equazione $F(x, y) = 0$ è l'unione degli assi cartesiani, dunque non è il supporto di una curva semplice vicino a $(0, 0)$.

Abbiamo ora il seguente fondamentale risultato, dovuto a Ulisse Dini:

Teorema 14.1.18. Sia $\Omega \subset \mathbb{R}^2$ un aperto e $F : \Omega \rightarrow \mathbb{R}$ una funzione con derivate parziali continue. Sia $(x_0, y_0) \in \Omega$ un punto tale che $F(x_0, y_0) = 0$ e $\text{grad}_{(x_0, y_0)}(F) \neq 0$. Allora vicino a (x_0, y_0) il luogo

$$Z = \{(x, y) \in \Omega : f(x, y) = 0\}$$

è il grafico di una funzione derivabile con derivata continua di una delle due coordinate in funzione dell'altra. Più esattamente se $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ si ha che vicino a (x_0, y_0) il luogo Z è il grafico di una funzione $y = Y(x)$ con

$$Y'(x) = -\frac{\partial F}{\partial x}(x, Y(x)) \cdot \left(\frac{\partial F}{\partial y}(x, Y(x)) \right)^{-1};$$

invece se $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$ si ha che vicino a (x_0, y_0) il luogo Z è il grafico di una funzione $x = X(y)$ con

$$X'(y) = -\frac{\partial F}{\partial y}(X(y), y) \cdot \left(\frac{\partial F}{\partial x}(X(y), y) \right)^{-1}.$$

Di questo teorema daremo una dimostrazione parziale, ma prima ne deduciamo il seguente:

Corollario 14.1.19. *Sia $\Omega \subset \mathbb{R}^2$ un aperto e $F : \Omega \rightarrow \mathbb{R}$ una funzione con derivate parziali continue. Se $\text{grad}_{(x,y)} F \neq 0$ per ogni $(x, y) \in \Omega$ allora gli insiemi di livello $\{(x, y) \in \Omega : F(x, y) = c\}$ al variare di $c \in \mathbb{R}$ sono tutti curve, e in ogni punto $(x, y) \in \Omega$ la retta tangente all'insieme di livello che passa per (x, y) è quella ortogonale a $\text{grad}_{(x,y)} F$.*

Dimostrazione. Applicando alla funzione $(x, y) \mapsto F(x, y) - c$ il Teorema 14.1.18 vediamo che l'insieme di livello c , se non vuoto, è localmente grafico di una coordinata in funzione dell'altra, dunque è una curva. A seconda di quale derivata parziale di F nel punto (x, y) sia non nulla, il vettore tangente alla curva nel punto è uno dei seguenti

$$\left(1, -\frac{\partial F}{\partial x}(x, y) \cdot \left(\frac{\partial F}{\partial y}(x, y) \right)^{-1} \right), \quad \left(-\frac{\partial F}{\partial y}(x, y) \cdot \left(\frac{\partial F}{\partial x}(x, y) \right)^{-1}, 1 \right),$$

dunque in ogni caso è ortogonale a $\text{grad}_{(x,y)} F = \left(\frac{\partial F}{\partial x}(x, y), \frac{\partial F}{\partial y}(x, y) \right)$. \square

Tornando alla dimostrazione del teorema del Dini, supponiamo ad esempio che $\frac{\partial F}{\partial y}(x_0, y_0) > 0$. Poiché F ha derivate parziali continue possiamo scegliere $\varepsilon, \delta > 0$ tali che $\frac{\partial F}{\partial y} > 0$ su $[x_0 - \varepsilon, x_0 + \varepsilon] \times [y_0 - \delta, y_0 + \delta]$. Esaminando la funzione $y \mapsto F(x_0, y)$ su $[y_0 - \delta, y_0 + \delta]$ vediamo che è strettamente crescente e nulla per $y = y_0$, dunque $F(x_0, y_0 - \delta) < 0 < F(x_0, y_0 + \delta)$. Poiché F è continua possiamo ora supporre, eventualmente prendendo un valore più piccolo di ε , che $F(x, y_0 - \delta) < 0 < F(x, y_0 + \delta)$ per ogni $x \in [x_0 - \varepsilon, x_0 + \varepsilon]$. Per ogni $x \in [x_0 - \varepsilon, x_0 + \varepsilon]$ esaminiamo ora la funzione $y = F(x, y)$ su $[y_0 - \delta, y_0 + \delta]$, osservando che è strettamente crescente, negativa nel primo estremo e nulla nel secondo; ciò comporta che essa si annulla in un unico punto $Y(x)$, il che fornisce la funzione $y = Y(x)$ di cui Z è il grafico vicino a (x_0, y_0) . È ora possibile, ma tecnicamente un po' complicato, verificare che tale Y è continua con derivata continua. Dando per buono questo fatto possiamo ricavare l'espressione di

$Y'(x)$ derivando (con la Proposizione 9.3.5) l'identità $F(x, Y(x)) = 0$, da cui otteniamo

$$\begin{aligned} \frac{\partial F}{\partial x}(x, Y(x)) + \frac{\partial F}{\partial y}(x, Y(x)) \cdot Y'(x) &= 0 \\ \Rightarrow Y'(x) &= -\frac{\partial F}{\partial x}(x, Y(x)) \cdot \left(\frac{\partial F}{\partial y}(x, Y(x)) \right)^{-1}. \end{aligned}$$

Dall'enunciato del Teorema 14.1.18 segue facilmente il:

Corollario 14.1.20. *Se F soddisfa le ipotesi del teorema del Dini e ha tutte le derivate parziali continue fino all'ordine k allora la funzione di cui l'insieme degli zeri di F è localmente grafico ha derivate continue fino all'ordine k .*

14.2 Curvatura di una curva piana

Il termine “curva” che abbiamo impiegato per gli oggetti di cui ci stiamo occupando appartiene anche al linguaggio naturale e indica un tratto non rettilineo di un percorso. In questo paragrafo, occupandoci esclusivamente delle curve piane, introdurremo una misura numerica di quanto punto per punto una curva sia *non rettilinea*.

14.2.1 Circonferenza osculatrice e curvatura

Abbiamo introdotto la retta tangente a una curva regolare $\alpha : [a, b] \rightarrow \mathbb{R}^2$ in un suo punto $\alpha(t_0)$ come l'unica retta avente contatto almeno duplice con α nel punto. Equivalentemente, possiamo vedere la tangente come il limite per t che tende a t_0 dell'unica retta passante per t e t_0 . La retta tangente fornisce un'approssimazione della curva al primo ordine, e in questo paragrafo scopriremo come approssimare la curva al secondo ordine, usando una circonferenza invece che una retta. L'idea geometrica è molto semplice: se vicino ad $\alpha(t_0)$ la curva non è rettilinea, presi valori distinti t_1, t_2, t_3 vicini a t_0 , avremo che $\alpha(t_1), \alpha(t_2), \alpha(t_3)$ sono punti del piano non allineati, dunque esiste un'unica circonferenza che li contiene. Facendo tendere t_1, t_2, t_3 contemporaneamente a t_0 tale circonferenza tenderà allora a una in cui i tre punti di contatto con α si sono fusi in un unico punto di contatto di ordine triplice. La formalizzazione di questa idea richiede però qualche passaggio analitico.

Iniziamo con un risultato valido in qualsiasi dimensione:

Lemma 14.2.1. *Se $v : [a, b] \rightarrow \mathbb{R}^n$ è una funzione derivabile e $\|v(t)\| = 1$ per ogni t allora $v'(t) \perp v(t)$ per ogni t .*

Dimostrazione. Derivando l'identità $\|v(t)\|^2 = 1$, ovvero $\sum_{j=1}^n (v_j(t))^2 = 1$, otteniamo $\sum_{j=1}^n 2v_j(t)v'_j(t) = 0$, ovvero $\langle v(t)|v'(t) \rangle = 0$. \square

Proposizione 14.2.2. *Sia $\alpha : [a, b] \rightarrow \mathbb{R}$ una curva regolare in parametro d'arco e derivabile due volte. Sia $t_0 \in [a, b]$ tale che $\alpha''(t_0) \neq 0$. Allora l'unica circonferenza avente contatto almeno triplice con α in $\alpha(t_0)$ è quella passante per $\alpha(t_0)$ e avente centro $\alpha(t_0) + \frac{\alpha''(t_0)}{\|\alpha''(t_0)\|^2}$.*

Dimostrazione. A meno di movimenti rigidi del piano possiamo supporre che $\alpha(t_0) = 0$ e $\alpha'(t_0) = (1, 0)$. Il Lemma 14.2.1 applicato a $v(t) = \alpha'(t)$ garantisce allora che $\alpha''(t_0) = (0, c)$ per qualche $c \in \mathbb{R}$ non nullo. Una circonferenza avente contatto almeno triplice con α in 0 certamente passa per 0 e ha in 0 retta tangente orizzontale, dunque ha centro $(0, r)$ e raggio $|r|$ per qualche $r \in \mathbb{R}$. Il contatto è poi almeno triplice se la funzione

$$f(t) = \|\alpha(t) - (0, r)\| - |r|$$

ha la derivata seconda nulla in t_0 , infatti le scelte fatte garantiscono già che $f(t_0) = f'(t_0) = 0$. Notiamo ora che

$$\begin{aligned} f(t) &= \sqrt{X(t)^2 + (Y(t) - r)^2} - |r| \\ f'(t) &= \frac{X(t)X'(t) + (Y(t) - r)Y'(t)}{\sqrt{X(t)^2 + (Y(t) - r)^2}} \\ f''(t) &= -\frac{(X(t)X'(t) + (Y(t) - r)Y'(t))^2}{\sqrt{(X(t)^2 + (Y(t) - r)^2)^3}} \\ &\quad + \frac{X(t)X''(t) + X'(t)^2 + (Y(t) - r)Y''(t) + Y'(t)^2}{\sqrt{X(t)^2 + (Y(t) - r)^2}}. \end{aligned}$$

Ne segue che $f''(t_0)$ può annullarsi solo se in $t = t_0$ si annulla

$$X(t)X''(t) + X'(t)^2 + (Y(t) - r)Y''(t) + Y'(t)^2,$$

e troviamo $0 + 1 + (-r)c + 0 = 0$, da cui $r = \frac{1}{c}$, che è precisamente quanto dovevamo dimostrare. \square

Definizione 14.2.3. Chiamiamo *circonferenza osculatrice* per una curva regolare in un suo punto, se esiste, la circonferenza avente contatto almeno triplice con la curva nel punto.

Osserviamo che la Proposizione 14.2.2 descrive la circonferenza osculatrice per una curva in parametro d'arco, ma il Teorema 14.1.12 garantisce che ogni curva regolare si può parametrizzare in parametro d'arco, dunque la definizione ha natura generale. Naturalmente però possono esistere punti nei quali la circonferenza osculatrice non esiste: per una curva α in parametro d'arco sono quelli in cui si annulla α'' ; per una curva qualsiasi sono quelli in cui α'' è multiplo di α' (ma verificarlo non è immediato, vedi anche la successiva Proposizione 14.2.7).

Per una circonferenza una buona misura numerica di quanto sia *curva* è il reciproco del raggio (infatti ad esempio a un raggio molto grande corrisponde un andamento quasi rettilineo). Ciò giustifica la:

Definizione 14.2.4. Chiamiamo *valore assoluto della curvatura* di una curva α in parametro d'arco in un suo punto $\alpha(t_0)$ la quantità $\|\alpha''(t_0)\|$.

Discuteremo tra poco come si possa raffinare la definizione precedente introducendo per la curvatura anche un segno. Prima di farlo però osserviamo che grazie al Teorema 14.1.12 possiamo estendere la nozione di curvatura a qualsiasi curva regolare, impiegando la riparametrizzazione in parametro d'arco; tuttavia esplicitare quest'ultima è tipicamente impossibile nella pratica, dunque avremo bisogno di metodi più efficaci.

14.2.2 Segno della curvatura e suo calcolo

Torniamo a considerare una curva $\alpha : [a, b] \rightarrow \mathbb{R}$ in parametro d'arco, considerandola come curva orientata, dunque vista soltanto a meno di riparametrizzazioni consistenti in una traslazione dell'intervallo di definizione. Se $t_0 \in [a, b]$ e $\alpha''(t_0) \neq 0$ sappiamo che tale vettore è ortogonale ad $\alpha'(t_0)$; possiamo dunque dare la seguente:

Definizione 14.2.5. Chiamiamo *curvatura* di α nel punto $\alpha(t_0)$ la quantità $\kappa(t_0)$ che vale 0 se $\alpha''(t_0) = 0$, vale $\|\alpha''(t_0)\|$ se $\alpha''(t_0) \neq 0$ e la rotazione di angolo $\frac{\pi}{2}$ che trasforma $\alpha'(t_0)$ in un multiplo positivo di $\alpha''(t_0)$ avviene in verso antiorario, mentre vale $-\|\alpha''(t_0)\|$ se $\alpha''(t_0) \neq 0$ e la rotazione di angolo $\frac{\pi}{2}$ che trasforma $\alpha'(t_0)$ in un multiplo positivo di $\alpha''(t_0)$ avviene in verso orario.

Geometricamente una curvatura positiva di α in un certo punto corrisponde al fatto che percorrendo α si sta svoltando in quel punto verso sinistra. Naturalmente una svolta a sinistra durante il viaggio di andata corrisponde a una svolta a destra nel viaggio di ritorno, dunque abbiamo la:

Osservazione 14.2.6. Invertendo l'orientazione di una curva la curvatura punto per punto cambia segno.

Illustriamo infine il metodo di calcolo della curvatura per una curva orientata espressa tramite una parametrizzazione qualsiasi, non necessariamente in parametro d'arco.

Proposizione 14.2.7. Se $\alpha : [a, b] \rightarrow \mathbb{R}^3$ è una curva regolare, la curvatura di α in un suo punto $\alpha(t)$ vale

$$\kappa(t) = \frac{\det \begin{pmatrix} \alpha'(t) & \alpha''(t) \end{pmatrix}}{\|\alpha'(t)\|^3}.$$

Anche se non diamo una dimostrazione completa di questa proposizione, possiamo almeno convincerci che la formula per $\kappa(t)$ è plausibile. Osserviamo intanto che $\det(\alpha'(t) \ \alpha''(t))$ coincide con $\det(\alpha'(t) \ v)$ dove v è la proiezione ortogonale di $\alpha''(t)$ su $\alpha'(t)^\perp$, dunque la formula al numeratore supplisce al fatto che se α non è in parametro d'arco non è detto che $\alpha''(t)$ sia ortogonale ad $\alpha'(t)$. Inoltre è chiaro che una formula per la curvatura deve contenere una normalizzazione rispetto ad $\|\alpha'(t)\|$, dato che è sempre possibile modificare una parametrizzazione semplicemente cambiando di un fattore costante la velocità di percorrenza; e possiamo vedere che l'esponente 3 di $\|\alpha'(t)\|$ al denominatore è quello giusto ponendo $\beta(s) = \alpha(\lambda \cdot s)$; abbiamo infatti (indicando $\lambda \cdot s$ con t)

$$\begin{aligned} \beta'(s) &= \lambda \cdot \alpha'(t), & \beta''(s) &= \lambda^2 \cdot \alpha''(t) \\ \Rightarrow \frac{\det \begin{pmatrix} \beta'(s) & \beta''(s) \end{pmatrix}}{\|\beta'(s)\|^3} &= \frac{\lambda^3 \cdot \det \begin{pmatrix} \alpha'(t) & \alpha''(t) \end{pmatrix}}{\lambda^3 \cdot \|\alpha'(t)\|^3} &= \frac{\det \begin{pmatrix} \alpha'(t) & \alpha''(t) \end{pmatrix}}{\|\alpha'(t)\|^3}. \end{aligned}$$

Esempio 14.2.8. Per concludere che $\kappa(t_0)$ è, a meno del segno, il valore $\|\alpha''(t_0)\|$, bisogna supporre che $\alpha'(t)$ valga 1 per ogni t , non basta supporre che $\|\alpha'(t_0)\| = 1$. Ad esempio per $\alpha(t) = (t + t^2, t^2)$ in $t_0 = 0$ si ha $\|\alpha'(0)\| = 1$ e $\alpha''(0) = (2, 2)$, dunque $\|\alpha''(0)\| = 2\sqrt{2}$, mentre $\kappa(0) = +2$.

A conclusione del paragrafo enunciamo un risultato del tutto ragionevole, anche se non immediato da dimostrare formalmente:

Teorema 14.2.9. Dati nel piano un punto P_0 e un vettore unitario v_0 e data una funzione continua $\kappa : [a, b] \rightarrow \mathbb{R}$ esiste una e una sola curva in parametro d'arco $\alpha : [a, b] \rightarrow \mathbb{R}^2$ tale che $\alpha(a) = P_0$ e $\alpha'(a) = v_0$ e avente punto per punto curvatura $\kappa(t)$.

In pratica questo risultato significa che *se so da dove parto e in quale direzione, so che andrò a velocità costante e so istante per istante quanto starò curvando verso sinistra o verso destra, allora so dove sarò in ogni istante.*

14.3 Curvatura e torsione di una curva nello spazio

In questo paragrafo estendiamo l'analisi geometrica delle curve piane a quelle nello spazio tridimensionale. Per comodità parlando di curve piane abbiamo spesso espresso i punti come vettori riga, ma converrà per il caso di curve nello spazio usare sempre vettori colonna. Inoltre eviteremo di usare t come parametro, poiché il simbolo t avrà un altro significato. Cominciamo con il seguente risultato (senza dimostrazione) che estende la Proposizione 14.2.2:

Proposizione 14.3.1. *Sia $\alpha : [a, b] \rightarrow \mathbb{R}^3$ una curva in parametro d'arco e sia $s_0 \in [a, b]$ tale che $\alpha''(s_0) \neq 0$. Allora:*

- *Esiste uno e un solo piano in \mathbb{R}^3 con il quale α ha in $\alpha(s_0)$ un contatto di ordine almeno triplice; tale piano è quello passante per $\alpha(s_0)$ con giacitura generata da $\alpha'(s_0)$ e $\alpha''(s_0)$;*
- *Su tale piano esiste una e una sola circonferenza con la quale α ha in s_0 un contatto di ordine almeno triplice; tale circonferenza è quella passante per $\alpha(s_0)$ con centro in $\alpha(s_0) + \frac{\alpha''(s_0)}{\|\alpha''(s_0)\|^2}$.*

Definizione 14.3.2. Chiameremo rispettivamente *piano osculatore* e *circonferenza osculatrice* per α in $\alpha(s_0)$ il piano e la circonferenza la cui esistenza è assicurata dalla proposizione precedente. Chiameremo inoltre *curvatura* di α in s_0 il numero $\kappa(s_0) = \|\alpha''(s_0)\|$.

Usando la Proposizione 14.2.7 non è difficile provare la seguente:

Proposizione 14.3.3. *La curvatura di una curva $\alpha : [a, b] \rightarrow \mathbb{R}^3$ in un suo punto $\alpha(s_0)$ coincide con la curvatura in $\alpha(s_0)$ della proiezione ortogonale di α sul suo piano osculatore in $\alpha(s_0)$.*

Ci proponiamo ora di introdurre una misura numerica della non-planarità di una curva α nello spazio. A questo fine introduciamo la seguente terminologia:

Definizione 14.3.4. Una curva $\alpha : [a, b] \rightarrow \mathbb{R}^3$ in parametro d'arco è detta *biregolare* se $\alpha''(s) \neq 0$ per ogni $s \in [a, b]$. Per una tale α chiamiamo:

- *vettore tangente* in $\alpha(s)$ il vettore $t(s) = \alpha'(s)$;

- *vettore normale* in $\alpha(s)$ il vettore $n(s) = \frac{\alpha''(s)}{\|\alpha''(s)\|}$;
- *vettore binormale* in $\alpha(s)$ il vettore $b(s) = t(s) \wedge n(s)$;
- *riferimento di Frénet* in $\alpha(s)$ la terna di vettori $(t(s), n(s), b(s))$.

Osserviamo che per costruzione $(t(s), n(s), b(s))$ è una base ortonormale positivamente orientata di \mathbb{R}^3 , ovvero $(t(s), n(s), b(s))$ è una matrice ortogonale con determinante +1. Inoltre la giacitura del piano osculatore per α in $\alpha(s)$, essendo generata da $\alpha'(s) = t(s)$ e $\alpha''(s) = \kappa(s)^{-1} \cdot n(s)$, è il piano ortogonale a $b(s)$. La variazione del piano osculatore è dunque misurata da $b'(s)$, per la quale abbiamo il fondamentale:

Teorema 14.3.5. *Se $\alpha : [a, b] \rightarrow \mathbb{R}^3$ è una curva in parametro d'arco e biregolare, esiste una funzione $\tau : [a, b] \rightarrow \mathbb{R}$ tale che $b'(s) = -\tau(s) \cdot n(s)$. Inoltre $n'(s) = -\kappa(s) \cdot t(s) + \tau(s) \cdot b(s)$.*

Dimostrazione. Per definizione abbiamo che $\|b'(s)\| = 1$ per ogni s , dunque il Lemma 14.2.1 garantisce che $b'(s)$ è ortogonale a $b(s)$. Inoltre derivando la relazione $b(s) = t(s) \wedge n(s)$ che definisce $b(s)$ otteniamo

$$\begin{aligned} b'(s) &= t'(s) \wedge n(s) + t(s) \wedge n'(s) \\ &= \kappa(s) \cdot n(s) \wedge n(s) + t(s) \wedge n'(s) \\ &= t(s) \wedge n'(s), \end{aligned}$$

il che comporta che $b'(s)$ è ortogonale anche a $t(s)$. Poiché la retta ortogonale a $t(s)$ e $b(s)$ è quella generata da $n(s)$, ne segue che $b'(s)$ è multiplo di $n(s)$, il che garantisce l'esistenza di $\tau(s)$. Per la seconda formula basta derivare la relazione $n(s) = b(s) \wedge t(s)$, ottenendo

$$\begin{aligned} n'(s) &= b'(s) \wedge t(s) + b(s) \wedge t'(s) \\ &= -\tau(s) \cdot n(s) \wedge t(s) + b(s) \wedge \kappa(s) \cdot n(s) \\ &= \tau(s) \cdot b(s) - \kappa(s) \cdot t(s). \end{aligned}$$

□

In forma matriciale il contenuto del teorema precedente è espresso dalle seguenti formule, dette di Frénet-Serret:

Corollario 14.3.6.

$$(t'(s), n'(s), b'(s)) = (t(s), n(s), b(s)) \cdot \begin{pmatrix} 0 & -\kappa(s) & 0 \\ \kappa(s) & 0 & -\tau(s) \\ 0 & \tau(s) & 0 \end{pmatrix}.$$

Poiché il piano osculatore è quello ortogonale al vettore binormale, e la torsione rappresenta la misura numerica della variazione del vettore binormale, otteniamo facilmente il:

Corollario 14.3.7. *Una curva nello spazio è contenuta in un piano se e soltanto se ha torsione identicamente nulla.*

Per migliorare la comprensione concettuale delle formule di variazione del riferimento di Frénet, osserviamo che esse sono una manifestazione particolare del seguente risultato generale:

Proposizione 14.3.8. *Se $M : [a, b] \rightarrow \mathcal{M}_{n \times n}(\mathbb{R})$ ha coefficienti punto per punto derivabili e $M(t)$ è ortogonale per ogni t , allora esiste $A : [a, b] \rightarrow \mathcal{M}_{n \times n}(\mathbb{R})$ tale che per ogni t la $A(t)$ è antisimmetrica e $M'(t) = M(t) \cdot A(t)$.*

Dimostrazione. L'ipotesi che $M(t)$ sia ortogonale per ogni t significa che

$${}^tM(t) \cdot M(t) = I_n$$

per ogni t . Possiamo ora derivare questa identità osservando che al prodotto righe per colonne tra matrici si applica la regola di Leibniz (infatti il prodotto è bilineare) e che la derivata della trasposta è la trasposta della derivata (infatti la trasposizione è lineare); otteniamo allora:

$${}^tM'(t) \cdot M(t) + {}^tM(t) \cdot M'(t) = 0.$$

Poniamo ora $A(t) = M^{-1}(t) \cdot M'(t)$, ovvero $A(t) = {}^tM(t) \cdot M'(t)$; la formula appena trovata garantisce precisamente che $A(t)$ è antisimmetrica, e per definizione si ha $M'(t) = M(t) \cdot A(t)$, e la dimostrazione è completa. \square

Ricordando che una matrice ortogonale definisce un movimento rigido, a partire dalla proposizione precedente si può anche dimostrare il fatto citato all'inizio del Capitolo 10, secondo cui una matrice antisimmetrica di sforzi applicata a un corpo materiale induce sul corpo un movimento rigido, senza deformazione.

14.3.1 Calcolo di curvatura e torsione

Come per le curve piane, esplicitare la riparametrizzazione in parametro d'arco di una data curva è tipicamente impossibile nella pratica, dunque servirà la seguente:

Proposizione 14.3.9. *Se $\alpha : [a, b] \rightarrow \mathbb{R}^3$ è una curva biregolare, la sua curvatura e la sua torsione nel punto $\alpha(s)$ valgono rispettivamente*

$$\begin{aligned}\kappa(s) &= \frac{\|\alpha'(s) \wedge \alpha''(s)\|}{\|\alpha'(s)\|^3}, \\ \tau(s) &= \frac{\langle \alpha'(s) \wedge \alpha''(s) | \alpha'''(s) \rangle}{\|\alpha'(s) \wedge \alpha''(s)\|^2}.\end{aligned}$$

Osservazione 14.3.10. La formula per $\kappa(s)$ estende quella data nella Proposizione 14.2.7, infatti se a due vettori $v, w \in \mathbb{R}^2$ si aggiunge una terza componente

nulla per pensarli come vettori di \mathbb{R}^3 si ha $v \wedge w = \begin{pmatrix} 0 \\ 0 \\ \det(v, w) \end{pmatrix}$.

Non dimostriamo la Proposizione 14.3.9, ma di nuovo verifichiamo almeno che essa è coerente con il fatto che curvatura e torsione restano ben definite rispetto a cambiamenti di parametrizzazione del tipo $\beta(s) = \alpha(\lambda \cdot s)$. Per la curvatura lo abbiamo già visto dopo l'enunciato della Proposizione 14.2.7; per la torsione abbiamo invece:

$$\frac{\langle \beta' \wedge \beta'' | \beta''' \rangle}{\|\beta' \wedge \beta''\|^2} = \frac{\langle (\lambda \cdot \alpha') \wedge (\lambda^2 \cdot \alpha'') | (\lambda^3 \cdot \alpha''') \rangle}{\|(\lambda \cdot \alpha') \wedge (\lambda^2 \cdot \alpha'')\|^2} = \frac{\langle \alpha' \wedge \alpha'' | \alpha''' \rangle}{\|\alpha' \wedge \alpha''\|^2}$$

dove β e tutte le sue derivate sono calcolate in s , mentre α e tutte le sue derivate sono calcolate in $\lambda \cdot s$.

Capitolo 15

Integrazione su curve

Nel capitolo precedente abbiamo imparato a integrare sulle curve le funzioni scalari, mentre in questo (limitandoci al caso piano) lo faremo per oggetti diversi, di natura più vettoriale che scalare (anche se useremo una simbologia diversa). Ne dedurremo una teoria analitica con molte conseguenze pratiche.

15.1 1-forme e loro integrali

Supponiamo di avere una curva $\alpha : [a, b] \rightarrow \Omega$, con $\Omega \subset \mathbb{R}^2$ aperto, e chiamiamo X e Y le componenti di α . È facile immaginare situazioni concrete nelle quali il “costo di percorrenza” di α puntualmente assume la forma

$$f(\alpha(t)) \cdot X'(t) + g(\alpha(t)) \cdot Y'(t)$$

con $f, g : \Omega \rightarrow \mathbb{R}$. Il “costo” totale della percorrenza di α risulta allora

$$\int_a^b (f(\alpha(t)) \cdot X'(t) + g(\alpha(t)) \cdot Y'(t)) dt.$$

(Ad esempio, se a percorrere α è una particella dotata di massa e di carica elettrica, e se in Ω agiscono orizzontalmente un campo elettrico di intensità variabile f e verticalmente un campo gravitazionale di intensità variabile g , allora questo integrale rappresenta il lavoro complessivo compiuto dalle due forze.) Nella formula precedente possiamo distribuire il fattore dt sui due addendi dell'integranda, ottenendo le scritte $X'(t) dt$ e $Y'(t) dt$, che rappresentano gli spostamenti infinitesimi lungo α nelle direzioni x e y , e che quindi è ragionevole riscrivere come dx e dy . Ciò motiva la seguente:

Definizione 15.1.1. Dato $\Omega \subset \mathbb{R}^2$ aperto chiamiamo 1-forma su Ω una scrittura del tipo

$$\omega(x, y) = f(x, y) dx + g(x, y) dy$$

dove $f, g : \Omega \rightarrow \mathbb{R}$ sono funzioni continue. Se $\alpha : [a, b] \rightarrow \Omega$ è una curva regolare, chiamiamo *integrale di ω su α* il valore

$$\int_{\alpha} \omega = \int_a^b (f(\alpha(t)) \cdot X'(t) + g(\alpha(t)) \cdot Y'(t)) dt$$

dove X e Y sono le componenti di α .

Notiamo che la definizione di integrale di una 1-forma su una curva si estende in modo evidente al caso di curve regolari a tratti, sommando gli integrali sui tratti regolari.

Lemma 15.1.2. Se β è ottenuta da α tramite un cambiamento di parametro allora $\int_{\beta} \omega = \int_{\alpha} \omega$ nel caso in cui il cambiamento preserva l'orientazione, mentre $\int_{\beta} \omega = -\int_{\alpha} \omega$ se la inverte.

Dimostrazione. Sia $\beta(s) = \alpha(\tau(s))$ con $\tau : [c, d] \rightarrow [a, b]$. Allora dalla formula di derivazione della funzione composta segue che

$$\int_{\beta} \omega = \int_c^d (f(\alpha(\tau(s))) \cdot X'(\tau(s)) + g(\alpha(\tau(s))) \cdot y'(\tau(s))) \cdot \tau'(s) ds$$

dove come sopra X e Y sono le componenti di α . Ricordiamo ora la regola di integrazione per sostituzione applicata alla nostra $\tau : [c, d] \rightarrow [a, b]$ monotona e a qualsiasi funzione continua $h : [a, b] \rightarrow \mathbb{R}$:

$$\int_c^d h(\tau(s)) \cdot |\tau'(s)| ds = \int_a^b h(t) dt.$$

Osserviamo infine che il cambiamento di parametro conserva l'orientazione se $\tau'(s) > 0$ per ogni s , mentre la inverte se $\tau'(s) < 0$ per ogni s . Combinando le due formule fornite con questo fatto otteniamo subito la conclusione. \square

15.2 Forme esatte

Data una funzione $U : \Omega \rightarrow \mathbb{R}$ con $\Omega \subset \mathbb{R}^2$ aperto, abbiamo definito il suo gradiente come il vettore che ha come componenti le sue derivate parziali. Definiamo ora una variante del gradiente e ne proviamo subito la proprietà fondamentale:

Definizione 15.2.1. Chiamiamo *differenziale* di $U : \Omega \rightarrow \mathbb{R}$ la 1-forma su Ω

$$dU(x, y) = \frac{\partial U}{\partial x}(x, y) dx + \frac{\partial U}{\partial y}(x, y) dy.$$

Proposizione 15.2.2. Se $\alpha : [a, b] \rightarrow \Omega$ è una curva regolare e $U : \Omega \rightarrow \mathbb{R}$ ammette derivate parziali continue allora

$$\int_{\alpha} dU = U(\alpha(b)) - U(\alpha(a)).$$

Dimostrazione. Grazie alla Proposizione 9.3.5 abbiamo che la derivata della funzione $t \mapsto U(\alpha(t))$, ovvero $t \mapsto U(X(t), Y(t))$, vale

$$\frac{\partial U}{\partial x}(\alpha(t)) \cdot X'(t) + \frac{\partial U}{\partial y}(\alpha(t)) \cdot Y'(t)$$

e questa è precisamente la funzione da integrare tra a e b per calcolare $\int_{\alpha} dU$, dunque la conclusione segue dal teorema fondamentale del calcolo integrale. \square

Da questa proposizione discende in particolare che l'integrale su una curva del differenziale di una funzione dipende soltanto dagli estremi della curva. Diamo ora un nome alle 1-forme che sono differenziali di funzioni e verifichiamo subito che vale anche il viceversa del fatto appena enunciato.

Definizione 15.2.3. Diciamo che una 1-forma ω su $\Omega \subset \mathbb{R}^2$ è *esatta* su Ω se esiste $U : \Omega \rightarrow \mathbb{R}$ tale che $\omega = dU$. In tal caso diciamo che U è un *potenziale* per ω .

Proposizione 15.2.4. Sia $\Omega \subset \mathbb{R}^2$ un aperto connesso. Una 1-forma su $\Omega \subset \mathbb{R}^2$ è esatta su Ω se e soltanto se per ogni curva regolare $\alpha : [a, b] \rightarrow \Omega$ l'integrale $\int_{\alpha} \omega$ dipende soltanto da $\alpha(a)$ e $\alpha(b)$.

Dimostrazione. Sappiamo già che se ω ammette un potenziale U allora

$$\int_{\alpha} \omega = U(\alpha(b)) - U(\alpha(a)).$$

Viceversa, supponiamo che per ogni α il valore di $\int_{\alpha} \omega$ dipenda soltanto da $\alpha(a)$ e $\alpha(b)$. Fissiamo un punto $(x_0, y_0) \in \Omega$ e osserviamo che grazie all'ipotesi possiamo definire una funzione $U : \Omega \rightarrow \mathbb{R}$ ponendo

$$U(x, y) = \int_{\alpha} \omega \quad \text{dove } \alpha \text{ è una qualsiasi curva avente primo} \\ \text{estremo } (x_0, y_0) \text{ e secondo estremo } (x, y).$$

A questo punto basta provare che U è un potenziale per ω , ovvero che se $\omega = f dx + g dy$ si ha $\frac{\partial U}{\partial x} = f$ e $\frac{\partial U}{\partial y} = g$. Ci limitiamo a dimostrare la prima uguaglianza (la seconda è analoga). Prendiamo dunque $(x, y) \in \Omega$ e proponiamoci di calcolare il limite per h che tende a 0 del rapporto incrementale

$$\frac{1}{h}(U(x+h, y) - U(x, y))$$

(che è definito per $|h|$ abbastanza piccolo). Calcoleremo questo limite soltanto per h che tende a 0 da destra (ma da sinistra è del tutto analogo). Fissiamo allora una qualsiasi curva α di primo estremo (x_0, y_0) e secondo estremo (x, y) , dunque per definizione $U(x, y) = \int_{\alpha} \omega$. Osserviamo quindi che una curva di primo estremo (x_0, y_0) e secondo estremo $(x+h, y)$ si ottiene percorrendo prima α e poi la curva $\beta(t) = (x+t, y)$ con $t \in [0, h]$, dunque

$$U(x+h, y) = \int_{\alpha} \omega + \int_{\beta} \omega.$$

Sostituendo nel rapporto incrementale otteniamo che esso vale

$$\frac{1}{h} \int_{\beta} \omega = \frac{1}{h} \int_0^h (f(x+t, y) \cdot 1 + g(x+t, y) \cdot 0) dt = \frac{1}{h} \int_0^h f(x+t, y) dt$$

e grazie al teorema fondamentale del calcolo integrale il limite di questo rapporto esiste e vale $f(x, y)$. \square

Corollario 15.2.5. *Sia $\Omega \subset \mathbb{R}^2$ un aperto connesso. Una 1-forma su $\Omega \subset \mathbb{R}^2$ è esatta su Ω se e soltanto se per ogni curva $\alpha : [a, b] \rightarrow \mathbb{R}$ chiusa e regolare a tratti si ha $\int_{\alpha} \omega = 0$.*

Dimostrazione. Se ω è esatta, ovvero $\omega = dU$, e $\alpha : [a, b] \rightarrow \Omega$ è chiusa, allora

$$\int_{\alpha} \omega = U(\alpha(b)) - U(\alpha(a))$$

si annulla poiché $\alpha(b) = \alpha(a)$. Se invece $\int_{\alpha} \omega = 0$ per ogni α chiusa, affermiamo che $\int_{\beta} \omega$ dipende solo dagli estremi di β per ogni curva β in Ω , da cui la conclusione segue grazie alla proposizione precedente. Per verificare la nostra affermazione dobbiamo provare che se β_1 e β_2 sono curve in Ω aventi i medesimi estremi allora $\int_{\beta_1} \omega = \int_{\beta_2} \omega$; a questo fine notiamo che percorrendo prima β_1 e poi la curva $-\beta_2$ ottenuta da β_2 invertendo l'orientazione otteniamo una curva α chiusa, dunque

$$0 = \int_{\alpha} \omega = \int_{\beta_1} \omega + \int_{-\beta_2} \omega = \int_{\beta_1} \omega - \int_{\beta_2} \omega$$

e il corollario è dimostrato. \square

15.2.1 Forme chiuse

Le condizioni per l'esattezza di una forma ω espresse dai due risultati precedenti sono naturalmente impossibili da verificare in pratica. Siccome invece l'esattezza è una caratteristica molto importante (in fisica, ad esempio, un campo di forze che espresso come 1-forma risulta esatto è detto *conservativo*, e gode di proprietà rilevanti), sorge in modo naturale l'esigenza di cercare condizioni più efficaci. Introduciamo a questo fine la seguente:

Definizione 15.2.6. Una 1-forma $\omega = f dx + g dy$ su un aperto $\Omega \subset \mathbb{R}^2$ si dice *chiusa* se $\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}$ su Ω .

Proposizione 15.2.7. Una 1-forma esatta è necessariamente chiusa.

Dimostrazione. Dire che ω è esatta su un aperto Ω significa dire che esiste un potenziale $U : \Omega \rightarrow \mathbb{R}$ tale che $\omega = dU$, ovvero che se $\omega = f dx + g dy$ allora $f = \frac{\partial U}{\partial x}$ e $g = \frac{\partial U}{\partial y}$. Ne deduciamo che

$$\frac{\partial g}{\partial x} = \frac{\partial^2 U}{\partial x \partial y}, \quad \frac{\partial f}{\partial y} = \frac{\partial^2 U}{\partial y \partial x}$$

e la conclusione è allora una conseguenza della Proposizione 9.3.7. \square

Verifichiamo ora che la condizione di chiusura è effettivamente soltanto necessaria per l'esattezza (come affermato dalla proposizione precedente), mentre *non è sufficiente*:

Esempio 15.2.8. Consideriamo su $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$ la 1-forma

$$\omega_0 = \frac{x \, dy - y \, dx}{x^2 + y^2}.$$

Posto $f(x, y) = -\frac{y}{x^2+y^2}$ e $g(x, y) = \frac{x}{x^2+y^2}$ abbiamo allora $\omega_0 = f \, dx + g \, dy$ su Ω ; inoltre

$$\frac{\partial g}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2},$$

dunque ω_0 è chiusa su Ω . Tuttavia se consideriamo la curva $\alpha : [0, 2\pi] \rightarrow \Omega$ data da $\alpha(\vartheta) = (\cos \vartheta, \sin \vartheta)$ abbiamo che α è chiusa e

$$\int_{\alpha} \omega_0 = \int_0^{2\pi} \frac{\cos \vartheta \cos \vartheta - \sin \vartheta (-\sin \vartheta)}{\cos^2 \vartheta + \sin^2 \vartheta} \, d\vartheta = \int_0^{2\pi} d\vartheta = 2\pi$$

dunque ω_0 non è esatta su Ω per il Corollario 15.2.5.

La forma ω_0 dell'esempio precedente ha una chiara interpretazione geometrica. Cominciamo osservando che ogni punto (x, y) di $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$ può essere espresso in coordinate polari (ρ, ϑ) come $(x, y) = (\rho \cos \vartheta, \rho \sin \vartheta)$. Qui $\rho = \sqrt{x^2 + y^2}$ è univocamente determinato, mentre ϑ lo è soltanto a meno di multipli interi di 2π (notiamo che questo è vero solo perché Ω non contiene $(0, 0)$, altrimenti non lo sarebbe). Più esattamente, ϑ non è ben definito, ma in ogni punto (x, y) può essere scelto come uno dei seguenti valori

$$\begin{aligned} \vartheta &= \arctan \frac{y}{x} + 2k\pi & \text{se } x > 0 \\ \vartheta &= \arctan \frac{y}{x} + \pi + 2k\pi & \text{se } x < 0 \\ \vartheta &= \operatorname{arccot} \frac{x}{y} + 2k\pi & \text{se } y > 0 \\ \vartheta &= \operatorname{arccot} \frac{x}{y} + \pi + 2k\pi & \text{se } y < 0 \end{aligned}$$

con $k \in \mathbb{Z}$. In altre parole, ϑ non è una funzione di (x, y) , ma può essere visto come una famiglia di funzioni che differiscono tra loro per costanti; ne

segue che, nonostante ϑ non sia ben definito, è ben definito il suo differenziale $d\vartheta$, di cui si trova la medesima espressione usando ciascuna delle precedenti espressioni di ϑ :

$$\begin{aligned}\frac{\partial}{\partial x} \left(\arctan \frac{y}{x} \right) &= \frac{\partial}{\partial x} \left(\operatorname{arccot} \frac{x}{y} \right) = -\frac{y}{x^2 + y^2}, \\ \frac{\partial}{\partial y} \left(\arctan \frac{y}{x} \right) &= \frac{\partial}{\partial y} \left(\operatorname{arccot} \frac{x}{y} \right) = \frac{x}{x^2 + y^2}.\end{aligned}$$

Abbiamo dunque scoperto che la misteriosa forma $\omega_0 = \frac{x dy - y dx}{x^2 + y^2}$ altro non è che $d\vartheta$. Integrando ω_0 su una curva α in $\mathbb{R}^2 \setminus \{(0, 0)\}$ stiamo allora punto per punto integrando la variazione della coordinata angolare, pertanto l'integrale complessivo uguaglia la variazione totale della coordinata angolare; ne deduciamo la:

Osservazione 15.2.9. Se $\omega_0 = \frac{x dy - y dx}{x^2 + y^2}$ e $\alpha : [a, b] \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$ è una curva chiusa allora $\int_{\alpha} \omega_0$ vale 2π volte il numero di avvolgimenti di α intorno a $(0, 0)$, contati positivamente se in verso antiorario e negativamente se in verso orario.

15.3 Formula di Gauss-Green

Nel paragrafo precedente è rimasta in sospeso la ricerca di una condizione *sufficiente* a garantire l'esattezza di una forma chiusa, e in questo paragrafo colmeremo questa lacuna. Per farlo iniziamo ricordando che una condizione *necessaria* per l'esattezza di $\omega = f dx + g dy$ è invece la chiusura, ovvero l'uguaglianza $\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}$. Introduciamo ora una nuova nozione:

Definizione 15.3.1. Chiamiamo *differenziale* di una 1-forma $\omega = f dx + g dy$ l'oggetto

$$d\omega = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx dy$$

(che può essere integrato sulla chiusura di un aperto limitato di \mathbb{R}^2).

Esempio 15.3.2. Se

$$\omega(x, y) = x^2 y^3 dx + \cos(xy) dy$$

abbiamo

$$d\omega = (-y \operatorname{sen}(xy) - 3x^2 y^2) dx dy,$$

dunque

$$\begin{aligned} \int_{[0,1] \times [0, \frac{\pi}{2}]} d\omega &= \int_0^{\pi/2} \left(\int_0^1 (-y \operatorname{sen}(xy) - 3x^2 y^2) dx \right) dy \\ &= \int_0^{\pi/2} (\cos y - y^2 - 1) dy = 1 - \frac{\pi}{2} - \frac{\pi^3}{24}. \end{aligned}$$

Data questa definizione abbiamo automaticamente che *una 1-forma è chiusa se e soltanto se ha differenziale nullo*, ma il differenziale si rivelerà utile anche al di là di questo fatto. Cominciamo con l'interpretare la definizione, spiegando come l'espressione di $d\omega$ sia quella che naturalmente si ottiene cercando di applicare a una 1-forma un'operazione di derivazione. Intanto partendo da $\omega = f dx + g dy$ è naturale pensare che sia valida la regola di Leibniz per la derivazione del prodotto, ovvero che

$$d\omega = df \cdot dx + f \cdot d(dx) + dg \cdot dy + g \cdot d(dy).$$

Abbiamo ora $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ e $dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy$; inoltre possiamo interpretare dx e dy come gli elementi di lunghezza nelle direzioni x e y , e osservare che essi sono indipendenti dal punto, pertanto $d(dx) = d(dy) = 0$. Giungiamo all'espressione

$$d\omega = \frac{\partial f}{\partial x} dx dx + \frac{\partial f}{\partial y} dy dx + \frac{\partial g}{\partial x} dx dy + \frac{\partial g}{\partial y} dy dy$$

nella quale possiamo interpretare $dx dx$, $dy dx$, $dx dy$ e $dy dy$ come gli elementi d'area orientati dati dai parallelogrammi aventi come lati gli elementi di lunghezza nelle opportune direzioni. Vediamo allora che $dx dx = dy dy = 0$, mentre $dy dx = -dx dy$, da cui riotteniamo l'espressione $d\omega = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx dy$.

15.3.1 Bordo di un insieme

Per introdurre il prossimo fondamentale risultato, detto *formula di Gauss-Green*, cominciamo ricordando l'uguaglianza

$$\int_{\alpha} dU = U(\alpha(b)) - U(\alpha(a))$$

valida per una curva $\alpha : [a, b] \rightarrow \Omega$ e una funzione $U : \Omega \rightarrow \mathbb{R}$. Possiamo interpretare il secondo membro dell'uguaglianza come l'integrale della funzione U calcolato sui punti $\alpha(b)$ e $\alpha(a)$, con il secondo contato negativamente. Poiché questi punti sono gli estremi di α , possiamo definire il *bordo* di α , indicato con $\partial\alpha$, come l'unione dei due punti $\alpha(b)$ e $\alpha(a)$, con il secondo contato negativamente, e riscrivere l'uguaglianza come

$$\int_{\alpha} dU = \int_{\partial\alpha} U.$$

La formula di Gauss-Green stabilisce un analogo di questa uguaglianza in una dimensione in più, ovvero quando la curva α viene sostituita da un sottoinsieme bidimensionale di \mathbb{R}^2 e la funzione U viene sostituita da una 1-forma. Per enunciare la formula dobbiamo però dare la:

Definizione 15.3.3. Sia $\Omega \subset \mathbb{R}^2$ la chiusura di un insieme aperto il cui bordo sia unione dei supporti di un insieme finito di curve regolari a tratti. Chiamiamo allora *bordo orientato* di Ω l'insieme di tali curve, ciascuna orientata in modo da lasciare Ω alla propria sinistra.

Esempio 15.3.4. Il bordo orientato del disco $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ è la circonferenza $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ orientata in verso antiorario.

Teorema 15.3.5. Se $\Omega \subset \mathbb{R}^2$ è limitato, è la chiusura di un insieme aperto, e ha bordo costituito da un insieme finito di curve regolari a tratti, e se ω è una 1-forma definita su un insieme aperto che contiene Ω , allora

$$\int_{\Omega} d\omega = \int_{\partial\Omega} \omega.$$

Nel caso in cui Ω è un rettangolo $[a, b] \times [c, d]$ la dimostrazione della formula di Gauss-Green è semplice: se $\omega = f dx + g dy$ abbiamo

$$\begin{aligned} \int_{\Omega} d\omega &= \int_{[a,b] \times [c,d]} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx dy \\ &= \int_c^d \left(\int_a^b \frac{\partial g}{\partial x} dx \right) dy - \int_a^b \left(\int_c^d \frac{\partial f}{\partial y} dy \right) dx \\ &= \int_c^d g(b, y) dy - \int_c^d g(a, y) dy - \int_a^b f(x, d) dx + \int_a^b f(x, c) dx. \end{aligned}$$

Osserviamo adesso che $\partial\Omega$ si può parametrizzare, rispettando l'orientazione, tramite le curve $\gamma_1, \gamma_2, -\gamma_3, -\gamma_4$, con

$$\begin{aligned}\gamma_1 &: [a, b] \rightarrow \mathbb{R}, & \gamma_1(x) &= (x, c), \\ \gamma_2 &: [c, d] \rightarrow \mathbb{R}, & \gamma_2(y) &= (b, y), \\ \gamma_3 &: [a, b] \rightarrow \mathbb{R}, & \gamma_3(x) &= (x, d), \\ \gamma_4 &: [c, d] \rightarrow \mathbb{R}, & \gamma_4(y) &= (a, y),\end{aligned}$$

da cui segue che

$$\int_{\partial\Omega} \omega = \int_a^b f(x, c) dx + \int_c^d g(b, y) dy - \int_a^b f(x, d) dx - \int_c^d g(a, y) dy$$

e la formula è dimostrata. Per insiemi Ω che non siano rettangoli è sufficiente approssimare Ω con un'unione di rettangoli e osservare che sommando l'integrale di ω sul bordo dei rettangoli tutti i contributi dei lati interni si cancellano, poiché ogni lato riceve orientazioni opposte dai due rettangoli a cui appartiene; si trova quindi l'integrale di ω su una spezzata con i lati paralleli agli assi cartesiani che approssima $\partial\Omega$, ed è un fatto che l'integrale di ω su tale spezzata approssima quello su $\partial\Omega$. (Si tratta però di un fatto non ovvio: osserviamo ad esempio che la lunghezza di Ω non è approssimata da quella della spezzata.)

Dalla formula di Gauss-Green deduciamo ora il seguente:

Corollario 15.3.6. *Se $\Omega \subset \mathbb{R}^2$ è limitato, è la chiusura di un insieme aperto, e ha bordo costituito da un insieme finito di curve regolari a tratti, allora la sua area misura*

$$\int_{\partial\Omega} x dy = - \int_{\partial\Omega} y dx.$$

Questo corollario afferma in particolare che l'area di un insieme si può calcolare eseguendo operazioni di misura esclusivamente sul suo bordo, il che ha notevolissime applicazioni pratiche.

15.3.2 Insiemi semplicemente connessi

Tornando allo scopo principale per cui abbiamo introdotto il differenziale di una 1-forma, cioè per lo studio della condizione di esattezza, abbiamo ora la:

Definizione 15.3.7. Un sottoinsieme aperto di \mathbb{R}^2 si dice *semplicemente connesso* se per ogni curva α semplice e chiusa contenuta in Ω il sottoinsieme limitato di \mathbb{R}^2 di cui α è il bordo (vedi il Teorema 14.1.15) è contenuto in Ω .

Informalmente, un insieme Ω è semplicemente connesso se è *privo di buchi*. Abbiamo ora finalmente la condizione che cercavamo:

Teorema 15.3.8. *Se $\Omega \subset \mathbb{R}^2$ è semplicemente connesso, allora ogni 1-forma chiusa su Ω è anche esatta su Ω .*

Dimostrazione. Sia ω una 1-forma chiusa su Ω ; per provare che ω è esatta su Ω , grazie al Corollario 15.2.5, basta verificare che $\int_{\alpha} \omega = 0$ per ogni curva α chiusa in ω . È in realtà possibile vedere che basta fare la verifica con α semplice e chiusa; ma in tal caso per ipotesi abbiamo che $\alpha = \partial A$ con $A \subset \Omega$, dunque ω è definita su A e per la formula di Gauss-Green abbiamo

$$\int_{\alpha} \omega = \int_A d\omega = \int_A 0 = 0$$

e il teorema è provato. □

Applicando il teorema precedente, bisogna prestare una certa attenzione: la condizione che Ω sia semplicemente connesso è *sufficiente* perché una forma ω chiusa su Ω sia anche esatta, ma non è necessaria. In altre parole, su un insieme non semplicemente connesso possono benissimo essere definite forme esatte. È tuttavia semplice dimostrare, usando la forma $\omega_0 = \frac{x dy - y dx}{x^2 + y^2}$ definita alla fine del paragrafo precedente, che su ogni insieme non semplicemente connesso sono definite forme chiuse non esatte.

Capitolo 16

Isometrie del piano

Nel Capitolo 9 abbiamo già incontrato la nozione di isometria (trasformazione che preserva le distanze) parlando di spazi vettoriali dotati di prodotti scalari. Abbiamo inoltre chiamato ortogonali le matrici che definiscono isometrie, e abbiamo esaminato alcune matrici ortogonali 2×2 (e anche $n \times n$ nel Capitolo 10). Affronteremo ora lo studio delle isometrie del piano euclideo in modo più sistematico, abbandonando l'ipotesi di linearità e adottando sia un punto di vista sintetico sia un punto di vista analitico.

16.1 Il gruppo delle isometrie

In questo paragrafo studiamo le isometrie da un punto di vista sintetico, ovvero senza impiegare coordinate.

16.1.1 Il piano euclideo e il gruppo delle isometrie

Indichiamo con \mathbb{E}^2 il piano euclideo, ovvero l'insieme \mathbb{R}^2 nel quale consideriamo come enti geometrici i punti e le rette, e fissiamo la consueta distanza d tra i punti e la consueta misura (in radianti) degli angoli formati da rette incidenti.

Definizione 16.1.1. Chiamiamo *isometria* di \mathbb{E}^2 una funzione $g : \mathbb{E}^2 \rightarrow \mathbb{E}^2$ invertibile e tale che $d(g(P), g(Q)) = d(P, Q)$ per ogni $P, Q \in \mathbb{E}^2$. Poniamo quindi

$$\text{Isom}(\mathbb{E}^2) = \{g : \mathbb{E}^2 \rightarrow \mathbb{E}^2, g \text{ isometria}\}.$$

Osservazione 16.1.2. Se $g : \mathbb{E}^2 \rightarrow \mathbb{E}^2$ soddisfa $d(g(P), g(Q)) = d(P, Q)$ per ogni $P, Q \in \mathbb{E}^2$ allora f è automaticamente iniettiva. In realtà si può provare

(lo faremo indirettamente nel Teorema 16.1.13) che f è automaticamente anche surgettiva, dunque la richiesta di invertibilità nelle definizioni è superflua.

Osservazione 16.1.3. Poiché le lunghezze dei lati di un triangolo ne determinano gli angoli, preservando le lunghezze dei segmenti ogni $f \in \text{Isom}(\mathbb{E}^2)$ preserva anche le ampiezze degli angoli.

Prima di procedere ricordiamo che un insieme è detto *gruppo* se è dotato di un'operazione binaria associativa per la quale esiste un elemento neutro e rispetto alla quale ogni elemento ha un inverso. Abbiamo allora la:

Proposizione 16.1.4. *Con l'operazione di composizione e come elemento neutro l'identità di \mathbb{E}^2 , l'insieme $\text{Isom}(\mathbb{E}^2)$ è un gruppo.*

16.1.2 Tipi di isometria

Elenchiamo ora alcune tipologie di isometria, che in realtà si riveleranno esaurire tutte le possibilità:

- La funzione identità $\text{id}_{\mathbb{E}^2}$;
- La traslazione τ_v di un vettore v ;
- La rotazione $\rho_{P,\vartheta}$ intorno a un punto P di angolo ϑ in verso antiorario;
- La riflessione σ_ℓ rispetto a una retta ℓ ;
- La glissoriflessione $\sigma_{\ell,v}$, definita come $\tau_v \circ \sigma_\ell$ dove ℓ è una retta e v è un vettore non nullo parallelo a ℓ .

Osservazione 16.1.5. Se v è un vettore non nullo parallelo a una retta ℓ allora si ha $\tau_v \circ \sigma_\ell = \sigma_\ell \circ \tau_v$.

Indichiamo ora con Γ l'insieme delle isometrie dei 5 tipi appena descritti. Come già annunciato si ha in realtà $\Gamma = \text{Isom}(\mathbb{E}^2)$, ma la dimostrazione richiede diversi passaggi geometrici (tutti peraltro di interesse indipendente).

16.1.3 Composizioni di isometrie

Iniziamo con la seguente:

Proposizione 16.1.6. *Se ℓ_1 e ℓ_2 sono rette distinte in \mathbb{E}^2 allora $\sigma_{\ell_1} \circ \sigma_{\ell_2}$ è:*

- τ_{2v} se ℓ_1 ed ℓ_2 sono parallele tra loro e v è un vettore che va da un punto di ℓ_2 a un punto di ℓ_1 perpendicolarmente a entrambe;

- $\rho_{P,2\vartheta}$ se ℓ_1 e ℓ_2 sono incidenti in un punto P e ϑ è l'angolo intorno a P da ℓ_2 a ℓ_1 in verso antiorario.

Dimostrazione. Supponiamo dapprima che ℓ_1 ed ℓ_2 siano parallele tra loro. Sia $Q \in \mathbb{E}^2$, sia r la retta ortogonale a ℓ_1 ed ℓ_2 passante per Q , e su r poniamo un sistema di riferimento con l'origine nel punto di intersezione con ℓ_2 , in modo che il punto di intersezione con ℓ_1 abbia coordinata $d > 0$. La tesi significa che $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(Q)$ appartiene a r (il che è evidente) e che se Q ha coordinata x su r allora $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(Q)$ ha coordinata $x+2d$. A questo fine basta notare che $\sigma_{\ell_2}(Q)$ ha coordinata $-x$, mentre $\sigma_{\ell_1}(\sigma_{\ell_2}(Q))$ ha coordinata $-(-x-d)+d = x+2d$.

Passiamo al caso in cui ℓ_1 e ℓ_2 sono incidenti in P , e definiamo ϑ come nell'enunciato. Poniamo sul piano un sistema di coordinate polari centrato in P in modo che ℓ_2 abbia coordinata angolare $\varphi = 0$ e ℓ_1 abbia coordinata angolare $\varphi = \vartheta$. Prendiamo un punto Q di \mathbb{E}^2 . È ovvio che $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(Q)$ ha la stessa distanza di Q da P , dunque per concludere basta vedere che se Q ha coordinata angolare ψ allora $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(Q)$ ha coordinata angolare $\psi + 2\vartheta$. E infatti $\sigma_{\ell_2}(Q)$ ha coordinata angolare $-\psi$, e $\sigma_{\ell_1}(\sigma_{\ell_2}(Q))$ ha coordinata angolare $-(-\psi - \vartheta) + \vartheta = \psi + 2\vartheta$. \square

Dal risultato precedente segue immediatamente che ogni traslazione o riflessione è esprimibile come la composizione di due riflessioni; ne deduciamo il:

Corollario 16.1.7. *Ogni elemento di Γ è composizione di al più tre riflessioni.*

Abbiamo ora la:

Proposizione 16.1.8.

- $\sigma_\ell \circ \tau_v$ è una riflessione se $v \perp \ell$, una glissoriflessione altrimenti;
- $\sigma_\ell \circ \rho_{C,\vartheta}$ è una riflessione se $C \in \ell$, una glissoriflessione altrimenti.

Dimostrazione. Per la prima affermazione, decomponiamo v come $w + u$ con $w \perp \ell$ e $u \parallel \ell$, e poniamo $s = \tau_{-w/2}(\ell)$. La Proposizione 16.1.6 comporta che $\sigma_\ell \circ \sigma_s = \tau_w$, dunque $\sigma_\ell \circ \tau_w = \sigma_s$, e

$$\sigma_\ell \circ \tau_v = \sigma_\ell \circ \tau_w \circ \tau_u = \sigma_s \circ \tau_u$$

è la riflessione σ_s se $u = 0$, cioè se $v \perp \ell$, e una glissoriflessione altrimenti (infatti $s \parallel \ell$, dunque $u \perp s$).

Passando alla seconda affermazione, supponiamo dapprima che C sia su ℓ , e poniamo $s = \rho_{C,-\vartheta/2}(\ell)$. La Proposizione 16.1.6 comporta che $\sigma_\ell \circ \sigma_s = \rho_{C,\vartheta}$,

dunque $\sigma_\ell \circ \rho_{C,\vartheta} = \sigma_s$. Nel caso invece in cui C è fuori da ℓ , chiamiamo P la proiezione ortogonale di C su ℓ e poniamo $s = \rho_{P,-\vartheta/2}(\ell)$. La Proposizione 16.1.6 comporta che $\sigma_\ell \circ \sigma_s = \rho_{P,\vartheta}$, dunque $\sigma_s = \sigma_\ell \circ \rho_{P,\vartheta}$ e

$$\sigma_\ell \circ \rho_{C,\vartheta} = \sigma_\ell \circ \rho_{P,\vartheta} \circ \rho_{P,-\vartheta} \circ \rho_{C,\vartheta} = \sigma_s \circ (\rho_{P,-\vartheta} \circ \rho_{C,\vartheta}).$$

Si può ora verificare (lo faremo in modo del tutto indipendente nella Proposizione 16.1.11) che $\rho_{P,-\vartheta} \circ \rho_{C,\vartheta}$ è la traslazione di vettore $-2v$, dove v è il vettore che va da P alla proiezione ortogonale Q di C su s . Poiché per costruzione P giace su s e $Q \neq P$ abbiamo che v è non nullo e parallelo a s , dunque $\sigma_s \circ \tau_{-2v}$ è una glissoriflessione, il che completa la dimostrazione. \square

Combinando le Proposizioni 16.1.6 e 16.1.8 ricaviamo facilmente il:

Corollario 16.1.9. Γ consiste di tutti e soli gli elementi di $\text{Isom}(\mathbb{E}^2)$ che sono composizione di al più tre riflessioni.

Prima di arrivare al risultato principale secondo cui Γ già coincide con $\text{Isom}(\mathbb{E}^2)$, ci servono ancora due risultati preliminari:

Lemma 16.1.10. Due isometrie di \mathbb{E}^2 che coincidono su tre punti non allineati coincidono su tutto \mathbb{E}^2 .

Dimostrazione. Osserviamo dapprima che le immagini tramite un'isometria di punti non allineati sono non allineate (infatti soddisfano in modo stretto le tre disuguaglianze triangolari). Basta allora notare che l'intersezione di tre circonferenze con centri non allineati è al più un punto, dunque ogni punto di \mathbb{E}^2 è univocamente determinato dalle sue distanze da tre punti non allineati. \square

Proposizione 16.1.11. Se $g_1, g_2 \in \text{Isom}(\mathbb{E}^2)$ sono traslazioni o rotazioni allora $g_1 \circ g_2$ è l'identità, una traslazione o una rotazione.

Dimostrazione. Distinguiamo le due possibilità sia per g_1 sia per g_2 . Se $g_1 = \tau_{v_1}$ e $g_2 = \tau_{v_2}$ chiaramente $g_1 \circ g_2 = \tau_{v_1+v_2}$. Se $g_1 = \rho_{C,\vartheta}$ e $g_2 = \tau_v$ poniamo $H = \tau_{-v/2}(C)$ e definiamo D come il punto tale che $C\hat{H}D$ è retto, mentre $H\hat{D}C$ misura $\vartheta/2$, come in Fig. 16.1. Notiamo che D è H se $\vartheta = \pi$, mentre giace da un lato o dall'altro della retta CH a seconda che $0 < \vartheta < \pi$ oppure $-\pi < \vartheta < 0$ (il caso in Fig. 16.1 è il primo). In Fig. 16.2 e Fig. 16.3 dimostriamo che il punto D rimane fisso rispetto a $\rho_{C,\vartheta} \circ \tau_v$, mentre H e C vengono ruotati intorno a D di angolo ϑ (tutti gli angoli segnati in figura misurano $\vartheta/2$). Grazie al Lemma 16.1.10 ciò comporta che $\rho_{C,\vartheta} \circ \tau_v = \rho_{D,\vartheta}$.

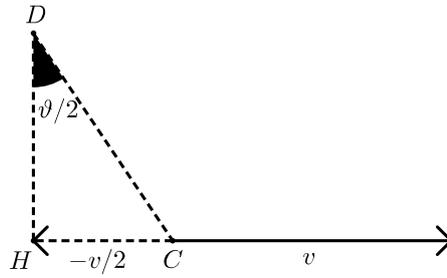


Figura 16.1: Composizione di una rotazione e di una traslazione

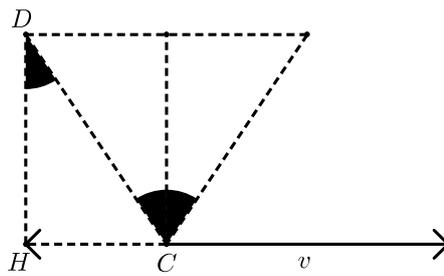


Figura 16.2: $\rho_{C,\vartheta} \circ \tau_v$ fissa D

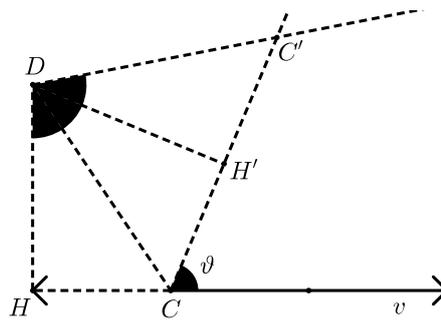


Figura 16.3: Immagini di C e H rispetto a $\rho_{C,\vartheta} \circ \tau_v$

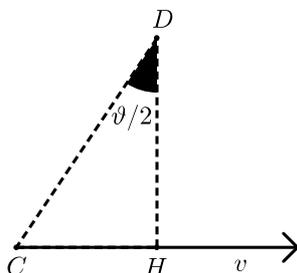


Figura 16.4: Composizione di una traslazione e di una rotazione

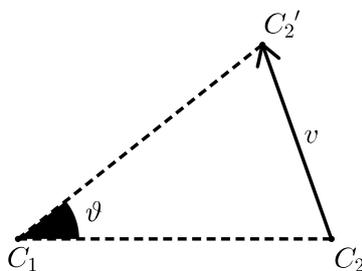


Figura 16.5: Composizione di due rotazioni con centri diversi e angoli opposti tra loro

Il caso $g_1 = \tau_v$ e $g_2 = \rho_{C,\vartheta}$ è simile (oppure si deduce dal precedente): se $H = \tau_{v/2}(C)$ e D è il punto tale che $C\hat{H}D$ è retto mentre $C\hat{D}H$ misura $\vartheta/2$, come in Fig. 16.4 si ha che $\tau_v \circ \rho_{C,\vartheta} = \rho_{D,\vartheta}$.

Veniamo al caso $g_1 = \rho_{C_1,\vartheta_1}$ e $g_2 = \rho_{C_2,\vartheta_2}$. Se $C_1 = C_2$ chiaramente $g_1 \circ g_2 = \rho_{C,\vartheta_1+\vartheta_2}$, con $C = C_1 = C_2$. Se $C_1 \neq C_2$ ma $\vartheta_2 = -\vartheta_1$, definiamo v come il vettore che va da C_2 a $\rho_{C_1,\vartheta}(C_2)$, come in Fig. 16.5, dove $\vartheta = \vartheta_1$. Ovviamente si ha $(\rho_{C_1,\vartheta_1} \circ \rho_{C_2,\vartheta_2})(C_2) = \tau_v(C_2)$, come evidente anche dalla stessa Fig. 16.5. La Fig. 16.6 e la Fig. 16.7 mostrano che $\rho_{C_1,\vartheta_1} \circ \rho_{C_2,\vartheta_2}$ agisce come τ_v anche su C_1 e su $P = \rho_{C_2,\vartheta}(C_1)$, dunque $\rho_{C_1,\vartheta_1} \circ \rho_{C_2,\vartheta_2} = \tau_v$ per il Lemma 16.1.10. Se infine $C_1 \neq C_2$ e $\vartheta_1 + \vartheta_2 \neq 0$ definiamo C come in Fig. 16.8, e affermiamo che $\rho_{C_1,\vartheta_1} \circ \rho_{C_2,\vartheta_2} = \rho_{C,\vartheta_1+\vartheta_2}$. La Fig. 16.9 mostra che C è in effetti mantenuto fisso, e con un argomento geometrico diretto si verifica che C_2 e (con più difficoltà) C_1 vengono ruotati intorno a C di angolo $\vartheta_1 + \vartheta_2$, onde la conclusione sempre grazie al Lemma 16.1.10. \square

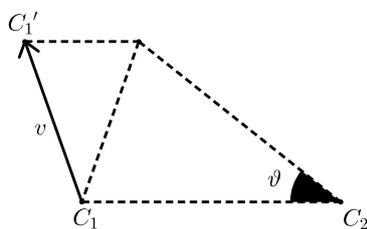


Figura 16.6: $\rho_{C_1, \vartheta_1} \circ \rho_{C_2, \vartheta_2}$ agisce come τ_v su C_1

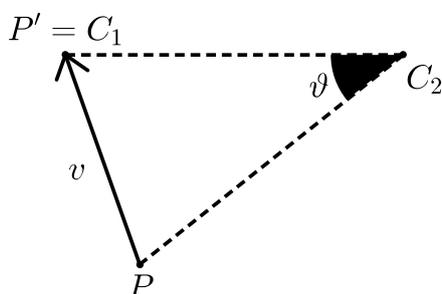


Figura 16.7: $\rho_{C_1, \vartheta_1} \circ \rho_{C_2, \vartheta_2}$ agisce come τ_v su P

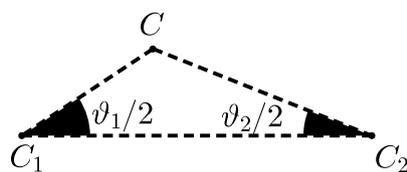


Figura 16.8: Composizione di due rotazioni di angoli con somma non nulla

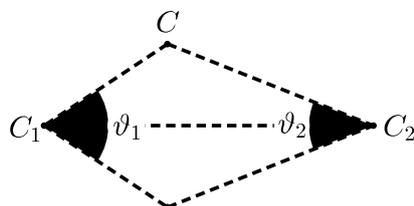


Figura 16.9: C viene trasformato in sé stesso

Corollario 16.1.12. Γ consiste di tutti e soli gli elementi di $\text{Isom}(\mathbb{E}^2)$ che sono composizione di riflessioni. In particolare la composizione di due elementi di Γ appartiene a Γ .

16.1.4 Il gruppo delle isometrie del piano

Siamo finalmente in grado di provare che le isometrie sopra descritte esauriscono tutte le possibilità.

Teorema 16.1.13. Sia $g: \mathbb{E}^2 \rightarrow \mathbb{E}^2$ tale che $d(g(P), g(Q)) = d(P, Q)$ per ogni $P, Q \in \mathbb{E}^2$. Allora $f \in \Gamma$. In particolare, Γ coincide con il gruppo $\text{Isom}(\mathbb{E}^2)$.

Dimostrazione. Per definizione Γ contiene le traslazioni, le rotazioni e le riflessioni. Il Corollario 16.1.12 comporta che contiene le inverse e le composizioni di trasformazioni che gli appartengono. Per concludere sarà allora sufficiente trovare $\gamma_1, \dots, \gamma_k \in \Gamma$ tali che $\gamma_k \circ \dots \circ \gamma_1 \circ f$ è l'identità di \mathbb{E}^2 , visto che allora $f = \gamma_k^{-1} \circ \dots \circ \gamma_1^{-1}$. A tal fine fissiamo in \mathbb{E}^2 tre punti P, Q, R tali che $d(P, Q) = d(P, R) = 1$ e $\widehat{QPR} = \frac{\pi}{2}$ in verso antiorario intorno a P . Se chiamiamo γ_1 la traslazione di vettore $-\overrightarrow{Pg(P)}$ e poniamo $g_1 = \gamma_1 \circ f$ abbiamo $g_1(P) = P$. Dunque $d(P, g_1(Q)) = d(P, Q) = 1$, pertanto se chiamiamo γ_2 la rotazione intorno a P di angolo $-\widehat{Q\hat{P}g_1(Q)}$ e poniamo $g_2 = \gamma_2 \circ g_1$ abbiamo $g_2(P) = P$ e $g_2(Q) = Q$. Abbiamo allora $d(P, g_2(R)) = d(P, R) = 1$ e $d(Q, g_2(R)) = d(Q, R) = \sqrt{2}$, da cui segue che $g_2(R)$ coincide con R o con il simmetrico di R rispetto alla retta ℓ che passa per P e Q . Definiamo ora γ_3 come l'identità di \mathbb{E}^2 nel primo caso e come σ_ℓ nel secondo, e poniamo $g_3 = \gamma_3 \circ g_2$, ovvero $g_3 = \gamma_3 \circ \gamma_2 \circ \gamma_1 \circ f$. Abbiamo allora che g_3 fissa tutti e tre i punti P, Q, R . Dal Lemma 16.1.10 segue che g_3 è l'identità di \mathbb{E}^2 , e la dimostrazione è completa. \square

Introduciamo ora una nozione ulteriore:

Definizione 16.1.14. Diremo che una terna ordinata di punti (O, A, B) di \mathbb{E}^2 è *positivamente orientata* se i punti non sono allineati e la rotazione intorno a O di angolo compreso tra 0 e π che sovrappone la semiretta OA alla semiretta OB avviene in verso antiorario intorno a O . Data $g \in \text{Isom}(\mathbb{E}^2)$, diremo che g *preserva l'orientazione* di \mathbb{E}^2 se manda terne di punti positivamente orientate in terne di punti positivamente orientate, altrimenti diremo che *inverte l'orientazione*.

Osservazione 16.1.15. Non è difficile vedere che se $g \in \text{Isom}(\mathbb{E}^2)$ manda una terna di punti positivamente orientata in una terna di punti positivamente orientata allora fa lo stesso con *qualsiasi* terna di punti positivamente orientata.

È immediato verificare che le riflessioni invertono l'orientazione, che la composizione di due isometrie che invertono l'orientazione la conserva, e che la composizione di un'isometria che inverte l'orientazione con una che la conserva (in qualsiasi ordine) inverte l'orientazione. Avendo provato che le isometrie di \mathbb{E}^2 sono tutte e sole quelle dei 5 tipi descritti (identità, traslazioni, rotazioni, riflessioni, glissoriflessioni) possiamo ora vedere che i 5 tipi sono distinti dal comportamento rispetto all'orientazione e da un'altra caratteristica geometrica molto trasparente, ovvero l'insieme dei punti fissi:

Tipo di isometria g	Punti fissi di g	g rispetto all'orientazione
Identità	\mathbb{E}^2	Conserva l'orientazione
Traslazione	\emptyset	Conserva l'orientazione
Rotazione	Un punto	Conserva l'orientazione
Riflessioni	Una retta	Inverte l'orientazione
Glissoriflessione	\emptyset	Inverte l'orientazione

16.2 Descrizione analitica delle isometrie

In questo paragrafo rivisiteremo quanto già visto in questo capitolo usando le coordinate piuttosto che il punto di vista della geometria sintetica. A questo fine fissiamo su \mathbb{E}^2 un sistema di riferimento cartesiano Oxy ortogonale, monometrico e positivamente orientato. Iniziamo con l'esprimere in coordinate le isometrie di tipo più semplice, con il seguente enunciato senza dimostrazione:

Proposizione 16.2.1.

- $\text{id}_{\mathbb{E}^2} \begin{pmatrix} x \\ y \end{pmatrix} = I_2 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{con } I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$
- $\tau_v \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + v$
- $\rho_{O,\vartheta} \begin{pmatrix} x \\ y \end{pmatrix} = R_\vartheta \cdot \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{con } R_\vartheta = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix};$
- $\rho \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \vartheta \begin{pmatrix} x \\ y \end{pmatrix} = R_\vartheta \cdot \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \right) + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$
 $= R_\vartheta \cdot \begin{pmatrix} x \\ y \end{pmatrix} + (I_2 - R_\vartheta) \cdot \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$

Osservazione 16.2.2. Un'isometria del tipo $g\begin{pmatrix} x \\ y \end{pmatrix} = R_\vartheta \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$ per $\vartheta \neq 2k\pi$ con $k \in \mathbb{Z}$ è sempre una rotazione di angolo ϑ ; il centro di rotazione $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ si trova imponendo $g\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, ovvero

$$\begin{cases} (1 - \cos \vartheta)x_0 + \sin \vartheta y_0 = a \\ -\sin \vartheta x_0 + (1 - \cos \vartheta)y_0 = b \end{cases}$$

e questo sistema ha soluzione unica perché la sua matrice incompleta ha determinante $(1 - \cos \vartheta)^2 + \sin^2 \vartheta = 2(1 - \cos \vartheta)$ che è non nullo per $\vartheta \neq 2k\pi$ con $k \in \mathbb{Z}$.

16.2.1 Riflessioni e glissoriflessioni

Veniamo ora a riflessioni e glissoriflessioni, cominciando con la:

Proposizione 16.2.3. *Se una retta ℓ passa per O e forma un angolo ϑ con l'asse Ox allora $\sigma_\ell\begin{pmatrix} x \\ y \end{pmatrix} = S_{2\vartheta} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$, con $S_\varphi = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$.*

Dimostrazione. Indicando con ℓ_0 l'asse Ox chiaramente $\sigma_{\ell_0}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$, e la formula è corretta. In generale $\sigma_\ell = \rho_{O,\vartheta} \circ \sigma_{\ell_0} \circ \rho_{O,-\vartheta}$, dunque si tratta dell'applicazione lineare associata alla matrice

$$\begin{aligned} & R_\vartheta \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot R_{-\vartheta} \\ &= \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{pmatrix} \\ &= \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{pmatrix} \cdot \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \vartheta - \sin^2 \vartheta & 2 \sin \vartheta \cos \vartheta \\ 2 \sin \vartheta \cos \vartheta & \sin^2 \vartheta - \cos^2 \vartheta \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\vartheta) & \sin(2\vartheta) \\ \sin(2\vartheta) & -\cos(2\vartheta) \end{pmatrix}. \end{aligned}$$

□

Osservazione 16.2.4. La dipendenza della matrice di σ_ℓ dalle funzioni goniometriche di 2ϑ e non di ϑ è del tutto logica, infatti le rette di inclinazione ϑ e $\vartheta + \pi$ coincidono, dunque danno luogo alla stessa riflessione.

Osservazione 16.2.5. Possiamo dimostrare la Proposizione 16.2.3 in altri due modi indipendenti:

- Usando le coordinate polari, possiamo osservare che σ_ℓ trasforma un punto di coordinate polari (ρ, ψ) nel punto di coordinate polari $(\rho, 2\vartheta - \psi)$, e dalle formule di addizione di seno e coseno segue facilmente la conclusione;
- Osservando che $\begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix} \in \ell$ e $\begin{pmatrix} \sin \vartheta \\ -\cos \vartheta \end{pmatrix} \in \ell^\perp$ deduciamo che la matrice M di σ_ℓ deve soddisfare

$$M \cdot \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix} = \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix}, \quad M \cdot \begin{pmatrix} \sin \vartheta \\ -\cos \vartheta \end{pmatrix} = \begin{pmatrix} -\sin \vartheta \\ \cos \vartheta \end{pmatrix}$$

$$\Rightarrow M = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \cdot \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{pmatrix}^{-1}$$

da cui di nuovo la conclusione con un semplice calcolo.

Teorema 16.2.6. *Le riflessioni rispetto a rette in \mathbb{E}^2 sono tutte e sole le trasformazioni affini del tipo*

$$g \begin{pmatrix} x \\ y \end{pmatrix} = S_\varphi \cdot \begin{pmatrix} x \\ y \end{pmatrix} + u \quad \text{con} \quad u \perp \begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{pmatrix}.$$

Una tale g è la riflessione rispetto alla retta ℓ contenente $\frac{1}{2}u$ e avente inclinazione $\varphi/2$ rispetto all'asse Ox .

Dimostrazione. Proviamo dapprima che tutte le riflessioni rispetto a rette hanno la forma indicata. Se ℓ ha inclinazione ϑ rispetto all'asse Ox , indicando con ℓ_0 la parallela a ℓ per O e con v il vettore che unisce O a ℓ in modo perpendicolare, sappiamo dalla Proposizione 16.1.6 che $\sigma_\ell \circ \sigma_{\ell_0} = \tau_{2v}$, dunque $\sigma_\ell = \tau_{2v} \circ \sigma_{\ell_0}$, e la conclusione segue subito. A questo punto è immediato verificare che ogni g come nell'enunciato è una riflessione, infatti quanto già provato comporta che g è effettivamente la riflessione rispetto alla retta ℓ descritta nella seconda parte dell'enunciato. \square

Osservazione 16.2.7. Dalla teoria sviluppata nel paragrafo precedente sappiamo che ogni isometria del tipo $g \begin{pmatrix} x \\ y \end{pmatrix} = S_\varphi \cdot \begin{pmatrix} x \\ y \end{pmatrix} + u$, essendo la composizione di una riflessione e di una traslazione, è una riflessione o una glissoriflessione. Per verificare che si tratti di una riflessione non conviene però verificare la condizione $u \perp \begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{pmatrix}$, bensì semplicemente verificare che g abbia punti fissi. In questo modo si trova peraltro direttamente l'asse di riflessione.

Esempio 16.2.8. Sia $g\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{25}\begin{pmatrix} -7 & 24 \\ 24 & 7 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \frac{4}{25}\begin{pmatrix} -4 \\ 3 \end{pmatrix}$. La matrice è quella di una riflessione, e imponendo $g\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ si trovano due equazioni entrambe equivalenti a $4x - 3y + 2 = 0$, dunque g è la riflessione rispetto alla retta avente tale equazione.

Teorema 16.2.9. *Le glissoriflessioni in \mathbb{E}^2 sono tutte e sole le trasformazioni affini del tipo*

$$g\begin{pmatrix} x \\ y \end{pmatrix} = S_\varphi \cdot \begin{pmatrix} x \\ y \end{pmatrix} + u \quad \text{con} \quad \left\langle u \mid \begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{pmatrix} \right\rangle \neq 0.$$

L'asse ℓ e il vettore w parallelo a ℓ tale che una tale g sia la glissoriflessione $\sigma_\ell \circ \tau_w$ si ottengono come segue:

- w è la proiezione ortogonale di u sulla retta generata da $\begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{pmatrix}$;
- ℓ è la retta contenente $\frac{1}{2}(u - w)$ di inclinazione $\varphi/2$ rispetto all'asse Ox .

Dimostrazione. Con w e ℓ come nella seconda parte dell'enunciato, il Teorema 16.2.6 comporta che la g della prima parte dell'enunciato è $\sigma_\ell \circ \tau_w$, e la conclusione segue subito. \square

Osservazione 16.2.10. Per provare che una trasformazione affine di \mathbb{E}^2 del tipo $g\begin{pmatrix} x \\ y \end{pmatrix} = S_\varphi \cdot \begin{pmatrix} x \\ y \end{pmatrix} + u$ è una glissoriflessione conviene provare che non ha punti fissi, individuare l'asse ℓ come l'unica retta che g manda in sé stessa, e trovare il vettore parallelo a ℓ tale che $g = \sigma_\ell \circ \tau_w$ come $w = g(P) - P$ con $P \in \ell$ scelto a piacimento.

Esempio 16.2.11. Se $g\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{13}\begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, la matrice di g è quella di una riflessione. Imponendo che un punto sia fissato da g si trova il sistema

$$\begin{cases} 5x - 12y + 13 = 13x \\ -12x - 5y + 13 = 13y \end{cases} \quad \text{ovvero} \quad \begin{cases} 8x + 12y = 13 \\ 12x + 18y = 13 \end{cases}$$

che è impossibile, dunque g è una glissoriflessione. Per trovare l'asse ℓ si osserva che è parallelo a quello della riflessione di matrice $\frac{1}{13}\begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix}$, che per

i calcoli già fatti ha equazione $2x + 3y = 0$. Bisogna allora trovare c tale che sostituendo x con $\frac{1}{13}(5x - 12y) + 1$ e y con $\frac{1}{13}(-12x - 5y) + 1$ nell'equazione $2x + 3y = c$ si ottenga un'equazione a essa equivalente. Con facili calcoli si ottiene $c = \frac{5}{2}$, dunque g ha asse $\ell : 4x + 6y = 5$. In particolare $\frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \ell$, e g lo trasforma in $\frac{1}{26} \begin{pmatrix} 19 \\ 9 \end{pmatrix}$, dunque il vettore di traslazione di g parallelo a ℓ è $\frac{1}{26} \begin{pmatrix} 19 \\ 9 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{13} \begin{pmatrix} 3 \\ -2 \end{pmatrix}$.

16.2.2 Composizione di isometrie descritte analiticamente

Tutti i risultati che abbiamo provato sulla composizione di isometrie si possono riottenere per via analitica. A titolo di esempio diamo una nuova dimostrazione di uno dei più impegnativi, ovvero il caso difficile del secondo punto della Proposizione 16.1.8:

Proposizione 16.2.12. *Se C non appartiene a ℓ allora $\sigma_\ell \circ \rho_{C,\vartheta}$ è la glissoriflessione $\sigma_s \circ \tau_{-2v}$, dove:*

- $s = \rho_{P,-\vartheta/2}(\ell)$, dove P è la proiezione ortogonale di C su ℓ ;
- v è il vettore che unisce P alla proiezione ortogonale di C su s .

Dimostrazione. Scegliamo un sistema di riferimento in cui P è l'origine O , l'asse ℓ è l'asse Ox , mentre $C = -\begin{pmatrix} 0 \\ k \end{pmatrix}$ con $k > 0$. Abbiamo allora

$$(\sigma_\ell \circ \rho_{C,\vartheta}) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \left(\begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \cdot \begin{pmatrix} x \\ y + k \end{pmatrix} - \begin{pmatrix} 0 \\ k \end{pmatrix} \right)$$

e dopo qualche passaggio otteniamo

$$(\sigma_\ell \circ \rho_{C,\vartheta}) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ -\sin \vartheta & -\cos \vartheta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -k \sin \vartheta \\ k(1 - \cos \vartheta) \end{pmatrix}.$$

La matrice di questa trasformazione affine è $S_{-\vartheta}$, che rappresenta la riflessione rispetto a una retta passante per O e di inclinazione $-\vartheta/2$, dunque proprio la retta s dell'enunciato. Resta da vedere che il vettore $w = \begin{pmatrix} -k \sin \vartheta \\ k(1 - \cos \vartheta) \end{pmatrix}$ coincide con $-2v$. A questo fine riscriviamo w come

$$w = k \begin{pmatrix} 2 \sin(-\vartheta/2) \cos(-\vartheta/2) \\ 2 \sin^2(-\vartheta/2) \end{pmatrix} = -2(-k \sin(-\vartheta/2)) \begin{pmatrix} \cos(-\vartheta/2) \\ \sin(-\vartheta/2) \end{pmatrix}$$

e osserviamo che $\begin{pmatrix} \cos(-\vartheta/2) \\ \sin(-\vartheta/2) \end{pmatrix}$ è il vettore direttore di s , da cui segue che la proiezione ortogonale v di C su s è proprio

$$(-k \sin(-\vartheta/2)) \begin{pmatrix} \cos(-\vartheta/2) \\ \sin(-\vartheta/2) \end{pmatrix}$$

e la dimostrazione è completa. \square

Capitolo 17

Pavimentazioni del piano

In questo capitolo applichiamo la teoria delle isometrie del piano per descrivere oggetti molto concreti: i modi in cui il piano può essere ricoperto usando piastrelle tutte uguali tra loro.

17.1 Pavimentazioni regolari

Supponiamo di avere una figura piana limitata P , come ad esempio un poligono (potremmo facilmente accettare che qualche porzione del bordo sia curva invece che dritta, ma non lo faremo). Intuitivamente, una *pavimentazione* del piano euclideo \mathbb{E}^2 con piastrella base P è una successione di figure piane P_1, P_2, P_3, \dots tale che:

- Ogni P_j è congruente a P , ovvero $P_j = g_j(P)$ per qualche $g_j \in \text{Isom}(\mathbb{E}^2)$;
- L'unione di tutte le P_j è tutto \mathbb{E}^2 ;
- Per $j \neq k$ si ha che P_j e P_k si incontrano al più in un vertice comune oppure un lato comune.

Osservazione 17.1.1. L'isometria g_j può invertire l'orientazione, dunque P può essere posata da entrambi i lati, cioè non ha un "sopra" e un "sotto".

Questa definizione non è in realtà del tutto soddisfacente, e dovremo modificarla, dato che include situazioni piuttosto innaturali. Informalmente, le pavimentazioni che ci interessano realmente sono quelle che rispondano a una certa *regola* di posatura, cioè per le quali, esaminando una porzione finita

sufficientemente grande, si possa con sicurezza sapere come la pavimentazione prosegue fino all'infinito. I prossimi esempi mostrano pavimentazioni che soddisfano le condizioni precedenti ma sono manifestamente “irregolari”.

Esempio 17.1.2. Partiamo dalla pavimentazione più semplice di \mathbb{E}^2 , data dai quadrati 1×1 con vertici nei punti di coordinate intere. Facciamo ora scorrere verticalmente ciascuna striscia di piastrelle $[n, n+1] \times \mathbb{R}$ di una quantità casuale, dipendente da n . Il risultato è una pavimentazione, ma una sua porzione finita non la determina tutta.

Esempio 17.1.3. Partiamo ora dalla pavimentazione con quadrati 2×2 con vertici nei punti di coordinate pari. Suddividiamo poi ciascuna piastrella in due rettangoli 2×1 , con un taglio verticale oppure orizzontale, in modo casuale dipendente dalla piastrella. Anche questa pavimentazione non è regolare.

Per formalizzare la regolarità di una pavimentazione useremo la seguente:

Definizione 17.1.4. Un sottoinsieme G di $\text{Isom}(\mathbb{E}^2)$ è detto *sottogruppo* se contiene la funzione identità e tutti gli inversi e i prodotti degli elementi che contiene.

Diamo dunque la definizione corretta:

Definizione 17.1.5. Se P è un poligono del piano diciamo che un sottogruppo G di $\text{Isom}(\mathbb{E}^2)$ definisce una *pavimentazione* di \mathbb{E}^2 con *piastrella base* P se l'unione dei $g(P)$ al variare di g in G copre \mathbb{E}^2 , e se per $g_1 \neq g_2$ in G si ha che l'intersezione di $g_1(P)$ e $g_2(P)$ consiste al più di un vertice comune oppure di un lato comune.

In questa definizione quello che conta realmente, più che P , è proprio G , che rappresenta la “regola di posatura”. Innanzitutto è facilissimo vedere che P può essere sostituito da qualsiasi altra piastrella $g(P)$ con $g \in G$ (e per questo nelle figure non indicheremo quale sia P). Inoltre P in alcuni casi può essere modificata senza cambiare G , e viceversa una stessa P può fungere da piastrella base per G diversi.

Esempio 17.1.6. La pavimentazione regolare in quadrati 1×1 con vertici nei punti di coordinate intere si ottiene prendendo $P = [0, 1] \times [0, 1]$, ma per G sono possibili scelte diverse. La più semplice è quella del sottogruppo

$$G_1 = \{\tau_{(n,m)} : n, m \in \mathbb{Z}\}$$

dato dalle traslazioni di vettori aventi entrambe le coordinate intere. Ma per tale G_1 possiamo anche modificare la piastrella base, ottenendo ad esempio

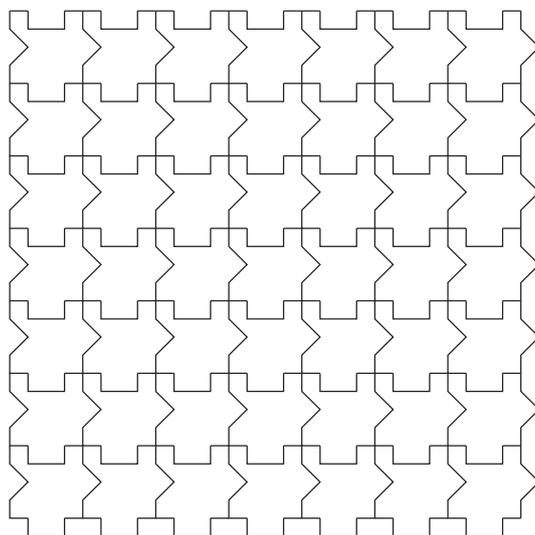


Figura 17.1: Pavimentazione avente come gruppo quello delle traslazioni di vettori con componenti intere

una pavimentazione come in Fig. 17.1. Osserviamo che questa figura è anche migliore di quella originale, dato che ci consente di “leggere” G_1 : nella figura originale non era affatto chiaro come si ottenessero l’uno dall’altro i quadrati, mentre ora per ogni coppia di piastrelle l’isometria che manda l’una nell’altra è chiaramente soltanto un’opportuna traslazione. Tornando alla pavimentazione in quadrati, possiamo ora verificare che con la stessa piastrella base $P = [0, 1] \times [0, 1]$ possiamo scegliere il sottogruppo anche in un altro modo. Precisamente, definiamo G_2 come il più piccolo sottogruppo di $\text{Isom}(\mathbb{E}^2)$ che contiene le 4 riflessioni rispetto ai lati di P , cioè l’insieme di tutte le isometrie ottenute componendo tali 4 riflessioni, in numero arbitrario e con ordine arbitrario. Notiamo che G_2 contiene certamente le rotazioni di angolo $\pm\pi/2$ intorno ai vertici di P , e le traslazioni di vettori $(\pm 2, 0)$ e $(0, \pm 2)$. Poiché ogni elemento di G_2 trasforma la pavimentazione regolare in quadrati 1×1 in sé stessa, ne deduciamo che anche G_2 può essere scelto come sottogruppo della pavimentazione. La figura con i soli quadrati non consente però di “leggere” G_2 : per farlo conviene introdurre in P una “decorazione”, come in Fig. 17.2.

Lo scopo del capitolo sarà quello di comprendere quali sottogruppi G possono essere impiegati per costruire una pavimentazione. Il risultato più importante sarà che, dopo opportune semplificazioni, esiste solo un numero finito di

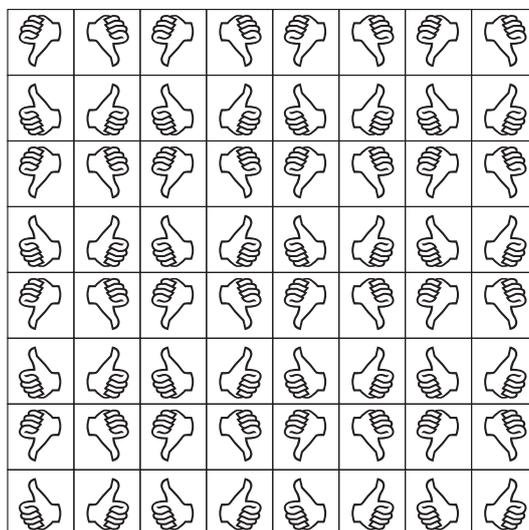


Figura 17.2: Una pavimentazione con piastrelle decorate in modo che il corrispondente sottogruppo di isometrie sia ben comprensibile

possibilità che si possono elencare agevolmente. Concludiamo il paragrafo con un risultato molto semplice:

Proposizione 17.1.7. *Se P e G definiscono una pavimentazione, allora le piastrelle della pavimentazione sono in corrispondenza biunivoca con G .*

Dimostrazione. Se a $g \in G$ si fa corrispondere la piastrella $g(P)$, per definizione si ha una funzione surgettiva da G nell'insieme delle piastrelle. Ma se $g_1 \neq g_2$ si ha che $g_1(P) \cap g_2(P)$ è al più un vertice comune oppure un lato comune, dunque in particolare $g_1(P) \neq g_2(P)$, dunque la corrispondenza è iniettiva. \square

17.2 Restrizioni sui sottogruppi delle pavimentazioni

Anche se non dimostreremo in modo del tutto completo il risultato di classificazione dei sottogruppi delle pavimentazioni, proveremo nelle prossime pagine che su un tale sottogruppo ci sono delle restrizioni, che renderanno il risultato di classificazione quanto meno plausibile. In tutto il paragrafo supponiamo che $G \subset \text{Isom}(\mathbb{E}^2)$ sia il sottogruppo di una pavimentazione.

Proposizione 17.2.1. G è discreto, ovvero non possono esistere $A, B \in \mathbb{E}^2$ e una successione g_1, g_2, g_3, \dots di elementi distinti di G tali che $g_n(A)$ tende a B per n che tende a $+\infty$.

Dimostrazione. Per assurdo: scelta una piastrella P che contiene A , arbitrariamente vicino a B ci sarebbero, per $N \in \mathbb{N}$ opportuno, le infinite piastrelle distinte

$$g_N(P), g_{N+1}(P), g_{N+2}(P), \dots$$

il che è impossibile. □

17.2.1 Restrizioni sugli elementi con punti fissi

Per qualsiasi punto A di \mathbb{E}^2 consideriamo ora l'insieme degli elementi di G che fissano A :

$$G_A = \{g \in G : g(A) = A\}.$$

È immediato vedere che anche G_A è un sottogruppo di $\text{Isom}(\mathbb{E}^2)$. Per enunciare le restrizioni che valgono su G_A diamo la seguente:

Definizione 17.2.2. Dato $A \in \mathbb{E}^2$ e $n \in \mathbb{N}$ con $n \geq 2$ chiamiamo:

- *Gruppo ciclico* di ordine n di centro A il sottogruppo di \mathbb{E}^2

$$C_{A,n} = \{\rho_{A,2\pi j/n} : j = 0, 1, \dots, n-1\};$$

- *Gruppo diedrale* di ordine $2n$ di centro A un sottogruppo del tipo

$$D_{A,n} = \{\rho_{A,2\pi j/n} : j = 0, 1, \dots, n-1\} \cup \{\sigma_{\ell_j} : j = 0, 1, \dots, n-1\}$$

dove $\ell_0, \ell_1, \dots, \ell_{n-1}$ sono rette passanti per A che formano ciascuna con la successiva un angolo di $\frac{\pi}{n}$.

Osservazione 17.2.3.

- $C_{A,n}$ è il più piccolo sottogruppo di $\text{Isom}(\mathbb{E}^2)$ che contiene $\rho_{A,2\pi/n}$;
- $D_{A,n}$ è il più piccolo sottogruppo di $\text{Isom}(\mathbb{E}^2)$ che contiene le riflessioni rispetto a due rette che formano in A un angolo di $\frac{\pi}{n}$.

Proposizione 17.2.4. Se A è un punto di \mathbb{E}^2 e appartiene a una piastrella P della pavimentazione avente sottogruppo G , allora G_A è uno dei seguenti:

- Se A è interno a P allora $G_A = \{\text{id}_{\mathbb{E}^2}\}$;

- Se A è su un lato e di P ma non è un vertice né il punto medio allora $G_A = \{\text{id}_{\mathbb{E}^2}\}$ oppure $G_A = \{\text{id}_{\mathbb{E}^2}, \sigma_\ell\}$ dove ℓ è la retta che contiene e (in particolare, G_A è del tipo $D_{A,1}$);
- Se A è il punto medio di un lato e di P allora $G_A = \{\text{id}_{\mathbb{E}^2}\}$ oppure $G_A = \{\text{id}_{\mathbb{E}^2}, \sigma_\ell\}$ oppure $G_A = C_{A,2}$;
- Se A è un vertice di P allora G_A è un gruppo ciclico $C_{A,n}$ oppure diedrale $D_{A,n}$ centrato in A .

Dimostrazione. Distinguiamo i casi come nell'enunciato.

- Se A è interno a P e $g \in G_A$ allora $g(P)$ è una piastrella che interseca P nel punto interno A , dunque $g(P) = P$, da cui $g = \text{id}_{\mathbb{E}^2}$.

Per tutti gli altri casi notiamo che un elemento g di G_A , grazie alla classificazione delle isometrie di \mathbb{E}^2 (si veda in particolare lo schema alla fine del Sottoparagrafo 16.1.4) è o l'identità, o una rotazione intorno ad A , o una riflessione rispetto a una retta passante per A .

- Se A è su un lato e di P ma non è un vertice allora la condizione che le piastrelle sono prive di punti interni comuni comporta che un elemento di G_A può essere solo l'identità oppure σ_ℓ oppure $\rho_{A,\pi}$; ma il fatto che piastrelle distinte hanno in comune al più un vertice o un lato comporta che G_A può essere $\rho_{A,\pi}$ solo se A è il punto medio di e , nel qual caso si vede immediatamente che G_A non può contenere sia σ_ℓ sia $\rho_{A,\pi}$, sempre grazie al fatto che piastrelle distinte non hanno punti interni comuni.
- Resta il caso in cui A è un vertice di P . Supponendo che P abbia angolo α in A , vediamo che $\{g(P) : g \in G_A\}$ sono piastrelle distinte della pavimentazione, tutte aventi vertice in A e angolo α in A , da cui segue intanto che G_A è finito e ha al più $\frac{2\pi}{\alpha}$ elementi. Distinguiamo allora due casi, a seconda che G_A contenga solo rotazioni intorno ad A oppure anche riflessioni rispetto a rette per A .

- Se G_A contiene solo rotazioni intorno ad A e G_A non è $\{\text{id}_{\mathbb{E}^2}\} = C_{A,1}$, allora possiamo prendere il minimo angolo positivo ϑ tale che $\rho_{A,\vartheta}$ appartiene a G_A , e $\vartheta < 2\pi$. Affermiamo intanto che $\vartheta = \frac{2\pi}{n}$ per qualche $n \in \mathbb{N}$; se così non fosse, nella successione $\vartheta, 2\vartheta, 3\vartheta, \dots$ che tende a $+\infty$ non comparirebbe 2π , dunque troveremmo $k \in \mathbb{N}$ con $k\vartheta < 2\pi < (k+1)\vartheta$, ma allora

$$G_A \ni \rho_{A,\vartheta}^{k+1} = \rho_{A,(k+1)\vartheta} = \rho_{A,(k+1)\vartheta-2\pi}, \quad 0 < (k+1)\vartheta - 2\pi < \vartheta$$

il che contraddice la minimalità di ϑ . Possiamo ora concludere questo caso provando che G_A consiste precisamente delle potenze di $\rho_{A,2\pi/n}$, ovvero che $G_A = C_{A,n}$. Se così non fosse esisterebbe φ tale che $\rho_{A,\varphi} \in G_A$ e $k \cdot \frac{2\pi}{n} < \varphi < (k+1) \cdot \frac{2\pi}{n}$ per qualche $k \in \mathbb{N}$, ma allora

$$G_A \ni \rho_{A,2\pi/n}^{k+1} \circ \rho_{A,\varphi}^{-1} = \rho_{A,(k+1)2\pi/n-\varphi}, \quad 0 < (k+1) \cdot \frac{2\pi}{n} - \varphi < \frac{2\pi}{n}$$

e di nuovo verrebbe contraddetta la minimalità di $\vartheta = \frac{2\pi}{n}$.

- Supponiamo invece che G_A contenga almeno una riflessione σ_ℓ . Allora o G_A è $\{\text{id}_{\mathbb{E}^2}, \sigma_\ell\}$, dunque è del tipo $D_{A,1}$, oppure contiene anche una rotazione (infatti se contiene una riflessione $\sigma_{\ell'}$ contiene la rotazione $\sigma_\ell \circ \sigma_{\ell'}$). Possiamo allora definire come sopra il minimo angolo positivo ϑ tale che $\rho_{A,\vartheta}$ appartiene a G_A . Lo stesso argomento di sopra mostra allora che $\vartheta = \frac{2\pi}{n}$ e che tutte le rotazioni in G_A sono potenze di $\rho_{A,\vartheta}$. Ora definiamo le rette $\ell_0, \ell_1, \dots, \ell_{n-1}$ ponendo $\ell_0 = \ell$ e in modo che ognuna formi con la successiva un angolo di $\frac{\pi}{n}$; sappiamo che $\sigma_{\ell_1} \circ \sigma_{\ell_0} = \rho_{A,2\pi/n}$, da cui segue che $\sigma_{\ell_1} \in G_A$, e analogamente $\sigma_{\ell_j} \in G_A$. Basta allora provare che G_A non può contenere alcuna riflessione oltre a $\sigma_{\ell_0}, \sigma_{\ell_1}, \dots, \sigma_{\ell_{n-1}}$: se contenesse σ_ℓ con $\ell \neq \ell_j$ per ogni j potremmo trovare j tale che ℓ e ℓ_j formano un angolo minore di $\frac{\pi}{n}$, ma allora $\sigma_\ell \circ \sigma_{\ell_j}$ sarebbe una rotazione in G_A di angolo minore di $\frac{2\pi}{n}$, e questo è assurdo.

□

Osservazione 17.2.5. Per concludere che G_A è sempre di tipo $C_{A,n}$ oppure $D_{A,n}$ non serve l'intera ipotesi che G sia il sottogruppo di una pavimentazione, è sufficiente supporre che G sia discreto. Infatti se G è discreto si vede facilmente che G_A è finito, e la conclusione segue dall'ultima parte della dimostrazione precedente. Tuttavia, per il caso in cui G è il gruppo di una pavimentazione, l'enunciato della Proposizione 17.2.4 contiene informazioni più specifiche sul tipo che G_A può avere a seconda della posizione di A rispetto a una piastrella della pavimentazione.

Il prossimo risultato migliora notevolmente l'ultimo punto della proposizione precedente:

Teorema 17.2.6. *Se A è un vertice di una piastrella P della pavimentazione avente sottogruppo G , allora G_A è un gruppo ciclico $C_{A,n}$ oppure diedrale $D_{A,n}$*

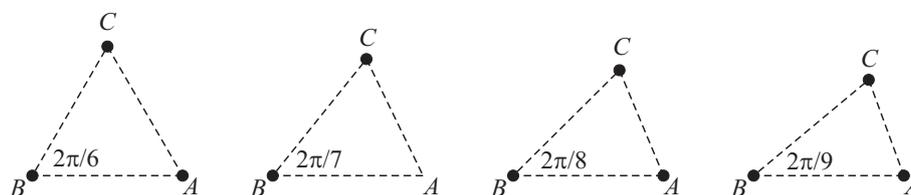


Figura 17.3: Il gruppo G_A non può contenere rotazioni di ordine maggiore di 6

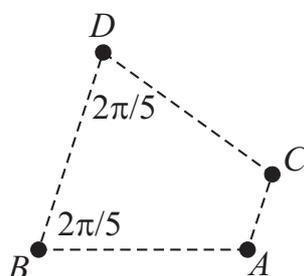


Figura 17.4: Il gruppo G_A non può contenere rotazioni di ordine 5

centrato in A con

$$n \in \{1, 2, 3, 4, 6\}.$$

Dimostrazione. Sappiamo già che G_A è del tipo $C_{A,n}$ oppure $D_{A,n}$, dunque basta vedere che $n \leq 6$ e $n \neq 5$. Per farlo osserviamo dapprima che G non può coincidere con G_A , poiché quest'ultimo è finito e P è limitata, dunque l'unione delle piastrelle $g(P)$ con $g \in G_A$ è limitata. Esaminiamo ora l'insieme di punti $X = \{g(A) : g \in G \setminus G_A\}$. Per definizione X non contiene A , e dal fatto che G è discreto segue che esiste in X un punto B avente minima distanza da A , cioè tale che $B \in X$ e $\overline{AB} \leq \overline{AC}$ per ogni $C \in X$. Sia ora $g \in G$ tale che $g(A) = B$ e notiamo che $\rho_{A,2\pi/n}$ appartiene a G , dunque anche $\rho_{B,2\pi/n} = g \circ \rho_{A,2\pi/n} \circ g^{-1}$ appartiene a G . Se $n \geq 7$ ponendo $C = \rho_{B,2\pi/n}(A)$ abbiamo che $C \in X$ e $\overline{AC} < \overline{AB}$, come suggerito in Fig. 17.3, il che contraddice la scelta di B . Per $n = 5$ poniamo invece $D = \rho_{B,2\pi/5}(A)$ e $C = \rho_{D,2\pi/5}(B)$ e abbiamo ancora $C \in X$ e $\overline{AC} < \overline{AB}$, come mostrato in Fig. 17.4; questo di nuovo contraddice la scelta di B , e la dimostrazione è completa. \square

17.2.2 Restrizioni sulle traslazioni

Come in precedenza manteniamo fissato un sottogruppo G di $\text{Isom}(\mathbb{E}^2)$ che sia associato a una pavimentazione. Poniamo ora

$$T = \{g \in G : g \text{ traslazione}\}$$

e osserviamo che si tratta ancora di un sottogruppo di $\text{Isom}(\mathbb{E}^2)$.

Proposizione 17.2.7. *Esistono vettori v_1 e v_2 con direzioni distinte tali che*

$$T = \{\tau_{n_1 v_1 + n_2 v_2} : n_1, n_2 \in \mathbb{Z}\}.$$

Dimostrazione. Affermiamo dapprima che G contiene traslazioni non banali in direzioni non parallele tra loro. Se G contiene solo traslazioni ciò è evidente. Altrimenti, supponiamo che G contenga una rotazione $\rho_{A_0, \vartheta}$ e scegliamo una piastrella P_0 di cui A_0 sia vertice o punto medio di un lato (vedi la Proposizione 17.2.4). Prendiamo quindi una piastrella $P_1 = g_1(P_0)$ disgiunta da P_0 e poniamo $A_1 = g_1(A_0)$. Abbiamo allora che $g \circ \rho_{A_0, \vartheta} \circ g^{-1}$ appartiene a G ed è $\rho_{A_1, \vartheta}$ oppure $\rho_{A_1, -\vartheta}$ a seconda che g_1 mantenga oppure inverta l'orientazione; ma G contiene gli inversi degli elementi che contiene, e concludiamo in ogni caso che $\rho_{A_1, -\vartheta}$ sta in G . Ne deduciamo che G contiene la traslazione $\tau_{v_1} = \rho_{A_0, \vartheta} \circ \rho_{A_1, -\vartheta}$, dove v_1 è il vettore che unisce A_1 a $\rho_{A_0, \vartheta}(A_1)$. Prendendo un'altra piastrella P_2 in un'altra direzione e ripetendo la costruzione troviamo poi facilmente che G contiene anche τ_{v_2} con v_2 linearmente indipendente da v_1 .

Supponiamo adesso che G non contenga rotazioni ma una glissoriflessione $\sigma_{\ell, w}$, dunque che contenga la traslazione τ_{v_1} con $v_1 = 2w$. Se prendiamo piastrelle P_1 e P_2 dalla stessa parte di ℓ ma a distanze diverse da ℓ , e prendiamo $g \in G$ tale che $g(P_1) = P_2$, possiamo avere i casi seguenti:

- Se $g = \tau_{v_2}$ allora v_2 è linearmente indipendente da v_1 , e abbiamo finito;
- Se g è una (glisso)riflessione con asse parallelo a ℓ , allora $g \circ \sigma_{\ell, w}$ è una traslazione τ_{v_2} con v_2 linearmente indipendente da v_1 , e di nuovo abbiamo finito;
- Se g è una (glisso)riflessione con asse non parallelo a ℓ , allora $g \circ \sigma_{\ell, w}$ è una rotazione oppure una traslazione τ_{v_2} con v_2 linearmente indipendente da v_1 , e ancora una volta abbiamo la conclusione, grazie anche al caso precedente.

Abbiamo infine il caso in cui G non contiene né rotazioni né glissoriflessioni. Ne segue che tutti gli assi di riflessione delle riflessioni di G sono paralleli tra

loro, e la conclusione che G contiene traslazioni non banali in direzioni non parallele tra loro è ora evidente.

Per concludere ricordiamo che G è discreto; ne segue che T non può contenere traslazioni di vettori non banali arbitrariamente corti, quindi possiamo scegliere $\tau_{v_1} \in T$ tale che $\|v_1\|$ sia positivo e minimale (il minimo viene assunto, ovvero è un minimo e non un estremo inferiore, sempre per il fatto che G è discreto). Successivamente possiamo scegliere $\tau_{v_2} \in T$ tale che $\|v_2\|$ sia minimale per le traslazioni in T che non siano nella direzione di v_1 . Se prendiamo il parallelogrammo Q di lati v_1 e v_2 vediamo ora che $\{\tau(Q) : \tau \in T\}$ è una pavimentazione del piano con sottogruppo T . Se per assurdo T non coincidesse con $\{\tau_{n_1 v_1 + n_2 v_2} : n_1, n_2 \in \mathbb{Z}\}$ troveremmo $\tau \in T$ che manda un vertice di questa pavimentazione in un punto che non è un vertice, e questo comporta facilmente una contraddizione alla definizione di v_1 oppure a quella di v_2 . \square

17.3 I 17 sottogruppi delle pavimentazioni

Nel paragrafo precedente abbiamo dimostrato che per un sottogruppo G di $\text{Isom}(\mathbb{E}^2)$ che corrisponda a una pavimentazione regolare ci sono forti restrizioni sia sui sottogruppi G_A degli elementi di G che fissano $A \in \mathbb{E}^2$, sia sul sottogruppo delle traslazioni di G . Queste restrizioni comportano che, *dopo avere opportunamente stabilito quali gruppi identificare tra loro*, per G esiste sono un numero finito (e precisamente 17) di possibilità.

17.3.1 Equivalenza di sottogruppi di pavimentazioni

Cominciamo a stabilire quali gruppi identificare tra loro. Supponiamo allora di averne due, G_1 e G_2 . Un primo caso in cui possiamo naturalmente ritenere che G_1 e G_2 definiscono *la stessa* pavimentazione è quello in cui

$$G_2 = \{\varphi \circ g_1 \circ \varphi^{-1} : g_1 \in G_1\} \quad (17.1)$$

con $\varphi \in \text{Isom}(\mathbb{E}^2)$, infatti in tal caso, detta P_1 una piastrella della pavimentazione di sottogruppo G_1 , e posto $P_2 = \varphi(P_1)$, si ha

$$\{g_2(P_2) : g_2 \in G_2\} = \{\varphi(g_1(P_1)) : g_1 \in G_1\}$$

cioè una pavimentazione di sottogruppo G_2 si ottiene applicando l'isometria φ a una pavimentazione di sottogruppo G_1 .

Un secondo caso simile è quello in cui vale la relazione (17.1) con φ omotetia, infatti in questa situazione le pavimentazioni a cui sono associati i sottogruppi G_1 e G_2 sono l'una un ingrandimento dell'altra.

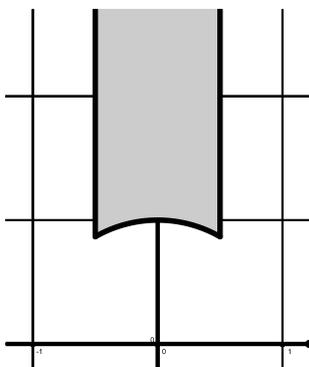


Figura 17.5: Un insieme che parametrizza di gruppi di traslazioni non equivalenti tra loro tramite similitudini

Combinando i due casi discussi si deduce che vanno identificati tra loro sottogruppi G_1 e G_2 per i quali vale la relazione (17.1) con φ similitudine. Ciò però ancora non basta, perché ad esempio restano infiniti sottogruppi di pavimentazioni del tipo $T = \{\tau_{n_1v_1+n_2v_2} : n_1, n_2 \in \mathbb{Z}\}$. Infatti tramite una similitudine ci si può ricondurre al caso in cui v_1 è il vettore di traslazione più corto di T e $v_1 = (1, 0)$, quindi si può scegliere v_2 come secondo vettore più corto, dunque $\|v_2\| \geq 1$, e avente prima coordinata in $[-\frac{1}{2}, \frac{1}{2})$, dunque v_2 appartenente all'insieme in Fig. 17.5, ma è possibile verificare che due elementi dell'insieme definiscono gruppi ottenibili l'uno dall'altro tramite una similitudine soltanto se entrambi sono sul bordo (e in posizione particolare uno rispetto all'altro).

Per poter identificare tra loro due gruppi del tipo $\{\tau_{n_1v_1+n_2v_2} : n_1, n_2 \in \mathbb{Z}\}$ (che chiaramente descrivono *la stessa* regola di posatura) dobbiamo considerare nella relazione (17.1) anche φ che modificano gli angoli; diamo allora (finalmente) la seguente:

Definizione 17.3.1. Diciamo che due sottogruppi G_1 e G_2 di pavimentazioni sono *equivalenti* tra loro se esiste una trasformazione affine $\varphi : \mathbb{E}^2 \rightarrow \mathbb{E}^2$ tale che

$$G_2 = \{\varphi \circ g_1 \circ \varphi^{-1} : g_1 \in G_1\}.$$

Osservazione 17.3.2. Dato G sottogruppo di $\text{Isom}(\mathbb{E}^2)$ e $\varphi : \mathbb{E}^2 \rightarrow \mathbb{E}^2$ trasformazione affine, tipicamente $\{\varphi \circ g \circ \varphi^{-1} : g \in G\}$ non è un sottogruppo di $\text{Isom}(\mathbb{E}^2)$. Dunque la definizione precedente si applica quasi sempre con φ isometria, ma in alcuni casi particolari capita di dover prendere φ solamente affine.

17.3.2 Quoziente del piano rispetto a un sottogruppo

Nell'esibire la lista dei 17 sottogruppi di pavimentazioni del piano euclideo forniremo una giustificazione almeno intuitiva del fatto che essi sono effettivamente *distinti* (o meglio, non equivalenti) basata sull'idea di prendere il quoziente del piano rispetto all'azione del sottogruppo. (In realtà, anche il fatto che ci sono *solo* le 17 possibilità che descriveremo si può dimostrare per la stessa via.) Cominciamo con il seguente:

Lemma 17.3.3. *Se G è il sottogruppo di $\text{Isom}(\mathbb{E}^2)$ di una pavimentazione con piastrella base P , allora la relazione \sim su \mathbb{E}^2 data da $B \sim A$ se esiste $g \in G$ tale che $B = g(A)$ è di equivalenza. Inoltre P interseca tutte le classi di equivalenza, e i suoi punti interni le intersecano in un punto solo.*

Dimostrazione. Riflessività, simmetria e transitività di \sim seguono facilmente dal fatto che G contiene l'identità, contiene gli inversi dei suoi elementi, e contiene le composizioni dei suoi elementi. Le ultime due affermazioni dell'enunciato discendono invece dal fatto che l'unione delle $g(P)$ al variare di $g \in G$ copre \mathbb{E}^2 , e che $g(P)$ incontra P al più sul bordo per $g \neq \text{id}_{\mathbb{E}^2}$. \square

Indicheremo con \mathbb{E}^2/G il quoziente \mathbb{E}^2/\sim di \mathbb{E}^2 rispetto alla relazione di equivalenza del lemma precedente. Dal lemma discende che \mathbb{E}^2/G si può descrivere come la piastrella base P con alcune identificazioni tra lati (o porzioni di lato), dunque è in ogni caso una superficie, eventualmente con bordo. Per mantenere traccia della *geometria* da cui la superficie \mathbb{E}^2/G proviene è però opportuno porre in evidenza su di essa i seguenti punti:

- Le classi di equivalenza $[A]$ con $A \in \mathbb{E}^2$ tale che $G_A = C_{A,n}$, detti *punti conici di ordine n* , che appartengono alla parte interna della superficie \mathbb{E}^2/G e sono indicati con un pallino vicino a cui sta scritto n ;
- Le classi di equivalenza $[A]$ con $A \in \mathbb{E}^2$ tale che $G_A = D_{A,n}$, detti *riflettori angolari di ordine n* , che appartengono al bordo della superficie \mathbb{E}^2/G e sono indicati con un pallino vicino a cui sta scritto n .

Indicheremo con $\Sigma(n_1, \dots, n_k)(m_1, \dots, m_h)$ una superficie Σ con punti conici di ordini n_1, \dots, n_k (all'interno) e riflettori angolari di ordini m_1, \dots, m_h (sul bordo).

17.3.3 La lista

Enunciamo ora formalmente il risultato di finitezza già più volte anticipato:

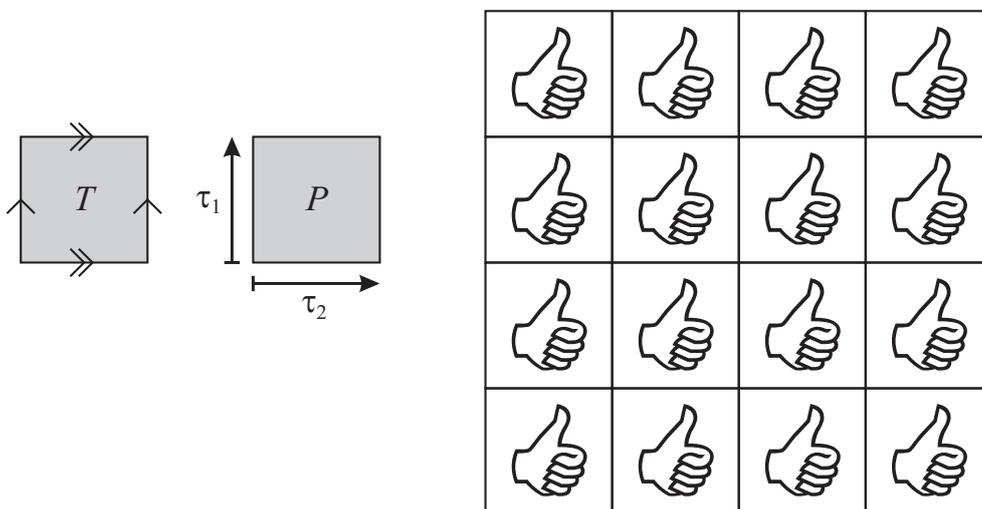
Teorema 17.3.4. *A meno della relazione di equivalenza definita nel Sottoparagrafo 17.3.1 esistono esattamente 17 sottogruppi di Isom(\mathbb{E}^2) di pavimentazioni.*

La dimostrazione si basa sulle Proposizioni 17.2.4 e 17.2.7 e sul Teorema 17.2.6, ma non è per nulla semplice. Elencheremo i 17 sottogruppi G a cui il teorema precedente si riferisce descrivendo:

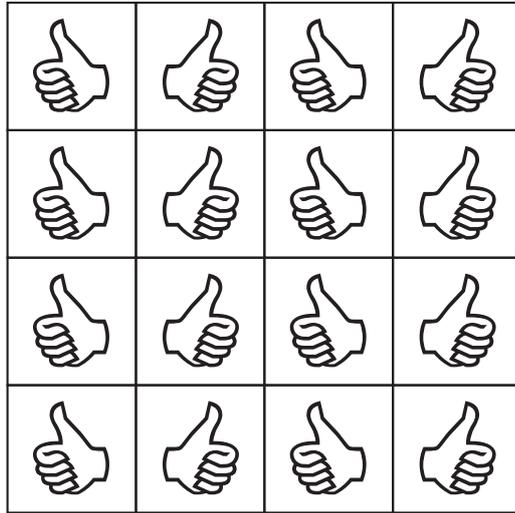
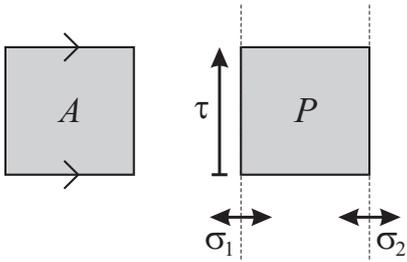
- La superficie $\mathbb{E}^2/G = \Sigma(n_1, \dots, n_k)(m_1, \dots, m_h)$ (talvolta con due raffigurazioni differenti);
- La piastrella base P della pavimentazione di sottogruppo G , insieme ad alcuni elementi di G che agiscono su P , e tramite i quali è possibile ottenere tutto G per composizione;
- Una porzione finita ma rappresentativa della pavimentazione di sottogruppo G .

Indicheremo con T il toro, con A l'anello, con K la bottiglia di Klein, con M la striscia di Möbius, con S la sfera, con D il disco e con \mathbb{P} il piano proiettivo.

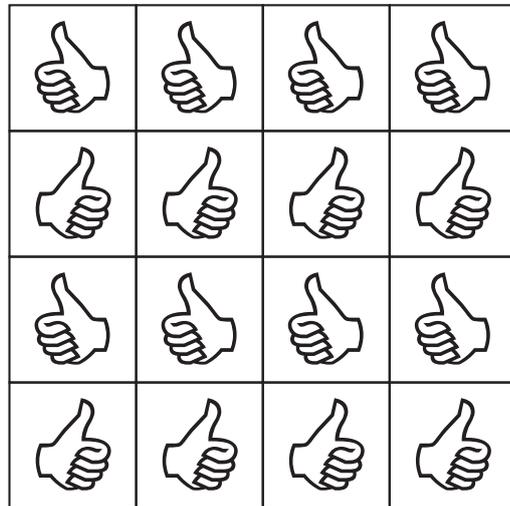
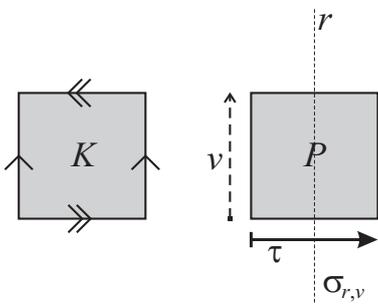
1. $\mathbb{E}^2/G = T$



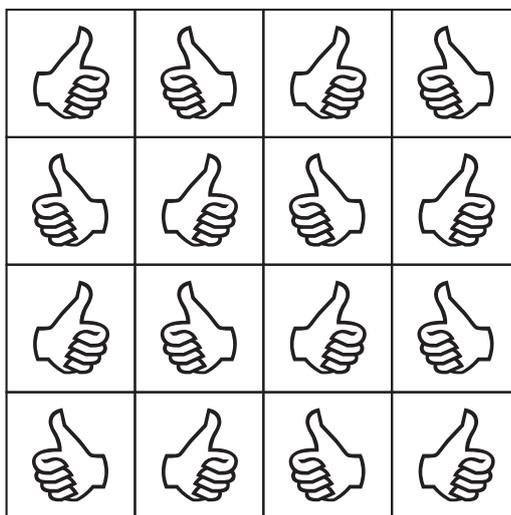
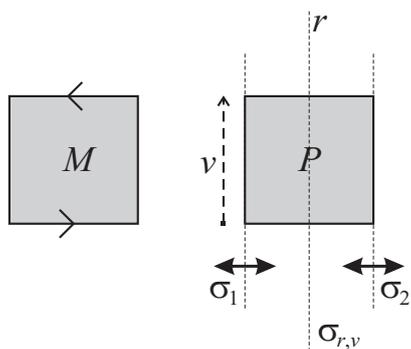
2. $\mathbb{E}^2/G = A$



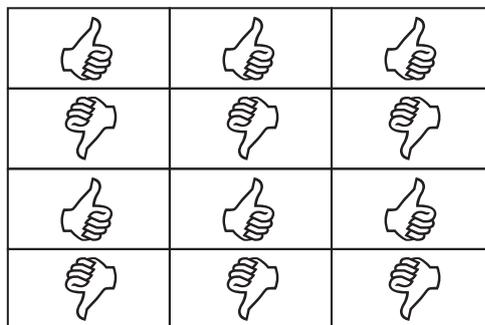
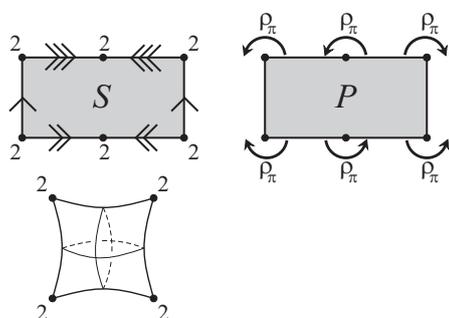
3. $\mathbb{E}^2/G = K$



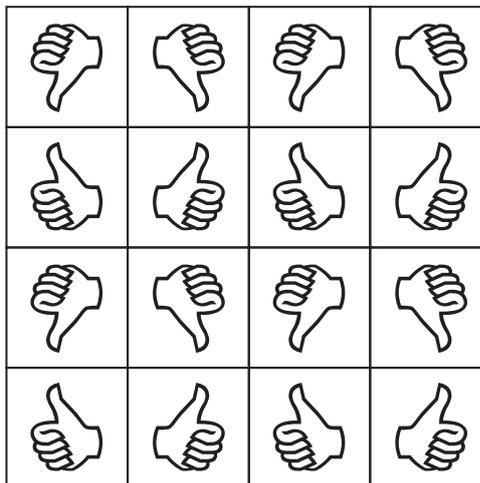
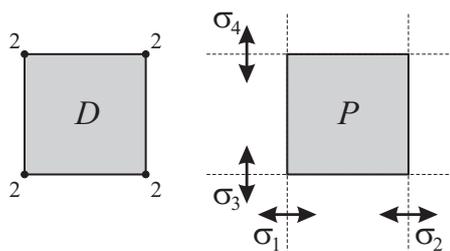
4. $\mathbb{E}^2/G = M$



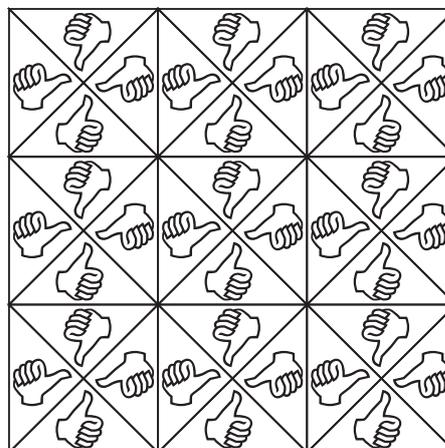
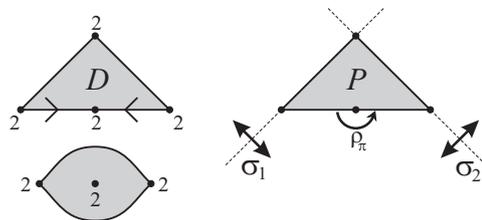
5. $\mathbb{E}^2/G = S(2, 2, 2, 2)(\)$



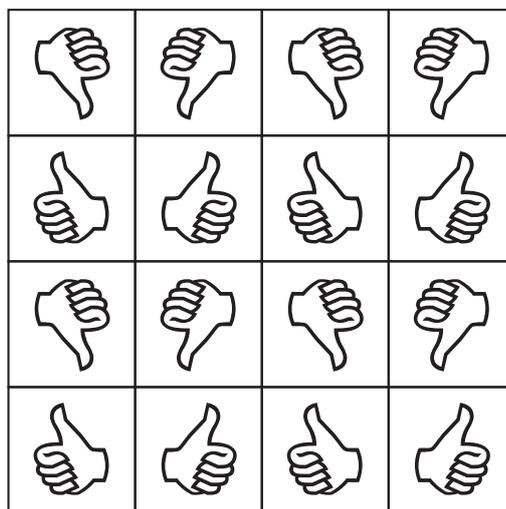
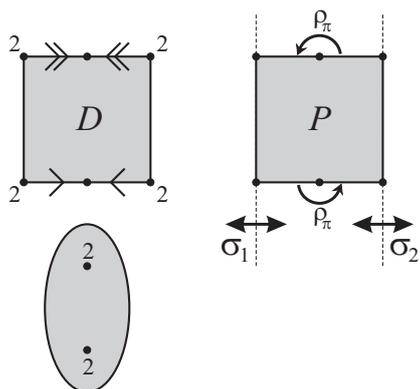
6. $\mathbb{E}^2/G = D(2, 2, 2, 2)$



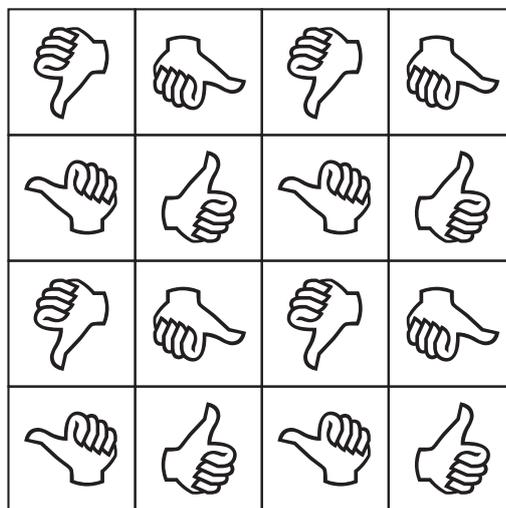
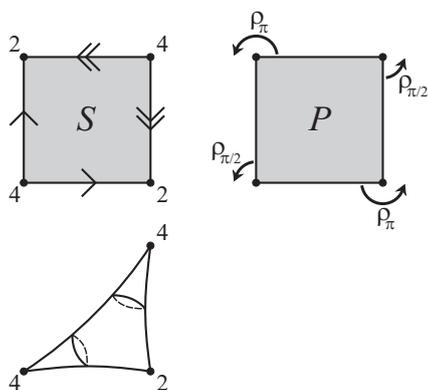
7. $\mathbb{E}^2/G = D(2)(2, 2)$



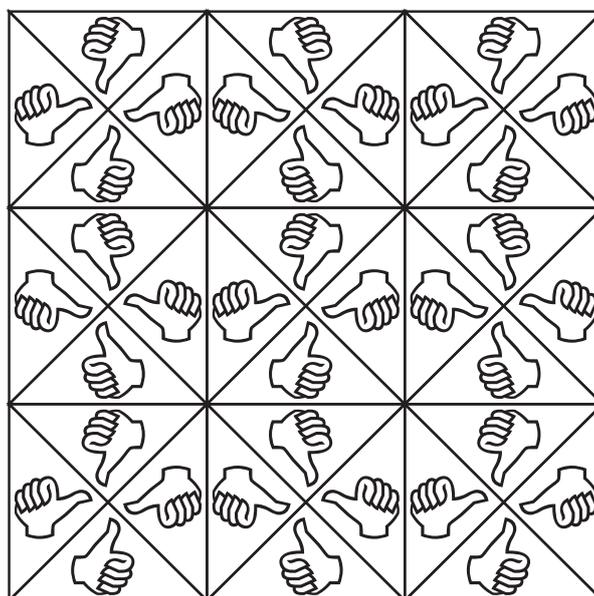
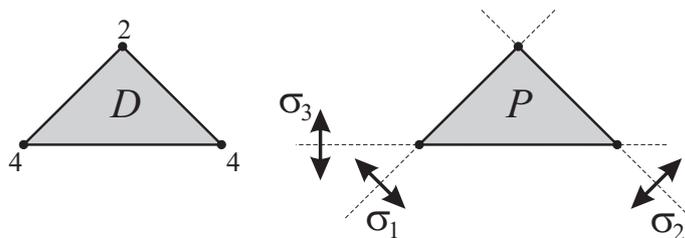
8. $\mathbb{E}^2/G = D(2, 2)(\)$



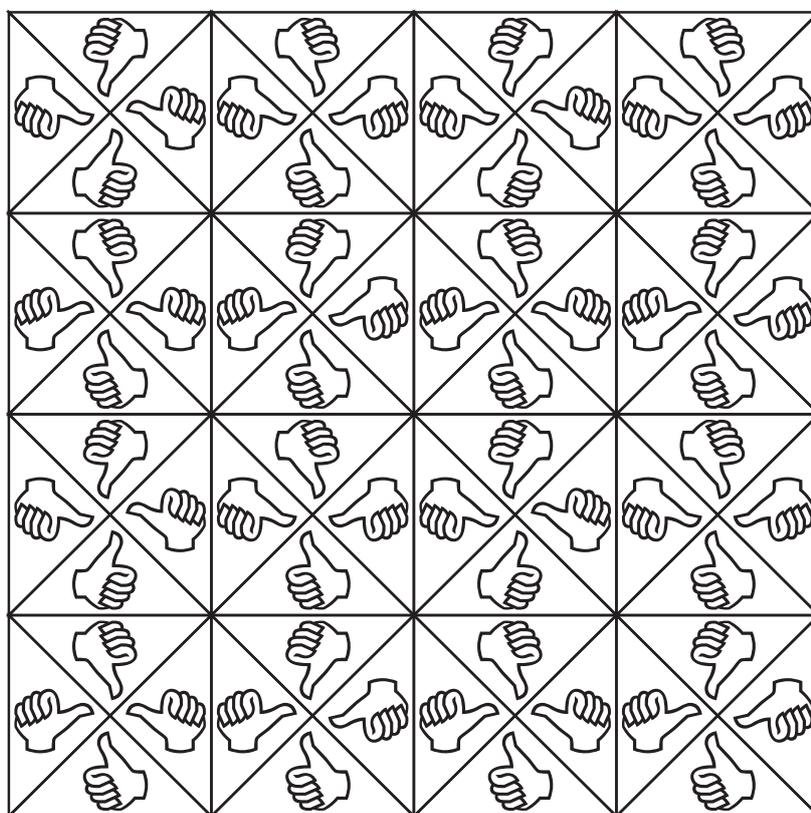
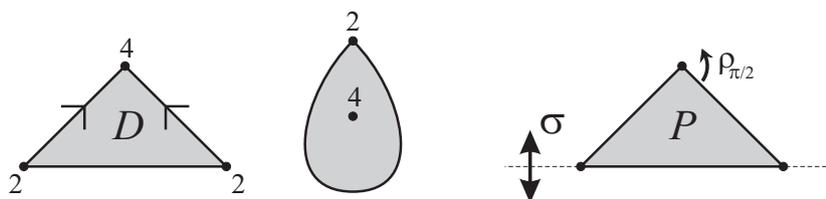
9. $\mathbb{E}^2/G = S(2, 4, 4)(2)$



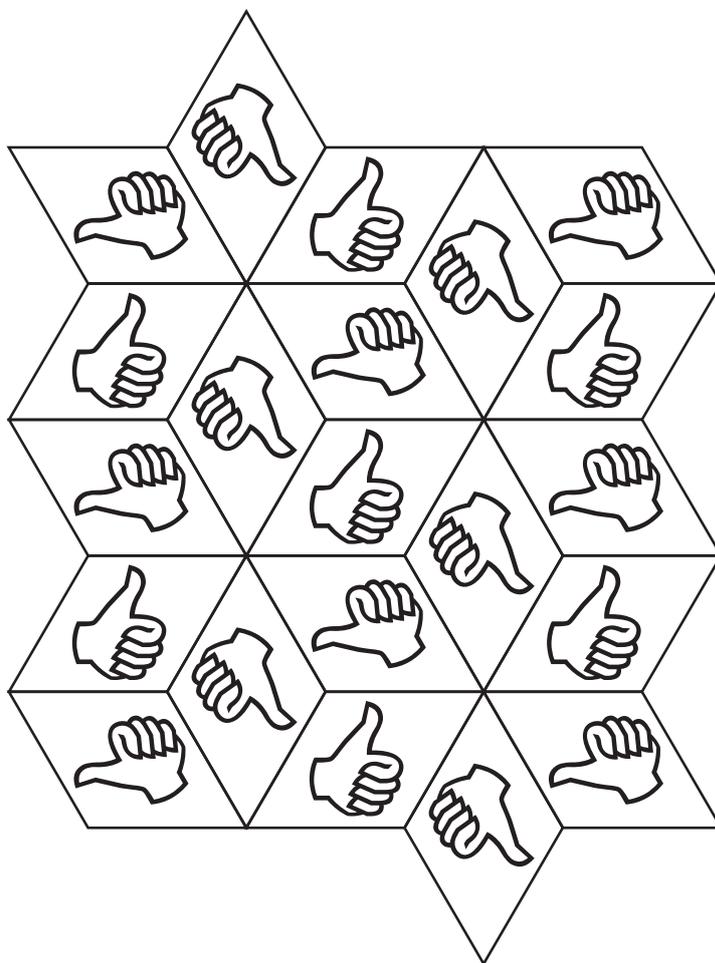
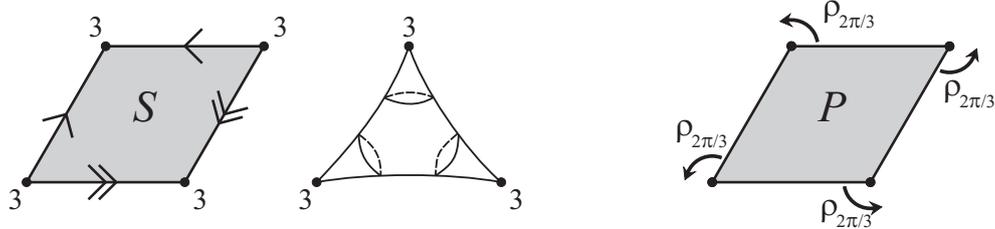
10. $\mathbb{E}^2/G = D(\) (2, 4, 4)$



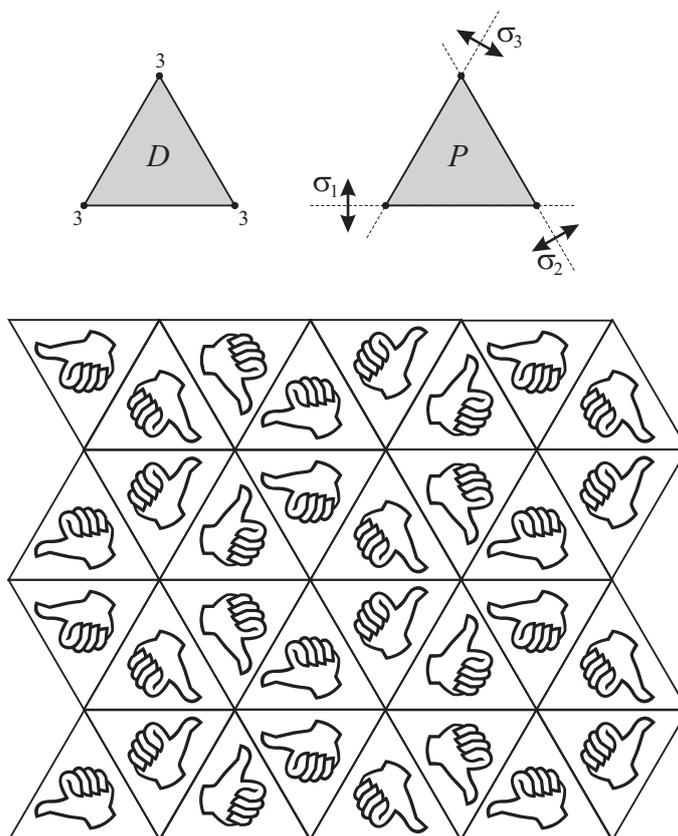
11. $\mathbb{E}^2/G = D(4)(2)$



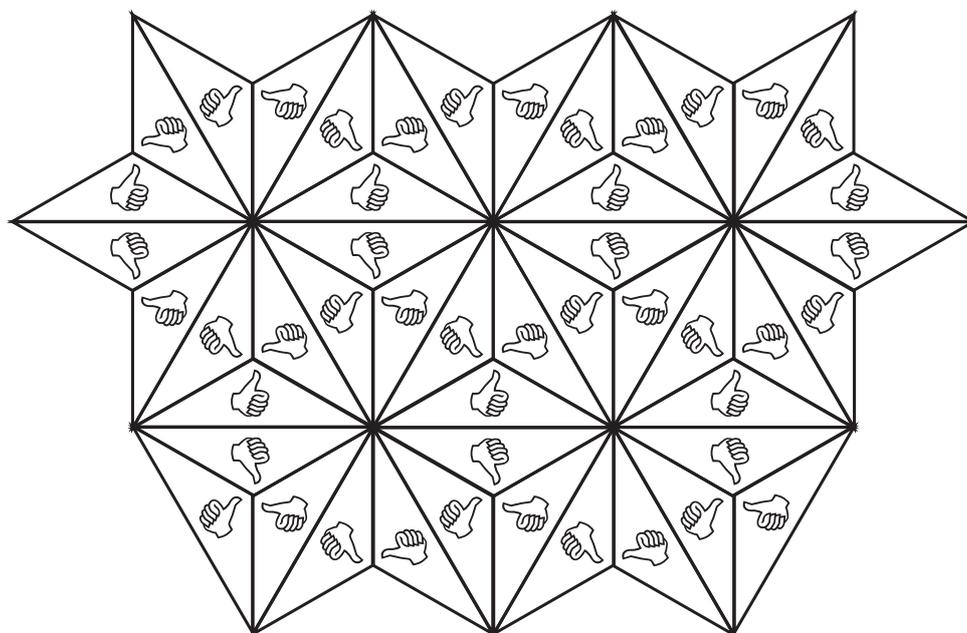
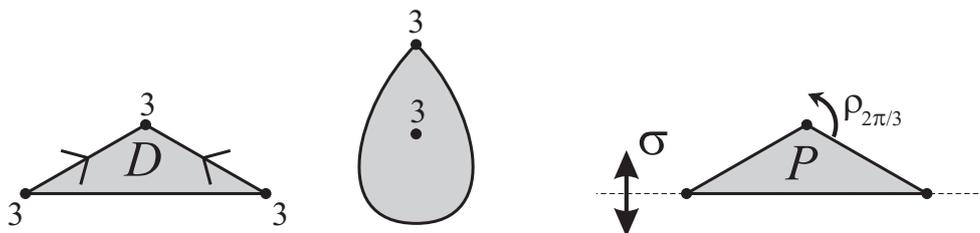
12. $\mathbb{E}^2/G = S(3, 3, 3)$



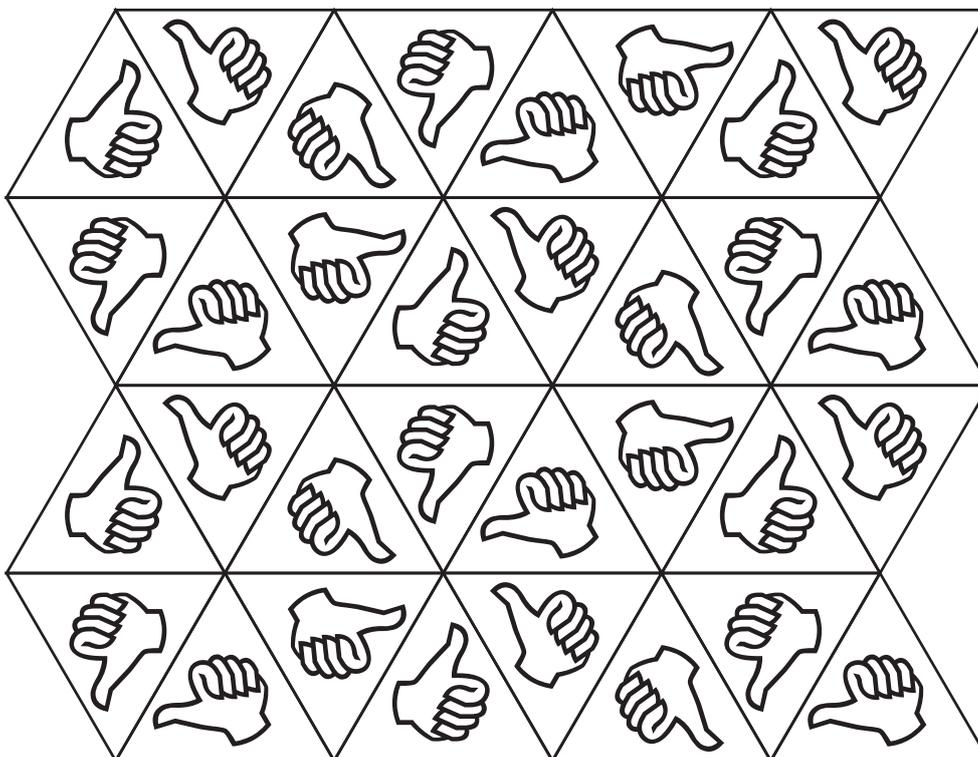
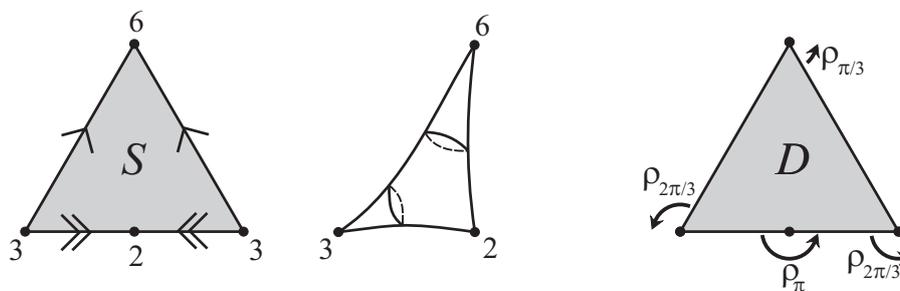
13. $\mathbb{E}^2/G = D(\cdot)(3, 3, 3)$



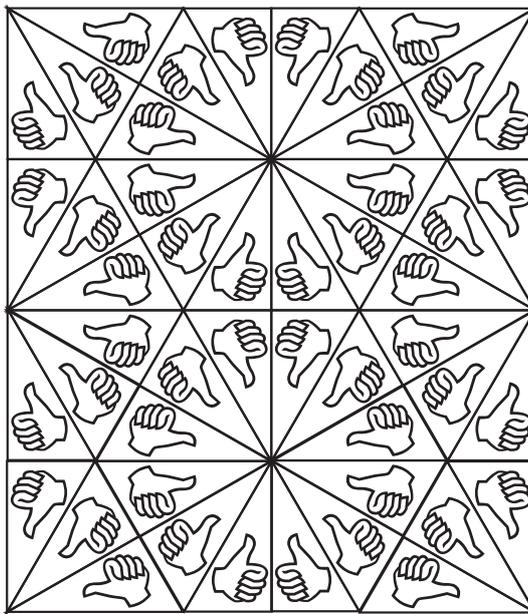
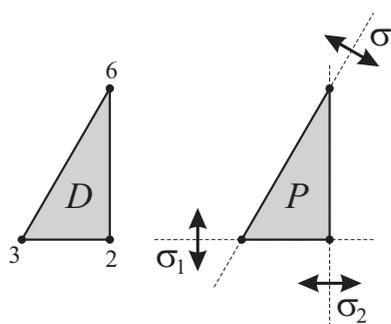
14. $\mathbb{E}^2/G = D(3)(3)$



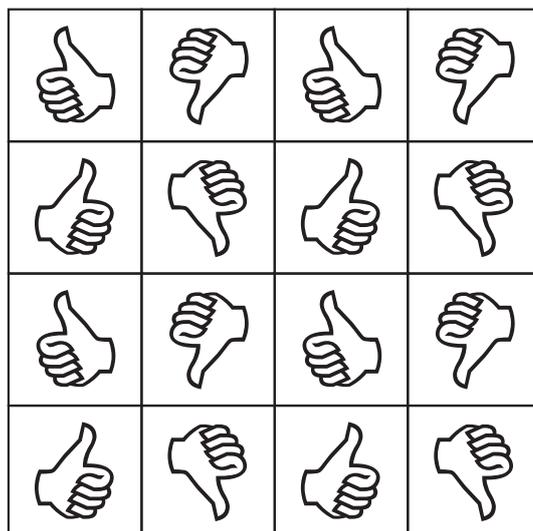
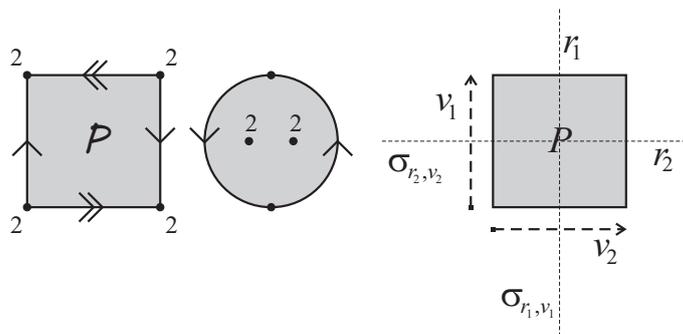
15. $\mathbb{E}^2/G = S(2,3,6)(\)$



16. $\mathbb{E}^2/G = D(\cdot)(2, 3, 6)$



17. $\mathbb{E}^2/G = \mathbb{P}(2, 2)(\)$



Indice analitico

- affine, cambiamento di coordinate —, 158
 affine, sottospazio —, 81
 affine, trasformazione —, 139
 aggiunta (di una matrice), 120
 A^{ij} , 75
 algebra, teorema fondamentale della —, 98
 algebrica, molteplicità — di un autovalore, 127
 angolo tra due rette, 103
 anti-lineare, applicazione —, 120
 antisimmetrica, matrice —, 53
 antisimmetrica, relazione —, 148
 aperto, sottoinsieme — di \mathbb{R}^n , 112, 177
 applicazione lineare, 47
 arco, parametro di —, 176
 area (con segno) del parallelogrammo, 71–72
 argomento (di un numero complesso), 96
 arrivo, base di — per la matrice di un'applicazione, 61
 associata, applicazione lineare — a una matrice, 54
 associata, matrice — a un'applicazione rispetto a basi date, 61
 associata, matrice — a un'equazione, 160
 associato, sottospazio vettoriale — a un sottospazio affine, 81
 assurdo, dimostrazione per, 21
 A^* , 120
 tA , 48
 autoaggiunta, applicazione —, 118, 121
 autovalore, 126
 autovettore, 126
 base, 35
 base canonica di $\mathcal{M}_{n \times m}(\mathbb{R})$, 37
 base canonica di \mathbb{R}^n , 36
 base canonica di $\mathbb{R}_{\leq d}[x]$, 37
 base ortogonale, 109
 base ortonormale, 110
 Bessel, disuguaglianza di —, 115, 121
 bilineare, applicazione —, 105
 Binet, teorema di —, 76
 binormale, vettore — a una curva nello spazio, 185
 biregolare, curva —, 184
 bordo di una curva, 197
 bordo orientato di un sottoinsieme di \mathbb{R}^2 , 197
 \mathbb{C} , 91–97
 cambiamento della matrice, 66
 cambiamento delle coordinate, 65
 cambiamento di base, 65–66
 cambiamento di coordinate affine, 158
 cambiamento di parametro, 173
 campo, 18
 canonica, base — di $\mathcal{M}_{n \times m}(\mathbb{R})$, 37
 canonica, base — di \mathbb{R}^n , 36

- canonica, base — di $\mathbb{R}_{\leq d}[x]$, 37
 canonico, prodotto scalare — di \mathbb{R}^n , 105
 canonico, prodotto scalare hermitiano — di \mathbb{C}^n , 120
 caratteristico, polinomio —, 126
 cartesiana, presentazione (o equazione) — di un piano nello spazio, 86–87, 116–118
 cartesiana, presentazione (o equazione) — di un sottospazio affine, 82
 cartesiana, presentazione (o equazione) — di una retta nel piano, 84, 116–118
 cartesiana, presentazione (o equazione) — di una retta nello spazio, 84–86, 116–118
 Cauchy-Schwarz, disuguaglianza di —, 107, 121
 chiusa, curva —, 171
 chiusa, forma —, 193
 ciclico, gruppo — di isometrie di $\text{Isom}(\mathbb{E}^2)$ di Kronecker, 56
 centrato in un punto, 219
 circonferenza, 139
 circonferenza osculatrice, 184
 $\mathbb{C}[z]$, 97
 classe di equivalenza, 148
 \mathbb{C} -lineare, applicazione —, 100
 \mathbb{C}^n , 100
 coefficiente di una matrice, 26, 100
 combinazione lineare, 32–33
 complesso, numero —, 91–97
 complesso, spazio vettoriale —, 99
 completamento (ad una base), 41–43
 completamento proiettivo, 160
 componente di un vettore numerico, 25
 conica degenerare, 145
 conica non degenerare, 142
 coniugata, matrice —, 125
 coniugato (di un numero complesso), 93
 connesso, sottoinsieme — di \mathbb{R}^n , 177
 conservativo, campo di forze —, 193
 coordinate di un vettore numerico, 100
 coordinate rispetto a una base, 36
 coordinate, cambiamento delle —, 65
 Cramer, teorema di —, 78
 curva chiusa, 171
 curva parametrizzata, 171
 curva regolare, 171
 curva semplice, 171
 curvatura di una curva nello spazio, 184
 curvatura di una curva piana, 180–182
 definita positiva, applicazione —, 105, 120
 degenerare, conica —, 145
 degenerare, conica non —, 142
 derivata parziale, 112
 derivata parziale seconda, 113
 derivazione di un polinomio, 48
 $\det(A)$, 71–77
 $\det_2(A)$, 71
 $\det_n(A)$, 73–76
 determinante di un'applicazione lineare, 127
 determinante di una matrice, 71–77
 $\det(f)$, 127
 $\det_3(A)$, 72
 diagonale principale, 48
 diagonalizzabile, matrice o applicazione lineare —, 125
 diedrale, gruppo — di isometrie di $\text{Isom}(\mathbb{E}^2)$ centrato in un punto, 219

- differenziale di una 1-forma, 195
 differenziale di una funzione, 191
 $\dim_{\mathbb{C}}(V)$, 100
 dimensione (di un sottospazio affine), 81
 dimensione complessa, 100
 dimensione di un sottospazio proiettivo, 157
 dimensione finita, 40
 dimensione reale, 39
 $\dim_{\mathbb{R}}(E)$, 81
 $\dim_{\mathbb{R}}(V)$, 39
 dipendenza lineare, 35
 diretta, somma —, 51
 discreto, sottogruppo — di $\text{Isom}(\mathbb{E}^2)$, 218
 distanza, 108
 distanza tra punti del piano, 103
 disuguaglianza di Bessel, 115, 121
 disuguaglianza di Cauchy-Schwarz, 107, 121
 disuguaglianza triangolare, 95, 108, 121
 divisione tra polinomi, 99
- \mathcal{E}_n , 36
 $\mathcal{E}_{n \times m}$, 37
 $e_i^{(n)}$, 36
 e_i , 36
 E_{ij} , 37
 ellisse, 139
 ellissoide, 164
 ellissoide proiettivo, 167
 equazioni cartesiane di un piano nello spazio, 86–87, 116–118
 equazioni cartesiane di un sottospazio affine, 82
 equazioni cartesiane di una retta nel piano, 84, 116–118
- equazioni cartesiane di una retta nello spazio, 84–86, 116–118
 equazioni parametriche di un piano nello spazio, 86–87, 116–118
 equazioni parametriche di un sottospazio affine, 82
 equazioni parametriche di una retta nel piano, 84, 116–118
 equazioni parametriche di una retta nello spazio, 84–86, 116–118
 equivalenti, sistemi lineari —, 70
 equivalenza di gruppi di pavimentazioni, 225
 equivalenza, classe di —, 148
 equivalenza, relazione di —, 148
 esatta, 1-forma —, 191
 esponenziale complessa, 97
 estrazione (di una base), 41–43
 euclideo, piano —, 201
- f_A , 54
 f^{-1} , 64
 $[f]_{\mathcal{B}}^{\mathcal{C}}$, 61
 Fibonacci, numeri di —, 11–13
 finita, dimensione —, 40
 forma (1-forma), 189
 formula di Grassmann, 44–45
 Frénet, riferimento di —, 185
 Frénet-Serret, formule di —, 185
 $\mathcal{F}(S, \mathbb{C})$, 100
 $\mathcal{F}(S, \mathbb{R})$, 26
 $\mathcal{F}(S, V)$, 27
- Gauss-Green, formula di —, 196–198
 generato, sottospazio —, 32
 generatori di un sottospazio, 32
 geometrica, molteplicità — di un autovalore, 127
 gradiente, 113, 132
 grado (di un polinomio), 19

- Gram-Schmidt, procedimento di ortogonalizzazione di —, 110–111, 121
- Grassmann, formula di —, 44–45
- Green, formula di Gauss—, 196–198
- gruppo, 202
- gruppo commutativo, 18
- hermitiana, applicazione —, 120
- hermitiana, matrice —, 122
- hermitiano, prodotto scalare —, 120
- hermitiano, prodotto scalare — canonico di \mathbb{C}^n , 120
- hessiana, matrice —, 132
- $\mathcal{L}_{\mathbb{R}}(V, W)$, 57
- i , 91
- identità, matrice —, 56
- Id_V , 52
- I_n , 56
- \Im , 93
- $\text{Im}(f)$, 49
- immaginaria, parte — di un numero complesso, 93
- immaginaria, unità —, 91
- immaginario puro, numero complesso —, 93
- immagine (di un'applicazione lineare), 49
- incognite (di una sistema), 67
- indice (di riga o colonna in una matrice), 26
- indipendenza lineare, 34
- induzione, principio di, 20
- infinito, punti all'— di una conica o quadrica, 160
- infinito, punto all'—, 154
- integrale di una 1-forma su una curva, 190
- integrale di una funzione scalare su una curva, 176
- intero, numero —, 15–16
- inversa (di un'applicazione), 64
- inverso di un numero complesso, 93, 94
- invertibile, applicazione —, 64
- iperbole, 140
- iperboloide ellittico o a due falde, 164
- iperboloide iperbolico o a una falda, 164
- iperboloide proiettivo, 167
- isometria, 118, 201
- isometria del piano euclideo, 137
- j , 91
- Jordan, teorema di —, 177
- κ , 182
- $\text{Ker}(f)$, 49
- Kronecker, δ di —, 56
- Laplace, formule di sviluppo di —, 75
- lineare, applicazione —, 47
- lineare, combinazione —, 32–33
- lineare, dipendenza —, 35
- lineare, indipendenza —, 34
- lineare, sistema —, 67
- lunghezza di una curva, 174
- Möbius, nastro di —, 151
- m.a., 127
- matrice (di un sistema), 67
- matrice antisimmetrica, 53
- matrice associata a un'applicazione rispetto a basi date, 61
- matrice complessa, 100
- matrice identità, 56
- matrice reale, 25–26
- matrice simmetrica, 53

- matrice, cambiamento della —, 66
 m.g., 127
 $\mathcal{M}_{n \times n}(\mathbb{C})$, 100
 $\mathcal{M}_{n \times m}(\mathbb{R})$, 25–26
 modulo (di un numero complesso), 94
 molteplicità algebrica di un autovalore, 127
 molteplicità della radice di un polinomio, 99
 molteplicità geometrica di un autovalore, 127
 monolatera, superficie —, 151
- \mathbb{N} , 15
 naturale, numero —, 15
 norma, 107, 121
 norma di un vettore in \mathbb{R}^n , 104
 normale, matrice —, 133
 normale, vettore — a una curva nello spazio, 185
 normalizzazione (di un vettore), 110
 noti, termini — (di un sistema lineare), 69
 nucleo (di un'applicazione lineare), 49
- omogenee, coordinate —, 156
 omogeneo, sistema lineare —, 69
 ordine, relazione di —, 148
 orientato, bordo — di un sottoinsieme di \mathbb{R}^2 , 197
 orientazione di una curva, 173
 orlata, 78
 orlati, teorema degli —, 78–80
 ortogonale, applicazione —, 118
 ortogonale, base —, 109
 ortogonale, proiezione —, 112–116, 121
 ortogonale, sistema — di vettori, 109
 ortogonali, rette — nel piano, 104
 ortogonali, vettori —, 109
 ortonormale, base —, 110
 ortonormale, sistema — di vettori, 109
 ortonormalizzazione, procedimento di — di Gram-Schmidt, 110–111, 121
 osculatore, piano —, 184
 osculatrice, circonferenza —, 180–182, 184
- \mathbb{P} , 149
 parabola, 140
 paraboloido ellittico, 164
 paraboloido iperbolico o a sella, 165
 paralleli, sottospazi affini —, 88
 parallelogrammo, area (con segno) del —, 71
 parallelogrammo, area del —, 72
 parametrica, presentazione (o equazione) — di un piano nello spazio, 86–87, 116–118
 parametrica, presentazione (o equazione) — di un sottospazio affine, 82
 parametrica, presentazione (o equazione) — di una retta nel piano, 84, 116–118
 parametrica, presentazione (o equazione) — di una retta nello spazio, 84–86, 116–118
 parametrizzata, curva —, 171
 parametro d'arco, 176
 parametro, cambiamento di —, 173
 parte immaginaria di un numero complesso, 93
 parte reale di un numero complesso, 93
 partenza, base di — per la matrice di un'applicazione, 61
 parziale seconda, derivata —, 113
 parziale, derivata —, 112

- pavimentazione (regolare), 215–216
 permutazione, 74–75
 perpendicolari, rette — nel piano, 104
 p_f , 126
 piano euclideo, 201
 piano nello spazio, 86–87, 116–118
 piano osculatore, 184
 piastrella base (di una pavimentazione), 216
 $\mathbb{P}^n(\mathbb{R})$, 149
 polare, forma — di un numero complesso, 96–97
 polinomio, 18–19
 polinomio caratteristico, 126
 polinomio complesso, 97
 polinomio, derivazione di un —, 48
 positiva, applicazione — definita, 105
 positiva, applicazione definita —, 120
 posto di una matrice, 25
 potenziale di una 1-forma, 191
 principale, diagonale —, 48
 procedimento di ortonormalizzazione di Gram-Schmidt, 110–111, 121
 prodotto di numeri complessi, 92
 prodotto per scalare, 23
 prodotto righe per colonne, 54–55
 prodotto scalare canonico di \mathbb{R}^n , 105
 prodotto scalare hermitiano, 120
 prodotto scalare hermitiano canonico di \mathbb{C}^n , 120
 prodotto scalare in \mathbb{R}^n , 104
 prodotto scalare su V , 105
 prodotto vettoriale, 117–118
 proiettivo, completamento —, 160
 proiettivo, sottospazio —, 156
 proiettivo, spazio —, 149
 proiezione nel quoziente, 148
 proiezione ortogonale, 112–116, 121
 punto conico, 226
 puro, numero complesso immaginario —, 93
 p_W , 112–116, 121
 \mathbb{Q} , 16–17
 quadrato, sistema lineare —, 69
 quadrica non degenera, 164
 quoziente, insieme —, 148
 \mathbb{R} , 17–18
 radici dell'unità, 98
 rango di un'applicazione o matrice, 68
 $\text{rank}(A)$, $\text{rank}(f)$, 68
 rappresentanti, insieme di — per una relazione di equivalenza, 148
 razionale, numero —, 16–17
 \mathfrak{R} , 93
 reale, numero —, 17–18
 reale, parte — di un numero complesso, 93
 regola di Sarrus, 72
 regolare, curva —, 171
 relazione su un insieme, 147
 resto della divisione tra polinomi, 99
 retta nel piano, 84, 116–118
 retta nello spazio, 84–86, 116–118
 retta tangente a una curva, 172
 riferimento di Frénet, 185
 riflessiva, relazione —, 147
 riflettore angolare, 226
 rigata, superficie —, 165
 \mathbb{R} -lineare, applicazione —, 100
 \mathbb{R}^n , 25
 Rouché-Capelli, teorema di —, 68–69
 Ruffini, regola di —, 99
 $\mathbb{R}[x]$, 18–19, 27
 $\mathbb{R}_{\leq d}[x]$, 29
 Sarrus, regola di —, 72
 scalare, 23

- scalare, prodotto — canonico di \mathbb{R}^n , 105
 scalare, prodotto — hermitiano, 120
 scalare, prodotto — hermitiano canonico di \mathbb{C}^n , 120
 scalare, prodotto — in \mathbb{R}^n , 104
 segnatura (di una permutazione), 74–75
 segno del volume di un solido, 72
 segno dell'area del parallelogrammo, 72
 segno di un numero reale, 48
 seminassi di un ellissoide, 164
 semplice, curva —, 171
 semplicemente connesso, sottoinsieme — di \mathbb{R}^2 , 198
 sesquilineare, applicazione —, 120
 sforzi, tensore degli —, 123
 sghembe, rette —, 88
 sgn, 48
 $\text{sgn}(\sigma)$, 74–75
 simile, matrice —, 125
 simmetrica, applicazione —, 105
 simmetrica, matrice —, 53
 simmetrica, relazione —, 147
 sistema lineare, 67
 sistema lineare omogeneo, 69
 sistema lineare quadrato, 69
 sistema lineare sottodeterminato, 69
 sistema lineare sovradeterminato, 69
 sistemi lineari equivalenti, 70
 somma di numeri complessi, 92
 somma di sottospazi, 43
 somma di sottospazi affini, 87–88
 somma di vettori, 23
 somma diretta, 51
 sottodeterminato, sistema lineare —, 69
 sottogruppo (del gruppo delle isometrie del piano), 216
 sottomatrice, 78
 sottospazio affine, 81
 sottospazio vettoriale, 27–28
 sovradeterminato, sistema lineare —, 69
 Span, 32
 spazio vettoriale complesso, 99
 spazio vettoriale, reale, 23–24
 spettrale, teorema —, 129–130
 stereografica, proiezione —, 155
 supporto di una curva, 173
 sviluppo di Laplace, formule di —, 75
 tangente, retta — a una curva, 172
 tangente, vettore — a una curva nello spazio, 184
 tensore degli sforzi, 123
 teorema degli orlati, , 78–80
 teorema di Binet, 76
 teorema di Cramer, 78
 teorema di Rouché-Capelli, 68–69
 teorema fondamentale dell'algebra, 98
 teorema spettrale, 129–130
 termini noti (di un sistema lineare), 69
 $\text{tr}(A)$, 48
 traccia di un'applicazione lineare, 127
 traccia di una matrice, 48
 transitiva, relazione —, 148
 trasposizione, 74
 trasposizione, trasposta, 48
 triangolare, disuguaglianza —, 95, 108, 121
 1-forma, 189
 unità immaginaria, 91
 unità, radici della —, 98
 unitaria, matrice —, 122
 unitario, vettore —, 109
 vettore, 23

vettore colonna, 25, 100
vettoriale, prodotto —, 117–118
vettoriale, sottospazio —, 27–28
vettoriale, spazio — reale, 23–24
volume (con segno) di un solido, 72
 $[v]_{\mathcal{B}}$, 36
 w^* , 120
 \mathbb{Z} , 15–16
 \bar{z} , 93
 \oplus , 51
 \perp , 104
 $\langle \cdot | \cdot \rangle_A$, 106, 121
 $\langle \cdot | \cdot \rangle_{\mathbb{R}^n}$, 104
 $\langle \cdot | \cdot \rangle_{\mathbb{C}^n}$, 120
 \wedge , 117–118